

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO RIO GRANDE
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM COMPUTACIONAL

**MODELO DE RECONSTRUÇÃO DE IMAGENS AFETADAS
POR VARIAÇÃO NA EXPOSIÇÃO BASEADO EM REDES
NEURAIS CONVOLUCIONAIS**

por

Cristiano Rafael Steffens

Tese para obtenção do Título de
Doutor em Modelagem Computacional

Rio Grande, julho – 2021

Cristiano Rafael Steffens

“MODELO DE RECONSTRUÇÃO DE IMAGENS AFETADAS POR VARIAÇÃO NA EXPOSIÇÃO BASEADO EM REDES NEURAIS CONVOLUCIONAIS”

Tese apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal do Rio Grande - FURG, como requisito parcial para obtenção do Grau de Doutor. Área concentração: Modelagem Computacional.

Aprovado em: 28 de julho de 2021.

BANCA EXAMINADORA

DocuSigned by:

Silvia Silva da Costa Botelho

536CB7A495C6434...

Prof.ª. Dra. Silvia Silva da Costa Botelho (Orientadora – FURG)

DocuSigned by:

Paulo Lilles Jorge Drews Junior

D211E40EAA4243D...

Prof. Dr. Paulo Lilles Jorge Drews Junior (Coorientador – FURG)

DocuSigned by:

Viviane Leite Dias de Mattos

5C845DBCE355428...

Prof.ª. Dra. Viviane Leite Dias de Mattos (FURG)

DocuSigned by:

Manuel Menezes de Oliveira Neto

B210D9982C74456...

Prof. Dr. Manuel Menezes de Oliveira Neto (UFRGS)

DocuSigned by:

Moacir Antonelli Ponti

28CF7F28C6014A5...

Prof. Dr. Moacir Antonelli Ponti (USP)

Ficha Catalográfica

S817m Steffens, Cristiano Rafael.

Modelo de reconstrução de imagens afetadas por variação na exposição baseado em redes neurais convolucionais / Cristiano Rafael Steffens. – 2021.

226 f.

Tese (doutorado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-Graduação em Modelagem Computacional, Rio Grande/RS, 2021.

Orientadora: Dra. Silvia Silva da Costa Botelho.

Coorientador: Dr. Paulo Lilles Jorge Drews Junior.

1. Restauração de Imagens Digitais 2. Restauração de imagens
3. Redes Neurais Convolucionais 4. Saturação 5. Fotografia
Computacional I. Botelho, Silvia Silva da Costa II. Drews Junior, Paulo
Lilles Jorge III. Título.

CDU 004.932

Catálogo na Fonte: Bibliotecário José Paulo dos Santos CRB 10/2344

“Toda a variedade, todo o charme e toda a beleza da vida é feita de luz e sombra.”
– Leo Tolstoy

AGRADECIMENTOS

À família.

À FURG pela infraestrutura e cuidado com as pessoas.

À Profa. Dra. Silvia Silva da Costa Botelho e ao Prof. Dr. Paulo Drews-Jr pela orientação.

À coordenação do Programa de Pós-graduação em Modelagem Computacional e professores.

Aos membros da banca pelas valorosas contribuições.

Às agências de fomento CAPES, CNPq e FAPERGS.

Aos colegas do Grupo de Automação e Robótica Inteligente - NAUTEC/FURG.

Aos professores e amigos do Centro de Ciências Computacionais - C3/FURG.

Aos colegas da Migrate Company pelo altruísmo e incentivo.

À SPROJECT Sistemas, Parque Tecnológico Oceantec e APL Marítimo.

Ao CREA-SC e aos amigos que ali fiz.

À Indra — Minsait Florianópolis pelo apoio.

Aos mais íntimos, que encorajaram, suportaram e comemoraram comigo as pequenas vitórias.

Ao fiel escudeiro Eng. Lucas Ricardo Vieira Messias por comprar as ideias boas e as absurdas, pela contribuição técnica, pela diligência, pela proatividade e sobretudo pela amizade.

Obrigado!

RESUMO

O trabalho apresenta um modelo de rede neural artificial para restauração de imagens danificadas por exposição inadequada, contemplando condições de subexposição e sobre-exposição. O problema abordado tem relevância em aplicações de visão computacional que envolvem obtenção de imagens em cenas onde a limitação do sensor ou arranjo óptico impedem que os detalhes da cena sejam adequadamente representados na imagem capturada. Em função da variabilidade de equipamentos e técnicas de fotografia disponíveis, da quantidade de circunstâncias não controladas que impactam o processo de aquisição de imagens opta-se pela modelagem baseada em redes neurais profundas. Nesta abordagem uma arquitetura de rede combinado com um procedimento de ajuste é capaz de convergir em um modelo a partir de dados pareados compostos por uma imagem com exposição inapropriada para a cena e uma imagem com a exposição adequada. A proposição desta modelagem leva em consideração os avanços recentes propiciados por redes convolucionais em problemas como segmentação semântica, transformação imagem-imagem e classificação de objetos em imagens, resultando em um modelo compacto e que pode ser incorporado como uma etapa de pré-processamento em aplicações de visão computacional. Com relação aos procedimentos técnicos, pode-se caracterizar a metodologia científica da pesquisa proposta em bibliográfica, descritiva e experimental. No que tange a avaliação dos resultados da pesquisa utilizam-se medidas de qualidade de imagem para avaliar a qualidade dos resultados produzido e métricas objetivas para avaliar o impacto desta em aplicações de visão computacional. Utiliza-se também a análise qualitativa para discutir os aspectos conceituais e explicitar o funcionamento do modelo, evidenciando o impacto das escolhas técnicas realizadas. Os resultados obtidos, tanto em termos de aprimoramento visual quanto na aplicação do modelo em problemas típicos da computação visual indicam que o modelo de rede neural convolucional proposto é capaz de melhorar imagens danificadas pela heterogeneidade de exposição, oferecendo ganho sobre métodos estado-da-técnica, tanto em conjuntos de dados simulados quanto em dados reais.

Palavras-chaves: Restauração de Imagens Digitais, Restauração de imagens, Redes Neurais Convolucionais, Saturação, Fotografia Computacional.

ABSTRACT

This work presents an artificial neural network model for the restoration of images damaged by underexposure and overexposure. The problem is relevant in computer vision applications that are applied in conditions where the limitation of the sensor or optical arrangement prevents the scene details from being adequately represented in the captured image. Due to uncontrolled conditions that impact the process, the variability of available equipment and photography techniques, and the feasibility of using paired datasets, we chose to address this problem using a deep learning based approach. The modeling takes into account the recent advances provided by convolutional networks in problems such as semantic segmentation, image-image transformation and classification of objects in images. Regarding the technical procedures, the scientific methodology of the proposed research can be characterized as bibliographical, descriptive and experimental. Regarding the evaluation of research results, a broad range of image quality metrics is used to assess the quality of the results produced by the model. We employ objective metrics to assess the impact that the use of the proposed model can have in computer vision applications. Qualitative analysis is also used to discuss conceptual aspects and to understand how the model works. The results obtained, measured by means of different image quality indicators, as well as by applying the model to typical problems of visual computing, indicate that the proposed convolutional neural network model is able to improve images damaged by exposure heterogeneity, offering gains over state-of-the-art methods, both on simulated and real data sets.

Palavras-chaves: Clipping, Image restoration, Image enhancement, Neural networks, Computational photography.

ÍNDICE

1	INTRODUÇÃO	21
1.1	Justificativa	22
1.2	Hipótese Científica	23
1.3	Objetivos	23
1.4	Organização do Texto	24
2	REVISÃO TEÓRICA E TRABALHOS RELACIONADOS	27
2.1	Restauração e Aprimoramento de Imagens (Métodos Clássicos)	27
2.1.1	Equalização de Histograma	27
2.1.2	Técnicas Baseadas em Constância de Cor (Retinex)	29
2.1.3	Técnicas baseadas em Fusão de Imagens	30
2.2	Técnicas baseadas em Redes Neurais	32
2.2.1	Considerações Gerais sobre Arquiteturas de Rede	32
2.2.2	Modelos para Restauração de Imagens Inapropriadamente Expostas	34
2.3	Modelos Diversos de Tradução Imagem-Imagem	49
2.4	Consolidação dos Trabalhos Relacionados	54
3	REQUISITOS E DEFINIÇÕES DE PROJETO	56
3.1	Requisitos da Arquitetura de Rede	57
3.2	Indicadores de Qualidade de Imagem	58
3.2.1	Métricas Cegas	58
3.2.2	Métricas Baseadas em Referência	59
3.3	Datasets	62
3.3.1	A6300 <i>Multi-Exposure Dataset</i> (real)	62
3.3.2	Cai2018 <i>Multi-Exposure Dataset</i> (real)	63
3.3.3	Dataset Multi-Exposição baseado em FiveK (sintético)	64
3.3.4	Dataset Multi-Exposição baseado HDR+ <i>Burst Photography Dataset</i> (sintético)	65
3.4	Validação a Nível de Aplicação	66
3.4.1	Reconhecimento de Imagens	68
3.4.2	Segmentação Semântica de Imagens	70
3.4.3	Procedimento	70
3.5	Visualização da Ativação Interna do Modelo	71
4	MODELO DE CORREÇÃO DE EXPOSIÇÃO AJUSTADO EM MEDIDAS DE SIMILARIDADE	73
4.1	Arquitetura da Rede	74
4.2	Bloco de Convoluções Dilatadas	78
4.3	Função Objetivo	80
4.4	Treinamento do Modelo	81
5	RESULTADOS E DISCUSSÃO	83

5.1	Considerações Gerais Sobre Indicadores de Qualidade de Imagem Aplicados . .	83
5.2	Resultados em Métricas de Qualidade de Imagem	89
5.2.1	Dataset Multi-Exposição baseado em FiveK (sintético)	90
5.2.2	HDR+ <i>Burst Photography Dataset</i> (sintético)	98
5.2.3	A6300 <i>Multi-Exposure Dataset</i> (real)	106
5.2.4	Cai2018 <i>Multi-Exposure Dataset</i> (real)	113
5.3	Outros Comparativos	120
5.4	Validação a Nível de Aplicação	123
5.4.1	Reconhecimento de Imagens	124
5.4.2	Segmentação Semântica de Imagens	125
5.4.3	Análise de Desempenho	129
5.5	Teste de Ablação	130
5.6	Visualização da Ativação Interna do Modelo	132
5.6.1	Aplicação em Imagem Sub-Exposta	132
5.6.2	Aplicação em Imagem Sobre-Exposta	136
5.6.3	Considerações sobre a Visualização	140
5.7	Limitações conhecidas	141
6	CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS	143
7	Apêndice	165
7.1	Teste de Correlação entre Indicadores de Qualidade de Imagem	166
7.2	Resultados para Dataset Multi-Exposição baseado em FiveK (sintético)	167
7.2.1	Testes de Normalidade	167
7.2.2	Média	174
7.2.3	Teste dos Postos Sinalizados de Wilcoxon	176
7.3	Resultados para HDR+ <i>Burst Photography Dataset</i> (sintético)	178
7.3.1	Testes de Normalidade	178
7.3.2	Média	185
7.3.3	Teste dos Postos Sinalizados de Wilcoxon	187
7.4	Resultados para A6300 <i>Multi-Exposure Dataset</i> (real)	189
7.4.1	Testes de Normalidade	189
7.4.2	Média	196
7.4.3	Teste dos Postos Sinalizados de Wilcoxon	198
7.5	Resultados para Cai2018 <i>Multi-Exposure Dataset</i> (real)	200
7.5.1	Testes de Normalidade	200
7.5.2	Média	207
7.5.3	Teste dos Postos Sinalizados de Wilcoxon	209
7.6	Avaliação do Impacto da Subexposição e Sobre-exposição em Aplicações de Reconhecimento de Imagens	211
7.7	Dataset Fivek	212
7.8	Dataset HDR+Burst	216

7.9 Dataset A6300 220

7.10 Dataset Cai2018 224

LISTA DE FIGURAS

2.1	Aplicação de diferentes métodos de Equalização de Histograma	29
2.2	Transformação de imagem baseado em Retinex Multiescala	30
2.3	Resultados obtidos com o modelo VP de Li et al. (2020b)	30
2.4	<i>Framework</i> de fusão de imagens de Ying et al. (2017b)	31
2.5	Arquitetura Convolutacional <i>Encoder-Decoder</i>	33
2.6	Visão geral da implementação da rede DeclipNet	34
2.7	Arquitetura da rede DeclipNet	35
2.8	Resultados obtidos com o modelo DeclipNet	35
2.9	Visão geral da implementação da rede SICE	36
2.10	Resultados obtidos com o modelo SICE	36
2.11	Visão geral da implementação da rede AgLLNet	37
2.12	Resultados obtidos com o modelo AgLLNet	38
2.13	Visão geral da implementação da rede LLIE-Net	38
2.14	Resultados obtidos com o modelo LLIE-Net	39
2.15	Visão geral da implementação da rede RDGAN	39
2.16	Resultados obtidos com o modelo RDGAN	40
2.17	Resultados obtidos com o modelo EnlightenGAN	41
2.18	Visão geral da implementação da rede KinD	41
2.19	Resultados obtidos com o modelo KinD	42
2.20	Visão geral da implementação de Ren et al. (2019)	43
2.21	Visão geral da arquitetura DRBN	43
2.22	Resultados obtidos com o modelo DRBN	44
2.23	Resultados obtidos com o modelo LLED-Net	45
2.24	Resultados obtidos com o modelo de Afifi et al. (2020)	46
2.25	Visão geral da arquitetura de Xiong et al. (2020)	47
2.26	Resultados obtidos com o modelo de Xiong et al. (2020)	47
2.27	Visão geral do <i>pipeline</i> SID. Fonte: Chen et al. (2018)	48
2.28	Comparativo de resultados para transformação de imagens em Chen et al. (2018)	49
2.29	Resultados da aplicação do modelo CAN	50
2.30	Esquemático do modelo DualGAN/ CycleGAN	51
2.31	Resultados da aplicação do modelo CycleGAN	52
2.32	Esquemático da arquitetura WESPE	52
2.33	Resultados da aplicação do modelo WESPE	53
3.1	Exemplo de conjunto de imagens que compõe o <i>dataset</i> A6300.	63
3.2	Exemplo de conjunto de imagens que compõe o <i>dataset</i> Cai2018.	64
3.3	Exemplo de conjunto de imagens sintéticas geradas a partir do <i>dataset</i> FiveK	65

3.4	Exemplo de conjunto de imagens sintéticas geradas a partir do <i>dataset</i> HDR+	66
4.1	Visão geral do modelo de rede convolucional de aprendizagem supervisionada para restauração de imagens inapropriadamente expostas	73
4.2	Arquitetura de rede convolucional de aprendizagem supervisionada para restauração de imagens inapropriadamente expostas	75
4.3	<i>Receptive field</i> do bloco convolucional proposto	79
4.4	Mapa de pesos baseado em níveis de intensidade utilizado no ajuste do modelo	81
5.1	Resultados da restauração de imagens subexpostas sintéticas	96
5.2	Resultados da restauração de imagens com sobre-exposição sintética	97
5.3	Resultados da restauração de imagens subexpostas sintéticas	104
5.4	Resultados da restauração de imagens com sobre-exposição sintética	105
5.5	Resultados da restauração de imagens subexpostas	111
5.6	Resultados da restauração de imagens sobre-expostas	112
5.7	Resultados da restauração de imagens noturnas	118
5.8	Resultados da restauração de imagens sobre-expostas	119
5.9	Impactos de exposição inapropriada na segmentação de instância	126
5.10	Impactos de exposição inapropriada na segmentação de instância (continuação)	127
5.11	Escala de cores utilizada para representação dos mapas de ativação	132
5.12	Imagem de entrada subexposta, saída do modelo e referência.	133
5.13	Primeiro bloco convolucional (imagem subexposta)	133
5.14	Blocos convolucionais dentro do <i>encoder</i> (imagem subexposta)	134
5.15	Blocos convolucionais dentro do <i>decoder</i> (imagem subexposta)	135
5.16	Fluxo paralelo ao <i>encoder-decoder</i> (imagem subexposta)	136
5.17	Três últimas camadas sequencias da rede (imagem subexposta)	136
5.18	Imagem de entrada sobre-exposta, saída do modelo e referência.	137
5.19	Primeiro bloco convolucional (imagem sobre-exposta)	137
5.20	Blocos convolucionais dentro do <i>encoder</i> (imagem sobre-exposta)	138
5.21	Blocos convolucionais dentro do <i>decoder</i> (imagem sobre-exposta)	139
5.22	Fluxo paralelo ao <i>encoder-decoder</i> (imagem sobre-exposta)	140
5.23	Três últimas camadas sequenciais da rede (imagem sobre-exposta)	140
7.1	Resultados qualitativos dos métodos relacionados em uma imagem sobre-exposta sintética	212
7.2	Resultados qualitativos dos métodos relacionados em uma imagem sobre-exposta sintética	213
7.3	Resultados qualitativos dos métodos relacionados em uma imagem subexposta sintética	214

7.4	Resultados qualitativos dos métodos relacionados em uma imagem subexposta sintética	215
7.5	Resultados qualitativos dos métodos relacionados em uma imagem sobreexposta sintética	216
7.6	Resultados qualitativos dos métodos relacionados em uma imagem sobreexposta sintética	217
7.7	Resultados qualitativos dos métodos relacionados em uma imagem subexposta sintética	218
7.8	Resultados qualitativos dos métodos relacionados em uma imagem subexposta sintética	219
7.9	Resultados qualitativos dos métodos relacionados em uma imagem sobreexposta sintética	220
7.10	Resultados qualitativos dos métodos relacionados em uma imagem subexposta real	221
7.11	Resultados qualitativos dos métodos relacionados em uma imagem sobreexposta real	222
7.12	Resultados qualitativos dos métodos relacionados em uma imagem subexposta real	223
7.13	Resultados qualitativos dos métodos relacionados em uma imagem sobreexposta sintética	224
7.14	Resultados qualitativos dos métodos relacionados em uma imagem sobreexposta real	225
7.15	Resultados qualitativos dos métodos relacionados em uma imagem subexposta real	226

LISTA DE TABELAS

2	Interpretação das medidas de qualidade de imagem aplicadas na avaliação.	62
3	Modelos de classificação considerados nos experimentos	70
4	Expansão do campo receptivo dentro da rede proposta através do fluxo <i>encoder</i>	78
5	ρ de Spearman (SRCC) para medidas de qualidade de imagem no <i>dataset</i> A6300	84
6	τ de Kendall para qualidade de imagens subexpostas do <i>Dataset</i> A6300 considerando resultados de restauração	86
7	Valor-p para o τ de Kendall para qualidade de imagens subexpostas do <i>Dataset</i> A6300 restauradas	86
8	τ de Kendall para medidas de qualidade de imagens sobre-expostas do <i>Dataset</i> A6300 considerando resultados de restauração	88
9	Valor-p para o τ de Kendall para medidas de qualidade de imagens sobre-expostas do <i>Dataset</i> A6300 restauradas	88
10	Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens subexpostas do <i>dataset</i> Fivek	91
11	Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens sobre-expostas do <i>dataset</i> Fivek	91
12	Mediana para restauração de imagens subexpostas geradas a partir do <i>dataset</i> <i>FiveK</i> (sintético)	93
13	Mediana para restauração de imagens sobre-expostas geradas a partir do <i>dataset</i> <i>FiveK</i> (sintético)	93
14	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do <i>dataset</i> <i>FiveK</i> (valor-p)	94
15	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do <i>dataset</i> <i>FiveK</i> (valor-p)	94
16	Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens subexpostas do dataset HDR+ <i>burst</i>	99
17	Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens sobre-expostas do dataset HDR+ <i>burst</i>	99
18	Mediana para restauração de imagens subexpostas geradas a partir do dataset HDR+ <i>burst</i> (sintético)	101
19	Mediana para restauração de imagens sobre-expostas geradas a partir do dataset HDR+ <i>burst</i> (sintético)	101
20	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do <i>dataset</i> HDR+ <i>burst</i> (valor-p)	102
21	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do <i>dataset</i> HDR+ <i>burst</i> (valor-p)	102

22	Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens subexpostas do <i>dataset</i> A6300	107
23	Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens sobre-expostas do <i>dataset</i> A6300	107
24	Mediana para restauração de imagens subexpostas do dataset A6300 (Steffens et al., 2018a) (real)	108
25	Mediana para restauração de imagens sobre-expostas do dataset A6300 (Steffens et al., 2018a) (real)	108
26	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do <i>dataset</i> A6300 <i>Multi-Exposure Dataset</i> (valor-p)	109
27	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do <i>dataset</i> A6300 <i>Multi-Exposure Dataset</i> (valor-p)	109
28	Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens subexpostas do <i>dataset</i> Cai et al. (2018)	114
29	Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens sobre-expostas do <i>dataset</i> Cai et al. (2018)	114
30	Mediana para restauração de imagens subexpostas do <i>dataset</i> Cai et al. (2018) (real)	115
31	Mediana para restauração de imagens sobre-expostas do <i>dataset</i> Cai et al. (2018) (real)	115
32	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do <i>dataset</i> Cai2018 <i>Multi-Exposure</i> (valor-p)	116
33	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do <i>dataset</i> Cai2018 <i>Multi-Exposure</i> (valor-p)	116
34	Comparativo entre métodos de restauração para imagens subexpostas	121
35	Comparativo entre métodos de restauração para imagens sobre-expostas	122
36	Acurácia Top-1 para o modelo VGG-16 (Simonyan & Zisserman, 2014) considerando imagens com diferentes níveis de subexposição e sobre-exposição simulada e restauradas pelo modelo de restauração proposto	125
37	Tempo médio de inferência em milissegundos para restauração de imagens utilizando o modelo proposto em milissegundos	130
38	Mediana para restauração de imagens subexpostas do <i>dataset</i> Cai et al. (2018) no teste de ablação	131
39	Mediana para restauração de imagens sobre-expostas do <i>dataset</i> Cai et al. (2018) no teste de ablação	132
G.40	Teste de correlação de Pearson (PCC) para medidas de qualidade de imagem no Dataset A6300 incluindo condições de sub e sobre-exposição	166
G.41	Resultados para o teste de normalidade para a saída dos modelos utilizando o <i>dataset</i> FiveK subexposto	167

G.42	Resultados para o teste de normalidade para a saída dos modelos utilizando o <i>dataset</i> FiveK sobre-exposto	170
G.43	Valores de media para restauração de imagens subexpostas do <i>dataset</i> FiveK	174
G.44	Valores de media para restauração de imagens sobre-expostas do <i>dataset</i> FiveK	175
G.45	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do <i>dataset</i> FiveK (estatística do teste)	176
G.46	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do <i>dataset</i> FiveK (estatística do teste)	177
G.47	Resultados para o teste de normalidade para a saída dos modelos utilizando o <i>dataset</i> HDR+ <i>burst</i> subexposto	178
G.48	Resultados para o teste de normalidade para a saída dos modelos utilizando o <i>dataset</i> HDR+ <i>burst</i> sobre-exposto	181
G.49	Valores de media para restauração de imagens subexpostas do <i>dataset</i> HDR+ <i>burst</i>	185
G.50	Valores de media para restauração de imagens sobre-expostas do <i>dataset</i> HDR+ <i>burst</i>	186
G.51	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do <i>dataset</i> HDR+ <i>burst</i> (estatística do teste)	187
G.52	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do <i>dataset</i> HDR+ <i>burst</i> (estatística do teste)	188
G.53	Resultados para o teste de normalidade para a saída dos modelos utilizando o <i>dataset</i> A6300 subexposto	189
G.54	Resultados para o teste de normalidade para a saída dos modelos utilizando o <i>dataset</i> A6300 sobre-exposto	192
G.55	Valores de média para restauração de imagens subexpostas do <i>dataset</i> A6300 <i>Multi-Exposure Dataset</i>	196
G.56	Valores de média para restauração de imagens sobre-exposta do <i>dataset</i> A6300 <i>Multi-Exposure Dataset</i>	197
G.57	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do <i>dataset</i> A6300 <i>Multi-Exposure Dataset</i> (estatística do teste)	198
G.58	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do <i>dataset</i> A6300 <i>Multi-Exposure Dataset</i> (estatística do teste)	199
G.59	Resultados para o teste de normalidade para a saída dos modelos utilizando o <i>dataset</i> Cai et al. (2018) subexposto	200
G.60	Resultados para o teste de normalidade para a saída dos modelos utilizando o <i>dataset</i> Cai et al. (2018) sobre-exposto	203

G.61	Valores de média para restauração de imagens subexpostas do <i>dataset</i> Cai2018 <i>Multi-Exposure Dataset</i>	207
G.62	Valores de média para restauração de imagens sobre-expostas do <i>dataset</i> Cai2018 <i>Multi-Exposure Dataset</i>	208
G.63	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do <i>dataset</i> Cai2018 <i>Multi-Exposure Dataset</i> (estatística do teste)	209
G.64	Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do <i>dataset</i> Cai2018 <i>Multi-Exposure Dataset</i> (estatística do teste)	210
G.65	Avaliação do impacto de distorções simuladas no desempenho de modelos de reconhecimento de imagens	211

LISTA DE SÍMBOLOS

A	Matriz
a	Vetor
a	Escalar
W	Matriz de pesos
D	Rede Discriminadora
G	Rede Geradora
X	Dados de Entrada (imagem ou lote de imagens)
Y	Dados de Referência (imagem ou lote de imagens)
$\hat{Y}$	Saída de modelo
ℓ	Camadas na rede neural
$\mathcal{L}$	Função Objetivo

Símbolos gregos

μ	Média
σ	Desvio Padrão
α	Constante Empírica
β	Constante Empírica
γ	Constante Empírica
λ	Constante Empírica

LISTA DE ABREVIATURAS

ABNT	Associação Brasileira de Normas Técnicas
AE	Erro Absoluto
ANN	<i>Rede Neural Artificial</i>
BIQI	<i>Blind Image Quality Index</i>
BM3D	<i>Block-matching and 3D filtering</i>
BN	Normalização por lote (<i>Batch Normalization</i>)
BNN	Bayesian neural network
BPDHE	<i>Brightness Preserving Dynamic Histogram Equalization</i>
BRISQUE	<i>Blind/Referenceless Image Spatial Quality Evaluator</i>
CAN	<i>Context Aggregation Network</i>
CDF	Função Distribuição Acumulada
CIEDE2000	Delta E (CIE 2000)
CLAHE	<i>Contrast Limited Adaptive Histogram Equalization</i>
CNN	Rede Neural Convolucional (<i>Convolutional neural network</i>)
DCT	Transformada Discreta de Cossenos
DHE	<i>Dynamic Histogram Equalization</i>
DNG	<i>Adobe Digital Negative</i>
DNN	Rede Neural Profunda
DSLR	<i>Digital Single-Lens Reflex</i>
DSSIM	Dissimilaridade Estrutural
ELU	<i>Exponential Linear Unit</i>
EV	Compensação de Exposição
FSIM	<i>Feature Similarity Index</i>
GAN	<i>Generative Adversarial Network</i>
GMSD	Gradient Magnitude Similarity Deviation
Grad-CAM	<i>Gradient-weighted Class Activation Mapping</i>
HDR	<i>High Dynamic Range</i>
i.i.d.	Independente e Identicamente Distribuído
IN	Normalização por Instância (<i>Instance Normalization</i>)
IoU	Razão Intersecção/União
JB	Teste de normalidade Jarque-Bera
JPEG	<i>Joint Photographic Experts Group</i> (padrão de compressão de imagens)
KL	Kullback–Leibler
LReLU	<i>Leaky Rectified Linear Unit</i>
MAE	Erro médio Absoluto
MOS	<i>Mean Opinion Score</i>
MSE	Erro médio quadrático
NN	Rede Neural (<i>Neural network</i>)

PCA	Análise de Componentes Principais (Principal Component Analysis)
PCA-NE	<i>Principal Component Analysis Noise Estimator</i>
PCC	Coefficiente de Correlação de Pearson
PDF	Função densidade de probabilidade
PSNR	Relação sinal-ruído de pico (<i>Peak Signal to Noise Ratio</i>)
QAC	<i>Quality-Aware Clustering</i>
RECO	<i>Polar Edge Coherence</i>
ReLU	<i>Rectified Linear Unit</i>
RL	Aprendizagem por Reforço (<i>Reinforcement learning</i>)
RMS	Valor Quadrático Médio (Root Mean Square)
RMSE	Erro médio Quadrático (<i>Root mean square error</i>)
RNN	Rede Neural Recorrente (<i>Recurrent neural network</i>)
SICE	<i>Single Image Contrast Enhancement</i>
SLR	<i>Single Lens Reflex</i>
SRCC	Coefficiente de Correlação de Postos de Spearman
SRGAN	<i>Super Resolution Generative Adversarial Network</i>
sRGB	<i>Standard Red Green Blue</i> (espaço de cor especificado pela IEC 61966-2-1)
SSIM	Métrica de Similaridade Estrutural entre Duas Imagens (<i>Structural SIMilarity</i>)
UIQI	<i>Universal Image Quality Index</i>
VIFP	<i>Visual Information Fidelity</i>

1 INTRODUÇÃO

Sezan & Tekalp (1990) definem restauração da imagem como o problema de estimar a imagem ideal a partir de sua renderização borrada e ruidosa. Mohapatra et al. (2014), por sua vez, tratam a restauração de imagens como a tentativa de melhorar a qualidade de uma imagem por meio do conhecimento do processo físico que levou à sua formação, com o objetivo de compensar ou desfazer defeitos que a degradam. Ainda segundo Mohapatra et al. (2014), a restauração de imagem difere do aprimoramento de imagem, uma vez que no segundo o objetivo do processamento é centrado na acentuação ou extração dos recursos da imagem, ao invés de focar na restauração das degradações.

Para Gonzalez & Woods (2009), o processamento digital de imagens destina-se primariamente à melhoria das informações visuais para a interpretação humana, bem como ao tratamento de dados de imagens para armazenamento, transmissão e representação, considerando a percepção automática através de computação visual. O presente trabalho tem a intenção de contribuir apresentando uma solução computacional baseada em aprendizagem profunda, dedicada à restauração de imagens digitais em que a representação das propriedades da cena é prejudicada pela parametrização fotograficamente incorreta ou por limitação do equipamento de aquisição. Assume-se que este processamento ocorre em momento posterior à aquisição, desconsiderando propriedades específicas do conjunto câmera-lente utilizado.

Imagens digitais fornecem uma representação da aparência visual de uma cena (Dawson-Howe, 2014). Mais precisamente, uma imagem é gerada pela combinação de uma fonte de iluminação e pela reflexão ou absorção de energia dessa fonte pelos elementos da cena cuja imagem está sendo gerada (Gonzalez & Woods, 2009). Para a aquisição de imagens digitais emprega-se, em geral, um arranjo de lentes ópticas que projetam a energia da cena em um sensor matricial. Em condições ideais, a resposta de cada sensor é proporcional à integral da energia luminosa projetada sobre a superfície do sensor. Um circuito eletrônico realiza a varredura das saídas do sensor e as converte um formato suportado pelos demais componente do sistema de aquisição de imagens.

Fabricantes de câmeras digitais tem se esforçado para produzir sistemas de aquisição mais fidedignos, concentrando esforços no aparato óptico, nos sensores e no processamento embarcado. Os detalhes sobre a implementação destas melhorias são, em geral, mantidos sob sigilo. Este trabalho não tem o propósito de cobrir os detalhes de implementação ou propor melhorias aos equipamentos e algoritmos utilizados para aquisição e armazenamento. Limita-se aqui em verificar os efeitos práticos de sua utilização e as dificuldades adicionais que a compressão com perdas impõe para a restauração de imagens e o seu impacto ao nível de aplicação.

Cenas com muito contraste (faixas dinâmicas que as câmeras não conseguem representar) representam um desafio para sistemas de aquisição de imagens, o mesmo é válido para cenas com pouco contraste (Afifi et al., 2020; Lv et al., 2018, 2021; Wang et al., 2020; Xiong et al., 2020). Imagens adquiridas a partir de câmeras convencionais, que operam no espectro da luz visível,

são comumente afetadas por artefatos e distorções decorrentes do excesso ou da falta de luz. A radiância da cena fora dos limites do sistema de aquisição resulta em subexposição e/ou sobre-exposição (*clipping*). Em fotografia digital, a subexposição é um fenômeno que ocorre quando o sensor da câmera é incapaz de capturar diferenças entre as partes mais escuras da imagem, fazendo com que detalhes sejam perceptíveis apenas nas áreas mais claras da cena fotografada. A subexposição pode ocorrer por diversos fatores, incluindo iluminação insuficiente, tempo de exposição muito curto ou abertura muito pequena do diafragma da lente. A sobre-exposição, por outro lado, ocorre quando o sensor recebe luz em excesso, saturando-o e tornando-no incapaz de diferenciar as partes mais claras da imagem. Neste caso, criam-se grandes regiões brancas onde não é possível visualizar qualquer objeto.

A subexposição e a sobre-exposição são características indesejáveis que prejudicam a performance de algoritmos empregados na visão computacional (Steffens et al., 2021). A exposição equivocada da cena pode ‘esconder’ detalhes muito claros ou muito escuros. Os algoritmos de visão computacional são tipicamente concebidos para funcionar com imagens que apresentam uma resposta linear a radiância correspondente à cena que está sendo imageada (Debevec & Malik, 2008). Neste sentido, admite-se que, quando os algoritmos são aplicados em imagens que apresentam distorções, seus resultados podem ser incorretos (Ai & Kwon, 2020; Singh & Parihar, 2020; Ignatov & Timofte, 2019; Jain & Raman, 2021).

A exposição inapropriada impacta, dentre outros, algoritmos para tarefas de classificação, segmentação e reconhecimento como os propostos em Redmon & Farhadi (2017), Hu et al. (2017) e Liu et al. (2016) sistemas de medição baseados em visão como os propostos em Shirmohammadi & Ferrero (2014), Malla et al. (2010); sistemas para reconhecimento de texto na cena como os apresentados por Luo et al. (2019), Xie et al. (2019) e Huang et al. (2020); robôs de serviço como os propostos em Lee et al. (2019b), Chi et al. (2018), Aravena et al. (2018) e Appuhamy & Madhusanka (2018); sistemas de vigilância inteligente baseada em vídeo em Simionelli & Quaglio (2015); Saini et al. (2016); e sistemas de reconhecimento voltados para veículos autônomos como os propostos em Arvind et al. (2018), Shi et al. (2017), Yang et al. (2017), Fairfield et al. (2016) e Kohli & Chadha (2019). Especialmente sobre veículos autônomos Kohli & Chadha (2019) apresentam um estudo de caso sobre acidente envolvendo direção autônoma noturna. Zhang et al. (2019a) e Wang et al. (2019b) também reportam impactos críticos causados por subexposição e sobre-exposição decorrentes de sombra e iluminação solar direta em sistemas de visão para veículos autônomos.

1.1 Justificativa

Por um lado, as técnicas de visão computacional atingiram um patamar onde podem garantir alta precisão e acurácia em diversos *benchmarks* (Wanner & Goldluecke, 2013; Russakovsky et al., 2015; Benenson et al., 2019). Entretanto, muitas destas técnicas concentram-se em aplicações que presumem iluminação controlada, homogênea, e superfícies com reflexão difusa, fatores que não refletem a realidade para aplicações em cenas ao ar livre. Este trabalho visa mi-

tigar os efeitos adversos da aquisição de imagens em condições desafiadoras através de modelos de aprendizagem profunda. A aprendizagem profunda vem sendo utilizada no tratamento de problemas desafiadores associados especialmente às ciências da vida, visão computacional, reconhecimento de voz, processamento de linguagem natural e veículos autônomos (LeCun et al., 2015). Os modelos de aprendizagem profunda utilizados atualmente são comumente baseados em redes neurais. Segundo Awodele & Jegede (2009), devido à sua habilidade de derivar informações com significância a partir de dados complexos ou incompletos, as redes neurais são adequadas para reconhecer padrões e encontrar tendências que são complexas demais para serem percebidas por humanos ou mesmo por técnicas de inteligência artificial tradicionais.

Para restauração e aprimoramento de imagens, o tratamento dos fenômenos ópticos, elétricos e algorítmicos depende de um conhecimento prévio de todas as variáveis envolvidas nos processos de aquisição, digitalização, compressão, transferência e armazenamento de imagens. Em virtude das condições de iluminação não controlada, da ampla gama de equipamentos de aquisição de imagens disponíveis e das propriedades dos diversos objetos que compõe uma cena, a modelagem através de redes neurais convolucionais permite que se obtenha um modelo robusto (Mangal et al., 2019; Hendrycks & Dietterich, 2019) e confiável para restauração de imagens afetadas pelas condições de captura. Além disso, vale ressaltar que as situações que levam à subexposição e à sobre-exposição podem ser reproduzidas por meio de simulação ou aquisição de dados no mundo real. Isto é importante para o ajuste supervisionado do modelo.

Sezan & Tekalp (1990) destaca que a restauração de imagens é um problema inverso mal-posto, ou seja, uma solução única pode não existir ou as soluções podem não depender continuamente dos dados. Dadas as condições e aplicações mencionadas na justificativa, evidencia-se a aplicabilidade de modelos baseados em aprendizado para lidar com o problema da restauração de imagens obtidas por câmeras digitais convencionais. Assim sendo, argumenta-se que a pesquisa pode tratar de forma eficaz o problema abordado, apresentar relevância científica ao trazer novos conhecimentos enquanto explora e estende o estado atual da evolução científica e, por fim, atender à necessidades fundamentais da comunidade acadêmica envolvida com a pesquisa de métodos de visão computacional e sistemas que se utilizam deste campo de pesquisa.

1.2 Hipótese Científica

É possível modelar uma rede neural capaz de restaurar imagens afetadas pela heterogeneidade de exposição, mantendo suas características dimensionais e informações estruturais relevantes para a visão computacional.

1.3 Objetivos

Tem-se como objetivo apresentar um modelo de rede neural artificial compacta capaz de restaurar imagens impactadas por exposição inapropriada obtidas a partir de câmeras digitais convencionais. Isso inclui preservar e aprimorar as características mais representativas para a

visão computacional, tais como definição, contraste, nitidez e correção de cor. Os resultados do estudo proposto são comparados quali-quantitativamente com outras alternativas atualmente disponibilizadas na literatura, baseadas ou não em aprendizado. Para a concretização deste objetivo geral, fazem-se pertinentes e necessários os seguintes objetivos específicos:

- Levantar e avaliar o estado da arte em restauração e aprimoramento de imagens digitais na literatura;
- Determinar o tipo, estrutura e modelo de rede que melhor se adéqua ao problema abordado;
- Testar os limites e a capacidade de generalização das redes neurais preexistentes;
- Apresentar um modelo de restauração para imagens subexpostas e sobre-expostas;
- Apresentar um modelo compacto, que possa ser integrado em aplicações de visão computacional contribuindo para sua robustez quanto à exposição inapropriada;
- Investigar e determinar métricas para avaliação de redes neurais aplicadas à restauração de imagens;
- Avaliar os resultados do modelo proposto;
- Aplicar, testar e validar em estudo de caso com aplicações de visão computacional;

1.4 Organização do Texto

No Capítulo 2 apresentam-se os conceitos fundamentais para o desenvolvimento deste trabalho, bem como elencam-se trabalhos estado da arte relacionados com o problema e a proposta apresentada. Métodos clássicos ou baseados em aprendizagem profunda que apresentem expressiva intersecção com o problema de pesquisa abordado no presente trabalho são discutidos.

No Capítulo 3 apresenta-se a metodologia utilizada para a realização da pesquisa. Na Seção 3.1 discutem-se os principais aspectos ligados à determinação da arquitetura de rede apresentando as premissas do modelo. Apresentam-se ainda, na Seção 3.2, as principais métricas e medidas de qualidade de imagem aplicadas ao problema de restauração de imagens. Na Seção 3.3 faz-se uma breve introdução aos conjuntos de dados utilizados no ajuste, validação e teste dos modelos de restauração de imagens afetadas por exposição inadequada. Apresentam-se dois *datasets* de imagens pareadas com subexposição e sobre-exposição real. Apresenta-se ainda dois *datasets* de imagens em condições ideais de exposição e uma metodologia para geração de saturação sintética de maneira controlada. Na Seção 3.4 apresentam-se dois problemas típicos de computação visual e robótica, nos quais o modelo de restauração de imagens é aplicado e testado. Por fim, na Seção 3.5 apresenta-se a técnica de visualização das ativações internas do modelo, utilizada para compreender o modelo.

No Capítulo 4 apresenta-se o modelo de rede neural para restauração. Apresentam-se os conceitos empregados e características da arquitetura da rede neural. Apresentam-se também as funções e procedimentos empregados para o ajuste do modelo. Expõe-se, desta forma, as propriedades que permitiram a obtenção de um modelo compacto e com boa capacidade de restauração de imagens.

No Capítulo 5 apresenta-se uma avaliação do modelo proposto considerando os resultados em quatro *datasets* distintos para condições de sub e sobre exposição. Os resultados são discutidos utilizando um conjunto de medidas de qualidade baseadas em referência, apontando que o modelo proposto atinge resultados significativamente melhores aos demais utilizados no comparativo. Faz-se uma avaliação qualitativa das imagens transformadas pelo modelo proposto, identificando-se as principais limitações e pontos de melhoria. Por fim, mostra-se como o modelo proposto pode impactar aplicações baseadas em computação visual. Apresenta-se também um estudo da estrutura e fluxo interno do modelo através da visualização por mapas de atenção.

O Capítulo 6 apresenta as considerações finais com relação ao trabalho desenvolvido e apresenta uma proposição de trabalho futuro. Faz-se uma avaliação do trabalho com relação aos objetivos e metodologia proposta, discutindo as principais contribuições alcançadas no período.

Uma parcela das ideias e resultados apresentados no presente texto tem intersecção com trabalhos publicados pelo autor em Steffens et al. (2018a), Steffens et al. (2018b), Steffens et al. (2017), Vaz-Jr et al. (2017), Hüttner et al. (2017), Vaz-Jr et al. (2018a), Vaz-Jr et al. (2018b), Steffens et al. (2019), Steffens et al. (2020a), Steffens et al. (2020b), Messias et al. (2020) e Steffens et al. (2021). Estes trabalhos apresentam resultados parciais para diferentes aspectos da restauração e avaliação da qualidade de imagens conforme os tópicos a seguir.

- *Deep Learning Based Exposure Correction for Image Exposure Correction with Application in Computer Vision for Robotics* (Steffens et al., 2018a) - Modelo de restauração de imagens inapropriadamente expostas baseado em redes geradora-adversária.
- *Análise Exploratória De Dados De Imagens Digitais Noturnas* (Steffens et al., 2018b) - Avaliação das propriedades estatísticas de imagens obtidas em condições de iluminação insuficiente.
- *Can Exposure, Noise and Compression Affect Image Recognition? An Assessment of the Impacts on State-of-the-Art ConvNets* (Steffens et al., 2019) - Validação dos impactos de exposição inapropriada, ruído e compressão em tarefas de visão computacional
- *CNN Based Image Restoration: Adjusting Ill-Exposed sRGB Images in Post-Processing* (Steffens et al., 2020a) - Modelo de restauração de imagens inapropriadamente expostas utilizando aprendizagem supervisionada.
- *A Pipelined Approach to Deal with Image Distortion in Computer Vision* (Steffens et al.,

2020b) - Proposição de modelo de restauração como etapa do pipeline de sistemas de visão computacional.

- *UCAN: A Learning-based Model to Enhance Poorly Exposed Images* (Messias et al., 2020) - Modelo de restauração de imagens inapropriadamente expostas utilizando aprendizagem supervisionada.
- *On Robustness of Robotic and Autonomous Systems Perception: An Assessment of Image Distortion on State-of-the-art Robotic Vision Model* (Steffens et al., 2021) - Investigação dos impactos de distorções de imagem em tarefas de robótica e sistemas autônomos.

2 REVISÃO TEÓRICA E TRABALHOS RELACIONADOS

A pesquisa apresentada teve o objetivo de criar um modelo de rede neural profunda capaz de minimizar os efeitos indesejáveis oriundos de exposição inadequada da cena no momento da aquisição da imagem. Busca-se fazer a restauração e reconstrução de imagens para aplicações de visão computacional e melhoria perceptual, incluindo características de cor, nitidez e estruturas. Este processamento é feito em etapa posterior à aquisição, quantização e compressão, o que inviabiliza e exclui dos trabalhos relacionados as abordagens que pressupõem acesso ao ajuste de tempo de exposição, ganho, curva de resposta do sensor, abertura da iris, e foco do conjunto óptico.

A subexposição e sobre-exposição geram efeitos adversos em toda a imagem, incorrendo em modificação dos valores de píxel, e fazendo com que não representem adequadamente a radiância da cena. Assim sendo, têm-se distintos problemas relevantes que precisam ser incorporados no modelo de restauração e aprimoramento, destacando-se o ajuste de nível de sinal em regiões que preservam algum dado e a interpolação do sinal para preenchimento de grandes regiões saturadas.

A estrutura deste capítulo é subdividida em duas frentes principais. Na Seção 2.1 abordam-se as principais técnicas voltadas ao processamento de imagens para restauração e aprimoramento. Estas técnicas são aplicadas no sentido de modificar a distribuição dos valores ou a preencher blocos da imagem. Na Seção 2.2 abordam-se os principais aspectos relacionados à morfologia, aplicação e treinamento de modelos de redes neurais convolucionais.

2.1 Restauração e Aprimoramento de Imagens (Métodos Clássicos)

Restauração de imagens digitais é um campo de engenharia que lida com métodos usados para recuperar uma cena original a partir de observações degradadas (Amudha et al., 2012). Definem-se como métodos clássicos aqueles que são baseados em transformações lineares diretas, ajuste de curva de intensidade ou modelos determinísticos. Os problemas relacionados ao processamento e aprimoramento de imagens digitais tem recebido expressiva atenção. Diversos autores buscaram categorizar as abordagens segundo suas propriedades teóricas, aplicação e características de implementação (Sezan & Tekalp, 1990; Amudha et al., 2012; Mohapatra et al., 2014; Narmadha et al., 2017). Neste sentido, apresenta-se uma revisão da literatura relacionada com a restauração de imagens subexpostas e sobre-expostas.

2.1.1 Equalização de Histograma

Um histograma de imagem é uma representação abstrata da frequência dos valores de intensidade na imagem. Nesta representação, são desconsideradas quaisquer informações sobre a posição dos píxeis. A equalização de histogramas é uma técnica clássica empregada com o

propósito de remapear valores de intensidade para uma distribuição específica. Em geral, níveis de intensidade bem distribuídos resultam em imagens com bom contraste (Wang et al., 2020). Em sua forma mais conhecida, a equalização é empregada com o objetivo de dar à imagem uma distribuição uniforme. Esta técnica é incapaz de extrapolar a partir de dados inexistentes, servindo apenas para ajuste dos níveis de intensidade já presentes na imagem original. O resultado esperado é uma melhoria no contraste da imagem.

A equalização, no formato clássico, busca remapear os valores de intensidade na imagem para uma nova distribuição de maneira global, no espaço de cor sRGB. Diversos autores apresentaram avanços que melhoram os resultados obtidos, levando em consideração particularidades de diferentes cenas. Dentre as técnicas baseadas em equalização de histogramas destacam-se os métodos *Brightness Preserving Dynamic Histogram Equalization* (BPDHE), de Ibrahim & Kong (2007); e *Dynamic Histogram Equalization* (DHE), de Abdullah-Al-Wadud et al. (2007).

Ibrahim & Kong (2007) argumenta que a forma clássica de equalização de histograma (*Global Histogram Equalization* - GHE) tende a introduzir deterioração visual desnecessária na imagem, especialmente na forma de saturação. Desta forma, o autor propõe BPDHE, um algoritmo desenhado com o objetivo de preservar a luminância média da imagem digital de entrada na imagem de saída. Inicialmente, o algoritmo suaviza o histograma de entrada com um filtro Gaussiano unidimensional. Na sequência, particiona o histograma suavizado com base em seus máximos locais. Em um terceiro passo, cada partição é atribuída a um novo intervalo dinâmico. Posteriormente, o processo de equalização do histograma é aplicado de independentemente para cada partição, com base nesse novo intervalo dinâmico. Por fim, como forma de compensar as mudanças na faixa dinâmica e no brilho médio da imagem, faz-se a normalização da imagem de saída para o brilho médio da entrada.

Assim como Ibrahim & Kong (2007), Abdullah-Al-Wadud et al. (2007) defendem que, embora excelente para melhoria de contraste, a técnica clássica de equalização de histograma causa efeitos colaterais como aparência desbotada, efeitos de tabuleiro de xadrez, ou artefatos indesejáveis. Estes efeitos colaterais são mais ou menos destacados dependendo da variação da distribuição de intensidade no histograma da imagem de entrada. *Dynamic Histogram Equalization* (DHE), proposto por Abdullah-Al-Wadud et al. (2007), é um algoritmo de equalização dinâmica que se propõe a realizar o aprimoramento de uma imagem sem causar perda de detalhes. Para tanto, a DHE particiona o histograma da imagem com base nos mínimos locais e os associa a intervalos de intensidade específicos para cada partição antes de equalizá-los separadamente. Essas partições ainda passam por um teste de reparticionamento para garantir a ausência de quaisquer partes dominantes. Em contraponto à equalização clássica, que considera a imagem como um todo e busca transformar a imagem para uma distribuição uniforme, o particionamento utilizado pelo DHE tem o propósito de trabalhar com sub-histogramas para realçar características locais.

A Figura 2.1 mostra os resultados da aplicação de métodos de melhoria de contraste baseados em equalização de histograma. Recomenda-se a visualização no formato digital. Destaca-se, além dos algoritmos BPDHE e DHE já elencados, o método de equalização adaptativa local

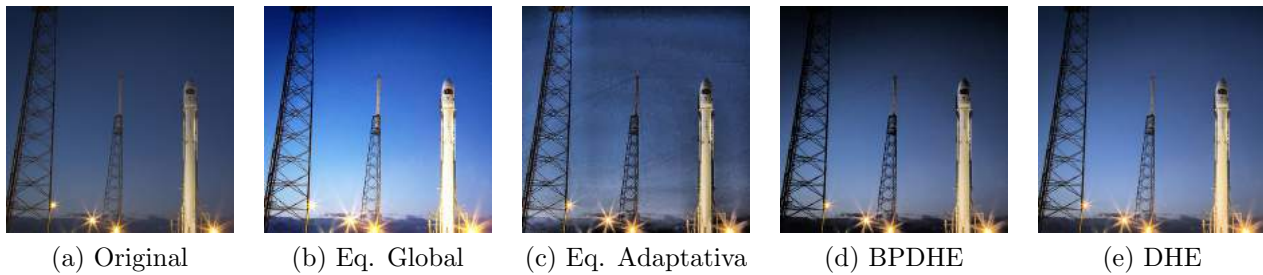


Figura 2.1: Aplicação de diferentes métodos de Equalização de Histograma

(CLAHE), de Zuiderveld (1994). Pode-se observar que a equalização global de histograma e o algoritmo DHE introduzem menos artefatos em forma de bloco na imagem, gerando uma saída onde as transições de intensidade são suaves. Identifica-se também que os métodos de equalização local são mais efetivos na preservação de detalhes da imagem, causando menos efeito de saturação. Apesar de sofrerem menos com efeitos adversos, estes algoritmos não resolvem completamente o problema. Nenhum dos algoritmos listados tem a capacidade de interpolar em regiões completamente saturadas da imagem de entrada.

2.1.2 Técnicas Baseadas em Constância de Cor (Retinex)

A constância de cor pode ser obtida estimando a cor da fonte de luz, seguida por uma transformação dos valores da imagem original usando essa estimativa de iluminação (Gijssen et al., 2011). A maior parcela dos algoritmos computacionais de constância de cor tem inspiração na teoria Retinex (contração de retina e córtex). Apresentada com profundidade em Land (1977), Retinex é a teoria da visão de cores humana proposta por Land (1977) para explicar sensações de cores em cenas reais. Através de experimentos de constância de cor, o autor foi capaz de identificar que a cor não se correlaciona com as respostas do receptor. Em cenas reais, o conteúdo da imagem inteira controla as aparências. Retinex é amplamente utilizado para identificar o processamento da imagem espacial responsável pela constância da cor (ou como o nome de algoritmos que imitam as interações espaciais da visão) para calcular a iluminação observadas em cenas complexas.

Em essência, os algoritmos de melhoria de imagens inspirados na Retinex buscam trabalhar de forma separada a cor e a iluminação (Forsyth & Ponce, 2015). Uma representação fiel da cena em imagens coloridas precisa combinar compressão da faixa dinâmica, consistência de cor e a luminosidade da reprodução tonal (Jobson et al., 1997). Dentre os métodos baseados em Retinex, destaca-se o algoritmo *Multiscale Retinex with Chromaticity Preservation* - MSRCR, de Petro et al. (2014), apresentado como uma extensão de métodos Retinex multiescala pre-existent. A Figura 2.2 apresenta os resultados atingidos pelo método em imagem subexposta (a, b) e imagem de cena com alta faixa dinâmica.

Com relação às limitações do algoritmo MSRCR, destaca-se que o modelo é altamente dependente dos parâmetros utilizados na transformação. Estes parâmetros precisam ser definidos individualmente por um observador humano. Conforme relatado pelos autores, a utilização



Figura 2.2: Aplicação de método de transformação de imagem baseado em Retinex Multiescala por intensidade. Fonte: Petro et al. (2014)

de parâmetros não-ideais pode levar a geração de artefatos de halo, perda de tonalidade da imagem, perda de detalhes na imagem, ou até mesmo inversão de cores.

Em proposta recente, Li et al. (2020b) apresentam o modelo VP (*Visual Perception*) para restauração de imagens obtidas em condições de pouca luz. Trata-se de um modelo que estende os conceitos da teoria Retinex. De acordo com os autores, o procedimento adotado por trabalhos anteriores baseados em Retinex realiza o aprimoramento estimando a intensidade da luz única, levando a problemas de fotossensibilidade visual desequilibrada e pouca adaptabilidade. Para resolver esses problemas, propõe-se o modelo de percepção visual que explora a relação entre fonte de luz e percepção visual para adquirir uma descrição matemática precisa da percepção visual. O modelo VP foi projetado para decompor a fonte de luz em intensidade de luz e distribuição espacial da luz, com a finalidade de descrever o processo de percepção de acordo com o sistema visual humano, produzindo uma estimativa de iluminação e refletância. A Figura 2.3 apresenta resultados gerados por este modelo.



Figura 2.3: Resultados obtidos com o modelo VP de Li et al. (2020b). Da esquerda para a direita: imagem com exposição inapropriada, imagem restaurada, imagem realçada utilizando equalização de histograma global. Fonte: Li et al. (2020b)

2.1.3 Técnicas baseadas em Fusão de Imagens

Técnicas baseadas em fusão de imagens são amplamente utilizadas para restauração. Dentre alguns trabalhos nesta linha pode-se elencar Toet (1992); Ancuti et al. (2012); Ancuti & Ancuti (2013); Ancuti et al. (2017); Ying et al. (2017b); Wang et al. (2018c); Vaz-Jr et al. (2018a).

Em comum, todas as abordagens baseadas em fusão de imagens se propõe a combinar diversos métodos de melhoria de contraste e redução de ruído através de pesos ponderados localmente, levando em consideração premissas específicas para o tipo de problema que se propõe a resolver.

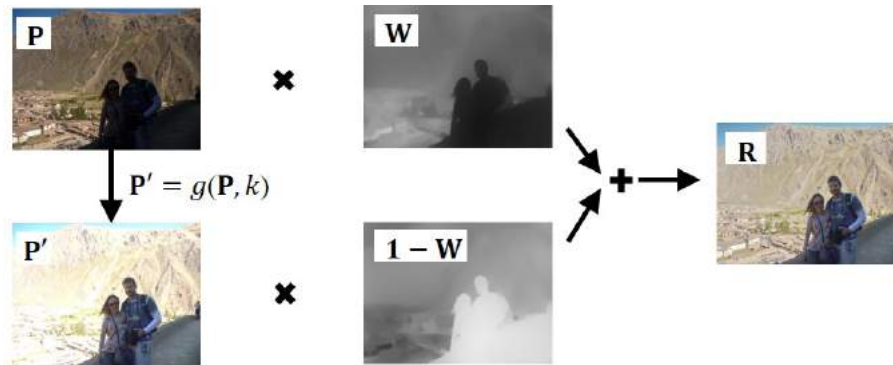


Figura 2.4: Abstração do *framework* de fusão de imagens para correção de subexposição utilizando fusão de imagens. Fonte: Ying et al. (2017b)

Destaca-se aqui a abordagem apresentada por Ying et al. (2017b) para tratamento de imagens obtidas em condições de pouca luz (cenas escuras). Ying et al. (2017b) argumentam que, embora muitas técnicas de aprimoramento de imagem tenham sido propostas para trabalhar esse problema, os métodos existentes inevitavelmente introduzem o contraste de maneira equivocada. O algoritmo proposto projeta inicialmente a matriz de peso para fusão de imagens usando técnicas de estimativa de iluminação. Uma abstração do algoritmo é apresentada na Figura 2.4. Dada uma imagem subexposta P , o algoritmo computa uma matriz de pesos W e uma imagem transformada P' . O resultado final R é obtido pelo somatório da multiplicação ponto a ponto entre as matrizes de imagem e peso.

A estimativa do pesos é componente chave, sendo projetada para que o algoritmo possa oferecer ganho relevante de contraste nas regiões muito escuras, ao mesmo tempo em que preserva o contraste em regiões apropriadamente expostas. Assume-se que a iluminação deve ser constante para regiões com estruturas similares. A partir desta estimativa o algoritmo é capaz de sintetizar imagens com várias exposições. Dentre estas imagens, escolhe-se a imagem sintética com boa exposição nas regiões onde a imagem original está subexposta. Finalmente, a imagem de entrada e a imagem sintética são fundidas de acordo com a matriz de peso para obter o resultado do aprimoramento. A aplicação do algoritmo de Ying et al. (2017b) é restrita à restauração do sinal presente na imagem. Este algoritmo não tem a capacidade de interpolar cores e texturas em regiões completamente saturadas.

Outra abordagem é apresentada por Chen et al. (2015b), que propõem um algoritmo de compensação de contraste para imagens obtidas em condições de pouca luz no espaço de cor RGB. Este algoritmo é baseado no modelo de percepção visual humana. Inicialmente, uma imagem colorida é transformada do espaço de cores RGB para o espaço de cores HSV (matiz, saturação e valor). O componente do vetor S é esticado linearmente para recuperar as informações de cores da imagem. O componente do H permanece inalterado. O componente V é utilizado para aumentar o brilho da imagem a fim de maximizar uma função de avaliação

da qualidade da imagem colorida. Por fim, faz-se a transformação inversa da imagem HSV para o espaço de cor RGB. Chen et al. (2015b) defende que a função a ser otimizada, levando em consideração entropia, contraste, fator hierárquico, função de vizinhança normalizada e largura de banda, é consistente com a qualidade percebida por observadores humanos.

Divergem do escopo da presente pesquisa as técnicas baseadas em fusão multi-espectral (Qi et al., 2013; Toet, 2005), as técnicas de fusão baseada na combinação de imagens noturnas/diurnas (Rao et al., 2010; Raskar et al., 2005) e as técnicas baseadas na combinação de múltiplas fotografias com diferentes tempos de exposição (Mertens et al., 2007). No primeiro caso, o modelo de restauração depende de uma entrada oriunda de um sensor adicional. No segundo caso, o modelo fica limitado ao imageamento prévio da cena para a qual se deseja restaura uma imagem, criando uma restrição temporal relevante. No terceiro caso, a aplicação depende do controle sobre o sensor no momento em que a imagem é adquirida, impactando ainda no tempo de aquisição, compressão, transmissão e processamento dos dados adquiridos.

2.2 Técnicas baseadas em Redes Neurais

Pouco exploradas no começo da década, Redes Neurais Artificiais com aplicações para restauração de imagem, remoção de ruído, *inpainting* e remoção de névoa passaram a atrair grande interesse da comunidade científica nos últimos anos. Dentre as aplicações desenvolvidas neste sentido pode se destacar: remoção do efeito de névoa (*de-haze*) (Cai et al., 2016; Ren et al., 2016; Gonçalves et al., 2018); super-resolução (Ledig et al., 2017; Lai et al., 2017); *de-clipping* (Cai et al., 2018); preenchimento de regiões faltantes (Pathak et al., 2016; Van Den Oord et al., 2016); restauração de visibilidade em fotografias noturnas (Chen et al., 2018; Yang et al., 2020; Guo et al., 2020; Ren et al., 2019; Lv et al., 2018; Li et al., 2020a; Lv et al., 2021; Wang et al., 2019a; Guo et al., 2019; Afifi et al., 2020; Xiong et al., 2020; Zhang et al., 2019b; Li et al., 2020b); e melhoria geral de qualidade de imagem (Ignatov et al., 2017, 2018; Gharbi et al., 2017; Huang et al., 2019; Moran et al., 2020; de Stoutz et al., 2018; Ignatov & Timofte, 2019).

2.2.1 Considerações Gerais sobre Arquiteturas de Rede

A maior parte dos modelos de transformação imagem-imagem supracitados baseiam-se em redes convolucionais com arquitetura *encoder-decoder*. A Figura 2.5 apresenta uma abstração de tal arquitetura. O *encoder* tem a função de extrair as propriedades mais importantes da imagem de entrada. Estruturalmente, o *encoder* consiste em várias camadas de convoluções seguidas, geralmente, por *Max-pooling* para seleção das *features* com maior ativação. Nestas arquiteturas, as primeiras camadas produzem poucos *feature maps* com resolução mais alta. Já nas camadas mais profundas são produzidos mais *feature maps* com resolução menor.

Já o *decoder* tem a função de utilizar os dados agregados pelo *encoder* e produzir uma nova imagem. Para tanto, o decodificador é composto por várias camadas de descompactação, geralmente implementadas por via de *upsampling* utilizando vizinhos mais próximos seguidos por convolução ou convolução transposta. A resolução das camadas aumenta e a quantidade

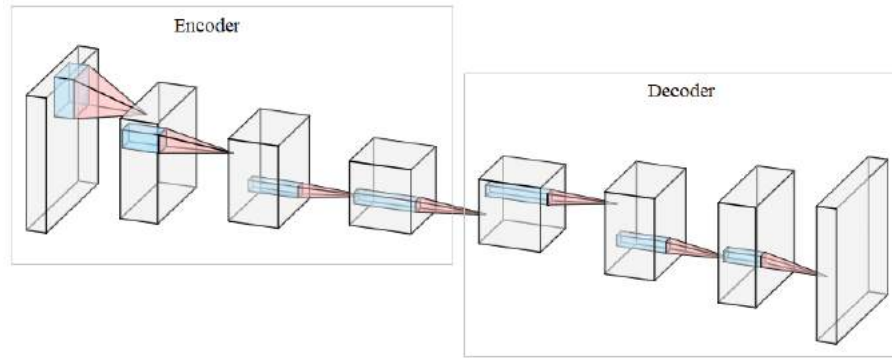


Figura 2.5: Arquitetura Convolutional *Encoder-Decoder*

de *feature maps* aumenta nas camadas mais profundas da rede, de forma que a última camada apresente as mesmas dimensões da imagem de entrada. Através destas transformações na dimensão da imagem as arquiteturas do tipo *encoder-decoder* atingem um campo receptivo expandido, mesmo com um número limitado de camadas.

U-Nets, são uma variação das arquiteturas *encoder-decoder* apresentada por Ronneberger et al. (2015). Nas U-Nets, camadas com a mesma resolução do *encoder* e do *decoder* são conectadas por meio de *skip connections*. Essas *skip connections* atuam como um caminho mais curto para transferir informações sobre a estrutura da imagem de entrada para as camadas finais da rede. Proposta inicialmente para tarefas de segmentação, a arquitetura U-Net tornaram-se um padrão para implementação de modelos de transformação imagem-imagem.

Alguns modelos de restauração utilizam ainda a arquitetura CAN (*Context Aggregation Network*). Esta arquitetura caracteriza-se pela utilização de um operador de convolução dilatada. Tais modelos vem sendo utilizados por permitirem a agregação de contexto multiescala. CAN's diferem de modelos *encoder-decoder* por dispensarem o uso de camadas de *down-scaling* e *up-scaling*, que modificam a resolução dos *feature maps* nas camadas internas da rede. Na sua forma usual todos os *feature maps* de uma CAN apresentam a mesma resolução, uma vez que as convoluções dilatadas permitem um aumento exponencial do campo receptivo sem impactar na resolução da saída produzida.

Em geral, modelos de redes neurais de transformação imagem-imagem empregam variações da arquitetura U-Net inicialmente proposto por Ronneberger et al. (2015), variando o número de parâmetros, funções de ativação internas, formas de concatenação entre camadas, e tratamentos para bordas externas da imagem conforme demanda da aplicação. Estes modelos tem sido amplamente empregados em realce e restauração de imagens (Jiang et al., 2019; Wang et al., 2018b; Honig & Werman, 2018; Cai et al., 2018; Lv et al., 2021; Guo et al., 2019; Xiong et al., 2020; Kwon et al., 2020; Zhang et al., 2019b; Chen et al., 2018; Wang et al., 2018b; Chen et al., 2018), segmentação (Zhang et al., 2020b, 2021) e geração/síntese de imagens Isola et al. (2017); Yi et al. (2017). Zhang et al. (2021) destacam que é notório que estes modelos tem excelente performance mas, para tanto, requerem uma quantidade maior de memória que outros modelos na literatura.

2.2.2 Modelos para Restauração de Imagens Inapropriadamente Expostas

O interesse da comunidade científica em modelos de correção de exposição e melhoria de contraste baseados em aprendizagem é recente e majoritariamente voltado para a restauração de imagens obtidas em condições de pouca luz. Discutem-se nesta Seção os trabalhos mais significativos relacionados ao tópico de pesquisa.

Honig & Werman (2018) introduzem o modelo DeclipNet. A arquitetura da rede proposta é composta por um gerador do tipo U-Net e uma Função Objetivo que combina de forma ponderada erro médio quadrático (MSE), perda perceptual (*perceptual loss*) e uma rede discriminadora adversária. Uma visão geral da proposta é apresentada na Figura 2.6. α , β e γ são coeficientes empíricos aplicados na combinação das três funções de perda que compõe a função objetivo. A rede é treinada em imagens truncadas, geradas a partir do *dataset* MS-COCO. A rede utiliza treinamento supervisionado com pares {imagem truncada, imagem original}. O objetivo da restauração é produzir resultados que pareçam naturais e plausíveis, dando pouca importância para o erro pixel-a-pixel.

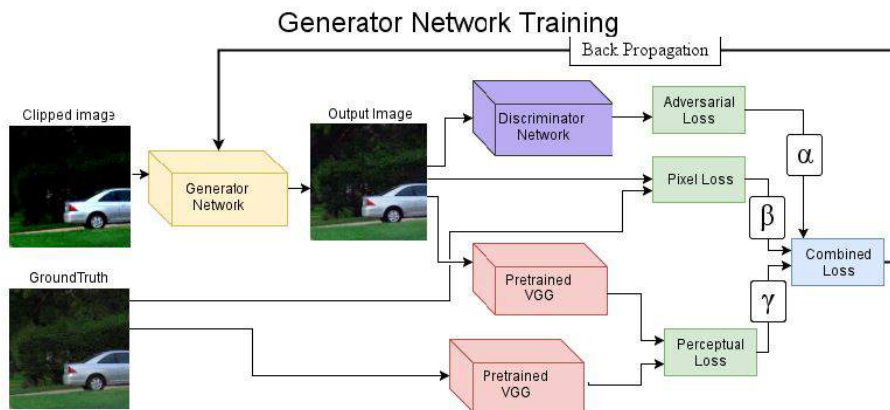


Figura 2.6: Visão geral da implementação da rede DeclipNet. Fonte: Honig & Werman (2018)

Com relação à estrutura de rede responsável pela transformação (gerador), os autores utilizam um modelo baseado em U-Net, conforme Figura 2.7. No *encoder* utilizam-se quatro blocos sequenciais no formato Convolução regular $3 \times 3 \rightarrow$ Ativação *Exponential Linear Unit* (ELU) $\rightarrow$ Convolução regular $3 \times 3 \rightarrow$ Ativação ELU $\rightarrow$ Normalização por Lotes. No *decoder* utilizam-se convoluções transpostas de 3×3 seguidas de Ativação ELU. Os autores não esclarecem qual a técnica utilizada para redução da resolução entre cada bloco do *encoder*.

Com relação à perda perceptual, Honig & Werman (2018) utilizam distância Euclidiana entre as *features* pré-treinadas de um modelo VGG16 (Simonyan & Zisserman, 2014). Para a rede adversária discriminadora, os autores utilizam a arquitetura DC-GAN apresentada por Radford et al. (2016).

Os resultados atingidos pela rede DeclipNet foram mensurados levando em consideração a opinião de 18 voluntários. Cada voluntário deveria escolher dentre 3 métodos de restauração aquele que produz resultados mais semelhantes à imagem referência (não truncada). A res-

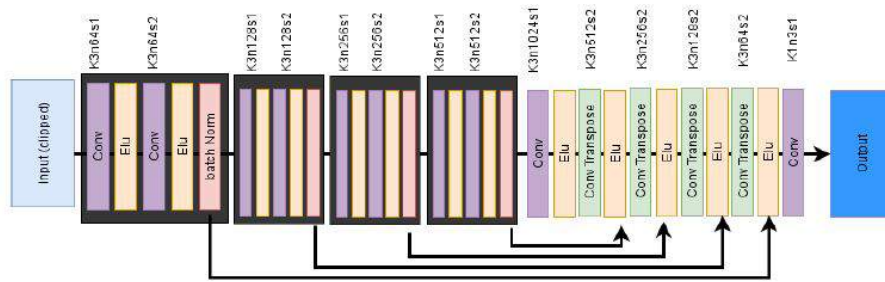


Figura 2.7: Arquitetura da rede DeclipNet . Fonte: Honig & Werman (2018)

tauração utilizando o modelo DeclipNet foi escolhida em 82,5% das vezes. Os autores argumentam que, em muitos casos, a imagem restaurada é indistinguível da imagem referência. Não são apresentados resultados quantitativos que contemplem a aplicação de métricas de qualidade de imagem.

A Figura 2.8 apresentam os resultados visuais fornecidos pela rede DeclipNet. Honig & Werman (2018) não discutem limitações sobre a resolução da imagem processada, quantidade de pesos treináveis ou número de iterações utilizadas para ajuste do modelo. A falta destes dados e a não disponibilização do modelo de forma pública impedem que os resultados sejam reproduzidos de forma confiável. Salienta-se ainda que todas as imagens utilizadas para avaliação são sintéticas, geradas a partir do dataset MS-COCO, não sendo apresentados resultados para aplicação em casos reais.



Figura 2.8: Resultados obtidos com o modelo DeclipNet. Da esquerda para a direita: destaque da região afetada, imagem de entrada, saída do modelo. Fonte: Honig & Werman (2018)

Cai et al. (2018) apresentam o modelo SICE - *Single Image Contrast Enhancer* - de aprimoramento de contraste composto por três sub-redes neurais convolucionais. A Figura 2.9 apresenta uma visão geral da implementação. No canto superior esquerdo tem-se a rede para aprimoramento de detalhes, composta por seis camadas de convoluções regulares 3×3 , seguidas por uma camada de convoluções 1×1 e uma camada residual de soma. No canto inferior

esquerdo, tem-se a rede para aprimoramento da luminância do tipo U-Net, que é composta por uma camada de convoluções regulares 9×9 com passo 2, uma camada de convoluções regulares 5×5 com passo 2, uma camada de convoluções regulares 3×3 com passo 1, uma camada de convolução transposta 3×3 com passo 1, uma camada de convolução transposta 5×5 com passo 2, uma camada de convolução transposta 9×9 com passo 2, e, por fim uma camada de convolução 1×1 na saída.

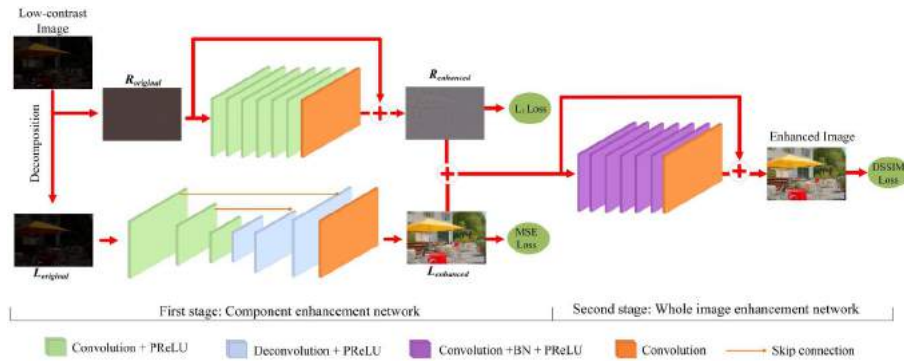


Figura 2.9: Visão geral da implementação da rede SICE. Fonte: Cai et al. (2018)

As saídas da rede para para aprimoramento de detalhes e da rede para para aprimoramento de luminância são combinadas e utilizadas como entrada para uma terceira rede, responsável por aprimorar a imagem como um todo. Esta terceira rede (lado direito da Figura 2.9) tem a mesma estrutura utilizada na rede a de aprimoramento de detalhes, exceto pelas camadas de normalização por lotes *Batch Normalization*. Todas as camadas convolucionais da rede são seguidas pela função de ativação *Parametric Rectified Linear Unit* (PReLU).

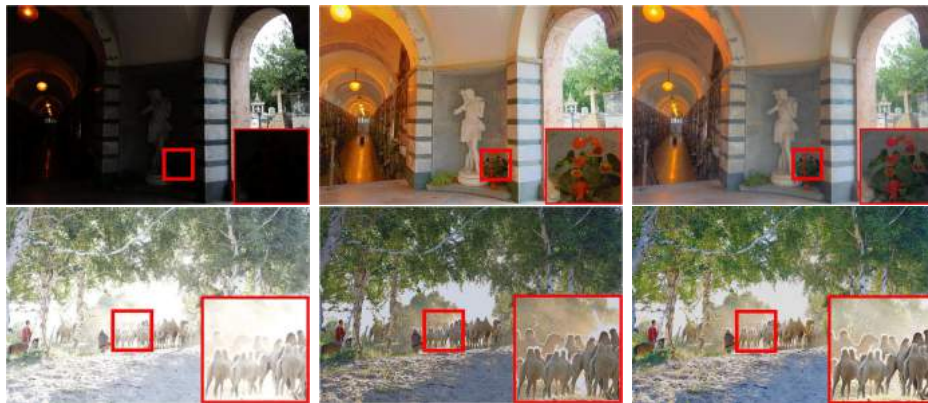


Figura 2.10: Resultados obtidos com o modelo SICE. Da esquerda para a direita: imagem com exposição inapropriada, imagem restaurada, imagem referência. Fonte: Cai et al. (2018)

A Figura 2.10 apresentam resultados obtidos pelo modelo proposto na restauração de imagens sub e sobre-expostas. Todo processo de treinamento e avaliação foi realizado utilizando um *dataset* com imagens em múltiplas exposições fotografado utilizando distintos equipamentos. Em imagens subexpostas com compensação de exposição EV−1, o método proposto por Cai et al. (2018) foi capaz de atingir uma Relação Sinal Ruído de Pico (PSNR) de 19,77 e um FSIM de 0,93 na restauração. Em imagens sobre-expostas com EV+1, o método alcançou

um PSNR de 20,21, com um FSIM de 0,93. A qualidade da restauração é limitada em regiões com saturação severa (em termos de área e intensidade). O *dataset* contendo 589 cenas com um total de 4.413 imagens é disponibilizado pelos autores. Uma descrição mais detalhada a respeito do conteúdo é apresentada na Seção 3.3.2.

Lv et al. (2021) apresentam o modelo AgLLNet para realce de imagens obtidas em condições de pouca iluminação, focando em fatores como cor, brilho, contraste, artefatos e ruído. Trata-se de uma evolução do modelo MBLLN (Lv et al., 2018) proposto pelos mesmos autores. A AgLLNet baseia-se em uma abordagem *end-to-end* com quatro sub-redes U-Net convolucionais. O modelo faz uso de dois mapas de atenção: o primeiro para guiar o aprimoramento da exposição e o segundo para guiar a supressão de ruído. O primeiro mapa de atenção distingue regiões subexpostas da imagem de regiões apropriadamente expostas. A segunda distingue entre ruídos do sensor e texturas reais. A partir destes guias, o modelo MBLLN pode trabalhar a entrada de forma adaptativa. O modelo incorpora ainda uma rede de reforço e aprimoramento de contraste para resolver a limitação de baixo contraste causada pela regressão. O modelo proposto tem aproximadamente 920.000 parâmetros treináveis.

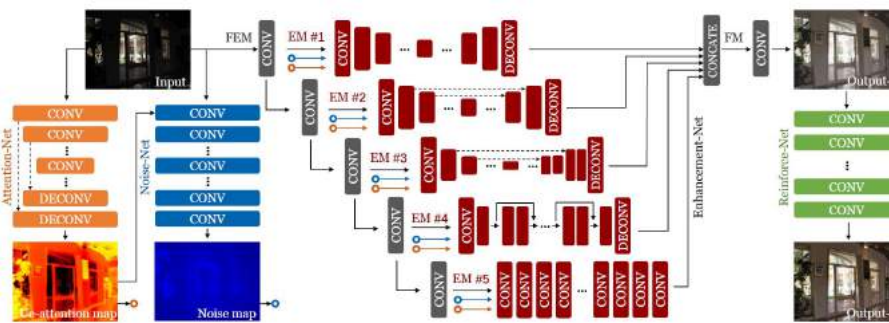


Figura 2.11: Visão geral da implementação da rede AgLLNet. Fonte: Lv et al. (2021)

Os resultados apresentados em Lv et al. (2021) mostram que o modelo proposto atinge uma PSNR de 25,24 e SSIM (Wang et al., 2004) de 0,94 em dados sintéticos sem ruído. Em dados sintéticos com ruído, o modelo atinge um PSNR de 20,84 e SSIM 0,82. Quando aplicado em dados reais do dataset LOL (Wei et al., 2018), a performance do modelo se mantém com PSNR de 20,84 e SSIM de 0,82. Os autores investigam também outras configurações, mostrando que o modelo sugerido oferece a melhor performance. Ressaltam ainda casos onde o modelo produz resultados não satisfatórios devido à textura e cor perdida pela compressão das imagens. Resultados visuais são apresentados na Figura 2.12.

Guo et al. (2019) propõem o modelo LLIE-Net de realce de imagens obtidas em condições de pouca iluminação. Trata-se de um modelo convolucional *end-to-end* inspirado em Retinex multiescala combinado com transformada *wavelet* discreta (*Discrete Wavelet Transformation - DWT*). Trata-se de um modelo em linha que compreende duas tarefas distintas: supressão de ruído e realce de imagem. A arquitetura da LLIE-Net, apresentada na Figura 2.13, é composta por 4 componentes: i) DWT e *Super-Resolution CNN* - SRCNN, destinadas à supressão de ruído e aprimoramento inicial da imagem; ii) Transformação logarítmica para geração de saídas com



Figura 2.12: Resultados obtidos com o modelo AgLLNet. Da esquerda para a direita: imagem com exposição inapropriada, imagem restaurada, imagem referência. Fonte: Lv et al. (2021)

diferentes níveis de brilho, seguida de convolução para ponderação; iii) Modelo U-Net para combinação de propriedades da imagem e; iv) Função de interpolação para produzir a imagem final a partir da combinação dos resultados das diversas transformações logarítmicas e das *features* extraídas pelo modelo U-Net. A função objetivo a ser minimizada durante treinamento é o erro médio absoluto (MAE).

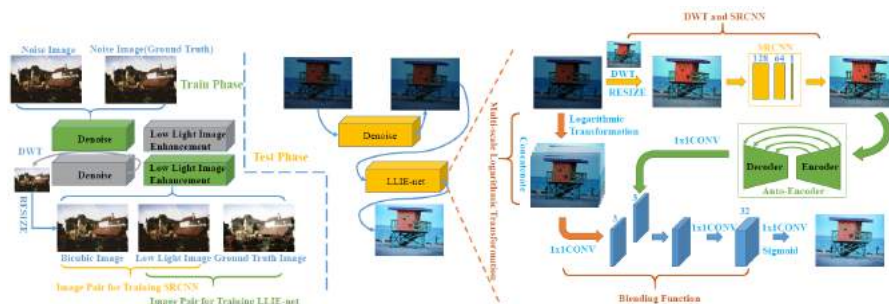


Figura 2.13: Visão geral da implementação da rede LLIE-Net. Fonte: Guo et al. (2019)

A Figura 2.14 apresenta os resultados obtidos pelo modelo LLIE-Net em dados sintéticos. Quantitativamente, Guo et al. (2019) apresenta uma validação utilizando as métricas PSNR e SSIM (Wang et al., 2004) em 1000 imagens sintéticas pareadas, apresentando PSNR de 23,68 e SSIM de 0,91. Apresenta-se ainda uma avaliação utilizando métricas cegas de qualidade de imagem em 64 imagens reais, mostrando um ILNIQE (Zhang et al., 2015) de 23,19 e SNM (Yeganeh & Wang, 2012) de 0,53. Em geral, o método produz resultados que se aproximam daqueles obtidos por modelos clássicos de equalização de histograma, Retinex e fusão de imagens.

Wang et al. (2019a), propõe o modelo RDGAN para aprimoramento de imagens subexpostas. A arquitetura do modelo é fortemente inspirada na teoria Retinex, sendo composta por duas sub-redes: a primeira faz a decomposição da imagem de entrada em componentes de refletância e iluminância, já a segunda faz a fusão dos componentes de reflectância e iluminância em uma imagem RGB realçada. Ambas as redes são baseadas em uma estrutura do tipo U-Net.

A rede de Decomposição Retinex (RD) utiliza como função objetivo uma combinação de erro



Figura 2.14: Resultados obtidos com o modelo LLIE-Net. Da esquerda para a direita: imagem com exposição inapropriada, imagem restaurada, imagem referência. Fonte: Guo et al. (2019)

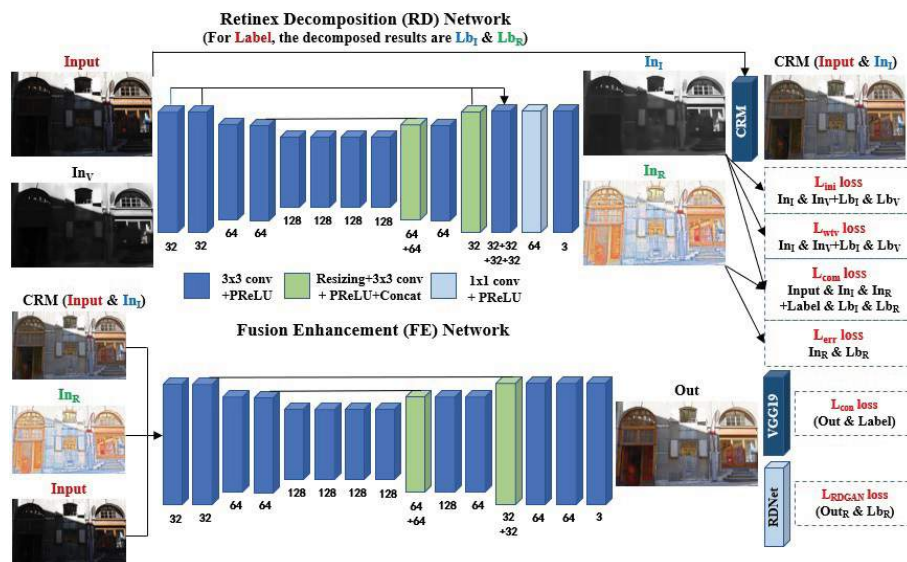


Figura 2.15: Visão geral da implementação da rede RDGAN. Fonte: Wang et al. (2019a)

médio quadrático, variância total, erro médio quadrático entre entrada e saída da rede e erro médio quadrático entre a luminância da imagem de entrada e a imagem de saída. Já a rede de fusão (FE) recebe como entrada a imagem de entrada original e as imagens correspondentes a refletância e iluminância estimadas pela RD. Como função de perda, utiliza-se uma combinação de *content-loss* sobre camadas intermediárias do modelo VGG-19 (Simonyan & Zisserman, 2014) e uma rede adversária operando sobre as componentes, voltada para aprimoramento de cor e detalhes.



Figura 2.16: Resultados obtidos com o modelo RDGAN. Da esquerda para a direita: imagem com exposição inapropriada, imagem restaurada, imagem referência. Fonte: Wang et al. (2019a)

Os resultados experimentais obtidos pelo modelo de Wang et al. (2019a) se mostram em linha com o estado-da-técnica no momento da publicação, atingindo PSNR de 22,34 e FSIMc (Zhang et al., 2011) de 0,95 em um conjunto de 53 imagens reais extraídos do dataset SICE Cai et al. (2018). Resultados visuais para o modelo são apresentados na Figura 2.16 Como efeitos adversos do realce utilizando RDGAN, os autores mencionam a amplificação de artefatos de compressão.

Jiang et al. (2019) apresentam os modelos EnlightenGAN e EnlightenGAN-N para restauração de imagens obtidas em condições de pouca luz. Trata-se de um modelo de aprendizagem não supervisionada e que dispensa dados pareados. Os autores alegam que tal abordagem faz com que o modelo generalize bem para imagens reais. A arquitetura do gerador é baseada em U-Net. O modelo emprega durante o treinamento um par de discriminadores: um discriminador global, que atua sobre a imagem como um todo e; e um discriminador PatchGAN que atua sobre blocos parciais da imagem. Com esta estratégia, busca-se produzir imagens globalmente consistentes e que preservem os detalhes existentes na imagem de entrada. EnlightenGAN emprega ainda uma camada de atenção que prioriza as regiões mais escuras da imagem. EnlightenGAN-N é uma variante da EnlightenGAN treinada com um outro conjunto de dados.

A Figura 2.17 apresenta alguns dos resultados gerados pelo modelo EnlightenGAN (Jiang et al., 2019). Observa-se que, apesar do surgimento de halos em bordas e reforço dos artefatos de compressão da imagem de entrada, o modelo apresenta resultados visualmente apelativos. Quantitativamente, os autores limitam a avaliação a métrica NIQE (Dominguez-Molina et al., 2003), mostrando que o modelo atinge resultados superiores aos apresentados por diversos modelos determinísticos e pelo modelo RetinexNet (Wei et al., 2018).

Zhang et al. (2019b) propõem o modelo KinD (*Kindling the Darkness*) para restauração



Figura 2.17: Resultados obtidos com o modelo EnlightenGAN. Da esquerda para a direita: imagem com exposição inapropriada e imagem restaurada. Fonte: Jiang et al. (2019)

de imagens obtidas em condições de pouca luz. Este modelo, assim como outros modelos contemporâneos, fundamenta-se na teoria Retinex e em uma estrutura de rede U-Net. O modelo é treinado com conjuntos pareados de imagens obtidas através de capturas com diferentes tempos de exposição. O modelo contém ainda uma função de mapeamento que permite a interferência do usuário na compensação de iluminação e um módulo para remoção de artefatos amplificados pela restauração da imagem.

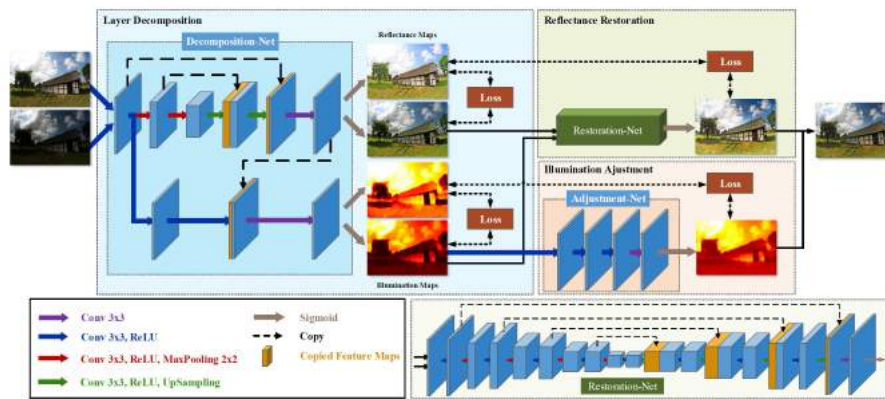


Figura 2.18: Visão geral da implementação da rede KinD. Dois fluxos correspondem à refletância e iluminação, respectivamente. Do ponto de vista funcional, a arquitetura pode ser dividida em três módulos, incluindo decomposição, restauração da refletância e ajuste de iluminação. Fonte: Zhang et al. (2019b)

A Figura 2.18 oferece uma visão geral do modelo KinD. Pode-se observar que o modelo é composto por 3 sub-redes: a primeira para a decomposição em iluminância e refletância, a segunda para restauração da refletância e uma terceira dedicada ao ajuste da iluminância. As funções objetivo utilizadas são direcionadas à preservação das características globais, através do erro médio quadrático, e à preservação de características estruturais e de alta frequência

utilizando SSIM (Wang et al., 2004).

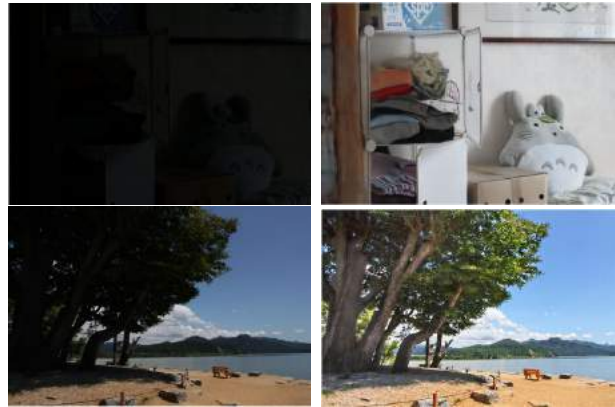


Figura 2.19: Resultados obtidos com o modelo KinD. Da esquerda para a direita: imagem com exposição inapropriada e imagem restaurada. Fonte: Zhang et al. (2019b)

Os resultados obtidos pelo modelo Kind, de Zhang et al. (2019b), utilizando as métricas PSNR, SSIM (Wang et al., 2004), LOE (Wang et al., 2013) e NIQE (Mittal et al., 2012) mostra resultados similares a outros modelos estado-da-técnica, atingindo PSNR de 20,86, SSIM de 0,80, e NIQE de 5,14. Os autores alegam ainda que o modelo é rápido, podendo restaurar uma imagem em resolução VGA (640px x 480px) em 50 milissegundos.

Ren et al. (2019) propõem um modelo de restauração de imagens obtidas sob condições de pouca iluminação. A rede consiste em dois fluxos distintos para aprender simultaneamente o conteúdo global e as estruturas salientes da imagem clara em uma rede unificada. O fluxo de conteúdo estima o conteúdo global da entrada de baixa luminosidade por meio de uma rede *encoder-decoder*. Os autores alegam que o *encoder* no fluxo de conteúdo tende a perder alguns detalhes da estrutura. Para compensar, é proposta uma rede neural recorrente (RNN) de variação espacial como um fluxo de borda para modelar detalhes de alta frequência, com a orientação de outro *auto-encoder*. A Figura 2.20 apresenta uma visão geral do modelo proposto por Ren et al. (2019). Como função objetivo o modelo combina erro médio quadrático, perda perceptual baseada em VGG (Simonyan & Zisserman, 2014) e uma rede adversária.

A avaliação dos resultados atingidos pelo modelo de Ren et al. (2019) em imagens é limitada a um pequeno conjunto de imagens sintéticas subexpostas. Nestas imagens, o comparativo mostra que o modelo se sobressai a outros voltados para o mesmo propósito. Uma avaliação mais detalhada é feita considerando o *dataset* DPED (Ignatov et al., 2017), onde o objetivo é a transformação imagem-imagem de dados obtidos a partir de câmeras de *smartphones* em imagens que se assemelhem às aquelas obtidas por câmeras DSLR. No *dataset* DPED, uma avaliação utilizando PSNR e SSIM (Wang et al., 2004) mostra que o desempenho é similar a outros modelos estado-da-técnica.

Yang et al. (2020) apresentam o modelo DRBN (*Deep Recursive Band Network*) para o tratamento de imagens degradadas por subexposição. O objetivo do modelo é recuperar uma representação linear de uma imagem de luz normal aprimorada utilizando imagens pareadas de pouca luz/luz normal, a fim de obter uma imagem realçada, recompondo as bandas fornecidas

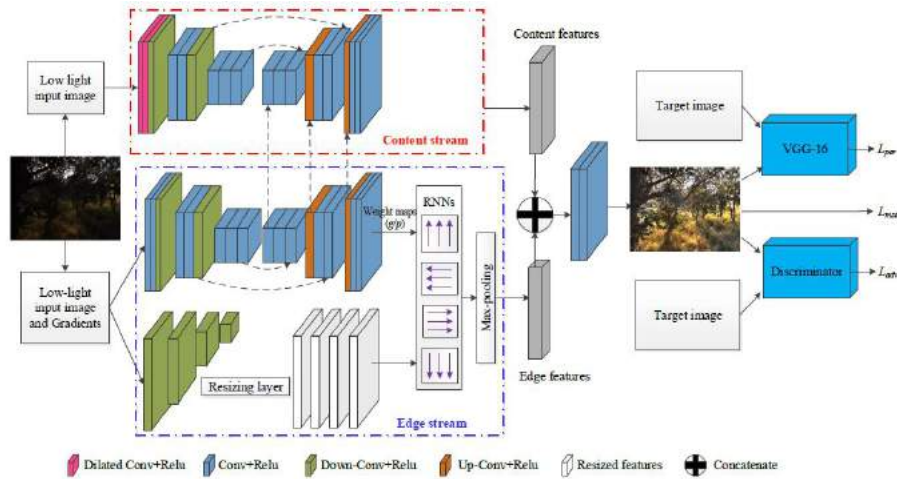


Figura 2.20: Visão geral da implementação de Ren et al. (2019). Fonte: Ren et al. (2019)

através de outra transformação linear aprendida com base em uma rede adversária com dados não pareados. Os autores alegam que a arquitetura é apropriada para treinar tanto com dados pareados e quanto não-pareados.

O arquétipo do modelo DRBN é apresentado na Figura 2.21. O modelo consiste em dois estágios: aprendizado da banda recursiva e recomposição de banda. Na aprendizagem de banda recursiva, por um lado, a rede é projetada para extrair uma série de informações globais e locais; por outro lado, a representação de banda extraída da imagem aprimorada no primeiro estágio do DRBN (aprendizado de banda recursiva) preenche a lacuna entre o conhecimento da restauração de dados pareados e a percepção. Na recomposição da banda, o modelo aprende a recompor a representação no sentido de ajustar com a ajuda desse design em dois estágios.

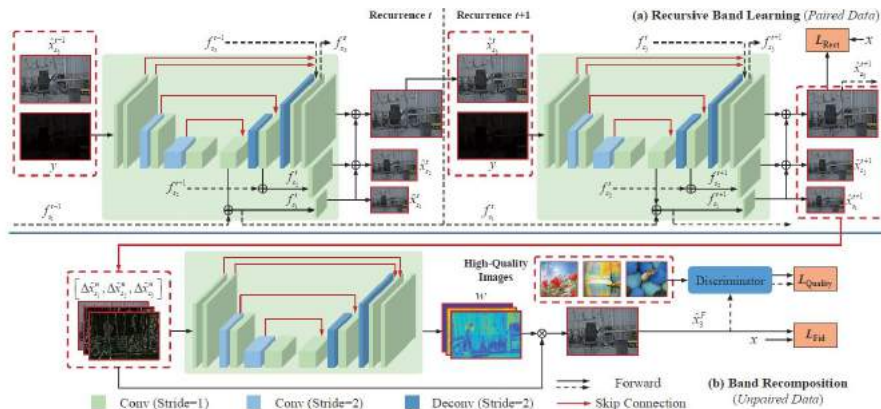


Figura 2.21: Visão geral da implementação da arquitetura DRBN. Fonte: Yang et al. (2020)

Os resultados do modelo DRBN reportados por Yang et al. (2020) se mostram, em termos de relação sinal ruído e similaridade estrutural, superiores aos modelos comparados, levando em consideração tanto modelos determinísticos quanto baseados em aprendizagem. Em *dataset* real de imagens obtidas sob condições de pouca luz o modelo atinge uma PSNR de 20,13 e SSIM de 0,82. Em geral, reporta-se um ganho expressivo de contraste e cor. Os autores defendem ainda que em alguns casos os resultados são perceptivelmente melhores que as imagens utilizadas



Figura 2.22: Resultados obtidos com o modelo DRBN. Da esquerda para a direita: imagem com exposição inapropriada, imagem restaurada, imagem referência. Fonte: Yang et al. (2020)

como referência.

Guo et al. (2020) apresentam o modelo Zero-DCE-Net (do inglês *Zero-Reference Deep Curve Estimation*). O modelo trata a tarefa de realce de luminosidade como uma tarefa de estimação de curva utilizando uma rede neural profunda do tipo U-net. O método treina uma rede leve e profunda, DCE-Net, para estimar curvas de píxel para o ajuste de faixa dinâmica da imagem. A estimativa de curva é projetada considerando a faixa de valor de píxel, a monotonicidade e a diferenciabilidade. O DCE se diferencia não requerer dados emparelhados ou não pareados durante o treinamento, o que é conseguido através de um conjunto de funções de perda sem referência cuidadosamente formuladas, que medem implicitamente a qualidade da melhoria e permitem o aprendizado utilizando gradiente descendente.

O modelo Zero-DCE-Net é projetado para encontrar as curvas que melhor realçam a imagem de entrada de forma iterativa. Para otimização, faz-se o uso de quatro funções objetivo: i) consistência espacial, calculada entre a imagem de saída e a imagem de entrada; ii) exposição, definida pelo erro médio absoluto entre a intensidade média de um bloco 16×16 px e sua correspondente ajustada pela *gray-level assumption* (Cepeda-Negrete & Sanchez-Yanez, 2013; Buchsbaum, 1980); iii) que busca corrigir possíveis desvios de cor baseada na *gray-level assumption* (Cepeda-Negrete & Sanchez-Yanez, 2013; Buchsbaum, 1980); e iv) suavidade de iluminação, que busca preservar a monotonicidade de píxeis em uma determinada região através do gradiente horizontal e vertical.

Os resultados atingidos pelo modelo Zero-DCE-Net (Guo et al., 2020) são discutidos utilizando tanto métricas de qualidade cegas quanto baseadas em referência. Um comparativo com diversos modelos evidencia que o modelo atinge performance em linha com o estado-da-técnica. Quando considerado o tempo de inferência, o modelo Zero-DCE-Net mostra-se expressivamente mais rápido. Guo et al. (2020) investiga ainda como o modelo impacta uma tarefa de reconhecimento facial em imagens noturnas, evidenciando melhora considerável na precisão e revocação.

Li et al. (2020a) apresentam o modelo LLED-Net. Trata-se de uma arquitetura baseada em U-Net, composta por *encoder-encoder* simétricos. No *encoder* utilizam-se somente camadas convolucionais com filtro 3×3 seguidas de ativação ReLU (*Rectified Linear Units*). Opta-se por não utilizar camadas de *pooling* para evitar perda de informações. No *decoder* utilizam-se camadas de deconvolução (convolução com um passo de entrada fracionário) seguido de ReLU. O treinamento é supervisionado requerendo, portanto, dados pareados. Como função objetivo

utiliza-se a métrica SSIM (Wang et al., 2004).



Figura 2.23: Resultados obtidos com o modelo LLED-Net. Da esquerda para a direita: imagem com exposição inapropriada, imagem restaurada, imagem realçada utilizando equalização de histograma global. Fonte: Li et al. (2020a)

Resultados visuais obtidos pelo modelo LLED-Net (Li et al., 2020a) são apresentados na Figura 2.23. O comparativo quantitativo apresentado leva em consideração um conjunto reduzido de imagens, nas quais o modelo se sobressai aos demais, atingindo um PSNR de 27.89 e um SSIM de 0.94. Cabe ressaltar, no entanto, que todos os modelos de restauração utilizados no comparativo são clássicos, havendo uma defasagem para o estado-da-técnica atual.

Afifi et al. (2020) propõem um modelo de restauração para imagens sRGB subexpostas e sobre-expostas. A abordagem formula o problema de correção de exposição como dois subproblemas principais, quais sejam, o realce de cor e o realce de detalhes. Trata-se de um modelo ponta-a-ponta desenhado para corrigir as informações globais de cores e refinar os detalhes da imagem. O modelo é composto por uma série de sub-redes do tipo U-Net, operando sobre a imagem de entrada em diferentes escalas, obtida através de pirâmide Laplaciana. A função objetivo do treinamento é dada pelo somatório de três erros: i) erro médio absoluto (MAE) entre a saída obtida e a saída esperada; ii) MAE piramidal, que busca manter a coerência entre as saídas de cada sub-rede; e iii) erro adversário, que emprega um discriminador com o objetivo de melhorar o realismo e apelo visual da saída gerada. Ao total, o modelo possui cerca de 7 milhões de parâmetros treináveis.

Além do modelo de rede, que é dependente de dados pareados, Afifi et al. (2020) também apresentam um novo conjunto de imagens renderizadas de raw-RGB para sRGB com diferentes configurações de exposição. Este *dataset* é construído tendo por base imagens brutas do dataset Five-K (Bychkovsky et al., 2011). Para cada imagem do *dataset* original, são geradas imagens que emulam subexposição e sobre-exposição em diferentes níveis (EVs -1.5, -, +0, +1, e +1.5). O *exposure value* zero indica as configurações de exposição originais. Como referência, utilizam-

se imagens retocadas por especialistas, conforme disponibilizadas no dataset Five-K original.



Figura 2.24: Resultados obtidos com o modelo de Afifi et al. (2020). Da esquerda para a direita: imagem com exposição inapropriada, imagem restaurada, imagem realçada utilizando equalização de histograma global. Fonte: Afifi et al. (2020)

Afifi et al. (2020) apresentam ainda avaliações e ablações sobre o método, incluindo comparações com o estado-da-técnica. O método alcança resultados em linha com métodos anteriores dedicados a imagens subexpostas e apresenta ganho significativo em imagens superexpostas. Em imagens sobre-expostas, o método atingiu PSNR de 19,19 e SSIM de 0,72, mostrando-se superior aos demais métodos utilizados no comparativo. Já em imagens subexpostas, o método atingiu PSNR de 19,37 e SSIM de 0,73. Afifi et al. (2020) investigara ainda a contribuição do erro adversário (discriminador), identificando que o mesmo diminui a qualidade dos resultados atingidos. Cabe destacar ainda que o modelo foi treinado e testado em imagens com resolução 512×512 pixels, tendo pouca robustez em regiões homogêneas de imagens de alta resolução.

Xiong et al. (2020) propõem um modelo não supervisionado para aprimoramento de imagens obtidas em condições de pouca luz utilizando redes neurais convolucionais desacopladas. O modelo de Xiong et al. (2020) é dividido em dois estágios: i) realce da iluminação e ii) supressão de ruído. Este modelo baseado em redes adversárias (GAN) de dois estágios para melhorar imagens reais de uma forma totalmente sem supervisão. Além de conjuntos de dados de referência convencionais, uma novo *dataset* é construído e usado para avaliar o desempenho do modelo.

Na Figura 2.25 pode-se observar a estrutura básica do modelo de Xiong et al. (2020), dividida em dois estágios. Pode-se observar a existência de dois geradores: o primeiro é utilizado para estimar um mapa de iluminação; o segundo, para produzir a imagem final com supressão de ruído a partir da imagem aprimorada no primeiro estágio (divisão ponto-a-ponto entre a original e o mapa de iluminação estimado), da imagem de entrada original e do mapa de iluminação estimado (máscara). Como artifício para expandir o campo receptivo do gerador, utiliza-se uma adaptação da rede piramidal proposta em Zhao et al. (2017). O discriminador utilizado é LSGAN (*Least Squares Generative Adversarial Networks*) que busca minimizar o χ^2 de Pearson.

Os resultados do modelo de Xiong et al. (2020) quando aplicado em *datasets* pareados apresentam resultados em linha com outros modelos contemporâneos, atingindo um PSNR de

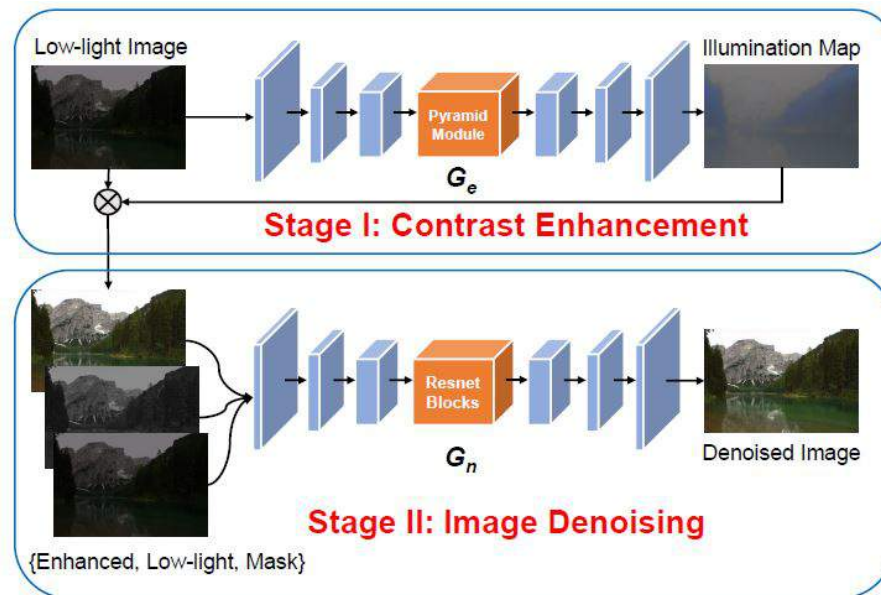


Figura 2.25: Visão geral da implementação da arquitetura de Xiong et al. (2020). Fonte: Xiong et al. (2020)



Figura 2.26: Resultados obtidos com o modelo de Xiong et al. (2020). Da esquerda para a direita: imagem com exposição inapropriada, imagem restaurada, imagem realçada utilizando equalização de histograma global. Fonte: Xiong et al. (2020)

19,78 e uma SSIM de 0,81. Apresenta-se ainda uma discussão baseada em dados do *dataset* LOL (Wei et al., 2018), evidenciando quem o modelo de Xiong et al. (2020) apresenta performance superior à Zhu et al. (2017) e Jiang et al. (2019). Resultados visuais do modelo podem ser observados na Figura 2.26, onde é possível identificar boa coerência de cores e leve borramento de regiões de alta frequência.

Por fim, Chen et al. (2018) apresentam o *dataset See-in-the-Dark* (SID), voltado para treinamento de modelos de restauração em imagens subexpostas noturnas. Esse *dataset* é composto por 5094 pares de imagens obtidos com duas câmeras:

- Conjunto 1: 2697 pares de imagens obtidas com câmera Sony $\alpha 7S$ II, com pares compostos de uma imagem subexposta e uma imagem apropriadamente exposta de ambientes com pouca luz. Os pares são obtidos variando-se o tempo de exposição do sensor.
- Conjunto 2: 2397 pares de imagens obtidas com câmera Fujifilm X-T2, com pares compostos de uma imagem subexposta e uma imagem apropriadamente exposta de ambientes com pouca luz. Os pares são obtidos variando-se o tempo de exposição do sensor.

Todas as imagens são disponibilizadas no formato bruto (RAW), preservando a totalidade dos dados da imagem tal como captada pelo sensor da câmera fotográfica. Com o objetivo de substituir o fluxo de processamento e compressão JPEG padrão utilizado por câmeras e aplicações de processamento de imagens, Chen et al. (2018) testam diferentes arquiteturas de redes convolucionais. A Figura 2.27 destaca como o modelo de restauração é aplicado com o objetivo de processar dados brutos. Para cada tipo de sensor, um novo treinamento do modelo é necessário.

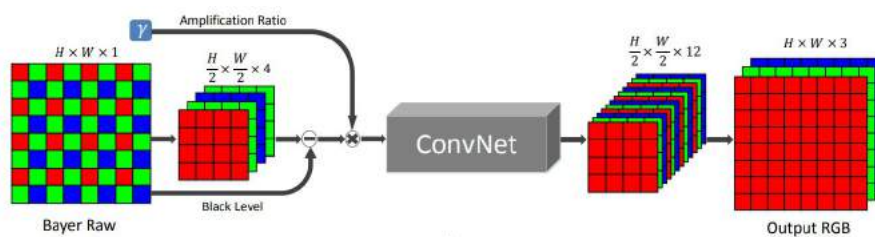


Figura 2.27: Visão geral do *pipeline* SID. Fonte: Chen et al. (2018)

Chen et al. (2018) testam diferentes arquiteturas de rede adaptadas para receber como entrada os dados brutos do sensor e fornecer como saída uma imagem no formato sRGB, em uma abordagem ponta-a-ponta. Os resultados obtidos indicam que a utilização de uma arquitetura U-Net Ronneberger et al. (2015), utilizando como função objetivo erro médio absoluto (MAE), apresenta desempenho superior aos demais modelos comparados. Os autores apresentam ainda testes substituindo a função objetivo por Dissimilaridade Estrutural (DSSIM) e erro médio quadrático (MSE), obtendo em ambas as situações resultados inferiores.

A Figura 2.28 apresenta os resultados da conversão de uma imagem crua para uma imagem no formato sRGB utilizando diferentes métodos. Verifica-se que a utilização de uma U-Net

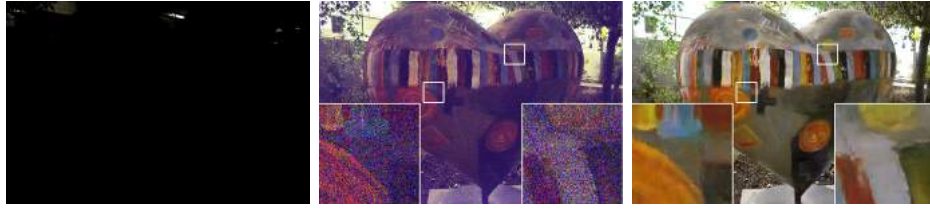


Figura 2.28: Comparativo de resultados para transformação de imagens em Chen et al. (2018). Da esquerda para a direita: saída do processamento utilizando os algoritmos embarcados da câmera, saída utilizando um fluxo de processamento clássico, saída utilizando U-Net para transformação. Fonte: Chen et al. (2018)

(Ronneberger et al., 2015) apresenta resultados superiores as demais alternativas apontadas. Embora não apresente contribuição metodológica ou teórica para a implementação de redes neurais convolucionais, é importante observar que os autores foram capazes de aplicar um modelo preexistente para um propósito até então não explorado. Cabe salientar que os resultados discutidos se limitam à aplicação do modelo de restauração em imagens RAW, nas quais todas as informações coletadas pelo sensor encontram-se preservadas.

2.3 Modelos Diversos de Tradução Imagem-Imagem

É usual que modelos de transformação imagem-imagem sejam aplicados com diferentes propósitos com mínima ou nenhuma modificação estrutural. Vários trabalhos recentes trabalham com arquiteturas de rede multipropósito. Esta Seção discute trabalhos que, muito embora sejam apenas tangencialmente relacionados com o trabalho proposto, oferecem contribuições metodológicas significativas ou ideias relevantes aplicadas aos modelos de redes convolucionais.

A maior parcela dos modelos de transformação imagem-imagem recentes é construída sobre uma arquitetura U-Net (Ronneberger et al., 2015). Diverge deste padrão o modelo *Context Aggregation Network* - CAN, proposto por Chen et al. (2017) com o intuito de substituir diversos operadores populares aplicados em melhoria de imagens por meio de modelos treinados. Chen et al. (2017) investiga diferentes arquiteturas de rede levando em consideração acurácia, tempo de execução e tamanho da rede. O modelo final é fortemente inspirado em Yu & Koltun (2015), proposto inicialmente para tarefas de segmentação semântica.

A arquitetura CAN de Chen et al. (2017) utilizam uma sequência de convoluções dilatadas (também conhecidas como *atrous convolutions*) como forma de agregar contexto. Neste modelo, todas as camadas intermediárias da rede operam com a mesma resolução da imagem de entrada, dispensando o uso de operadores de *pooling* ou deconvolução. A arquitetura da rede CAN dispensa ainda o uso de *skip connections*, reduzindo expressivamente os requisitos de memória necessários para sua utilização depois do treinamento. A primeira camada da rede utiliza convoluções sem dilatação. A segunda camada utiliza uma taxa de dilatação 2^1 . A terceira camada utiliza uma taxa de dilatação 2^2 . A dilatação é aumentada exponencialmente até a sexta camada, onde a taxa de dilatação atinge 2^6 . Desta forma, a arquitetura permite agregar informações globais da imagem de maneira gradual, minimizando significativamente a

quantidade de recursos computacionais necessários para o processamento de imagens em alta resolução.

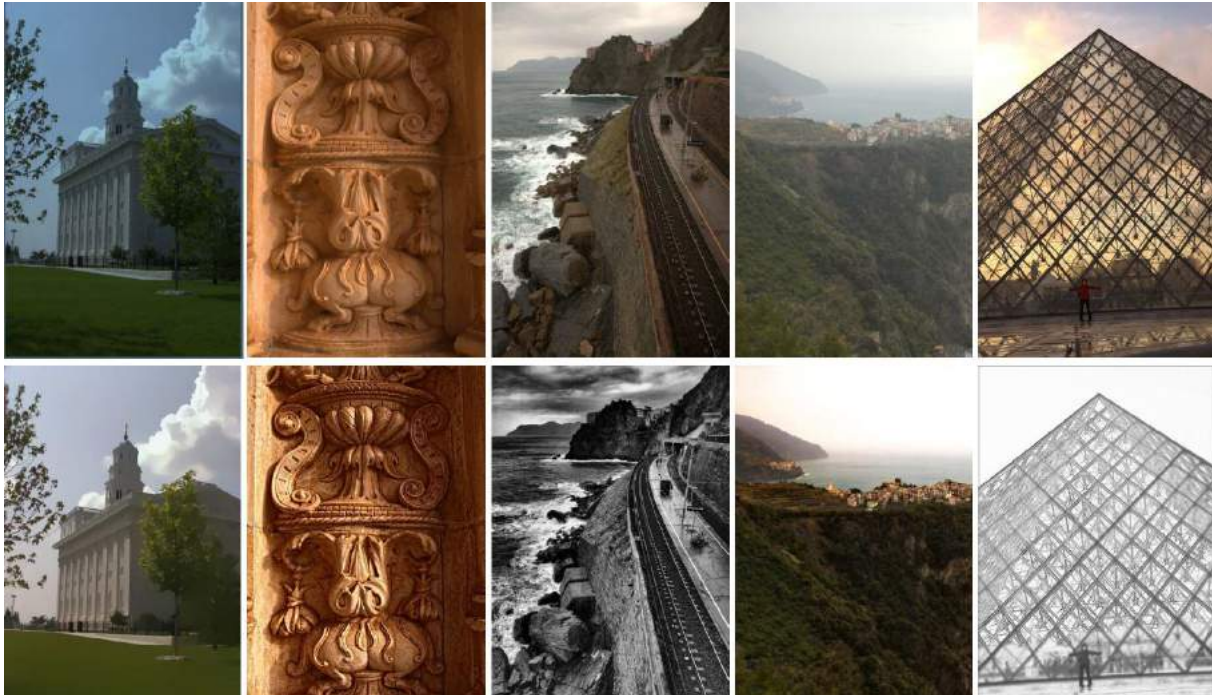


Figura 2.29: Resultados da aplicação do modelo CAN para diferentes aplicações de transformação imagem-imagem. Da esquerda para a direita: L_0 smoothing, Multiscale Tone Mapping, Estilo fotográfico, dehazing, e Pencil drawing. Fonte: Chen et al. (2017)

A Figura 2.29 apresenta os resultados obtidos na replicação de diferentes operadores empregados para transformação de imagens. Todo o processamento é realizado no espaço de cor sRGB, utilizando dados pareados. Utiliza-se como função objetivo o erro médio quadrático (MSE) e o otimizador Adam, de Kingma & Ba (2014). Cada modelo é ajustado individualmente de acordo com a aplicação. Os resultados são avaliados quantitativamente utilizando as medidas PSNR e SSIM, evidenciando que a arquitetura supera outros métodos utilizados no comparativo.

Isola et al. (2017) propõem a arquitetura de rede Pix2Pix, baseada em redes adversárias condicionais (*GANs*). Trata-se de uma arquitetura de rede de propósito geral para problemas de tradução de imagem-imagem. Tal arquitetura aprende não apenas o mapeamento da imagem de entrada para a imagem de saída, mas também uma função de perda para treinar esse mapeamento. Isso possibilita aplicar a mesma abordagem genérica a problemas que tradicionalmente exigiram formulações de perda muito diferentes. A rede geradora é do tipo U-Net, e a função objetivo combina de forma ponderada erro médio absoluto (MAE) e a saída de uma rede discriminadora. Isola et al. (2017) demonstra que a abordagem é eficaz na síntese de fotos a partir de mapas de rótulos, na reconstrução de objetos a partir de mapas de borda e na colorização de imagens, entre outras tarefas.

Zhu et al. (2017) apresentam CycleGAN, uma nova metodologia de treinamento de redes adversárias que dispensa o uso de dados pareados. Esta arquitetura tem o objetivo de tradu-

zir uma imagem de um domínio de origem para um domínio de destino, mesmo que a saída esperada não seja conhecida. Assim como a arquitetura Pix2Pix, rede geradora na CycleGAN é uma variação do modelo U-Net. Resultados qualitativos apresentados pelo autor, em várias tarefas onde os dados de treinamento emparelhados não existem, incluindo transferência de estilo, transfiguração de objeto, aprimoramento de imagem, mostram o potencial da alternativa. Comparações quantitativas em relação a métodos anteriores demonstram a superioridade desta abordagem em tarefas de tradução que envolvem alterações de cor e textura, nas quais o método geralmente é bem-sucedido.

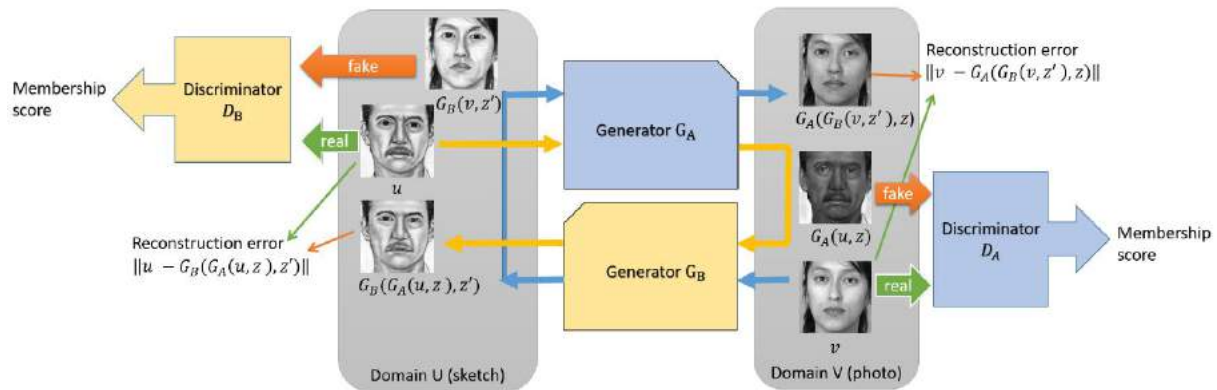


Figura 2.30: Esquemático do modelo DualGAN/ CycleGAN. Para aprender a mapear imagens entre dois domínios distintos o modelo emprega dois conjuntos gerador-discriminador operando de maneira cíclica. Fonte: Yi et al. (2017)

O modelo CycleGAN (Zhu et al., 2017) é equivalente ao modelo DualGAN (Yi et al., 2017), apresentado na Figura 2.30. Intuitivamente, para mapear imagens entre dois domínios distintos, o modelo emprega dois conjuntos gerador-discriminador operando de maneira cíclica. Isto é, uma entrada no domínio U é fornecida para o gerador G_A , que a transforma em uma imagem no domínio V , sendo posteriormente utilizada como entrada no gerador G_B , onde é transformada novamente para seu domínio original. A função objetivo avalia então o erro médio absoluto (MAE) entre a imagem de entrada e a imagem de saída. O ciclo inverso ocorre com as entradas no domínio V . Duas redes discriminadoras completam a função objetivo. Ao final do treinamento, o modelo é capaz de converter imagens entre os dois domínios.

A Figura 2.31 apresenta os resultados obtidos com o modelo CycleGAN para modificação da profundidade de campo de imagens. Nesta aplicação, o objetivo é preservar a visibilidade dos objetos mais próximos da câmera e criar um efeito de borrramento nos objetos mais distantes da cena, reproduzindo os efeitos atingidos com uma câmera DSLR. Os resultados evidenciam que o método proposto consegue mimetizar de forma convincente as características desejadas na imagem de saída.

Ignatov et al. (2018) apresentam a arquitetura *Weakly Supervised Photo Enhancer* (WESPE). Esta abordagem dispensa o uso de dados pareados, necessitando apenas de dois conjuntos de imagens caracterizadas pela transformação que o modelo deve aprender. A rede utiliza três redes de transformação de imagem (G e F), uma rede discriminadora para textura (D_t), uma



Figura 2.31: Resultados da aplicação do modelo CycleGAN para modificação da profundidade de campo em imagens. Lado a lado: entrada e imagem transformada. Fonte: Zhu et al. (2017)

rede discriminadora para (D_c) para cor, e uma rede VGG19 para perda perceptual. A Figura 2.32 apresenta uma visão geral do esquema de treinamento proposto.

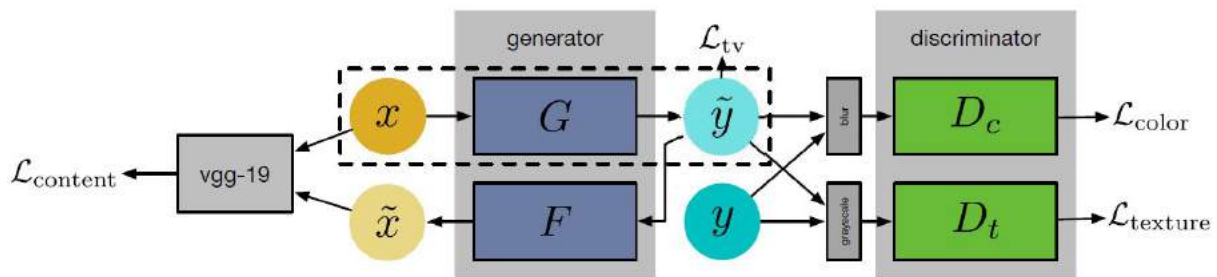


Figura 2.32: Esquemático da arquitetura WESPE. Utilizam-se três redes de transformação de imagem (G e F), uma rede discriminadora para textura (D_t), uma rede discriminadora para cor (D_c) para cor, e uma rede VGG19 para perda de conteúdo. Fonte: Ignatov et al. (2018)

Na WESPE os geradores G e F tem uma estrutura na qual todo fluxo de dados ocorre na resolução original da imagem de entrada. Os geradores são compostos por uma camada convolucional 9×9 , quatro blocos sequenciais na forma Convolução $3 \times 3 \rightarrow$ Normalização por Lote $\rightarrow$ Convolução $3 \times 3 \rightarrow$ Normalização por Lote $\rightarrow$ Camada residual, duas camadas de convolução 3×3 e uma camada convolucional 9×9 na camada de saída.

Uma importante contribuição de Ignatov et al. (2018) são as funções de perda implementadas com o objetivo de otimizar realce de imagem independente de dados pareados. Para tanto, os autores introduzem três redes secundárias treináveis, sendo uma voltada para a otimização de textura, uma para otimização de cor e a terceira, na forma de gerador, voltada a converter a imagem transformada de volta à imagem original. A entrada do modelo é comparada com a saída gerada pelo segundo utilizando a diferença de *features* do modelo VGG19.

A Figura 2.33 apresenta os resultados da aplicação do modelo WESPE na transformação de imagens obtidas através de câmeras de baixa qualidade em imagens que imitam aquelas obtidas por câmeras de alta qualidade. Os resultados quantitativos apresentados por Ignatov et al. (2018) indicam que as imagens restauradas pelo método atingem PSNR entre 20,66 e 22,01 em



Figura 2.33: Resultados da aplicação do modelo WESPE para realce de imagem. Saídas na segunda linha. Fonte: Ignatov et al. (2018)

diferentes *datasets*. Uma avaliação considerando similaridade estrutural SSIM traz resultados entre 0,92 e 0,94 nestes datasets. Um estudo conduzido com observadores humanos indica que a restauração utilizando WESPE é preferida em relação aos demais métodos comparados.

Observa-se que a transformação imagem-imagem é útil em aplicações com distintos propósitos. Em geral, os modelos de transformação compartilham aspectos estruturais e metodológicos. Esta intersecção e reaproveitamento entre diversos modelos é natural, uma vez que restauração de cor e textura, combinadas com uma imagem de alta qualidade percebida, são objetivos comuns a todas as aplicações. Apesar das diferenças de propósito, estes trabalhos se fazem relevantes por propor arquiteturas e funções objetivo que podem ser incorporados ao presente trabalho.

O problema de remoção de névoa (*dehazing*) difere do problema abordado no presente trabalho, pois se preocupa com a distorção na imagem em decorrência da distância entre o objeto na cena e a câmera. Assim sendo, busca estimar profundidade e transmissão do meio participativo para fazer o ajuste dos níveis de intensidade na imagem e restaurar a visibilidade. Cai et al. (2016) utilizam uma CNN para estimar um valor de transmissão baseado em recortes de tamanho 16×16 de imagens coloridas (RGB), estimando transmissão constante para todos píxeis e todos canais de cor do recorte processado, o que pode levar ao surgimento de artefatos nas bordas e na coloração da imagem. Para restaurar uma imagem utilizando este método, os autores propõe a aplicação do *guided filter* (He et al., 2013) na saída da CNN utilizada. Gonçalves et al. (2018) propõem uma arquitetura de rede melhorada com o mesmo propósito, incorporando *skip connections*, e a concatenação com a imagem de entrada antes da última camada. Gonçalves et al. (2018) utiliza ainda o *guided filter* como uma parte da arquitetura de CNN, fazendo com que o impacto deste filtro para manter a estrutura seja considerado no ajuste de pesos durante o treinamento.

O problema de super-resolução difere da pesquisa proposta no sentido de que busca trabalhar a escala da imagem em termos espaciais. Ledig et al. (2017) apresentam *Super-Resolution Generative Adversarial Network* (SRGAN), um modelo de rede neural deliberadamente inspi-

rado no modelo ResNet (He et al., 2016a), com o propósito de escalar uma imagem em até quatro vezes, enquanto busca manter um aspecto realístico. Para tanto, Ledig et al. (2017) propõe uma nova função objetivo, que leva em consideração i) o erro médio quadrático (MSE) píxel-a-píxel entre a imagem referência e a saída da rede neural; ii) o erro perceptual (*perceptual loss*), conforme proposto por Johnson et al. (2016); e uma rede neural adversária conforme introduzida por Bengio & Courville (2016).

Também voltado para o problema de super-resolução, Lai et al. (2017) introduzem *Laplacian Pyramid Super-Resolution Network* (LapSRN), um modelo de super-resolução progressivo, onde o refinamento é feito em etapas sucessivas através de um *framework* de pirâmide Laplaciana. Em cada nível da pirâmide, o modelo recebe dados em baixa qualidade e busca retornar uma imagem refinada. Para aumento da dimensionalidade, os autores fazem uso de convolução transposta. O uso da pirâmide laplaciana dentro da rede permite que a esta obtenha melhor definição em regiões de alta frequência na imagem. Também neste sentido, os autores argumentam que a substituição do erro médio quadrático como função objetivo, conhecido por produzir imagens suavizadas, pela função de penalidade Charbonnier (uma versão diferenciável do Erro Médio Absoluto). Na arquitetura da LapSRN, cada nível de super-resolução é penalizado de acordo com a resolução da imagem.

A incorporação de técnicas, métodos e arquiteturas apresentadas nos modelos de transformação imagem-imagem elencados serviram de base para a o desenvolvimento do método proposto, contribuindo para o aprimoramento e aumento de qualidade da restauração obtida. O modelo proposto e discutido na Seção 4 incorpora, a exemplo de vários dos modelos relacionados, uma arquitetura do tipo U-Net (Ronneberger et al., 2015), camadas de convolução dilatada (Chen et al., 2017). Avaliou-se também a viabilidade de uma função objetivo composta por erro píxel-a-píxel e perda adversária utilizando PatchGAN (Isola et al., 2017; Radford et al., 2016). Da mesma forma, considera-se que avanços recentes obtidos em áreas relacionadas podem ser utilizadas para minimizar as principais limitações do modelo proposto no presente trabalho.

2.4 Consolidação dos Trabalhos Relacionados

Apresenta-se um resumo consolidando os principais aspectos dos métodos convencionais de processamento de imagens e modelos baseados em aprendizagem profunda que são utilizados no restante deste trabalho. Para cada entrada, apresenta-se uma síntese das características. Destaca-se que, a nível de aplicação, a grande maioria dos trabalhos relacionados tem foco na restauração de imagens noturnas ou obtidas em condições de pouca luz. As aplicações em imagens sobre-expostas são menos frequentes.

- U-Net ou FCN com *skip connections*: Honig & Werman (2018); Cai et al. (2018); Lv et al. (2021); Guo et al. (2019); Xu et al. (2020); Xiong et al. (2020); Kwon et al. (2020); Zhang et al. (2019b); Jiang et al. (2019); Chen et al. (2018); Wang et al. (2018b); Lv

- et al. (2018); Afifi et al. (2020); Zhang et al. (2020a); Guo et al. (2020); Li et al. (2021); Atoum et al. (2020); Wang et al. (2019a);
- Utilizam GAN para ajuste: Honig & Werman (2018); Xiong et al. (2020); Jiang et al. (2019); Wang et al. (2019a); Hu et al. (2018); Afifi et al. (2020);
 - Disponibilizam informações sobre tamanho do modelo: Lv et al. (2021); Guo et al. (2020); Li et al. (2021); Wei et al. (2018); Lv et al. (2018); Afifi et al. (2020);
 - Embasados na teoria Retinex: Guo et al. (2019); Liang et al. (2020); Zhang et al. (2019b); Wang et al. (2019a); Wei et al. (2018); Hao et al. (2021);
 - São desenhados para trabalhar com imagens RGB: Honig & Werman (2018); Cai et al. (2018); Lv et al. (2021); Guo et al. (2019); Zhang et al. (2020a); Xu et al. (2020); Xiong et al. (2020); Kwon et al. (2020); Liang et al. (2020); Zhang et al. (2019b); Guo et al. (2020); Li et al. (2021); Atoum et al. (2020); Jiang et al. (2019); Wang et al. (2019a); Wei et al. (2018); Wang et al. (2018b); Lv et al. (2018); Afifi et al. (2020); Yang et al. (2020);
 - São desenhados para operar com dados brutos (RAW): Xu et al. (2020); Wang et al. (2019c); Chen et al. (2018); Hu et al. (2018);
 - São desenhados para condições de iluminação insuficiente: Lv et al. (2021); Guo et al. (2019); Zhang et al. (2020a); Xu et al. (2020); Xiong et al. (2020); Kwon et al. (2020); Liang et al. (2020); Zhang et al. (2019b); Guo et al. (2020); Li et al. (2021); Atoum et al. (2020); Jiang et al. (2019); Wang et al. (2019a); Wei et al. (2018); Chen et al. (2018); Lv et al. (2018); Yang et al. (2020);
 - São desenhados para condições de sobre-exposição: Cai et al. (2018); Honig & Werman (2018); Afifi et al. (2020);
 - Não necessitam de dados pareados para ajuste: Zhang et al. (2020a); Guo et al. (2020); Li et al. (2021); Jiang et al. (2019);
 - São avaliados utilizando dados pareados (PSNR, SSIM, FSIM, VIFP): Lv et al. (2021); Guo et al. (2019); Zhang et al. (2020a); Xu et al. (2020); Xiong et al. (2020); Liang et al. (2020); Wang et al. (2019c); Zhang et al. (2019b); Guo et al. (2020); Li et al. (2021); Atoum et al. (2020); Wang et al. (2019a); Chen et al. (2018); Lv et al. (2018); Afifi et al. (2020); Yang et al. (2020);

Com relação às arquiteturas de rede empregadas, destaca-se a popularidade dos modelos baseados em U-Net, seguidos por outros modelos totalmente convolucionais (FCN) que empregam algum tipo de atalho (*Skip Connections*) entre as camadas do início e do final da rede. Apenas 6 modelos baseados em aprendizagem profunda utilizam uma estrutura geradora-adversária (GAN) para treinamento. Apenas 6 autores disponibilizam informações sobre a quantidade de parâmetros do modelo.

3 REQUISITOS E DEFINIÇÕES DE PROJETO

Na seção 3.1 deste capítulo discutem-se os requisitos, definições e balizadores para o ajuste do modelo. Na seção 3.2 apresenta-se uma visão geral das principais métricas aplicadas para avaliação de qualidade de imagens. Na Seção 3.3 apresentam-se os quatro datasets utilizados para ajustar o modelo e avaliar os resultados da restauração. Na seção 3.4 são apresentados os procedimentos para validação do modelo proposto a nível de aplicação. Por fim, na Seção 3.5 apresenta-se a técnica de visualização das ativações internas do modelo, utilizada para compreender e explicar o funcionamento do modelo.

Inicialmente, busca-se conceituar os principais aspectos e balizadores para o desenvolvimento do trabalho. Para tanto, a Seção 3.3 concentra-se na definição da arquitetura da rede para que seja adequada ao problema que se pretende mitigar. Nesta etapa é importante a determinação da representação dos dados, da estrutura da rede, da função objetivo (aquilo que se deseja minimizar) e a determinação inicial dos hiperparâmetros. Da mesma forma, apresentam-se as métricas e conjuntos de dados utilizados para ajuste e validação do modelo e os procedimentos para validação do modelo proposto.

A avaliação preliminar das métricas de qualidade e medidas de similaridade entre imagens compõe, juntamente com a definição da arquitetura e aprendizagem da rede, um importante instrumento para a pesquisa. A importância destas é manifestada em dois aspectos. No sentido tradicional, as métricas e medidas são uma forma objetiva, direta e não ambígua de mensurar a qualidade do modelo proposto e estabelecer um comparativo com os demais estados-da-técnica. Não menos importante, conhecidas as características de cada uma das métricas de qualidade e medidas similaridade, pode-se incorporá-las ao procedimento de treinamento da rede, de forma que propriedades específicas de estrutura e cor da imagem sejam restauradas de forma adequada. Por fim, entender as métricas permite ainda a seleção apropriada das técnicas estatísticas.

Modelos de redes neurais artificiais requerem *datasets* amplos e diversos. Uma vez que a obtenção destes dados é uma tarefa onerosa e passível de interpretações diversas, apresentam-se na Seção 3.3 os principais conjuntos de dados reais e sintéticos utilizados no desenvolvimento do trabalho. Não obstante, os resultados obtidos nesta pesquisa são constrictos aos tipos de dados, cenários, equipamentos e técnicas de processamento de imagens aplicadas em cada um dos *datasets* utilizados na avaliação do modelo.

Para além da avaliação a nível de métricas e medidas de similaridade de imagens, faz-se também uma validação dos resultados obtidos pelo modelo de restauração de imagens em aplicações comuns de visão computacional. A validação em nível de aplicação é relevante no sentido de contribuir para a completude dos objetivos. Na Seção 3.4 descrevem-se os procedimentos, conjuntos de dados, modelos de reconhecimento e segmentação utilizados para verificar o impacto do modelo de restauração proposto em atividades importantes para a robótica e automação baseada em visão.

Por fim, verificada a natureza caixa-preta das redes neurais convolucionais, faz-se relevante

a utilização de técnicas para visualização das ativações internas do modelo. Estas técnicas são concebidas para explicar as redes neurais a partir de seus resultados e dos parâmetros aprendidos. A visualização permite verificar se a rede está utilizando as partes corretas da imagem para realizar a restauração corrigir possíveis equívocos embutidos.

3.1 Requisitos da Arquitetura de Rede

A arquitetura adequada para uma rede neural artificial deve refletir os dados de entrada e de saída, estando intrinsecamente ligada com o uso esperado do modelo. As restrições de espaço e performance também se fazem relevantes na definição da arquitetura de rede, otimizadores, função objetivo a ser minimizada e nos procedimentos empregados no ajuste do modelo. Nesta seção apresentam-se as definições iniciais e balizadores considerados na elaboração do modelo de restauração de imagens proposto.

Alinhado com os objetivos da pesquisa, tem-se como requisitos que o modelo proposto precisa apresentar desempenho satisfatório em imagens de média e alta-resolução (maiores que 300×300 píxeis) no espaço de cor sRGB, 8 bits por canal e que já tenham passado por algum tipo de compressão. O modelo deve ser capaz de restaurar imagens mesmo que estas tenham passado por compressão com perdas, condição que é comum à maioria das aplicações baseadas em computação visual. O modelo deve ser capaz de fazer a inferência de vários quadros por segundo, sendo passível de aplicação em sistemas de visão computacional baseados em vídeo que requerem processamento em tempo real.

Para o propósito desta pesquisa, o objetivo da rede é sempre minimizar as medidas de erro ou dissimilaridade. O algoritmo de *backpropagation* pode ser considerado a pedra fundamental das redes neurais artificiais e aprendizagem profunda (Bengio & Courville, 2016). O *backpropagation* é implementado em duas etapas. Na primeira, propagação ou *forward pass*, entradas são passadas através da rede e as previsões de saída obtidas. Na segunda, retropropagação ou *backward pass*, calcula-se o gradiente para aplicar recursivamente a regra da cadeia para atualizar os pesos em nossa rede. Este processo é, em geral, repetido até que um determinado valor de erro seja atingido ou um número fixo de iterações tenham sido executadas.

Os pesos das redes neurais artificiais são ajustados iterativamente durante o treinamento do modelo. Ao método de atualização de pesos dá-se o nome otimizador. Utiliza-se nesta pesquisa uma estratégia de otimização por mini-lotes com o otimizador Adam (Kingma & Ba, 2014). Nesta estratégia, um sub-conjunto de imagens do *dataset* é extraída por amostragem aleatória simples com substituição. O estudo das técnicas de amostragem para aprendizagem profunda tem sido amplamente explorado (Qian & Klabjan, 2020; Li et al., 2014; Perrone et al., 2019). A presente pesquisa limita-se a aplicar amostragem simples em virtude das propriedades conhecidas nos dados utilizados para o ajuste e da viabilidade de execução.

A utilização de mini-lotes durante o treinamento do modelo é apropriada considerando a necessidade de processamento de imagens em média e alta resolução. Modelos convolucionais e, especialmente, modelos que fazem uso de *skip-connections*, requerem uma quantidade signi-

ficativa de memória para o processamento de cada imagem de entrada, inviabilizando o ajuste do modelo sobre o conjunto de dados de treinamento como um todo em uma única iteração.

Na tarefa de predição densa píxel-a-píxel de imagens, costuma-se empregar modelos de redes neurais convolucionais. Em imagens, a informação é representada pela espacialidade. Considerando que o espaço desempenha o papel mais importante, modelos convolucionais se mostram mais adequados (em contraposição à modelos de rede densos ou recorrentes). Em decorrência, modelos para classificação, segmentação, geração, *inpainting* e restauração de imagens se beneficiam do uso de estruturas convolucionais. Redes neurais convolucionais se tornaram o padrão para implementação de modelos cujo dado de entrada é imagem.

Propõe-se uma arquitetura de rede neural artificial destinada à restauração de imagens sRGB obtidas a partir de câmeras digitais convencionais, impactadas por exposição inapropriada. Objetiva-se que o modelo seja capaz de preservar as características mais representativas para a visão computacional como definição, contraste, nitidez e cor. Deseja-se um modelo ponta-a-ponta eficiente, que minimize o uso de memória e capacidade de processamento requerido, sem sacrificar de maneira significativa a qualidade dos resultados produzidos.

3.2 Indicadores de Qualidade de Imagem

Os avanços oferecidos pelas Redes Neuras Convolucionais (CNN) aplicadas para transformação e melhoria de imagem trouxeram à tona uma maior preocupação com a forma de mensurar a qualidade de uma imagem. Este tema tem ganho significativa relevância na comunidade acadêmica, uma vez que diversos autores argumentam que existe inconsistência entre a qualidade percebida pelo olho humano e os valores fornecidos por métricas e medidas de qualidade. Entre outros, neste tópico destacam-se os trabalhos de Johnson et al. (2016), Blau & Michaeli (2018), Blau et al. (2018), Egiazarian et al. (2018) e Prashnani et al. (2018).

De forma geral, os indicadores de qualidade de imagem podem ser categorizados em:

- Medidas e métricas de qualidade baseadas em referência: neste caso, assume-se a existência de dados pareados (uma imagem referência e uma imagem sobre a qual se busca inferir a qualidade);
- Medidas e métricas cegas: neste caso, assume-se a existência de um modelo capaz de discernir entre uma imagem boa e uma imagem ruim.

3.2.1 Métricas Cegas

Em um estudo feito no âmbito do presente trabalho, buscou-se identificar a viabilidade da utilização de métricas cegas para avaliação de imagens obtidas sob distintas condições de iluminação e exposição. Os resultados completos desta avaliação foram publicados em Steffens et al. (2017). Esta avaliação levou em consideração conjuntos de imagens de resolução 320×240 px, obtidos sob condições controladas de iluminação e exposição. O estudo avaliou as seguintes medidas cegas: Metric-Q (Zhu & Milanfar, 2010); PCA-NE (Pyatykh et al., 2013);

BRISQUE (Mittal et al., 2012) e QAC (Xue et al., 2013). Metric-Q e PCA-NE são voltadas para a avaliação de ruído. BRISQUE e QAC tem escopo de avaliação mais aberto, baseando-se em estatísticas da imagem.

Sobretudo, verificou-se que as métricas cegas avaliadas apresentam alguma relação com a qualidade percebida da imagem. Apesar disso, considerou-se que oferecem avaliação insuficiente sobre a qualidade percebida, sendo suscetíveis à mudanças de escala. Identificou-se ainda que a indicação de qualidade oferecida tanto pela Metric-Q quanto BRISQUE tem relação direta com o brilho da imagem (ou seja, privilegiam luminância em detrimento de textura). Desta forma, considera-se que as métricas cegas não são uma opção viável para avaliar a restauração de imagens subexpostas ou sobre-expostas. Ressalta-se ainda que os resultados observados na aplicação dessas medidas mostraram-se inconsistentes, avaliando imagens sub e sobre-expostas como apresentando qualidade superior às obtidas com exposição inadequada.

3.2.2 Métricas Baseadas em Referência

Com relação às métricas baseadas em referência, diversos autores têm apontado para a inconsistência entre a qualidade medida e a qualidade percebida. Egiazarian et al. (2018) estuda o problema da avaliação de qualidade visual de imagens que passaram por remoção de ruído, com ênfase em imagens com baixo contraste e textura semelhante à de ruído. Nestas condições, a remoção de ruído geralmente resulta também em perda ou suavização dos detalhes da imagem. Para o estudo, Egiazarian et al. (2018) utilizaram 75 imagens sem ruído e 300 imagens com ruído Gaussiano removido por meio do algoritmo BM3D (Dabov et al., 2007). Os autores identificam que o coeficiente de correlação de Spearman entre PSNR e MOS é próximo de zero. Os autores avaliam, entre outras métricas, SSIM (Wang et al., 2004), FSIM (Zhang et al., 2011), GMSD (Xue et al., 2014) e verificam que nenhuma métrica apresenta correlação forte ou muito forte com a avaliação humana.

Blau & Michaeli (2018) busca provar matematicamente a contraposição entre qualidade percebida e distorção para aplicações de super-resolução, ou seja, algoritmos que buscam gerar uma imagem em alta resolução a partir de uma única imagem em baixa resolução. Para tanto, Blau & Michaeli (2018) compara os resultados fornecidos pela métrica cega proposta em Ma et al. (2017), com um conjunto de medidas de distorção baseadas em referência. Baseados neste comparativo, Blau & Michaeli (2018) demonstram que, ao contrário do que se poderia esperar, para os trabalhos recentes no tópico existe uma relação inversa entre a qualidade percebida e a medida de distorção entre uma imagem reescalada e sua referência.

Em virtude das inconsistências apresentadas entre a qualidade percebida por avaliadores humanos e das avaliações fornecidas por métricas e medidas de qualidade com referência, Blau & Michaeli (2018) argumenta que as alternativas recentes utilizando redes adversárias (GANs) fornecem uma maneira mais adequada de avaliar o limite de distorção da percepção. Os resultados do estudo dão suporte teórico e evidência empírica das vantagens das GANs na restauração de imagens, evidenciando a contribuição de melhorias introduzidas por redes com SRGAN

(Ledig et al., 2017), SICE-CNN (Lai et al., 2017), Context-Encoder (Pathak et al., 2016).

Apesar da controvérsia em relação ao tema, métricas e medidas de qualidade de imagem baseadas em referência ainda são amplamente utilizadas, especialmente por oferecerem uma forma determinística de comparar os resultados do processamento com a saída esperada. Dentre as métricas baseadas em referência destacam-se:

- Relação Sinal-Ruído de Pico – PSNR;
- Erro médio quadrático – MSE;
- Erro médio absoluto – MAE;
- *Structural SIMilarity* – SSIM de Wang et al. (2004);
- Canny Intersecção sobre União – (Canny, 1986);
- Diferença de histogramas (256 *bins*);
- *Gradient Magnitude Similarity Deviation* – GMSD de Xue et al. (2014);
- *Visual Information Fidelity* – VIFP de Sheikh & Bovik (2004);
- *Feature Similarity Index* em escala de cinza – FSIM de Zhang et al. (2011);
- *Feature Similarity Index* com cromaticidade – FSIMc de Zhang et al. (2011);
- *Polar Edge Coherence* – RECO de Baroncini et al. (2009);
- Delta E (CIE 2000) de Sharma et al. (2005);

A relação sinal-ruído de pico (PSNR), calculada em função do erro quadrático médio (MSE), tem sido tradicionalmente empregada como uma métrica de qualidade consistente para avaliação de imagens (Huynh-Thu & Ghanbari, 2008). A métrica define a relação entre a máxima energia de um sinal e o ruído que afeta sua representação fidedigna. Embora sua aplicação seja mais adequada para a avaliação de algoritmos de compressão, a PSNR é com frequência empregada para verificação da qualidade resultante de métodos voltados para restauração (Gharbi et al., 2017; Ren et al., 2016; Cai et al., 2016), super-resolução (Wang et al., 2018a; Ledig et al., 2017), remoção de ruído e demais tarefas que se beneficiam da comparação entre uma imagem distorcida e sua referência.

O erro médio quadrático pixel-a-pixel (MSE) e o erro médio absoluto (MAE) são medidas amplamente utilizadas para avaliar modelos de regressão. Ambas as medidas fornecem um indicativo de fácil interpretação sobre a qualidade dos resultados. MSE e MAE diferem entre si na forma como tratam a distribuição do erro. MSE penaliza mais os erros que são mais distantes do valor esperado. Já MAE pondera todos os valores de erro de forma equivalente. Essas medidas são adequadas para mensurar a qualidade de restauração em partes de baixa frequência da imagem (regiões de intensidade homogênea).

A Similaridade estrutural (SSIM), de Wang et al. (2004), é outra métrica frequentemente empregada para avaliar o nível de distorção entre uma imagem e sua referência. De forma distinta ao que ocorre com MSE e MAE, a SSIM considera variância e covariância entre valores de intensidade para blocos 3×3 das imagens. Esta medida combina estimadores de luminância, contraste e estrutura. O valor da métrica é limitado ao intervalo $[0; 1]$, onde 1 indica que as imagens são idênticas.

Como forma de avaliar o impacto da restauração, considera-se para este trabalho duas aplicações de detecção de borda. Segundo Gonzalez & Woods (2009), os pixels de borda ocorrem em pontos onde a intensidade de uma função imagem muda abruptamente, e as bordas (ou segmentos de borda) são conjuntos de pixels de borda conexos. Para a identificação dos pixels de borda aplica-se o operador de gradiente de Sobel. Para identificação de bordas aplica-se o método do Canny (1986). Os detectores de borda são métodos de processamento de imagem local desenvolvidos para detectar os pixels da borda. Utiliza-se como medida de similaridade a fração da união pela intersecção dos pixels destacados. Um valor mais próximo de 1 indica que as bordas da imagem restaurada e da imagem referência apresentam maior intersecção. Estas medidas são interessantes por oferecerem uma avaliação da restauração proposta em uma aplicação muito comum dentro da visão computacional.

Também com maior ênfase em similaridade de bordas e gradientes, faz-se uso das medidas GMSD (Xue et al., 2014) e RECO (Baroncini et al., 2009). A GMSD trabalha com o conceito de similaridade de magnitude de gradiente pixel-a-pixel, recebendo uma imagem distorcida e uma imagem referência como entrada. A medida leva ainda em consideração o desvio padrão do gradiente. Valores menores de GMSD indicam que as imagens são mais similares, no sentido de que as regiões com transição forte de nível de intensidade foram preservadas.

As métricas VIFP (Sheikh & Bovik, 2004), FSIM e FSIMc (Zhang et al., 2011) são utilizadas para uma comparação entre duas imagens em um espaço de *features*. FSIM opera sobre imagens em escala de cinza explorando o fato de que o sistema visual humano responde a uma imagem principalmente de acordo com suas características primitivas como texturas e bordas, especificamente a congruência de fase e a magnitude do gradiente. FSIMc estende a proposta da métrica FSIM agregando informações de cromaticidade no espaço de cores YIQ (com modulação de amplitude em quadratura).

Utiliza-se também no comparativo a métrica CIEDE 2000, de Sharma et al. (2005). CIEDE 2000 é uma métrica voltada especialmente para a avaliação da diferença de cor entre duas imagens. A medida pode produzir valores no intervalo $[0; 100]$, onde valores mais baixos indicam uma melhor preservação das cores. A Tabela 2 sintetiza as principais características das medidas de avaliação de similaridade entre a imagem restaurada e a imagem referência, servindo como um guia de interpretação dos resultados obtidos.

Combinadas, as métricas e medidas permitem aferir a qualidade das restaurações em termos de definição, contraste, nitidez e correção de cor. Métricas como MSE, PSNR, MAE e CIEDE 2000 são capazes de avaliar cor e a correção global da imagem. Medidas como intersecção de bordas com Canny e Sobel, GMSD conseguem aferir a qualidade do modelo de restauração em

Tabela 2: Interpretação das medidas de qualidade de imagem aplicadas na avaliação.

Medida	Referência	Mínimo	Máximo	Interpretação
PSNR	N.A.	0	∞	Maior é melhor
MSE	N.A.	0	1	Menor é melhor
MAE	N.A.	0	1	Menor é melhor
SSIM	Wang et al. (2004)	0	1	Maior é melhor
Sobel IoU	N.A.	0	1	Maior é melhor
Canny IoU	N.A.	0	1	Maior é melhor
Dif. Hist.	N.A.	0	∞	Menor é melhor
GMSD	Xue et al. (2014)	0	∞	Menor é melhor
VIFP	Sheikh & Bovik (2004)	0	∞	Mais próximo de 1 é melhor
FSIM	Zhang et al. (2011)	0	1	Maior é melhor
FSIMc	Zhang et al. (2011)	0	1	Maior é melhor
RECO	Baroncini et al. (2009)	- ∞	∞	Mais próximo de 1 é melhor
CIEDE 2000	Sharma et al. (2005)	0	100	Menor é melhor

termos de contraste, observado nas transições de intensidade entre partes da cena que diferem por sua luminância ou densidade óptica (contraste de borda). Quanto mais contrastada for a fronteira entre uma área escura e outra mais clara, maior será a acuidade e com ela a nitidez percebida na imagem (Gonzalez & Woods, 2009). SSIM combina correção de cor e contraste, uma vez que é calculada em pequenos blocos da imagem. Por fim, FSIM e VIFP complementam esta avaliação ao fazer uma validação ao utilizar extração de características como uma etapa da avaliação de similaridade entre imagens.

3.3 Datasets

Entre as principais limitações para o desenvolvimento de abordagens baseadas em redes neurais convolucionais está a necessidade de *datasets* amplos e com grande variabilidade. Para o desenvolvimento deste trabalho, faz-se o uso de quatro distintos datasets (dois reais e dois sintéticos). Desta forma, é possível avaliar a aplicabilidade do modelo tanto para condições de sub e sobre-exposição simulados, quanto para condições reais, onde existe a interferência do meio, da óptica, da eletrônica utilizada para a aquisição e dos algoritmos de compressão.

3.3.1 A6300 *Multi-Exposure Dataset* (real)

O dataset *A6300 Multi-Exposure Dataset*, produzido no decorrer do presente trabalho e inicialmente apresentado em Steffens et al. (2018a), é composto por 116 cenas. Cada cena é representada por um conjunto de 4 imagens: uma imagem apropriadamente exposta utilizando uma única fotografia, uma imagem subexposta, uma imagem sobre-exposta e uma composição das anteriores utilizando o método de *Tone Mapping* de Mertens et al. (2007). Todas as imagens são adquiridas utilizando uma câmera digital Sony $\alpha 6300$. As imagens sub e sobre-expostas são obtidas a partir de compensação de exposição com prioridade de abertura, com valores de exposição (EV) no intervalo EV -0.7 – EV +0.7.

Todas as imagens são arquivadas utilizando o compressão JPEG com perdas. A compressão JPEG cria desafios adicionais para o processo de restauração de imagens sub e sobre expostas,



Figura 3.1: Exemplo de conjunto de imagens que compõe o *dataset* A6300.

destacando-se: i) perda de nitidez e definição próximo a regiões de alto contraste em decorrência da aproximação por cossenos; ii) artefatos de blocos decorrentes do modelo de processamento em que cada bloco de 8×8 é processado separadamente, resultando em artefatos visíveis, especialmente quando as taxas de compressão altas são utilizadas; iii) perda de detalhes de cor decorrente da compressão dos canais de cromaticidade. Apesar destas limitações da compressão JPEG, este formato ainda é o mais utilizado para fins de armazenamento e transmissão de imagens, sendo implementado como alternativa padrão na maior parte das câmeras comercialmente disponíveis.

O dataset gerado contempla imagens de cenas em ambientes fechados (móveis, eletrodomésticos, pessoas, arquiteturas feitas pelo homem e iluminação artificial) e cenas ao ar livre (árvores, nuvens, gramados, veículos, edificações e iluminação natural). Para fins de utilização como dados pareados, todas as imagens neste dataset foram selecionadas manualmente, sendo removidas cenas que apresentam variação visível do conteúdo em função da dinâmica da cena em si. Incluem-se nestes casos cenas com veículos em movimento rápido, monitores e televisores, borrões e deformações, e imagens fora de foco. A Figura 3.1 apresenta uma das cenas que compõem este *dataset*.

3.3.2 Cai2018 *Multi-Exposure Dataset* (real)

Inicialmente apresentado por Cai et al. (2018), este *dataset* é composto por 589 cenas. Para cada cena são disponibilizados de 3 a 18 fotos em baixo contraste com diferentes níveis de exposição e uma referência gerada através da composição de imagens usando o método de Mertens et al. (2007). Para adquirir as imagens sub e sobre-expostas, os autores utilizam compensação de exposição $EV \pm \{0.5, 0.7, 1.0, 2.0, 3.0\}$. Assim como Steffens et al. (2018a) este dataset contém imagens de ambiente interior e imagens obtidas ao ar livre, todas comprimidas utilizando o formato JPEG. Cai et al. (2018) utilizam sete modelos de câmera, incluindo Sony α 7RII, Sony NEX-5N, Canon EOS-5D Mark II, Canon EOS-750D, Nikon D810, Nikon D7100 e iPhone 6s.

A Figura 3.2 apresenta uma das cenas que compõe o *dataset* de Cai et al. (2018). Destaque-se a existência de perda severa de informação em decorrência de saturação causada por exposição inapropriada. Os efeitos de sub e sobre-exposição resultam em blocos grandes em que pouca ou nenhuma informação é preservada. Isso se reflete em perda de cor, textura e gradiente de maneira geral, além de modificar a representação da imagem no domínio da frequência.

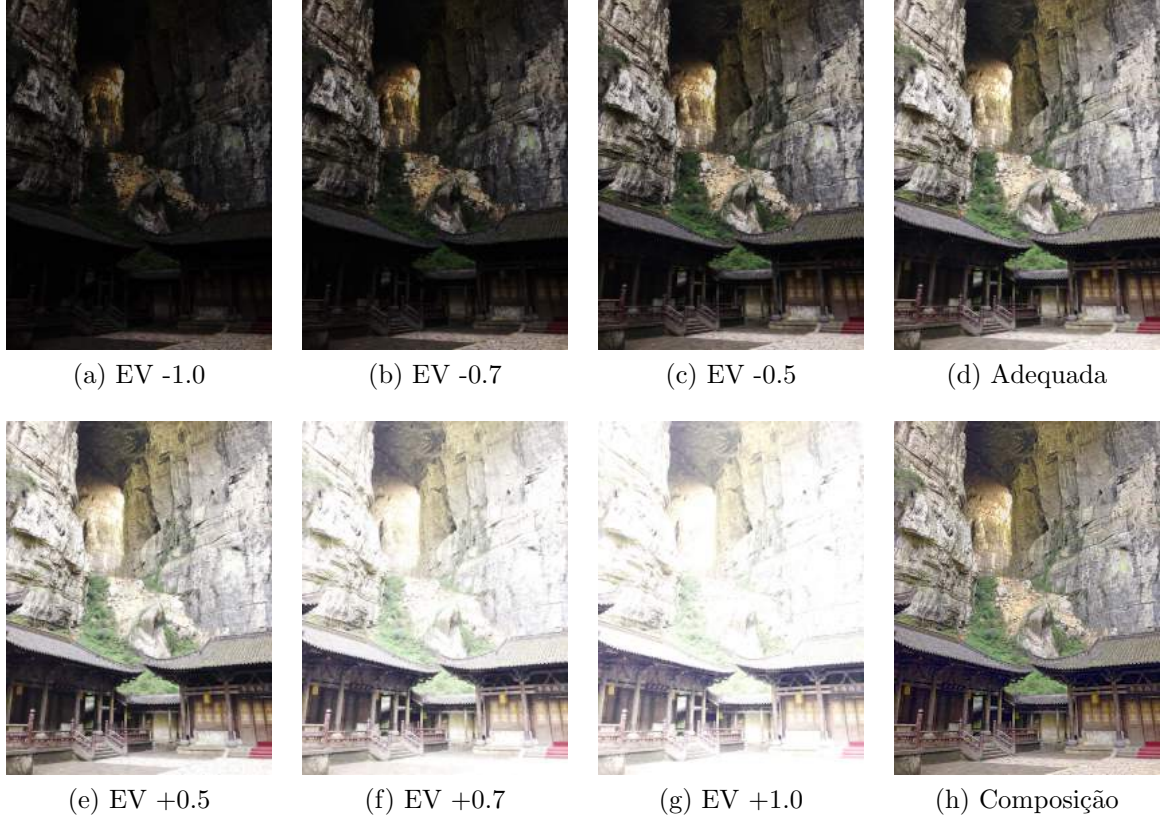


Figura 3.2: Exemplo de conjunto de imagens que compõe o *dataset* Cai2018.

3.3.3 Dataset Multi-Exposição baseado em FiveK (sintético)

O *dataset* MIT-Adobe FiveK (Bychkovsky et al., 2011) é composto por 5000 cenas obtidas através de câmera SLR (*single-lens reflex*) por vários fotógrafos. Estas imagens são disponibilizadas no formato não comprimido DNG (Adobe *Digital Negative*), portanto, não apresentam qualquer artefato decorrente de compressão. Toda informação registrada pelo sensor da câmera é preservada. No entanto, este dataset não contém imagens pareadas e requer pré-processamento para que possa ser utilizado para treinamento e avaliação do modelo de restauração de imagens imprópriamente expostas.

Para treinamento, os dados são convertidos do formato DNG para uma representação sRGB padrão. A imagem no formato sRGB é utilizada como referência. A partir desta imagem, geram-se as imagens danificadas, truncando os valores de acordo com um percentil predeterminado. Esta operação é feita através da Equação 3.1, em que I é a imagem referência, C é a imagem truncada resultante, P_{LT} e P_{HT} são os valores de percentil que definem respectivamente os limiares inferior e superior para o truncamento. Para os experimentos conduzidos, trabalha-se com valores de LT e HT no conjunto $\{25,15,10,5\}$ aplicados arbitrariamente.

$$C_{ij} = \begin{cases} P_{LT}, & I_{ij} \leq P_{LT} \\ I_{ij}, & P_{LT} \leq I_{ij} \leq P_{HT} \\ P_{HT}, & I_{ij} \geq P_{HT} \end{cases} \quad (3.1)$$

A operação de truncamento, para reprodução dos efeitos de *hard clipping*, é então seguida

por normalização min-max (Equação 3.2) de forma a estender os valores da imagem truncada para todo intervalo de representação.

$$\hat{I} = \frac{C_{ij} - \min(C)}{\max(C) - \min(C)}. \quad (3.2)$$

Nota-se que esta abordagem apresenta algumas limitações com relação à utilização de dados reais. Apesar de reproduzir com alguma fidelidade visual os efeitos de saturação e subexposição, esta abordagem é incapaz, por exemplo, de reproduzir o efeito de *blooming*. Este efeito é percebido em câmeras reais, se apresentando na imagem capturada como uma distorção que emana feixes de luz que se estendem das bordas de áreas claras em uma imagem, sendo decorrente do arranjo óptico e do vazamento de carga entre elementos da matriz do sensor (Hasinoff, 2014). Apesar das limitações, acredita-se que a transformação é capaz de reproduzir de forma aceitável os principais efeitos adversos da exposição inapropriada, conforme pode ser visualizado na Figura 3.3.

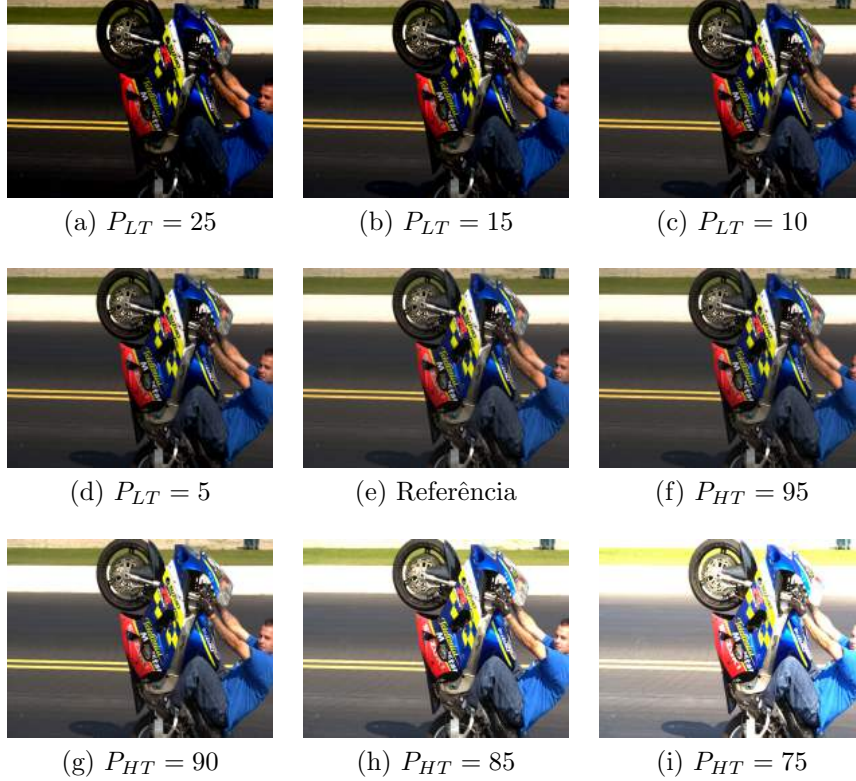


Figura 3.3: Exemplo de conjunto de imagens sintéticas geradas a partir do *dataset* FiveK, apresentando o impacto causado pelas operações de truncamento e normalização por percentil.

3.3.4 Dataset Multi-Exposição baseado HDR+ *Burst Photography Dataset* (sintético)

Inicialmente apresentado por Hasinoff et al. (2016), este *dataset* compreende 3640 cenas. Para cada cena são disponibilizadas sequências de imagens em diferentes exposições obtidas por câmeras de *smartphones*. Para cada cena são disponibilizadas entre 2 e 10 imagens, totalizando

28461 imagens no *dataset* como um todo. Para cada cena, é disponibilizado também o resultado do alinhamento das imagens (necessário para compensar movimento da câmera em relação à cena) e um resultado final da composição de imagens obtido através de método proposto pelos autores. No presente trabalho, em particular, utiliza-se a imagem JPEG gerada pelo método, com saturação e subexposição sintéticos através do truncamento dos valores de intensidade, de maneira análoga àquela descrita na Seção 3.3.4. Assim como para os dados do FiveK, trabalha-se com a danificação de até 25% dos pixels da imagem referência, gerando 8 imagens danificadas pra cada imagem referência.

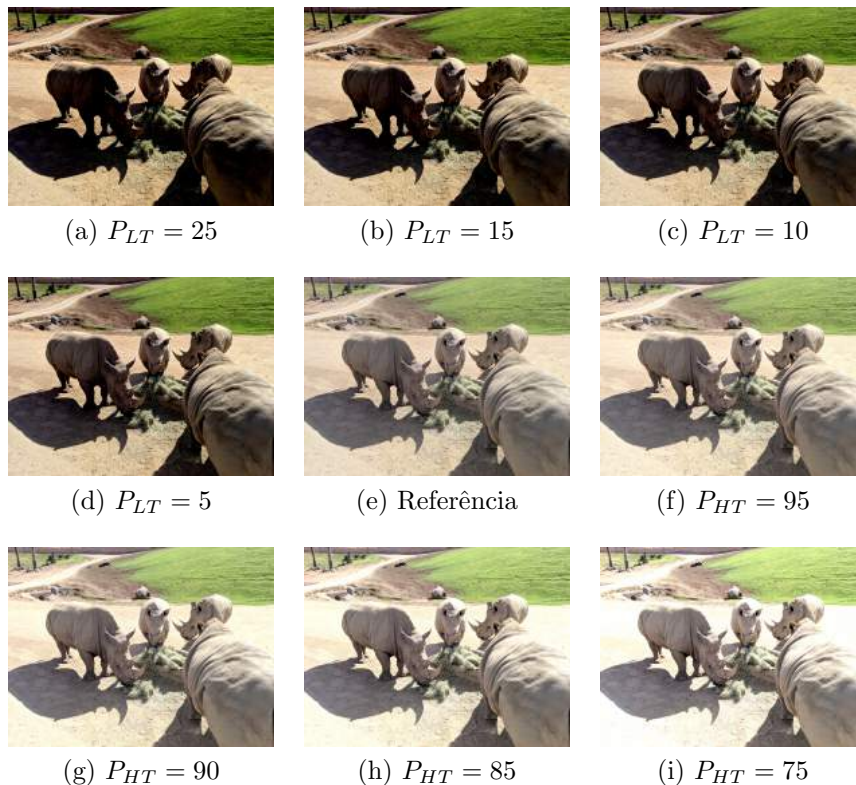


Figura 3.4: Exemplo de conjunto de imagens sintéticas geradas a partir do *dataset* HDR+, apresentando o impacto causado pelas operações de truncamento e normalização por percentil.

A Figura 3.4 apresenta um conjunto de imagens com sub e sobre-exposição sintética gerado a partir de uma imagem do *dataset* HDR+. P_{LT} e P_{HT} indicam os percentis utilizados na Equação 3.1. Na aplicação verifica-se que a transformação aplicada prejudica de maneira severa a visibilidade de regiões muito claras e muito escuras, emulando os efeitos de mudança de cromaticidade e textura típicos de fotografias obtidas com parâmetros inapropriados de exposição.

3.4 Validação a Nível de Aplicação

Além dos aspectos estéticos, a presença de regiões subexpostas e sobre-expostas impacta também no desempenho de diversas aplicações baseadas em visão computacional. Para investi-

gar o impacto destes fenômenos em aplicações recentes, propõe-se um arcabouço conceitual de avaliação que reproduz condições de imagens mal expostas passíveis na visão computacional. Faz-se uma avaliação rigorosa da robustez de vários modelos de reconhecimento de imagem e investiga-se seu desempenho sob distintas distorções de imagem. Propõe-se uma abordagem baseada em pipeline para mitigar os efeitos adversos das distorções de imagem, incluindo o modelo de restauração proposto como uma etapa de pré-processamento de imagem que visa estimar a exposição adequada. Além disso, explora-se os impactos das distorções de imagem na tarefa de segmentação, uma tarefa que desempenha um papel fundamental na navegação autônoma, prevenção de obstáculos, seleção de objetos e outras tarefas de robótica.

Esta validação do modelo a nível de aplicação, tem valor considerando que a visão computacional tornou-se um catalisador da implementação de sistemas automatizados e robóticos que dependem da percepção do ambiente. Entre as aplicações práticas dessas associações de hardware-software estão sistemas de vigilância biométrica (Ito et al., 2017), inspeções visuais automatizadas (Molina et al., 2018; Soares et al., 2017), rastreamento de objetos (Sánchez-Ramírez et al., 2020; Voigtlaender et al., 2019; Zhang et al., 2019a), mapeamento de ambiente (Diane et al., 2019), os robôs domésticos e assistivos (Iocchi et al., 2015; Piyathilaka & Kodagoda, 2015), robótica de campo (Weber et al., 2018; Weis et al., 2017) e carros autônomos (Kohli & Chadha, 2019; Chen et al., 2015a). Além disso, a percepção baseada na visão também se mostrou valiosa em diversos sistemas robóticos e autônomos, sendo usada para servovisão (Young et al., 2020), detecção de obstáculos em veículos não tripulados (Drews-Jr et al., 2016; Gaya et al., 2016; Du et al., 2018; Ma et al., 2019), localização e mapeamento (Ha et al., 2018), navegação (Teso-Fz-Betoño et al., 2020), estimativa de distância (Gao et al., 2019), fechamento de *loop* (Qiu et al., 2018) e manipulação de garra de robô (Qian et al., 2020; Liu et al., 2020, 2018; Jia et al., 2020).

Os módulos de percepção visual empregados nas tarefas anteriormente mencionadas são, em geral, implementados utilizando redes neurais convolucionais (CNNs). Uma parte significativa destes módulos é baseada em contribuições apresentadas em Redes Neurais Convolucionais (CNN) destinadas ao reconhecimento de imagens, como DenseNet (Huang et al., 2017), Inception-v3 (Szegedy et al., 2016), Inception-v4 e Inception-ResNet-v2 (Szegedy et al., 2017), MobileNetV1 (Howard et al., 2017), MobileNet-v2 (Sandler et al., 2018), NASNet (Zoph et al., 2018), NASNetMobile (Zoph et al., 2018), ResNet (He et al., 2016a) ResNet-v2 (He et al., 2016b), ResNeXt (Xie et al., 2017), VGG (Simonyan & Zisserman, 2014), e Xception (Chollet, 2017). Portanto, considera-se que esta tarefa da visão computacional oferece um importante caso de estudo, uma vez que é uma base comum para diversas aplicações.

Apesar da tarefa de reconhecimento de imagem ser amplamente explorada e se tratar de uma das áreas mais maduras da visão computacional, uma limitação que muitas vezes passa despercebida diz respeito à robustez dos modelos de reconhecimento em imagens que não exibem exposição adequada. Considerando que tais circunstâncias mostram-se comuns em qualquer sistema baseado em visão, seus efeitos sobre o desempenho das previsões finais não foram inspecionados meticulosamente. A pesquisa sobre este tema tornou-se cada vez mais relevante

desde o acidente com um carro autônomo em desenvolvimento pela Uber Technologies Inc., que foi minuciosamente abordado por Kohli & Chadha (2019). Cenas de alto contraste e sobre-exposição devido a sombras e forte luz solar também foram relatados como críticos para carros autônomos por Zhang et al. (2019a) e Wang et al. (2019b).

3.4.1 Reconhecimento de Imagens

A fim de avaliar a resiliência de diversos modelos de reconhecimento de imagem, propõe-se o uso de conjuntos de imagens gerados sinteticamente em condições superexposição e subexposição. Todos os modelos de classificação são utilizados usados com conjuntos idênticos de pesos e condições de entrada formas de entrada fornecidas por seus autores. Os modelos foram previamente preparados para estarem adequados ao *ImageNet ILSRVC Challenge* Russakovsky et al. (2015).

Classificar objetos apresentados em uma imagem é o objetivo dos modelos de classificação de imagens. Com o passar dos anos, o Deep Learning tornou-se a forma padrão de resolver problemas de classificação de imagens. O Desafio ILSRVC fez com que a arquitetura dos modelos mudasse para atingir a melhor precisão de classificação, dentro do número de parâmetros na rede.

VGG Proposto por Simonyan & Zisserman (2014), o modelo VGG alcançou o primeiro e o segundo lugares no ILSVRC-2014, principalmente porque o aumento da profundidade da rede (*i.e.* mais camadas empilhadas) em combinação com pequenos (3×3) filtros de convolução, acaba produzindo uma grande melhoria em relação aos métodos experimentados anteriormente. Enquanto o VGG requer uma grande quantidade de recursos computacionais por causa de sua grande largura de camadas convolucionais, o VGG garantiu seu lugar como um dos sistemas mais amplamente usados para extração de recursos em aplicativos de perda de percepção, estilo e contexto. Ele foi integrado às estratégias de treinamento de vários modelos de aprendizado profundo de tradução de imagem para imagem.

ResNet Proposto por He et al. (2016a), o modelo ResNet obteve o primeiro lugar no ILSVRC-2015. Os autores reformulam explicitamente as camadas como funções residuais de aprendizagem em relação às entradas da camada, em vez de funções de aprendizagem não referenciadas (módulo residual). Ao fazer isso, o modelo é capaz de evitar tanto o problema do gradiente de desaparecimento, quanto o problema da degradação na otimização. Em termos de estrutura, o modelo é composto principalmente de 3×3 convoluções e camadas de *pooling* médias.

Inception-v3 Proposto por Szegedy et al. (2016), o modelo Inception-v3 apresenta convoluções fatoradas e regularização agressiva, aumentando a eficiência da rede e melhorando a precisão. O módulo Iniciação usa convoluções de tamanhos diferentes para capturar detalhes em escalas variadas (5×5 , 3×3 , 1×1).

Inception-ResNet-v2 Proposto por Szegedy et al. (2017), o Inception-ResNet-v2 mostra que combinar uma arquitetura simplificada uniforme de início com conexões residuais e mais módulos de início do que Szegedy et al. (2016), acelera o treinamento e obtém melhores resultados de precisão.

DenseNet Proposto por Huang et al. (2017), DenseNet é um modelo baseado em He et al. (2016a) onde cada camada obtém entradas adicionais de todas as camadas anteriores e passa em seus próprios mapas de recursos para todas as camadas subsequentes em vez do elemento-sábio adição entre a entrada e a saída de uma camada.

Xception Proposta por Chollet (2017), a rede Xception é inspirada nas arquiteturas iniciais com conexões saltadas e convoluções separáveis modificadas em termos de profundidade como uma melhoria.

MobileNetV1 Proposto por Howard et al. (2017), MobileNetV1 é um CNN com convoluções separáveis em profundidade entre as camadas de convoluções regulares. Consequentemente, os parâmetros e as operações de multiplicação-adição são consideravelmente reduzidos, o que é adequado para dispositivos móveis ou quaisquer dispositivos com baixo poder computacional.

MobileNetV2 Proposta por Sandler et al. (2018), a rede MobileNetV2 tem um bloco residual com uma passada de 1 e um bloco com uma passada de 2 para *downsizing*, superando Howard et al. (2017).

NASNet Proposto por Zoph et al. (2018), NASNet é um modelo que utiliza as informações adquiridas em um pequeno conjunto de dados sobre um grande buscando a melhor camada convolucional no primeiro. Os autores também propõem a técnica de regularização **Scheduled Drop Path**, que melhora significativamente a generalização do modelo.

A Tabela 3 fornece detalhes sobre datas de lançamento, tamanho da rede, resolução da imagem de entrada e precisão de validação Top-1 obtida por cada modelo, de acordo com os relatórios oficiais ¹. Devido ao pré-processamento, otimização da estrutura de aprendizado profundo e precisão dos pontos flutuantes utilizados durante a inferência, a precisão real pode variar. Para evitar a interferência dessas variáveis, reexecuta-se a inferência no conjunto de validação original nas mesmas condições de todas as outras imagens distorcidas. No entanto, a metodologia de avaliação também pode ser aplicada a qualquer modelo e conjunto de dados para tarefas de reconhecimento de imagem ou segmentação.

¹Resultados atualizados do estado da arte estão disponíveis em <https://paperswithcode.com/sota/image-classification-on-imagenet>. Os vencedores do desafio para cada edição podem ser encontrados em <http://www.image-net.org/challenges/LSVRC/2017/results>, <http://www.image-net.org/challenges/LSVRC/2016/results>, <http://www.image-net.org/challenges/LSVRC/2015/results> e <http://www.image-net.org/challenges/LSVRC/2014/results>.

Tabela 3: Modelos de classificação considerados nos experimentos. Precisão de validação Top-1 de acordo com os relatórios oficiais.

Modelo	Ano	Tamanho	Parâmetros	Top-1	Resolução (px.)
VGG-16 (Simonyan & Zisserman, 2014)	2014	528 MB	138,357,544	0,71	224
ResNet50 (He et al., 2016a)	2016	98 MB	25,636,712	0,75	224
Inception-v3 (Szegedy et al., 2016)	2016	92 MB	23,851,784	0,78	299
Inception-ResNet-v2 (Szegedy et al., 2017)	2017	215 MB	55,873,736	0,80	299
DenseNet201 (Huang et al., 2017)	2017	80 MB	20,242,984	0,77	224
Xception (Chollet, 2017)	2017	88 MB	22,910,480	0,79	299
MobileNetV1 (Howard et al., 2017)	2017	16 MB	4,253,864	0,70	224
MobileNetV2 (Sandler et al., 2018)	2018	14 MB	3,538,984	0,71	224
NASNetLarge (Zoph et al., 2018)	2018	343 MB	88,949,818	0,83	331
NASNetMobile (Zoph et al., 2018)	2018	23 MB	5,326,716	0,74	224

3.4.2 Segmentação Semântica de Imagens

Mask-RCNN Proposto por He et al. (2017), Mask-RCNN é um modelo para segmentação de instância (*i.e.* para encontrar instâncias de um objeto contável na cena). Seu objetivo é distinguir cada instância de cada objeto na imagem no nível do pixel. Este modelo é baseado em FPN (*Feature Pyramid Networks for Object Detection*) (Lin et al., 2017) o que lhe permite boa detecção de objetos em diferentes escalas. O modelo Mask-RCNN é processado em duas etapas: primeiro, o modelo gera propostas sobre as regiões onde pode haver um objeto; em seguida, o modelo identifica a classe do objeto, cria uma caixa delimitadora e gera uma máscara em nível de pixel do objeto com base na proposta do primeiro estágio.

3.4.3 Procedimento

Primeiro, avaliam-se os impactos das distorções de imagem na aplicação de reconhecimento de objetos. Esta avaliação é realizada considerando que este campo de pesquisa maduro e que os modelos de reconhecimento definiram as direções no desenvolvimento de outras tarefas de visão e são, portanto, uma boa medida do desempenho geral da percepção da visão para a robótica.

A avaliação da robustez das redes de reconhecimento de imagem para distorção comum de imagens foi realizada, com métricas calculadas no subconjunto de validação ILSRVC *ImageNet Challenge*.

1. As imagens são carregadas individualmente utilizando a biblioteca *Python Scikit-Image* van der Walt et al. (2014), tipo de dados inteiro sem sinal de 8 bits. Todos os arquivos são armazenados no formato JPEG compactado, com variações no tamanho e proporção da imagem.
2. A imagem de entrada passa por um processo de distorção para replicar os efeitos de subexposição e sobre-exposição.
3. As imagens foram redimensionadas e recortadas de acordo com as restrições aceitas pelo

modelo, momento em que é utilizada uma interpolação *spline* de primeira ordem. O filtro gaussiano em $\sigma = \frac{s-1}{2}$ foi usado como um método *anti-aliasing* para reduzir a escala das imagens, s sendo o fator de escala. Se o modelo assim o exigir, outras transformações de imagem específicas foram feitas para adaptar a entrada e a representação dos dados.

4. A imagem finalmente está pronta para ser inserida como entrada no sistema e, uma vez que a inferência prossiga, os resultados são armazenados para posterior avaliação.
5. Esta avaliação julga algumas das métricas de avaliação mais populares: Acurácia Top-1, Acurácia Top-5 e F1-Score. Levando em consideração o número de instâncias verdadeiras para cada rótulo, uma média ponderada com apenas os resultados Top-1 é usada para Precisão, Recuperação e Pontuação F1.

Em seguida, testa-se um *pipeline* alternativo de visão computacional. A fim de minimizar os impactos indesejáveis da exposição incorreta modifica-se o *pipeline* tradicional, introduzindo uma etapa de restauração de imagem, que opera no espaço de cores RGB, após a imagem já ter sido compactada e transmitida. A restauração ocorre imediatamente antes do algoritmo de reconhecimento de objeto de forma que o *pipeline* possa ser facilmente aplicado a outras tarefas de visão sem a necessidade de adaptações adicionais ou personalização de hardware.

Finalmente, a fim de investigar se os resultados da avaliação na tarefa de reconhecimento de objetos também se verificam a aplicações distintas de percepção, investiga-se como a exposição inadequada afeta a segmentação de instâncias. Os modelos de segmentação são frequentemente aplicados para tarefas como seleção de objetos, detecção de obstáculos, navegação autônoma e interação humano-robô. Esses sistemas geralmente compartilham os mesmos blocos de construção básicos que se mostraram bem-sucedidos com modelos de reconhecimento de objeto baseados em CNN Zhang et al. (2019a); Ha et al. (2018); Teso-Fz-Betoño et al. (2020); Qiu et al. (2018); Liu et al. (2018); Jia et al. (2020); Ma et al. (2019); Du et al. (2018). Como esses modelos contam com o mesmo conjunto de técnicas e arquiteturas populares, espera-se que eles mostrem os mesmos pontos fortes e fracos dos modelos CNN de reconhecimento de objetos.

3.5 Visualização da Ativação Interna do Modelo

Uma das maiores ressalvas aos modelos de restauração baseados em redes neurais artificiais diz respeito ao seu modo de operação ‘caixa-preta’. Uma vez que estes modelos apresentam milhares de parâmetros treináveis, a visualização dos componentes internos da rede é uma importante ferramenta para compreensão e identificação de oportunidades de melhoria no que diz respeito à arquitetura da rede.

Apesar de serem amplamente utilizadas na resolução de vários problemas da restauração e realce de imagens, a compreensão de como os modelos baseados em CNN funcionam e como os milhares de parâmetros se relacionam para chegar à resposta final ainda é limitada. A visualização das regiões de entrada que são importantes para as previsões do modelo através

de uma explicação visual oferece uma descrição intuitiva sobre a operação da arquitetura, permitindo identificar quais as informações da imagem de entrada que o modelo considera mais relevantes. Um melhor entendimento sobre a operação interna do um modelo específico ajuda na previsão de como ele vai se comportar em situações inesperadas, além de contribuir para o desenvolvimento de soluções mais eficientes.

Utiliza-se para efeitos de visualização o método Mapeamento de Ativação de Classes Ponderadas por Gradiente (Grad-CAM, do inglês *Gradient-weighted Class Activation Mapping*), de Selvaraju et al. (2017). Este método permite observar, através de mapas de atenção, quais as propriedades da imagem de entrada são levadas em consideração em cada camada da rede para, ao final produzir uma imagem restaurada. Considera-se este tipo de visualização adequado para a compreensão da arquitetura proposta.

4 MODELO DE CORREÇÃO DE EXPOSIÇÃO AJUSTADO EM MEDIDAS DE SIMILARIDADE

Neste trabalho propõe-se uma arquitetura de rede neural artificial, capaz de restaurar imagens impactadas por exposição inapropriada, obtidas a partir de câmeras digitais convencionais, preservando as características mais representativas para a visão computacional como definição, contraste, nitidez e cor. A rede recebe uma imagem sRGB como entrada e produz uma imagem sRGB de dimensões idênticas na saída. Tem-se como meta gerar um modelo ponta-a-ponta eficiente, que minimize o uso de memória e capacidade de processamento requerido, sem sacrificar de maneira significativa a qualidade dos resultados produzidos.

Apresenta-se neste capítulo um modelo de correção de exposição de imagem no espaço de cor sRGB. Tratando-se de uma tarefa de transformação imagem-imagem opta-se por um modelo convolucional em detrimento de modelos densos – adequados para dados estruturados – e modelos recorrentes – adequados para dados sequenciais. Busca-se um modelo que produza melhoria visual perceptível, especialmente em termos de correção de cor, preenchimento de grandes regiões e qualidade de textura em regiões saturadas.

A arquitetura proposta utiliza camadas convolucionais com dilatação ampla para de agregação de contexto, camadas convolucionais treináveis para *down-scaling* e *up-scaling* dentro da rede, e um número reduzido de *skip connections*, de modo a diminuir os requisitos de memória durante o treinamento. Nesta arquitetura inclui-se um bloco de convoluções dilatadas, como forma de minimizar a quantidade de parâmetros do modelo ao limitar a quantidade de filtros disponíveis e forçar, através da arquitetura, que todas as camadas internas da rede tenham acesso a um campo receptivo amplo. O modelo é ajustado utilizando uma função objetivo que explora as características de imagens sRGB inapropriadamente expostas.

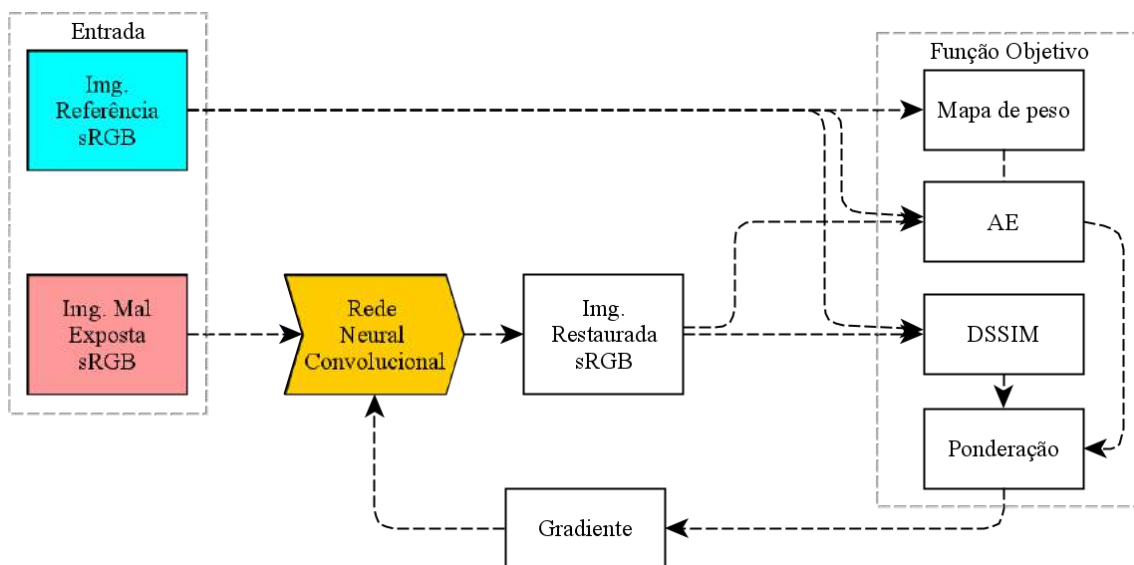


Figura 4.1: Visão geral do modelo de rede convolucional de aprendizagem supervisionada para restauração de imagens inapropriadamente expostas, durante processo de treinamento

Uma visão geral do modelo proposto é apresentada na Figura 4.1. O ajuste dos parâmetros do modelo de restauração é feito utilizando-se medidas de similaridade entre a saída do modelo e a imagem referência correspondente. O treinamento deste modelo requer a existência de dados pareados no *dataset*. O processo de aprendizagem é totalmente supervisionado.

4.1 Arquitetura da Rede

A arquitetura de rede é apresentada na Figura 4.2. A base da rede segue um padrão ‘U-Net’. Apresentada pela primeira vez por Ronneberger et al. (2015), a U-Net é baseada em uma estrutura *encoder-decoder*, aprimorada com o uso de *skip-connections*. Nas arquiteturas *encoder-decoder*, as primeiras camadas da rede produzem poucos *feature maps* com resolução alta. À medida que a profundidade aumenta, a estrutura prevê mais *feature maps*, com resoluções menores. O *encoder* é composto por sucessivas camadas com redução espacial e aumento no número de *feature maps*. O *decoder*, por outro lado, é composto por sucessivas camadas com aumento de resolução e diminuição na quantidade de *feature maps*. Desta forma, as últimas camadas possuem as mesmas dimensões da entrada. Em geral, nas arquiteturas *encoder-decoder* a ampliação espacial dos *feature maps* é realizada através de convoluções transpostas (*transposed convolutions* ou *deconvolution layers*). Arquiteturas *encoder-decoder* favorecem a criação de um campo receptivo (*receptive field*) amplo, propriedade que permite que as camadas profundas da rede tenham acesso à informação contextual oferecida por uma área ampla da imagem de entrada.

Arquiteturas *encoder-decoder*, no entanto, apresentam limitações relevantes para aplicações voltadas para restauração e realce de imagens. Estas limitações são decorrentes, em maior proporção, da perda de informação causada pela redução espacial da imagem dentro das camadas da rede. Dada somente uma abstração da imagem (*feature maps* ao final do *encoder*), recuperar os detalhes da imagem original torna-se um problema subdeterminado (Mao et al., 2016). U-Nets tratam este problema ao introduzir atalhos (*skip connections*) entre camadas do *encoder* e do *decoder* que apresentam as mesmas dimensões. Desta forma, melhoram o fluxo do gradiente dentro da rede.

Ronneberger et al. (2015) é conhecido como o primeiro trabalho a explorar o uso de atalhos entre camadas da rede para fins de predição densa (transformação de imagens). O modelo de Ronneberger et al. (2015) destinava-se originalmente a resolver o problema de segmentação de imagens médicas. As vantagens da utilização de atalhos dentro da arquitetura foram posteriormente evidenciadas para outras aplicações, destacando-se He et al. (2016a), com o modelo ResNet destinado à tarefa de classificação, e Huang et al. (2017), com o modelo DenseNet voltado para reconhecimento de objetos. Atualmente, a maioria dos modelos estado da arte em transformação de imagens fazem o uso de arquiteturas inspiradas nestes modelos.

Para expandir o campo receptivo do modelo de rede neural artificial, Yu & Koltun (2015) propõe uma abordagem distinta ao uso de uma estrutura *encoder-decoder*, também voltado para problemas de predição densa. Yu & Koltun (2015) introduzem o conceito de rede de

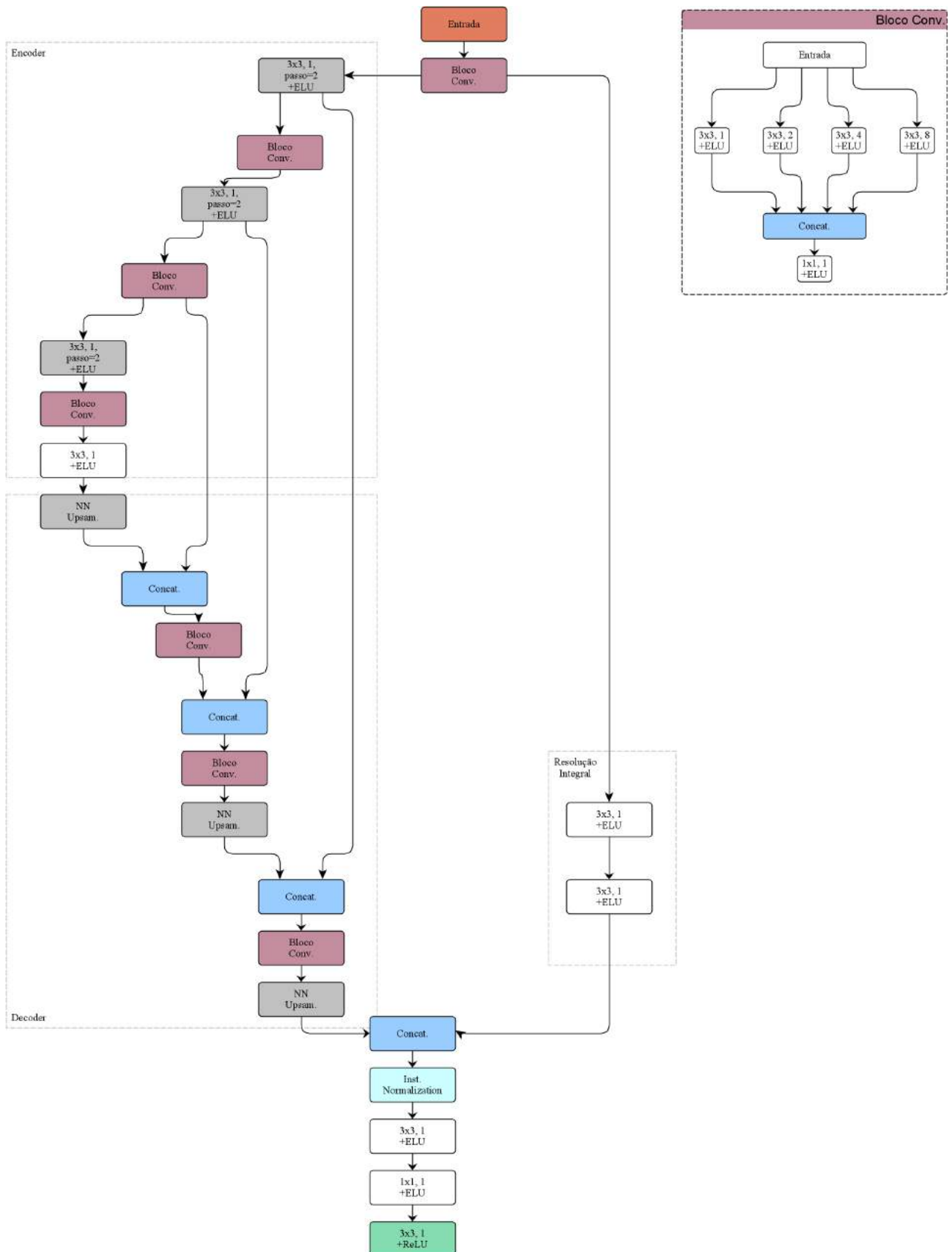


Figura 4.2: Arquitetura de rede convolucional de aprendizagem supervisionada para restauração de imagens inapropriadamente expostas. As camadas são assim definidas: tamanho do *kernel* de convolução, taxa de dilatação, tamanho do passo e função de ativação

agregação de contexto (CAN, do inglês *Context Aggregation Network*). Na CAN, as camadas de redução espacial dos *feature maps*, utilizando convolução com passo > 1 ou camadas de votação, é preterida em favor de convoluções dilatadas. O uso de convoluções dilatadas permite expansão exponencial do campo receptivo sem incorrer em perda de resolução ou cobertura (Yu & Koltun, 2015).

A estrutura da rede proposta alia propriedades das arquiteturas U-net e CAN, de forma a minimizar o número de parâmetros treináveis e a profundidade da rede. O primeiro fluxo se inspira em ideias apresentadas em Milletari et al. (2016); Ronneberger et al. (2015); Drozdal et al. (2016); Isola et al. (2017); Pathak et al. (2016), ou seja, uma arquitetura *encoder-decoder*, aprimorada por meio de *skip connections*. Esta arquitetura *encoder-decoder* é construída sob a suposição de que o *encoder* é capaz de aprender as informações distribuídas espacialmente e codificá-las em uma representação que pode ser posteriormente decodificada para uma imagem que corresponda à saída desejada da rede. O uso desse design, em que a imagem é reduzida (*down-sampling*) e depois expandida (*up-sampling*) permite que esse segmento da rede atue como um observador global, uma vez que torna o campo receptivo significativamente mais amplo usando menos convoluções sequenciais.

Os estágios de *down-sampling* e *up-sampling* podem ser construídos usando abordagens distintas. Uma abordagem popular, empregada entre outros por Gharbi et al. (2017), Milletari et al. (2016), Ronneberger et al. (2015) e Isola et al. (2017), consiste na utilização de uma sequência de camadas convolucionais com passo > 1 . Outra abordagem comum para *down-sampling*, usada principalmente em redes de classificação como Szegedy et al. (2017), é usar o operador de votação *max-pooling*, uma função *down-sampling* não linear que retorna o máximo de um conjunto de valores. Operadores de *Max-Pooling* somente permitem a passagem dos *features* com maior ativação, levando ao descarte da informação dos demais *features* dentro do filtro. Embora operações *Max-Pooling* sejam excelentes para problemas de classificação e demandem menos poder computacional, observa-se que para problemas de transformação imagem-imagem é interessante que a maior quantidade possível de informação na entrada seja preservada. Desta forma, opta-se por camadas convolucionais treináveis, para as quais o procedimento de treinamento pode determinar a melhor combinação de *features* em detrimento de outras alternativas de *pooling* para *down-sampling*.

Para o estágio de *up-sampling* no decodificador, as opções variam da camada convolucional transposta (também conhecida como deconvolução), como usado em Milletari et al. (2016); Isola et al. (2017); *Up-sampling* Bilateral Profundo, como proposto em Chen et al. (2016) e Gharbi et al. (2017); a *up-sampling* por vizinhos mais próximos (repetição de valores) seguido de convolução, como usado em Ronneberger et al. (2015). Com base nos resultados apresentados por Odena et al. (2016), que mostraram redução significativa de artefatos de tabuleiro ao utilizar *up-sampling* por vizinhos mais próximos seguido por convolução, opta-se por aplicar essa combinação.

Ainda na Figura 4.2, um segundo fluxo, sem redimensionamento, ajuda a transportar informações por um caminho mais curto, ajudando a rede a reter características importantes

da estrutura da imagem. Embora, em geral, a profundidade melhore o desempenho de uma rede, ela também dificulta o treinamento baseado em gradiente, já que as redes mais profundas tendem a ser mais não-lineares (Hochreiter et al., 2001). A importância das conexões atalho (também conhecidas como residuais) no problema de transformação imagem-imagem foi evidenciada em Drozdal et al. (2016), mostrando uma melhora significativa com relação às redes que não aplicam esta técnica. Os *feature maps* são unificados através de concatenação seguida por convolução.

Os dois fluxos da rede são combinados em um bloco convolucional que atua como mecanismo de ponderação para as *features* concatenadas através de 32 filtros. Por fim, a camada de saída da rede é uma convolução regular de 3×3 com 3 filtros (uma para cada canal de cor RGB), seguida por uma função de ativação ReLU para restringir as saídas aos limites válidos de representação, no intervalo $[0; 1]$. Exceto pela camada de saída, todas as camadas convolucionais são compostas por 32 filtros. Dentro do bloco convolucional, utiliza-se 8 filtros por tipo de núcleo de convolução. A rede geradora tem ao total cerca de 381 mil pesos treináveis, valor pequeno se comparado aos modelos estado-da-arte aplicados com propósito semelhante.

Exceto na última camada da rede, todas as operações de convolução são seguidas por uma função de ativação Unidade Linear Exponencial (ELU, do inglês *Exponential Linear Unit*). A função de ativação ELU (Equação 4.1) foi utilizada pela primeira vez por Clevert et al. (2015), que argumentam que esta função de ativação proporciona uma aceleração no aprendizado e leva a uma precisão maior em comparação com ReLU (*Rectified Linear Units*, com resposta linear para números positivos e zero para números negativos) e Leaky ReLU (uma versão melhorada das ReLUs em que entradas negativas produzem saídas negativas atenuadas). Para entradas não negativas a ativação ELU funciona como uma função identidade. Para entradas negativas, o resultado é uma curva suave até a saída ser igual a $-\alpha$ (empregou-se $\alpha = 1$ de forma empírica). Clevert et al. (2015) defende ainda que os valores negativos dos ELUs permitem que eles empurrem a ativação média da unidade para mais perto de zero. A contribuição desta função de ativação para acelerar o processo de treinamento e a acurácia geral de modelos convolucionais foi evidenciada empiricamente por Hendrycks & Gimpel (2016).

$$f(x) = \begin{cases} x, & x \geq 0, \\ \alpha(e^x - 1), & x < 0. \end{cases} \quad (4.1)$$

Utiliza-se ainda um *Instance Normalization* (Ulyanov et al., 2016). Nesta técnica de normalização, média e variância são calculadas individualmente para cada canal do *feature map* em cada imagem de entrada considerando ambas as dimensões espaciais. Desta forma, a normalização por instância atua como uma forma de atuar sobre o contraste da imagem. A contribuição desse tipo de normalização já foi evidenciada em Pan et al. (2018). A normalização por instância se difere da normalização por lotes por não necessitar dos dados de uma população. Desta forma o modelo pode utilizar os mesmos parâmetros para treinamento e inferência.

4.2 Bloco de Convoluções Dilatadas

O Bloco de Convoluções Dilatadas é um arranjo proposto com o objetivo de: i) trabalhar com um grande campo receptivo; ii) limitar por *design* o número de filtros de convolução necessários em cada camada e evitar que o treinamento do modelo resulte em filtros repetidos; e iii) reduzir a quantidade de operações aritméticas necessárias. Embora os fundamentos que justificam a aplicação de convoluções dilatadas sejam amplamente conhecidos na área de pesquisa, não são conhecidas abordagens anteriores que as tenham aplicado na forma apresentada neste trabalho.

A restauração de imagens, assim como outras tarefas de predição densa, exige raciocínio contextual multiescala em combinação com a saída na resolução original (Yu & Koltun, 2015). Nesse sentido, as convoluções dilatadas são capazes de fornecer campos receptivos grandes sem a necessidade de filtros grandes (com muitos parâmetros treináveis), dispensando o empilhamento de camadas convolucionais (redes muito profundas). As convoluções dilatadas também podem capturar uma representação hierárquica de um espaço de entrada muito maior do que as convoluções padrão, permitindo que sejam dimensionadas para grandes tamanhos de contexto (Gupta & Rush, 2017).

O bloco convolucional inclui quatro camadas paralelas de convoluções dilatadas 3×3 , com taxas de dilatação variando de 2^0 a 2^3 . Assim, cada bloco convolucional é capaz de agregar *features* de uma região 19×19 no espaço de entrada utilizando apenas 9 pesos treináveis para cada filtro. As saídas de todas as convoluções dilatadas são concatenadas e passam então por uma camada de convolução 1×1 , que atua de forma a adicionar não linearidade e ponderar os valores das convoluções dilatadas, atribuindo maior ou menor importância de acordo com sua contribuição para minimizar a função objetivo.

A Figura 4.3 mostra o campo receptivo do bloco convolucional proposto. Aplicadas de forma sucessiva e intercaladas com camadas de redução na resolução da imagem, o bloco de convoluções dilatadas permite que a restauração de características locais tenha consistência global. A Tabela 4 mostra a expansão do campo receptivo nas primeiras camadas da rede (*encoder*).

Tabela 4: Expansão do campo receptivo dentro da rede proposta através do fluxo *encoder*. Verifica-se que em 7 camadas a arquitetura da rede chega a um *receptive field* de 255×255 pixels da imagem de entrada

Camada	Tamanho do Filtro ²	Passo	Dilatação	Escala de Saída	<i>Receptive Field</i>
Bloco Convolucional	17	1	8	1	17
Conv. com passo 2	3	2	1	1/2	19
Bloco Convolucional	17	1	8	1/2	51
Conv. com passo 2	3	2	1	1/4	55
Bloco Convolucional	17	1	8	1/4	119
Conv. com passo 2	3	2	1	1/8	127
Bloco Convolucional	17	1	8	1/8	255

²Para o bloco convolucional o filtro é esparsos, cobrindo 31 de 289 (17^2) *features* de entrada.

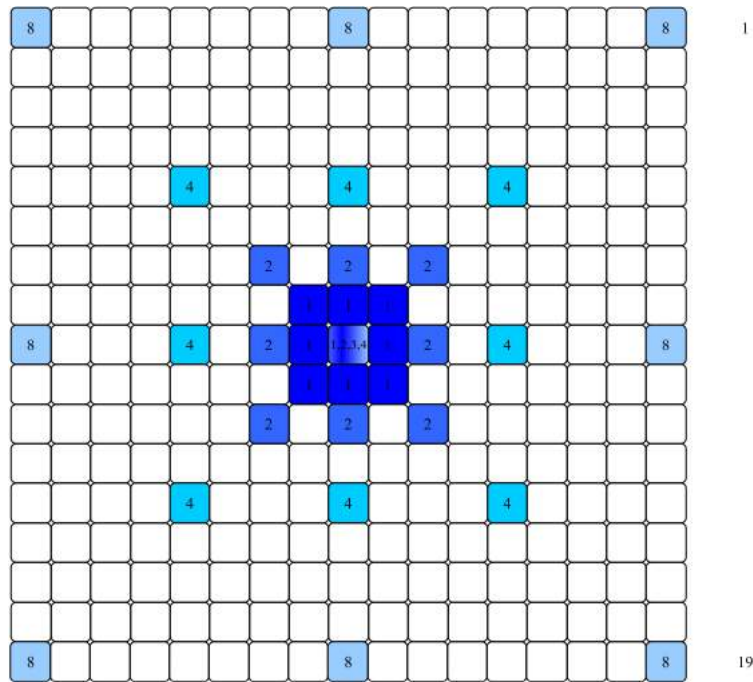


Figura 4.3: *Receptive field* do bloco convolucional proposto. Os números na imagem indicam a taxa de dilatação

A expansão do *receptive field* tem valor, especialmente quando se considera imagens em resolução média e alta, para as quais esta alternativa colabora no sentido de agregar informação global da imagem. A estruturação em camadas paralelas, em contraposição ao modelo de camadas sequenciais como proposto por Chen et al. (2017) e Yu & Koltun (2015), propicia um caminho sem perdas para as *features*, em que cada camada de convolução tem acesso aos mesmos dados de entrada. Por outro lado, esta abordagem aumenta a quantidade de memória necessária em tempo de inferência, uma vez que o resultado do processamento das camadas iniciais da rede precisa ser armazenado por mais tempo.

Por fim, as convoluções 1×1 introduzidos por Lin et al. (2013), tem sido aplicadas com sucesso como uma alternativa, de baixo custo computacional, para aprofundar os modelos sem impactar de maneira relevante a quantidade de parâmetros e adicionar não linearidade à rede. Utiliza-se convoluções de 1×1 em cada bloco como forma de impor a seleção constante de *features*, garantindo assim que a rede seja capaz de codificar os filtros mais significativos e descartar aqueles que não contribuem para gerar melhoria. Em vez de um bloco de pixels, como ocorre com convoluções de dimensões maiores, as convoluções 1×1 consideram um único píxel na imagem (ou do *feature map* na entrada), em todos os canais u , agindo como uma rede densa (completamente conectada) aplicada em cada posição.

Combinando convoluções 1×1 , convoluções 3×3 regulares e convoluções dilatadas 3×3 , o bloco proposto permite cobrir um grande campo receptivo e ainda preservar a localidade. Dentro de um campo receptivo, quanto mais próximo um píxel estiver do centro, mais ele contribuirá para o cálculo das *features* de saída. Através do bloco de convoluções dilatadas, a localidade é imposta pela arquitetura da rede.

Uma vez que o bloco convolucional é aplicado repetidamente em toda a rede, cada *feature map* nas camadas mais profundas é capaz de acessar *features* codificados de uma área ampla na imagem de entrada. Isso fornece à rede a capacidade de construir *features* globais relevantes, dispensando etapas adicionais de processamento para garantir a consistência global na transformação imagem-imagem. Um campo receptivo amplo permite que o modelo trabalhe com propriedades implícitas da imagem completa (como distribuição de luminância ou tipo de cena). Cada convolução no bloco utiliza $\frac{u}{4}$ filtros, reduzindo o número de *features* da entrada, o que força a rede a minimizar filtros que não fornecem contribuição significativa.

4.3 Função Objetivo

Utiliza-se uma função de erro customizada para enfatizar regiões da imagem mais próximas ao limite do sensor e portanto, mais propensas a sofrer os efeitos adversos da exposição inadequada. Essa função objetivo combina, de forma ponderada, dissimilaridade estrutural (DSSIM) e erro médio absoluto entre píxel-a-píxel.

DSSIM é uma medida baseada no índice de similaridade estrutural entre duas imagens SSIM, de Wang et al. (2004). Naturalmente, sua utilização só é viável quando uma imagem danificada e a sua referência são conhecidas, requerendo dados pareados. O índice SSIM considera a degradação da imagem como uma mudança percebida nas informações estruturais, incluindo ainda termos de compensação da luminância e do contraste. A SSIM assume que a informação estrutural é dada pela interdependência entre os píxeis, especialmente quando estão espacialmente próximos. Essas dependências carregam informações importantes sobre a estrutura dos objetos na cena. Com relação a luminância, Wang et al. (2004) defendem que as distorções da imagem tendem a ser menos visíveis em regiões brilhantes, enquanto o contraste é se reflete em distorções menos visíveis onde há textura na imagem.

O índice SSIM é calculado considerando blocos 3×3 das imagens. O SSIM entre uma imagem distorcida a e sua referência b é dado pela Equação 4.2, em que μ é a média aritmética dos valores de intensidade para os píxeis no bloco, σ^2 é a variância, σ_{ab} é a covariância, c_1 e c_2 são variáveis utilizadas para estabilizar a divisão quando os valores do denominador são muito pequenos.

$$\text{SSIM}(a, b) = \frac{(2\mu_a\mu_b + c_1)(2\sigma_{ab} + c_2)}{(\mu_a^2 + \mu_b^2 + c_1)(\sigma_a^2 + \sigma_b^2 + c_2)}. \quad (4.2)$$

A partir do índice SSIM, a medida DSSIM é dada pela Equação 4.3. Os valores de saída são no intervalo $[0; 0,5]$:

$$\text{DSSIM}(a, b) = \frac{1 - \text{SSIM}(a, b)}{2}. \quad (4.3)$$

Embora o DSSIM forneça uma boa avaliação da similaridade entre duas imagens, o índice não consegue avaliar os valores de píxeis na posição exata. Portanto, complementa-se a função objetivo com o erro absoluto (AE) entre valores de píxel da saída do modelo.



(a) Imagem referência

(b) Mapa de pesos

Figura 4.4: Mapa de pesos baseado em níveis de intensidade utilizado no ajuste do modelo. Pontos mais claros no mapa de pesos indicam maior atenção

De forma a ponderar a retropropagação do erro, atribui-se maior peso à regiões mais propensas à saturação ou subexposição. Para tanto, propõe-se uma matriz de pesos W computada a partir da imagem referência b . Assumindo-se que as imagens estão representadas no intervalo $[0; 1]$, valores próximos dos limites tem ponderação maior que valores no centro da escala. A Figura 4.4 apresenta um caso de aplicação do mapa de pesos colorizado mostrando como regiões muito claras e regiões muito escuras na imagem de entrada, onde os efeitos adversos da exposição inapropriada se manifestam de forma mais acentuada, são ponderadas para ter maior atenção no ajuste do modelo.

$$AE(a, b) = |a - b|. \quad (4.4)$$

$$W = |b - 0,5|. \quad (4.5)$$

Ao final, a função objetivo é dada por:

$$\mathcal{L}(a, b) = \lambda W \circ AE(a, b) + (1 - \lambda) DSSIM(a, b), \quad (4.6)$$

onde $\lambda = 0.2$ é uma constante empírica utilizada para compensar a diferença de escala entre as duas funções de erro.

4.4 Treinamento do Modelo

O modelo proposto é ajustado e testado em quatro conjuntos de imagens, previamente descritas na Seção 3.3. Em todos os casos utiliza-se 70% do *dataset* para treinamento e a parte remanescente para testes e validação. As amostras utilizadas para cada etapa são selecionadas aleatoriamente. Utiliza-se o otimizador Adam, de Kingma & Ba (2014), com os hiperparâmetros padrão. A atualização do pesos é feita em minilotes de 8 imagens com resolução variada (devido à restrições impostas pelo *hardware* utilizado para o treinamento). Todos os dados utilizados

para treinamento são pareados.

Todos os pesos da rede são inicializados utilizando o método de Glorot & Bengio (2010), com distribuição normal $\mathcal{N}(\mu, \sigma)$, sendo $\mu = 0$ e $\sigma = \frac{2}{features_{entrada} + features_{saida}}$. O treinamento para imagens com subexposição e sobre-exposição é realizado de maneira separada, resultando em um modelo específico para restauração de imagens subexpostas e um modelo específico para imagens sobre-expostas.

Inicialmente, ajusta-se os modelos para o *dataset* multiexposição sintético baseado em FiveK (apresentado na Seção 3.3.3). O ajuste do modelo para os demais *datasets* é feito utilizando como ponto de partida os pesos pré-ajustados para os dados FiveK. Iniciar o treinamento a partir de um modelo pré-ajustado permite que a rede alcance um valor ótimo em um curto espaço de tempo. Afora isso, a utilização de dados sintéticos expande significativamente a quantidade de dados disponíveis para treinamento.

O ajuste baseado em gradiente é um processo iterativo. Neste sentido, faz-se útil o uso de um critério de parada para identificar quando o treinamento parou de surtir resultados. Para o modelo proposto faz-se a validação através do erro médio quadrático (MSE). O treinamento é encerrado uma vez que sejam processados 300 minilotes de imagens sem surtir ganhos maiores que 10^{-5} . Critério idêntico é aplicado às demais arquiteturas utilizadas no comparativo.

5 RESULTADOS E DISCUSSÃO

Neste Capítulo discutem-se os resultados da pesquisa e do modelo proposto. Primeiro, na Seção 5.1 apresentam-se uma avaliação das métricas e medidas de qualidade de imagem quando aplicadas, identificando-se as suas características quando aplicadas em imagens subexpostas e sobre-expostas. Em seguida, na Seção 5.2 apresentam-se os resultados detalhados obtidos pelo modelo proposto em quatro datasets. Posteriormente, na Seção 5.3 apresenta-se um quadro comparativo sumarizado de outros trabalhos relacionados na literatura. Depois disso, na Seção 5.4 apresenta-se uma avaliação dos impactos da exposição inapropriada de imagens em aplicações de visão computacional comuns em tarefas de robótica e automação, evidenciando a viabilidade da aplicação do modelo nestas situações. Na Seção 5.6 faz-se uma investigação do funcionamento do modelo de restauração proposto sob a perspectiva das ativações internas, mostrando como cada camada interna da rede contribui para a restauração. Finalmente, na Seção 5.7 discutem-se as limitações conhecidas do modelo.

5.1 Considerações Gerais Sobre Indicadores de Qualidade de Imagem Aplicados

Conforme discussão apresentada na Seção 3.2, a aplicabilidade e coerência entre diversas medidas de qualidade de imagem é um tópico de pesquisa que ainda levanta questionamentos. Não foram encontrados trabalhos destinados à avaliação de imagens obtidas sob condições impróprias de exposição, tema abordado no presente trabalho. Desta forma, conduziu-se um estudo de correlação entre estas métricas quando aplicadas na comparação de imagens para o *dataset* A6300. Uma descrição detalhada deste dataset é apresentada na Seção 3.3.1. Este dataset é composto por conjuntos contendo uma imagem apropriadamente exposta, uma imagem subexposta, uma imagem sobre-exposta e uma composição das anteriores. Todas as imagens são comprimidas utilizando JPEG. Para o comparativo, utilizando medidas baseadas em referência empregou-se dados pareados na forma {subexposta, apropriadamente exposta} e {sobre-exposta, apropriadamente exposta}.

Fez-se inicialmente uma avaliação da normalidade dos dados utilizando os testes de Jarque-Bera (Jarque & Bera, 1980), Shapiro-Wilk (Shapiro & Wilk, 1965) e Anderson-Darling (Stephens, 1974) com nível de significância $\alpha = 0.05$. Os testes de normalidade indicaram que a maior parte das medidas testadas não segue uma distribuição normal. Assim sendo, apresentam-se os resultados utilizando o teste de correlação de Spearman, mais adequada para o conjunto de dados que não normal. Os resultados utilizando o Coeficiente de Correlação de Pearson são disponibilizados no Apêndice 7.1.

Na Tabela 5 considera-se a estatística não-paramétrica de Spearman. Verifica-se uma correlação exata entre PSNR, MSE e MAE. As três medidas apresentam ainda uma correlação muito forte com GMSD e Delta E (CIE 2000). Identifica-se ainda que Delta E (CIE 2000) apre-

Tabela 5: ρ de Spearman (SRCC) para medidas de qualidade de imagem no *dataset* A6300 (232 amostras pareadas, das quais 116 contém imagens subexpostas e 116 contém imagens sobre-expostas). Os resultados indicam que a maior parcela das métricas avaliadas apresenta entre si correlação forte ou muito forte. Métricas populares como PSNR, MSE, MAE, SSIM e FSIM não apresentam correlação forte apenas com as medidas VIFP e RECO

	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Dif. Hist.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
PSNR	1.00	-1.00	-1.00	0.64	0.67	0.68	-0.75	-0.94	-0.29	0.77	0.78	-0.43	-0.97
MSE	-1.00	1.00	1.00	-0.64	-0.67	-0.68	0.75	0.94	0.29	-0.77	-0.78	0.43	0.97
MAE	-1.00	1.00	1.00	-0.63	-0.65	-0.66	0.76	0.93	0.28	-0.75	-0.75	0.42	0.97
SSIM	0.64	-0.64	-0.63	1.00	0.87	0.77	-0.65	-0.68	-0.69	0.84	0.84	-0.75	-0.73
Sobel IoU	0.67	-0.67	-0.65	0.87	1.00	0.79	-0.61	-0.71	-0.60	0.87	0.87	-0.72	-0.71
Canny IoU	0.68	-0.68	-0.66	0.77	0.79	1.00	-0.63	-0.69	-0.54	0.76	0.76	-0.66	-0.72
Dif. Hist.	-0.75	0.75	0.76	-0.65	-0.61	-0.63	1.00	0.70	0.39	-0.62	-0.62	0.50	0.74
GMSD	-0.94	0.94	0.93	-0.68	-0.71	-0.69	0.70	1.00	0.33	-0.82	-0.82	0.47	0.93
VIFP	-0.29	0.29	0.28	-0.69	-0.60	-0.54	0.39	0.33	1.00	-0.42	-0.42	0.93	0.40
FSIM	0.77	-0.77	-0.75	0.84	0.87	0.76	-0.62	-0.82	-0.42	1.00	1.00	-0.55	-0.79
FSIMc	0.78	-0.78	-0.75	0.84	0.87	0.76	-0.62	-0.82	-0.42	1.00	1.00	-0.55	-0.80
RECO	-0.43	0.43	0.42	-0.75	-0.72	-0.66	0.50	0.47	0.93	-0.55	-0.55	1.00	0.52
CIEDE 2000	-0.97	0.97	0.97	-0.73	-0.71	-0.72	0.74	0.93	0.40	-0.79	-0.80	0.52	1.00

senta correlação forte ou muito forte com onze das demais medidas. FSIM e FSIMc apresentam correlação exata entre si, e correlação forte com outras 9 das outras métricas. GMSD apresenta correlação muito forte com PSNR, MSE e Delta E (CIE 2000). PSNR e SSIM, duas medidas de qualidade bastante populares para descrever a qualidade de algoritmos de restauração, apresentam entre si uma correlação forte. PSNR e SSIM apresentam também uma correlação forte ou muito forte com as demais medidas, exceto VIFP e RECO.

Observa-se ainda que nos dados testados a maioria das medidas apresenta correlação forte ou muito forte com as demais. Nestes casos, observa-se valores-p menores do que o nível de significância de 0,05, indicando também que os coeficientes de correlação são significativos. Destoam desta avaliação as medidas VIFP e RECO, que apresentam correlação fraca ou moderada com as demais medidas, enquanto apresentam entre si correlação muito forte. VIFP e RECO apresentam entre si correlação significativa ($\alpha < 0,05$). Observa-se ainda que Delta E (CIE 2000), de Sharma et al. (2005), é a medida que apresenta maior correlação com as demais. Delta E (CIE 2000) apresenta correlação muito forte com PSNR, MSE, MAE e GMSD. Delta-E apresenta ainda correlação forte com SSIM, Diferença de histograma, e Intersecção sobre União dos operadores Canny e Sobel.

Em uma segunda análise, explora-se a correlação das métricas quando empregadas na avaliação de imagens já restauradas pelos métodos de restauração. Considera-se aqui as restauração feitas pelo método proposto, U-Net de Ronneberger et al. (2015), CAN de Chen et al. (2017), Abdullah-Al-Wadud et al. (2007), Ying et al. (2017b), Fu et al. (2015), AMSR de Lee et al. (2013), Petro et al. (2014), Dong et al. (2011) e Ying et al. (2017c) em um conjunto de 50 imagens extraídas por amostragem aleatória simples sem reposição do *dataset* a6300 (descrito na Seção 3.3.1). Faz-se uma análise individualizada para casos de sub e sobre-exposição.

Nesta segunda análise utilizou-se o Tau-b bicaudal de Kendall (Kendall, 1938, 1945) pelo método assintótico, com ajustes para empates. Segundo Newson (2002) e Croux & Dehon (2010) o teste de Kendall é mais robusto e mais eficiente que o teste de correlação de Spearman, sendo menos sensível a pontos fora da curva e assimetria. Ressalta-se que, para o conjunto testado, os resultados dos testes são semelhantes.

As Tabelas 6 e 7 apresentam respectivamente a estatística do teste de correlação de Kendall e o valor-p correspondente para um conjunto de 500 imagens obtidas a partir da restauração de 50 imagens subexpostas com 10 diferentes métodos. Se o valor-p está abaixo do nível de significância $\alpha = 0,05$ rejeita-se a hipótese nula de que as medidas são estatisticamente independentes. Observa-se que a hipótese nula é rejeitada para a maior parcela das métricas. São exceções as correlações da métrica VIFP com as métricas SSIM, intersecção de bordas por Sobel e Canny, FSIM e FSIMc. Também são exceções as correlações da métrica RECO com as métricas FSIM e FSIMc.

Com relação ao valor da estatística, observa-se que neste cenário que a maioria das métricas embora rejeitem a hipótese de que são estatisticamente independentes, apresentam correlação fraca ou muito fraca. Por outro lado, observa-se que MSE apresenta correlação muito forte com PSNR, MAE e CIEDE 200. MAE apresenta correlação muito forte com MSE e CIEDE 2000.

Tabela 6: τ de Kendall para qualidade de imagens subexpostas do *Dataset* A6300 considerando resultados de restauração

	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Dif. Hist.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
PSNR	1.000	-0.942	-0.893	0.655	0.287	0.212	-0.473	-0.599	-0.293	0.341	0.340	-0.350	-0.863
MSE	-0.942	1.000	0.935	-0.697	-0.315	-0.253	0.478	0.635	0.251	-0.387	-0.386	0.303	0.911
MAE	-0.893	0.935	1.000	-0.665	-0.295	-0.224	0.488	0.594	0.276	-0.351	-0.349	0.313	0.916
SSIM	0.655	-0.697	-0.665	1.000	0.439	0.383	-0.316	-0.670	-0.105	0.573	0.580	-0.180	-0.692
Sobel IoU	0.287	-0.315	-0.295	0.439	1.000	0.419	-0.145	-0.408	-0.012	0.555	0.558	-0.107	-0.316
Canny IoU	0.212	-0.253	-0.224	0.383	0.419	1.000	-0.147	-0.410	-0.033	0.514	0.517	-0.089	-0.245
Dif. Hist.	-0.473	0.478	0.488	-0.316	-0.145	-0.147	1.000	0.349	0.223	-0.140	-0.138	0.228	0.478
GMSD	-0.599	0.635	0.594	-0.670	-0.408	-0.410	0.349	1.000	0.059	-0.604	-0.599	0.121	0.597
VIFP	-0.293	0.251	0.276	-0.105	-0.012	-0.033	0.223	0.059	1.000	0.027	0.027	0.715	0.271
FSIM	0.341	-0.387	-0.351	0.573	0.555	0.514	-0.140	-0.604	0.027	1.000	0.960	-0.055	-0.367
FSIMc	0.340	-0.386	-0.349	0.580	0.558	0.517	-0.138	-0.599	0.027	0.960	1.000	-0.055	-0.370
RECO	-0.350	0.303	0.313	-0.180	-0.107	-0.089	0.228	0.121	0.715	-0.055	-0.055	1.000	0.316
CIEDE 2000	-0.863	0.911	0.916	-0.692	-0.316	-0.245	0.478	0.597	0.271	-0.367	-0.370	0.316	1.000

Tabela 7: Valor-p para o τ de Kendall para qualidade de imagens subexpostas do *Dataset* A6300 restauradas

	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Dif. Hist.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
PSNR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MSE	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MAE	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SSIM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sobel IoU	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.684	0.000	0.000	0.000	0.000
Canny IoU	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.247	0.000	0.000	0.002	0.000
Dif. Hist.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GMSD	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.040	0.000	0.000	0.000	0.000
VIFP	0.000	0.000	0.000	0.000	0.684	0.247	0.000	0.040	0.000	0.343	0.351	0.000	0.000
FSIM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.343	0.000	0.000	0.054	0.000
FSIMc	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.351	0.000	0.000	0.053	0.000
RECO	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.054	0.053	0.000	0.000
CIEDE 2000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

FSIM e FSIMc também apresentam entre si correlação muito forte. GMSD apresenta correlação moderada com PSNR, MSE, MAE, SSIM, intersecção de borda por Sobel e Canny, FSIM, FSIMc e CIEDE 200. SSIM apresenta correlação moderada com PSNR, MSE, MAE, intersecção de borda por Sobel, GMSD, FSIM, FSIMc e CIEDE 2000 mostrando-se uma boa métrica para a avaliação de qualidade de forma geral. Observa-se ainda que FSIM e intersecções de borda tem correlação moderada com várias das métricas avaliadas. RECO apresenta correlação forte apenas com VIFP.

As Tabelas 8 e 9 apresentam respectivamente a estatística do teste de correlação de Kendall e o valor-p correspondente para um conjunto de 500 imagens obtidas a partir da restauração de 50 imagens sobre-expostas com 10 diferentes métodos. Rejeita-se a hipótese nula de que as medidas são estatisticamente independentes quando o valor-p está abaixo do nível de significância $\alpha = 0,05$. Considerando somente as imagens sobre-expostas identifica-se que somente a métrica VIFP não rejeita a hipótese nula quando testada a correlação com as demais métricas testadas (exceto por VIFP e RECO).

Assim como ocorre com as imagens subexpostas, observa-se que em imagens sobre-expostas a maioria das métricas rejeita a hipótese de que são estatisticamente independentes. No entanto, destaca-se que entre as 80 combinações possíveis 47 apresentam correlação fraca ou muito fraca, 22 apresentam correlação moderada, 6 apresentam correlação forte e 3 apresentam correlação muito forte. Dentre as correlações muito fortes, destacam-se as correlações entre MAE, MSE e PSNR e a correlação entre FSIM e FSIMc. Novamente, em condições de sobre-exposição identifica-se que as métricas RECO e FSIM apresentam correlação forte entre si, mas pouca correlação com as demais.

Observa-se, a partir da avaliação das medidas de qualidade de imagem para amostras pareadas, uma coerência na identificação da similaridade para aplicação em imagens sub e sobre-expostas. Cabe ressaltar que os valores apresentados podem variar entre diferentes datasets. Observado o fato de que estas métricas foram propostas ao longo de mais de uma década, por diferentes equipes, utilizando dados e formas de avaliação distintas, pode-se concluir a partir do estudo realizado que a aplicação conjunta pode oferecer fortes indícios sobre a qualidade final das imagens restauradas.

No remanescente do texto, os resultados são apresentados e discutidos levando em consideração as medidas e testes mais adequados para distribuições não normais. Faz-se uma distinção entre os casos de imagens sub e sobre-expostas, avaliando os resultados dos modelos de forma separada. Esta separação é adequada uma vez que alguns dos algoritmos de restauração comparados foram desenvolvidos com foco em restauração somente de imagens subexpostas, resultando em performance ruim quando aplicados em imagens sobre-expostas. Os Apêndices 7.2.2, 7.3.2, 7.4.2 e 7.5.2 complementam os resultados aqui discutidos, apresentando os valores de média para imagens subexpostas e sobre-expostas.

Tabela 8: τ de Kendall para medidas de qualidade de imagens sobre-expostas do *Dataset* A6300 considerando resultados de restauração

	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Dif. Hist.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
PSNR	1.000	-0.925	-0.880	0.590	0.220	0.163	-0.485	-0.676	-0.009	0.296	0.286	-0.166	-0.817
MSE	-0.925	1.000	0.952	-0.640	-0.227	-0.176	0.502	0.734	-0.002	-0.334	-0.323	0.150	0.884
MAE	-0.880	0.952	1.000	-0.619	-0.208	-0.154	0.506	0.705	-0.005	-0.303	-0.291	0.146	0.883
SSIM	0.590	-0.640	-0.619	1.000	0.326	0.343	-0.331	-0.637	0.038	0.537	0.528	-0.107	-0.661
Sobel IoU	0.220	-0.227	-0.208	0.326	1.000	0.494	-0.219	-0.321	-0.061	0.430	0.433	-0.229	-0.217
Canny IoU	0.163	-0.176	-0.154	0.343	0.494	1.000	-0.154	-0.283	-0.027	0.492	0.492	-0.179	-0.168
Dif. Hist.	-0.485	0.502	0.506	-0.331	-0.219	-0.154	1.000	0.474	0.025	-0.209	-0.191	0.149	0.471
GMSD	-0.676	0.734	0.705	-0.637	-0.321	-0.283	0.474	1.000	-0.046	-0.468	-0.449	0.110	0.683
VIFP	-0.009	-0.002	-0.005	0.038	-0.061	-0.027	0.025	-0.046	1.000	0.065	0.071	0.620	0.003
FSIM	0.296	-0.334	-0.303	0.537	0.430	0.492	-0.209	-0.468	0.065	1.000	0.946	-0.099	-0.312
FSIMc	0.286	-0.323	-0.291	0.528	0.433	0.492	-0.191	-0.449	0.071	0.946	1.000	-0.100	-0.306
RECO	-0.166	0.150	0.146	-0.107	-0.229	-0.179	0.149	0.110	0.620	-0.099	-0.100	1.000	0.143
CIEDE 2000	-0.817	0.884	0.883	-0.661	-0.217	-0.168	0.471	0.683	0.003	-0.312	-0.306	0.143	1.000

Tabela 9: Valor-p para o τ de Kendall para medidas de qualidade de imagens sobre-expostas do *Dataset* A6300 restauradas

	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Dif. Hist.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
PSNR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.740	0.000	0.000	0.000	0.000
MSE	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.952	0.000	0.000	0.000	0.000
MAE	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.863	0.000	0.000	0.000	0.000
SSIM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.186	0.000	0.000	0.000	0.000
Sobel IoU	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.031	0.000	0.000	0.000	0.000
Canny IoU	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.351	0.000	0.000	0.000	0.000
Dif. Hist.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.378	0.000	0.000	0.000	0.000
GMSD	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.110	0.000	0.000	0.000	0.000
VIFP	0.740	0.952	0.863	0.186	0.031	0.351	0.378	0.110	0.000	0.023	0.013	0.000	0.904
FSIM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.023	0.000	0.000	0.001	0.000
FSIMc	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.013	0.000	0.000	0.001	0.000
RECO	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000
CIEDE 2000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.904	0.000	0.000	0.000	0.000

5.2 Resultados em Métricas de Qualidade de Imagem

Nesta seção, apresenta-se um comparativo do método proposto com os demais métodos de aprimoramento de imagem da literatura. Todos os resultados apresentados são sobre dados reservados para teste. A avaliação quantitativa inclui várias medições de qualidade de imagem, incluindo os clássicos relação sinal-ruído (PSNR), erro médio absoluto (MAE) píxel-a-píxel, erro médio quadrático (MSE) e similaridade estrutural (SSIM) (Wang et al., 2004), bem como os menos populares *Gradient Magnitude Similarity deviation* (GMSD) (Xue et al., 2014), intersecção de Sobel sobre união e diferença de histograma.

A maioria das técnicas aplicadas na análise estatística de dados são baseadas em modelos teóricos que pressupõe distribuição normal. Diante disso, a avaliação da normalidade da distribuição dos dados é primordial para a adequada descrição da amostra e sua análise inferencial. Verificando-se que os dados não seguem uma distribuição normal, o uso de técnicas estatísticas que tem normalidade como pressuposto incorre no enviesamento dos parâmetros e da inferência dos testes.

De forma geral, uma primeira avaliação da normalidade de um conjunto de dados pode ser realizada através da visualização de seu histograma e diagrama de caixas, identificando grandes assimetrias, descontinuidades de dados e picos multimodais. Como etapa seguinte, com o intuito de verificar a normalidade dos dados, pode-se aplicar o teste de Jarque-Bera (JB) (Jarque & Bera, 1980). O teste de Jarque-Bera verifica como o coeficiente de curtose e o coeficiente de assimetria de um conjunto de dados se ajustam aos de uma distribuição normal. Quanto mais próximo de 0 for a estatística JB, maior a probabilidade da distribuição ser normal. O teste tem como hipótese nula H_0 a normalidade. Desta forma, se o valor-p for menor do que um determinado nível de significância, rejeita-se a normalidade.

O teste JB de normalidade leva em consideração os valores de assimetria e curtose, que representam aspectos ligados à forma do histograma: desviado para a esquerda/direita (simetria) ou apiculado/achatado (curtose). Jarque-Bera utiliza como parâmetros os coeficientes de curtose 3 e assimetria 0. Desvios muito grandes, como, por exemplo, uma curtose acima de 4 e assimetria acima de 1, invalidam a avaliação dos erros-padrão e intervalos de confiança.

Neste trabalho, utiliza-se ainda o teste Anderson-Darling (Stephens, 1976) para verificar se a função densidade de probabilidade (FDP) dos dados segue uma distribuição normal. O teste tem como hipótese nula H_0 que a amostra é extraída de uma população que pertence a uma distribuição normal. Os valores críticos dependem da distribuição para a qual o teste é aplicado. Utiliza-se ainda o teste de Shapiro-Wilk (Shapiro & Wilk, 1965), que tem como hipótese nula H_0 a normalidade.

Os testes de hipótese para verificação da normalidade aplicados apontam resultados mistos. Uma avaliação inicial da normalidade dos dados utilizando os testes de Jarque-Bera (Jarque & Bera, 1980), Shapiro-Wilk (Shapiro & Wilk, 1965) e Anderson-Darling Stephens (1974) com nível de significância $\alpha = 0.05$ é apresentada nos Apêndices 7.2.1, 7.3.1, 7.4.1 e 7.5.1. Para a avaliação, os dados foram agrupados por *dataset* de imagens utilizados no comparativo, método

de restauração e indicador de qualidade de imagem.

As estatísticas paramétricas são sempre preferíveis aos métodos não-paramétricos por serem mais poderosas. Nos métodos não-paramétricos há perda de informação pois os dados são trabalhados em forma de *ranking*, perdendo a magnitude das observações. No entanto, identificada a não-normalidade dos dados, opta-se por utilizar, para fins de comparação, testes não-paramétricos, uma vez que os mesmos não supõem uma distribuição específica para a população.

5.2.1 Dataset Multi-Exposição baseado em FiveK (sintético)

Apresentam-se nesta Seção os resultados obtidos pelo método proposto e o comparativo com outros métodos aplicados em imagens subexpostas ou sobre-expostas sintéticas geradas a partir do *dataset* MIT-Adobe FiveK (Bychkovsky et al., 2011). Uma parcela significativa dos métodos utilizados no comparativo tem execução lenta (apresentam complexidade computacional não linear e/ou não exploram paralelismo de processamento). Por uma questão de viabilidade, não são utilizados todos os dados do conjunto de teste. Todos os valores apresentados são computados sobre uma amostra de 50 imagens selecionadas de maneira aleatória, sem substituição entre os dados reservados para teste. Todas as imagens processadas são processadas com valor de píxel no intervalo $[0; 1]$. Todos os métodos processam as mesmas imagens. Para modelos baseados em aprendizagem profunda, de forma a garantir a isonomia, o ajuste foi feito utilizando a mesma estratégia empregada no treinamento do modelo proposto (ver Seção 4.4) variando apenas o tamanho dos mini-lotes de forma que o processo fosse exequível no *hardware* utilizado. Os dados apresentados a seguir foram obtidos considerando 50 amostras aleatórias simples extraídas do *dataset* sem repetição.

Inicialmente, apresenta-se o resumo estatístico para os resultados obtidos pelo método proposto. A Tabela 10 apresenta resumidamente os resultados obtidos pelo método proposto quando aplicado para correção de imagens que apresentam subexposição. Os indicadores de qualidade são calculados sempre entre a saída do modelo de restauração avaliado e a imagem referência (não alterada). A tabela apresenta informações de média e mediana (50%). Em distribuições simétricas, os valores de ambas as estatísticas tendem a ser muito próximos. A média, no entanto, é mais afetada por valores extremos, sejam eles muito altos ou muito baixos. A comparação da média com a mediana fornece ainda uma medida da assimetria da distribuição.

Para imagens com subexposição, verifica-se que PSNR apresenta mediana e média altos e bastante próximos. Em termos de variabilidade, o intervalo interquartil (75% – 25%) baixo indica que os valores observados tendem a ficar próximos da medida de tendência central. Com relação às demais medidas de qualidade utilizadas observa-se que estas ficam bastante próximas do limite desejável, apresentando também pouca dispersão. Observa-se ainda que os valores de FSIM e FSIMc apresentam alguma discrepância, com FSIM apresentando valores mais altos.

Ainda em relação ao resumo estatístico apresentado, considerando imagens sobre-expostas, verifica-se, de maneira geral, que os resultados obtidos na restauração de imagens sobre-expostas

Tabela 10: Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens subexpostas do *dataset* Fivek

	PSNR ↑	MSE ↓	MAE ↓	SSIM ↑	Sobel IoU ↑	Canny IoU ↑	Hist. Diff. ↓	GMSD ↓	VIFP ↑	FSIM ↑	FSIMc ↑	RECO	CIEDE 2000 ↓
Média	24,675	0,004	0,053	0,938	0,851	0,711	5,585	0,711	0,783	0,980	0,978	0,939	6,599
D. Padrão	2,766	0,005	0,021	0,045	0,072	0,116	1,858	0,523	0,129	0,013	0,014	0,120	1,660
Mínimo	14,806	0,001	0,023	0,724	0,622	0,246	2,698	0,153	0,358	0,943	0,939	0,560	4,564
25%	23,312	0,002	0,040	0,934	0,826	0,668	4,597	0,426	0,710	0,975	0,972	0,911	5,671
Mediana	24,670	0,003	0,050	0,951	0,866	0,728	5,151	0,541	0,811	0,987	0,985	0,962	6,306
75%	26,447	0,005	0,059	0,962	0,901	0,799	6,269	0,875	0,873	0,990	0,988	1,014	7,076
Máximo	30,171	0,033	0,157	0,983	0,944	0,900	11,668	3,265	0,976	0,993	0,991	1,124	15,127

Tabela 11: Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens sobre-expostas do *dataset* Fivek

	PSNR ↑	MSE ↓	MAE ↓	SSIM ↑	Sobel IoU ↑	Canny IoU ↑	Hist. Diff. ↓	GMSD ↓	VIFP ↑	FSIM ↑	FSIMc ↑	RECO	CIEDE 2000 ↓
Média	20,164	0,015	0,082	0,891	0,724	0,487	5,757	3,624	0,762	0,926	0,919	0,758	9,445
D. Padrão	4,011	0,015	0,046	0,109	0,133	0,148	2,300	3,131	0,267	0,055	0,056	0,237	3,702
Mínimo	11,165	0,001	0,024	0,377	0,161	0,150	2,444	0,522	0,261	0,713	0,705	0,271	4,869
25%	17,305	0,005	0,050	0,860	0,675	0,398	3,764	1,483	0,580	0,904	0,894	0,568	6,543
Mediana	20,956	0,008	0,065	0,928	0,748	0,509	5,323	2,923	0,747	0,944	0,935	0,762	8,383
75%	22,852	0,019	0,108	0,955	0,817	0,593	6,976	4,352	0,950	0,963	0,958	0,890	11,828
Máximo	29,596	0,076	0,251	0,983	0,885	0,787	11,136	15,147	1,451	0,980	0,976	1,316	22,439

(Tabela 11) são inferiores aos obtidos na restauração de imagens subexpostas (Tabela 10). Os resultados apresentam também maior dispersão em torno da tendência central.

As medidas de dissimilaridade ou diferença, como MSE, MAE, diferença de histogramas, GMSD e CIEDE 2000, seguem o mesmo padrão. Os valores para estas medidas quando o método é aplicado para imagens sobre-expostas são mais altos do que aqueles observados quando o modelo de restauração é aplicado em imagens subexpostas indicando que o modelo apresenta melhor performance quando aplicado em imagens subexpostas. O desvio padrão amostral e intervalo interquartil apresentam uma dispersão maior em torno da tendência central, indicando que a qualidade da restauração obtida tem maior dispersão, sendo, portanto, menos confiável. As medidas VIFP e RECO refletem o mesmo padrão com valores mais distantes de 1 para as imagens sobre-expostas.

As Tabelas 12 e 13 apresentam um comparativo com outros métodos de restauração da literatura. Para cada medida de qualidade são identificados os três métodos de restauração que apresentam melhor performance, avaliados pela mediana. Compara-se o modelo proposto aos modelos baseados em redes neurais U-Net (Ronneberger et al., 2015) e CAN-24 (Chen et al., 2017). Compara-se também às abordagens clássicas (que não empregam aprendizado profundo) de aprimoramento de imagens de Abdullah-Al-Wadud et al. (2007), Dong et al. (2011), Lee et al. (2013), Petro et al. (2014), Fu et al. (2015), Ying et al. (2017b) e Ying et al. (2017c). Incluem-se ainda no comparativo as imagens não tratadas, de forma a permitir uma observação do ganho proporcionado pela aplicação dos modelos de restauração.

Para medidas que atribuem grande importância à regiões de borda ou alto gradiente (Sobel IoU, Canny IoU, GMSD, RECO) verifica-se que a maioria dos métodos de restauração comparados acaba piorando os resultados. Isto é mais evidente para imagens subexpostas, uma vez que somente o modelo proposto, o modelo de Ronneberger et al. (2015) e o modelo de Chen et al. (2017) apresentaram resultados melhores do que a imagem não processada, segundo estas medidas. Pode-se dizer que, para aplicações que dependem de detecção de bordas, linhas ou primitivas, o pré-processamento de imagens utilizando os métodos clássicos poderia, de fato, degradar o resultado final.

De forma geral observa-se que, tanto para imagens sub quanto sobre-expostas, os modelos baseados em redes neurais convolucionais superam os modelos baseados em processamento de imagens clássico. A Tabela 12 mostra ainda que os métodos de Ying et al. (2017b), Fu et al. (2015) e Petro et al. (2014) apresentam resultados semelhantes em termos de PSNR e SSIM. Apesar disto, os métodos baseados em CNN oferecem notável vantagem. Condição semelhante é observada entre as medidas que atribuem grande relevância à correção de cor, MSE, MAE, FSIMc e CIEDE 2000.

Para as imagens sobre-expostas, observa-se que todos os métodos apresentam performance inferior à observada em imagens subexpostas. Somente o método de Lee et al. (2013) apresenta comportamento distinto, com performance superior em algumas métricas. Os resultados mostram ainda que o modelo proposto supera com alguma margem os modelos de redes convolucionais anteriores Ronneberger et al. (2015); Chen et al. (2017) treinadas com as mesmas

Tabela 12: Mediana para restauração de imagens subexpostas geradas a partir do *dataset FiveK* (sintético). Os três melhores resultados para cada medida de qualidade são sublinhados

Método	PSNR↑	MSE↓	MAE↓	SSIM↑	Sobel IoU↑	Canny IoU↑	Hist. Diff.↓	GMSD↓	VIFP↑	FSIM↑	FSIMc↑	RECO	CIEDE 2000↓
M. Proposto	<u>24,670</u>	<u>0,003</u>	<u>0,050</u>	<u>0,951</u>	<u>0,866</u>	<u>0,728</u>	<u>5,151</u>	<u>0,541</u>	<u>0,811</u>	<u>0,987</u>	<u>0,985</u>	<u>0,962</u>	<u>6,306</u>
Ronneberger et al. (2015)	<u>26,118</u>	<u>0,002</u>	<u>0,042</u>	<u>0,913</u>	0,766	<u>0,653</u>	<u>5,219</u>	<u>0,589</u>	<u>0,977</u>	<u>0,982</u>	<u>0,980</u>	<u>0,916</u>	<u>5,575</u>
Chen et al. (2017)	<u>24,453</u>	<u>0,004</u>	<u>0,048</u>	<u>0,932</u>	<u>0,817</u>	0,642	6,053	1,045	<u>0,918</u>	0,975	0,968	<u>1,009</u>	8,931
Abdullah-Al-Wadud et al. (2007)	18,089	0,016	0,100	<u>0,787</u>	0,616	0,323	6,329	4,088	0,404	0,906	0,900	0,565	11,615
Ying et al. (2017b)	19,636	0,011	0,084	0,813	0,661	0,437	5,769	3,625	0,465	0,921	0,917	0,486	9,027
Fu et al. (2015)	19,814	0,010	0,084	0,807	0,676	0,400	5,686	3,098	0,434	0,927	0,924	0,550	9,627
Lee et al. (2013)	7,452	0,180	0,373	0,017	0,020	0,001	11,719	24,798	0,625	0,567	0,552	-11,788	33,413
Petro et al. (2014)	19,148	0,012	0,096	<u>0,777</u>	0,731	0,561	<u>5,622</u>	3,151	0,616	0,946	0,931	0,582	10,266
Dong et al. (2011)	16,524	0,023	0,113	<u>0,728</u>	0,533	0,293	6,592	6,256	0,321	0,848	0,843	0,384	12,682
Ying et al. (2017c)	15,630	0,027	0,133	<u>0,753</u>	0,589	0,315	6,645	8,938	0,318	0,863	0,853	0,387	14,288
Não Tratada	19,648	0,011	0,093	<u>0,778</u>	<u>0,851</u>	<u>0,718</u>	5,854	<u>0,836</u>	0,801	<u>0,980</u>	<u>0,978</u>	0,802	<u>8,030</u>

Tabela 13: Mediana para restauração de imagens sobre-expostas geradas a partir do *dataset FiveK* (sintético)

Método	PSNR↑	MSE↓	MAE↓	SSIM↑	Sobel IoU↑	Canny IoU↑	Hist. Diff.↓	GMSD↓	VIFP↑	FSIM↑	FSIMc↑	RECO	CIEDE 2000↓
M. Proposto	<u>20,956</u>	<u>0,008</u>	<u>0,065</u>	<u>0,928</u>	<u>0,748</u>	<u>0,509</u>	<u>5,323</u>	<u>2,923</u>	<u>0,747</u>	<u>0,944</u>	<u>0,935</u>	<u>0,762</u>	<u>8,383</u>
Ronneberger et al. (2015)	<u>18,936</u>	<u>0,013</u>	<u>0,083</u>	<u>0,857</u>	<u>0,632</u>	<u>0,463</u>	<u>5,395</u>	<u>3,986</u>	<u>0,895</u>	<u>0,921</u>	<u>0,913</u>	<u>1,158</u>	<u>9,890</u>
Chen et al. (2017)	<u>20,225</u>	<u>0,010</u>	<u>0,073</u>	<u>0,914</u>	0,719	<u>0,479</u>	<u>5,179</u>	<u>2,267</u>	<u>0,993</u>	<u>0,947</u>	<u>0,939</u>	<u>1,237</u>	<u>9,267</u>
Abdullah-Al-Wadud et al. (2007)	12,704	0,054	0,181	<u>0,754</u>	0,538	0,267	6,312	11,706	0,364	0,813	0,808	0,355	16,993
Ying et al. (2017b)	10,105	0,098	0,278	<u>0,715</u>	0,586	0,320	8,230	8,269	0,474	0,868	0,862	0,636	26,456
Fu et al. (2015)	10,928	0,081	0,246	<u>0,757</u>	0,608	0,309	7,479	7,400	0,429	0,867	0,861	0,627	25,017
Lee et al. (2013)	15,522	0,028	0,123	0,838	0,563	0,235	6,688	8,530	1,636	0,861	0,856	1,912	12,859
Petro et al. (2014)	13,551	0,044	0,167	<u>0,794</u>	0,568	0,342	6,303	8,543	0,390	0,851	0,839	0,425	16,853
Dong et al. (2011)	16,672	0,085	0,262	<u>0,697</u>	0,503	0,276	7,792	9,109	0,386	0,814	0,804	0,440	26,128
Ying et al. (2017c)	9,794	0,105	0,306	<u>0,681</u>	0,564	0,301	8,456	9,021	0,463	0,863	0,856	0,654	29,297
Não Tratada	12,280	0,059	0,191	<u>0,823</u>	<u>0,638</u>	0,385	6,855	6,229	0,452	0,893	0,888	0,584	20,056

Tabela 14: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do *dataset* FiveK (valor-p)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	0,606	0,208	0,612	0,000	0,000	0,000	0,546	0,017	0,000	0,000	0,000	0,184	0,191
Chen et al. (2017)	0,022	0,004	0,026	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,016	0,000
Abdullah-Al-Wadud et al. (2007)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017b)	0,000	0,000	0,000	0,000	0,000	0,000	0,572	0,000	0,000	0,000	0,000	0,000	0,000
Fu et al. (2015)	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000
Lee et al. (2013)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,988	0,000	0,000	0,019	0,000
Petro et al. (2014)	0,000	0,000	0,000	0,000	0,000	0,000	0,454	0,000	0,000	0,000	0,000	0,000	0,000
Dong et al. (2011)	0,000	0,000	0,000	0,000	0,000	0,000	0,007	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017c)	0,000	0,000	0,000	0,000	0,000	0,000	0,008	0,000	0,000	0,000	0,000	0,000	0,000
Não Tratada	0,001	0,000	0,000	0,000	0,060	0,626	0,043	0,001	0,052	0,000	0,000	0,000	0,004

Tabela 15: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do *dataset* FiveK (valor-p)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	0,138	0,443	0,194	0,000	0,000	0,000	0,478	0,010	0,000	0,000	0,000	0,000	0,124
Chen et al. (2017)	0,599	0,138	0,592	0,783	0,004	0,017	0,349	0,241	0,000	0,012	0,016	0,000	0,357
Abdullah-Al-Wadud et al. (2007)	0,000	0,000	0,000	0,000	0,000	0,000	0,004	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017b)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000
Fu et al. (2015)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Lee et al. (2013)	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000
Petro et al. (2014)	0,000	0,000	0,000	0,000	0,000	0,000	0,005	0,000	0,000	0,000	0,000	0,000	0,000
Dong et al. (2011)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017c)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,011	0,000
Não Tratada	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000

condições.

Ainda dentro da análise numérica, foram computadas as estatísticas não paramétricas Qui-Quadrado de Friedman (Friedman, 1937, 1939) e Teste de Postos Sinalizados de Wilcoxon bicaudal (Wilcoxon, 1992). O teste de Friedman, adequado para a comparação de três grupos ou mais dependentes por meio de uma variável qualitativa ordinal ou quantitativa sem distribuição normal, tem como hipótese nula que as medianas populacionais são iguais. O processamento se dá da seguinte forma: para cada métrica verifica-se se existe diferença significativa entre os resultados da restauração. No Qui-Quadrado de Friedman todas as estatísticas tiveram nível de significância inferior a 10^3 evidenciando que as diferenças são estatisticamente significativas.

Uma vez que se verifica que existe uma diferença entre os resultados obtidos por cada um dos métodos de restauração, faz-se uma análise pareada dos modelos utilizando o Teste de Postos Sinalizados de Wilcoxon. Este teste substitui o teste t de Student quando os dados não apresentam distribuição normal. O Teste de Postos Sinalizados de Wilcoxon verifica se o tratamento A (imagem restaurada utilizando o modelo de rede neural proposto) produz valores maiores do que o tratamento B (imagem restaurada utilizando o modelo comparado).

A Tabela 14 apresenta o valor-p para o teste de Wilcoxon para imagens subexpostas. Utilizando-se um nível de significância $\alpha = 0,05$ pode-se dizer que existe diferença estatisticamente significativa entre os resultados produzidos pelo método proposto e a imagem não tratada em todas as métricas avaliadas, exceto pela intersecção dos operadores Canny e Sobel e pela métrica VIFP. No caso dos operadores de borda Canny e Sobel, esta é uma propriedade que indica que as transições de intensidade estão sendo preservadas e que os efeitos indesejáveis de deslocamento ou atenuação de borda são pouco impactantes na restauração modelo proposto. Em uma análise por métrica, verifica-se que a hipótese nula do teste de Wilcoxon não é rejeitada para 3 dos modelos de restauração comparada, indicando que estes métodos podem produzir distribuições de intensidade similares aos do modelo proposto. Observa-se ainda que o modelo U-net³ (Ronneberger et al., 2015) apresenta resultados similares (diferença se deve ao acaso, não rejeita a hipótese nula) aos do modelo proposto nas métricas PSNR, MSE, SSIM, RECO e CIEDE 2000.

A Tabela 15 apresenta o valor-p para o teste de Wilcoxon para imagens sobre-expostas. Nesta condição observa-se que a hipótese nula é rejeitada para todas as métricas em todos os comparativos exceto pelos modelos CAN (Chen et al., 2017) e U-Net (Ronneberger et al., 2015). Neste rejeita-se a hipótese nula de que os dados pertencem a populações diferentes nas métricas PSNR, MSE, MAE, diferença de histogramas e CIEDE 2000. O valor da estatística para o Teste de Postos Sinalizados de Wilcoxon é disponibilizado no Apêndice 7.2.3.

Qualitativamente, a Figura 5.1 apresenta as saídas de diferentes modelos para uma imagem subexposta. Nota-se uma melhora significativa na visibilidade dos elementos da cena, na restauração da textura e na re-coloração. Mesmo em regiões nas quais todos os três canais são

³U-Net (Ronneberger et al., 2015) e CAN (Chen et al., 2017) são modelos baseados em aprendizagem profunda. Os modelos foram ajustados utilizando o mesmo conjunto de dados e procedimentos empregados para o modelo proposto.

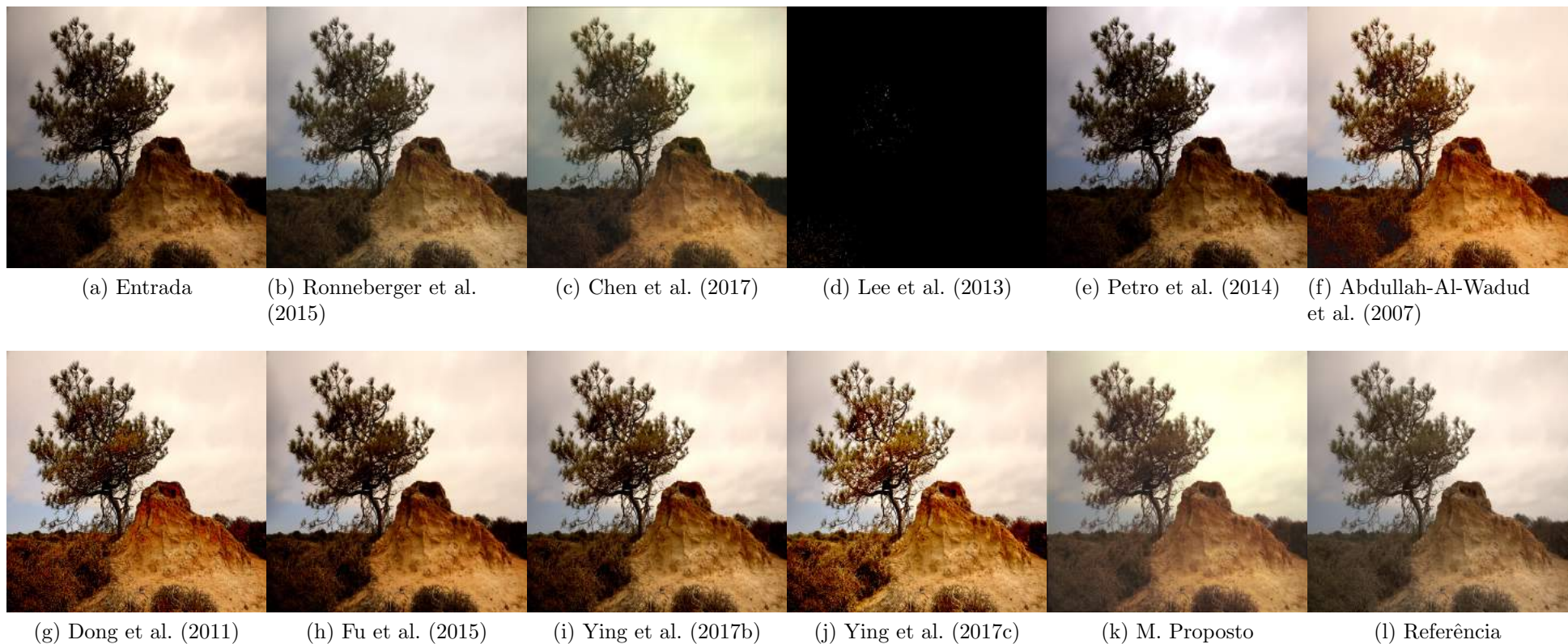


Figura 5.1: Resultados da restauração atingida por diferentes métodos para restauração de subexposição sintética. A cena apresenta alto contraste, com regiões muito escuras na parte inferior e regiões muito claras na parte superior. O modelo proposto apresenta cores mais próximas da imagem referência, preservando a textura e visibilidade dos elementos da cena

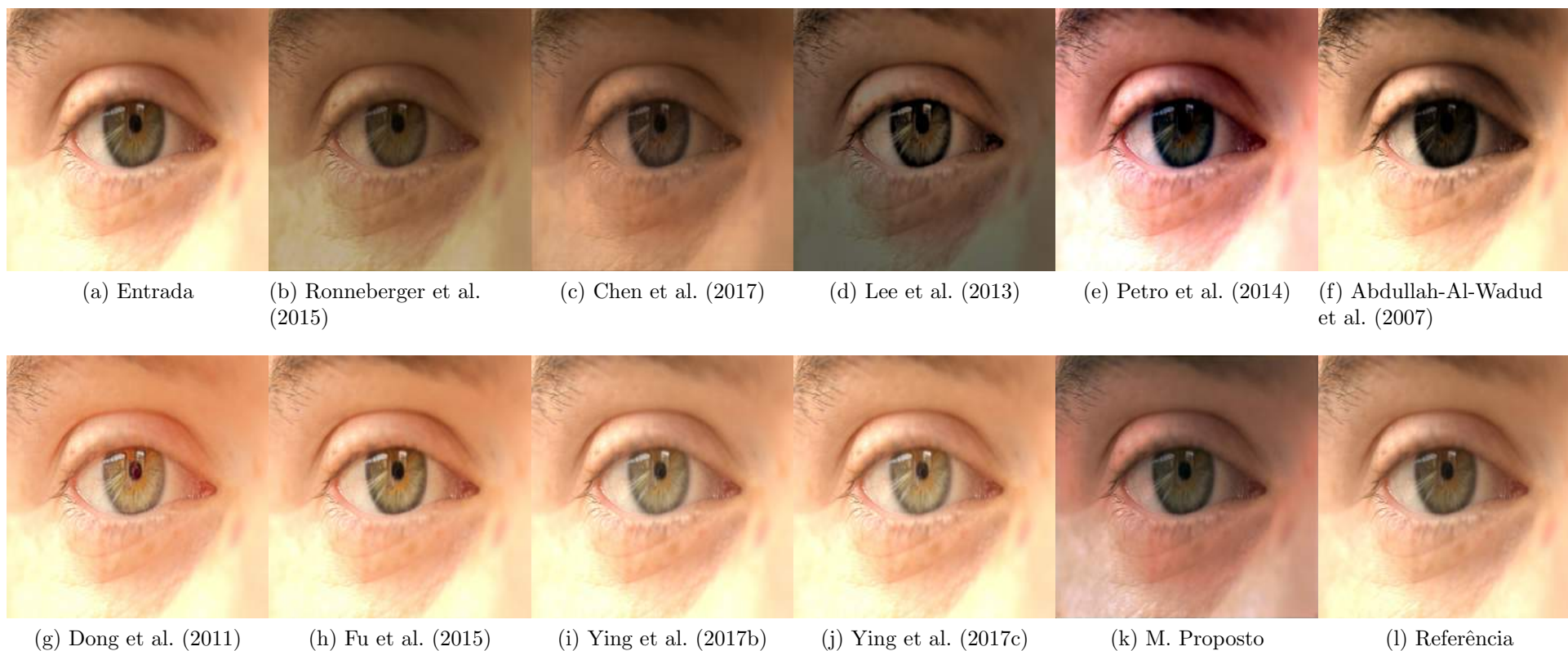


Figura 5.2: Resultados da restauração atingida por diferentes métodos para restauração na sobre-exposição sintética. Cena apresenta alto contraste, com regiões escuras na parte central e regiões muito claros na parte inferior. O modelo proposto resulta em cores distintas da imagem referência, porém preserva a textura e visibilidade dos elementos da cena, além de manter um aspecto mais natural que as demais restaurações

afetados por subexposição, observa-se que o modelo proposto consegue restaurar a suavidade da superfície. Uma comparação qualitativa mais abrangente com os métodos de referência pode ser encontrada no material suplementar (Apêndice 7.7).

A Figura 5.2 apresenta um comparativo do modelo proposto com os demais modelos da literatura na restauração de uma imagem sobre-exposta sintética. A imagem apresenta alto contraste, com pixels muito escuros na parte central e regiões muito claros na parte inferior. Para esta cena, os melhores resultados de restauração são atingidos utilizando modelos baseados em aprendizagem. Entre as técnicas clássicas, verifica-se um padrão de transformação de cor e preservação de blocos saturados, efeitos indesejáveis para a aplicação em questão. O modelo proposto resulta em cores distintas da imagem referência, no entanto preserva a textura e visibilidade dos elementos da cena. Salienta-se também que a imagem referência apresenta pontos de próximos da saturação na parte inferior fazendo com que o resultado atingido pelo modelo proposto ofereça maior visibilidade dos detalhes.

O modelo proposto mantém um aspecto mais natural que as demais restaurações, mostrando-se mais agradável aos olhos que a imagem referência. Este resultado pode ser atribuído ao processo de aprendizagem, no qual o modelo tem acesso à milhares de imagens de referência, aprendendo a identificar características inerentes às imagens apropriadamente expostas. Ao utilizar treinamento completamente supervisionado, o modelo aprende a maximizar o resultado médio. É importante ressaltar que a qualidade percebida por humanos em uma imagem está associada com as características únicas do indivíduo que as está observando e, portanto, não existe uma única medida de qualidade que possa representá-la. Em alguns casos, a imagem referência utilizada para a validação quantitativa dos resultados pode ser diferente da imagem que um observador humano considerasse ideal.

5.2.2 HDR+ *Burst Photography Dataset* (sintético)

Discutem-se nesta Seção os resultados da aplicação do método proposto sobre o conjunto de imagens sintéticas geradas a partir do *dataset* HDR+*burst* (Hasinoff et al., 2016). O procedimento de transformação para imagens utilizadas na avaliação é descrito na Seção 3.3.4. Este conjunto de dados se diferencia do anterior uma vez que utiliza imagens resultantes de um processo de composição a partir de múltiplas fotografias com distintos tempos de exposição, comprimidos no formato JPEG.

Além dos efeitos adversos de amostragem, quantização e *clipping* inerentes à aquisição de imagens, o processo de composição e compressão introduz artefatos de imagem. Estes artefatos são, em geral, relacionados ao alinhamento de múltiplas imagens da mesma cena, transformação de cores e aparência não natural, inconsistência decorrente de iluminação não homogênea e artefatos de blocos resultantes do algoritmo de compressão aplicado. Os dados apresentados a seguir foram obtidos considerando 50 amostras aleatórias simples extraídas do *dataset* sem repetição.

Um resumo estatístico dos resultados atingidos pelo método é disponibilizado nas Tabelas

Tabela 16: Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens subexpostas do dataset HDR+*burst*

	PSNR↑	MSE↓	MAE↓	SSIM↑	Sobel IoU ↑	Canny IoU ↑	Hist. Diff. ↓	GMSD↓	VIFP↑	FSIM↑	FSIMc↑	RECO	CIEDE 2000 ↓
Média	28,090	0,004	0,038	0,951	0,873	0,757	4.836	0.788	0,915	0,980	0,978	0,919	4,738
D. Padrão	4,412	0,012	0,038	0,049	0,079	0,095	1.791	1.440	0,131	0,023	0,023	0,102	3,171
Mínimo	10,796	0,000	0,016	0,684	0,431	0,328	2.686	0.205	0,334	0,835	0,831	0,472	2,486
25%	26,533	0,001	0,021	0,944	0,854	0,717	3.320	0.321	0,875	0,979	0,978	0,879	3,336
Mediana	29,007	0,001	0,029	0,967	0,899	0,779	4.248	0.433	0,940	0,985	0,983	0,935	3,838
75%	31,494	0,002	0,040	0,976	0,918	0,814	5.992	0.705	0,986	0,989	0,988	0,975	4,934
Máximo	33,096	0,083	0,262	0,987	0,958	0,873	10.084	10.233	1,094	0,994	0,993	1,114	23,021

Tabela 17: Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens sobre-expostas do dataset HDR+*burst*

	PSNR↑	MSE↓	MAE↓	SSIM↑	Sobel IoU ↑	Canny IoU ↑	Hist. Diff. ↓	GMSD↓	VIFP↑	FSIM↑	FSIMc↑	RECO	CIEDE 2000 ↓
Média	22,424	0,008	0,063	0,939	0,804	0,644	5,795	2,005	1,208	0,964	0,956	1,298	8,621
D. Padrão	3,441	0,011	0,041	0,031	0,079	0,092	1,947	1,537	0,312	0,018	0,018	0,284	2,775
Mínimo	11,198	0,002	0,030	0,853	0,508	0,362	2,879	0,311	0,581	0,896	0,889	0,594	4,630
25%	20,358	0,004	0,041	0,922	0,776	0,616	4,155	0,882	0,968	0,959	0,951	1,118	7,157
Mediana	23,116	0,005	0,048	0,945	0,817	0,654	5,767	1,474	1,157	0,966	0,959	1,271	8,114
75%	24,191	0,009	0,072	0,962	0,868	0,704	6,994	2,608	1,421	0,975	0,968	1,488	9,092
Máximo	28,107	0,076	0,270	0,984	0,909	0,809	11,024	7,065	1,959	0,988	0,983	1,861	24,143

16 e 17. Na restauração de imagens subexpostas observa-se assimetria nos resultados, com valores de média e mediana divergindo entre si. Para as medidas PSNR, SSIM, Sobel IoU, Canny IoU, VIFP, FSIM e FSIMc uma mediana mais elevada que a média indica que os valores na parte de baixo da distribuição são mais distantes da tendência central que os valores no topo da distribuição. Ou seja, existe uma assimetria na distribuição que faz com que a maior parte dos valores esteja concentrada acima da média. De fato, esta assimetria fica mais evidente se observados os valores do primeiro e terceiro quartil. Para as medidas em que valores menores correspondem a um aumento de similaridade entre uma imagem restaurada e sua referência, observa-se comportamento similar. Nestes casos, o valor da mediana apresenta-se inferior ao valor da média.

Ao passo que o modelo apresenta performance média levemente superior neste *dataset* se comparado ao dataset FiveK, percebe-se nestes resultados um aumento da variabilidade e maior erro nas medidas voltadas para cor. Este comportamento tem sentido, na medida em que se leva em consideração o efeito da compressão em blocos realizada no padrão JPEG. Esta compressão frequentemente resulta na atenuação de gradiente da imagem dentro do bloco e consequente perda de informação. Uma vez que os valores de intensidade no píxel são perdidos, o método tem mais dificuldade em interpolar a partir de sua vizinhança.

A Tabela 18 apresenta o comparativo da restauração utilizando o método proposto com os demais. Na medida PSNR observa-se uma ampla vantagem do método proposto. O mesmo ocorre com MSE, MAE, SSIM, GMSD e CIEDE 2000. Merecem também destaque os modelos baseados em redes convolucionais propostos por Ronneberger et al. (2015) e Chen et al. (2017). Entre os métodos baseados em processamento de imagens clássico, destaca-se Petro et al. (2014) apresentando bons resultados nas medidas Sobel IoU e diferença de histogramas.

Vale notar que, para as imagens subexpostas, a transformação aplicada na geração de imagens subexpostas resulta em um erro médio muito pequeno, deixando pouco espaço para os métodos de aprimoramento de imagem. As medidas relacionadas à cor e gradiente refletem esta afirmação, visto que em oito das medidas testadas, o conjunto de imagens não processadas compõe a relação dos três melhores. Assim como reportado anteriormente, a aplicação de métodos de restauração leva, em vários casos, à uma piora das condições da imagem.

A Tabela 19 apresenta o comparativo para imagens sobre-expostas. Verifica-se aqui um maior espaço para melhoria, uma vez que as deformações aplicadas na imagem sobre-exposta sintética resultaram em uma diferença mais relevante com relação à imagem referência. Para este conjunto de dados, a imagem não processada compõe a lista dos três melhores somente na avaliação da intersecção sobre união dos filtros de borda.

Assim como nos demais cenários avaliados, os modelos de restauração baseados em redes neurais convolucionais apresentaram desempenho superior aos modelos baseados em técnicas de processamento de imagem clássicas. Dentre os modelos de restauração comparados, o modelo proposto apresenta o maior PSNR, SSIM, Sobel IoU, canny IoU, FSIM e FSIMc. O modelo proposto apresenta ainda menor MSE, MAE e GMSD. VIFP e RECO indicam também que as imagens restauradas pelo modelo são as que mais se aproximam das imagens referência. A

Tabela 18: Mediana para restauração de imagens subexpostas geradas a partir do dataset HDR+*burst* (sintético). Os três melhores resultados para cada medida de qualidade são sublinhados

Método	PSNR $\uparrow$	MSE $\downarrow$	MAE $\downarrow$	SSIM $\uparrow$	Sobel IoU $\uparrow$	Canny IoU $\uparrow$	Hist. Diff. $\downarrow$	GMSD $\downarrow$	VIFP $\uparrow$	FSIM $\uparrow$	FSIMc $\uparrow$	RECO	CIEDE 2000 $\downarrow$
M. Proposto	<u>29,007</u>	<u>0,001</u>	<u>0,029</u>	<u>0,967</u>	<u>0,899</u>	<u>0,779</u>	<u>4,248</u>	<u>0,433</u>	<u>0,940</u>	<u>0,985</u>	<u>0,983</u>	<u>0,935</u>	<u>3,838</u>
Ronneberger et al. (2015)	<u>22,757</u>	<u>0,005</u>	<u>0,055</u>	<u>0,864</u>	<u>0,723</u>	<u>0,658</u>	<u>4,345</u>	<u>0,729</u>	<u>1,017</u>	<u>0,968</u>	<u>0,966</u>	<u>1,059</u>	<u>5,977</u>
Chen et al. (2017)	21,482	0,007	<u>0,070</u>	<u>0,908</u>	0,764	0,600	5,756	1,093	<u>1,025</u>	0,959	0,951	<u>1,112</u>	8,740
Abdullah-Al-Wadud et al. (2007)	19,201	0,012	0,093	<u>0,857</u>	0,693	0,440	6,325	2,849	0,501	0,929	0,925	0,710	9,495
Ying et al. (2017b)	20,896	0,008	0,073	0,878	0,721	0,524	5,305	3,360	0,541	0,944	0,942	0,561	7,490
Fu et al. (2015)	20,763	0,008	0,076	<u>0,866</u>	0,715	0,475	5,580	3,058	0,510	0,942	0,940	0,600	7,953
Lee et al. (2013)	7,022	0,199	0,380	<u>0,019</u>	0,017	0,000	11,719	30,025	0,897	0,441	0,429	-63,905	32,674
Petro et al. (2014)	21,196	0,008	0,072	0,860	<u>0,788</u>	0,623	4,879	1,682	0,637	0,962	0,951	0,665	8,032
Dong et al. (2011)	18,273	0,015	0,092	0,797	0,632	0,387	6,167	5,464	0,439	0,887	0,880	0,477	9,093
Ying et al. (2017c)	14,973	0,032	0,148	0,778	0,631	0,364	6,529	11,636	0,367	0,862	0,858	0,436	14,567
Não Tratada	<u>22,459</u>	<u>0,006</u>	0,069	0,859	<u>0,878</u>	<u>0,776</u>	4,962	<u>0,566</u>	0,868	<u>0,978</u>	<u>0,977</u>	0,897	<u>6,036</u>

Tabela 19: Mediana para restauração de imagens sobre-expostas geradas a partir do dataset HDR+*burst* (sintético). Os três melhores resultados para cada medida de qualidade são sublinhados

Método	PSNR $\uparrow$	MSE $\downarrow$	MAE $\downarrow$	SSIM $\uparrow$	Sobel IoU $\uparrow$	Canny IoU $\uparrow$	Hist. Diff. $\downarrow$	GMSD $\downarrow$	VIFP $\uparrow$	FSIM $\uparrow$	FSIMc $\uparrow$	RECO	CIEDE 2000 $\downarrow$
M. Proposto	23,116	<u>0,005</u>	<u>0,048</u>	<u>0,945</u>	<u>0,817</u>	<u>0,654</u>	5,767	1,474	1,157	0,966	0,959	1,271	8,114
Ronneberger et al. (2015)	<u>20,895</u>	<u>0,008</u>	<u>0,064</u>	<u>0,854</u>	0,644	0,543	5,398	2,689	<u>1,362</u>	<u>0,951</u>	<u>0,946</u>	<u>1,274</u>	<u>7,503</u>
Chen et al. (2017)	<u>20,714</u>	<u>0,008</u>	<u>0,066</u>	<u>0,920</u>	<u>0,766</u>	<u>0,582</u>	<u>5,493</u>	<u>1,808</u>	<u>1,183</u>	<u>0,954</u>	<u>0,937</u>	<u>1,312</u>	<u>9,880</u>
Abdullah-Al-Wadud et al. (2007)	13,099	0,049	0,191	0,799	0,660	0,439	6,758	7,763	0,462	0,893	0,888	0,469	17,148
Ying et al. (2017b)	11,181	0,076	0,251	0,738	0,665	0,451	7,669	7,709	0,503	0,890	0,884	0,530	21,799
Fu et al. (2015)	11,928	0,064	0,230	0,786	0,685	0,438	7,011	7,488	0,482	0,902	0,891	0,528	20,092
Lee et al. (2013)	14,937	0,032	0,141	0,754	0,440	0,199	7,923	9,315	1,719	0,800	0,798	3,029	14,452
Petro et al. (2014)	13,993	0,040	0,164	<u>0,854</u>	0,716	0,557	6,423	6,163	0,508	0,922	0,918	0,492	13,954
Dong et al. (2011)	12,972	0,068	0,231	<u>0,722</u>	0,602	0,385	7,585	9,108	0,433	0,844	0,837	0,406	20,719
Ying et al. (2017c)	9,818	0,104	0,299	0,690	0,624	0,391	7,528	9,634	0,459	0,857	0,849	0,533	26,633
Não Tratada	13,918	0,041	0,163	0,853	<u>0,734</u>	<u>0,564</u>	6,521	5,589	0,534	0,928	0,921	0,540	13,928

Tabela 20: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do *dataset* HDR+*burst* (valor-p)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	0,000	0,000	0,000	0,000	0,000	0,000	0,457	0,000	0,000	0,000	0,000	0,000	0,000
Chen et al. (2017)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,000	0,000	0,000	0,000
Abdullah-Al-Wadud et al. (2007)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017b)	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000
Fu et al. (2015)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Lee et al. (2013)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,851	0,000	0,000	0,000	0,000
Petro et al. (2014)	0,000	0,000	0,000	0,000	0,000	0,000	0,012	0,000	0,000	0,000	0,000	0,000	0,000
Dong et al. (2011)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017c)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Não Tratada	0,000	0,000	0,000	0,000	0,004	0,299	0,000	0,010	0,000	0,000	0,000	0,000	0,000

Tabela 21: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do *dataset* HDR+*burst* (valor-p)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	0,000	0,000	0,000	0,000	0,000	0,000	0,318	0,000	0,001	0,000	0,000	0,277	0,041
Chen et al. (2017)	0,000	0,003	0,010	0,000	0,000	0,000	0,008	0,000	0,008	0,000	0,000	0,131	0,000
Abdullah-Al-Wadud et al. (2007)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017b)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Fu et al. (2015)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Lee et al. (2013)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Petro et al. (2014)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Dong et al. (2011)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017c)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Não Tratada	0,000	0,000	0,000	0,000	0,000	0,005	0,001	0,000	0,000	0,000	0,000	0,000	0,000

métrica CIEDE 2000, por outro lado, aponta que a restauração feita através do modelo de U-Net, adaptado de Ronneberger et al. (2015), produz os melhores resultados em termos de acurácia de cor.

Assim como para o *dataset* Adobe-MIT FiveK fez-se a verificação dos níveis de significância utilizando as estatísticas não paramétricas Qui-Quadrado de Friedman (Friedman, 1937, 1939) e Teste de Postos Sinalizados de Wilcoxon bicaudal (Wilcoxon, 1992). Seguiu-se o mesmo procedimento reportado na Seção 5.2.1. Novamente o teste de Friedman indica que as diferenças entre os tratamentos são estatisticamente significativas. O teste de Wilcoxon em imagens subexpostas, apresentado na Tabela 20 indicam que não existe diferença significativa na intersecção de Canny entre a imagem não tratada e a imagem não processada. Esta é uma propriedade desejável, uma vez que indica que o modelo preservou as transições abruptas de intensidade na imagem de entrada e gerou resultados similares aos da imagem referência.

De forma geral, para imagens subexpostas, ao combinar os resultados apresentados nas Tabelas 18 com os testes de significância apresentados na tabela 20 observa-se que o modelo proposto produz resultados melhores e estatisticamente significantes em todas as métricas, exceto diferença de histogramas com o modelo U-Net (Ronneberger et al., 2015) e VIFP com a técnica de AMSR de Lee et al. (2013). Nas imagens sobre-expostas, combinando as Tabelas 19 e 21 também observam-se diferenças estatisticamente significantes. Nesta condição, somente na métrica RECO verifica-se que a hipótese nula é rejeitada para o comparativo com os modelos baseados em redes convolucionais Ronneberger et al. (2015) e Chen et al. (2017). As estatísticas do teste de Wilcoxon são apresentadas no Apêndice 7.3.3.

A Figura 5.3 apresenta os resultados para restauração de imagem subexposta obtida com diferentes métodos. Dentre os modelos baseados em redes convolucionais, os resultados do modelo proposto superam os obtidos com a arquitetura de rede de Ronneberger et al. (2015) e Chen et al. (2017). Ronneberger et al. (2015) e Chen et al. (2017) mostram expressivo borramento nas regiões de borda e atenuação geral da cor. Os métodos baseados em processamento de imagens clássicos resultam em cores mais vibrantes, mas aumentam também o ruído.

A Figura 5.4, apresenta os resultados da restauração de imagem severamente saturada gerada a partir do *dataset* HDR+*burst*. Na região central da imagem de entrada, observam-se vários blocos com saturação total dos valores de intensidade. Os modelos de restauração clássicos (Lee et al., 2013; Petro et al., 2014; Abdullah-Al-Wadud et al., 2007; Dong et al., 2011; Fu et al., 2015; Ying et al., 2017b,c) mostram-se insuficientes para o tratamento da saturação. Dentre os modelos baseados em aprendizagem, o modelo proposto se sobressai aos demais por apresentar maior correção de cor, menor incidência de artefatos de bloco (especialmente com relação ao modelo CAN-24 de Chen et al. (2017)) e melhor definição em regiões de borda. Estatísticas adicionais podem ser encontradas no Apêndice 7.3.2.

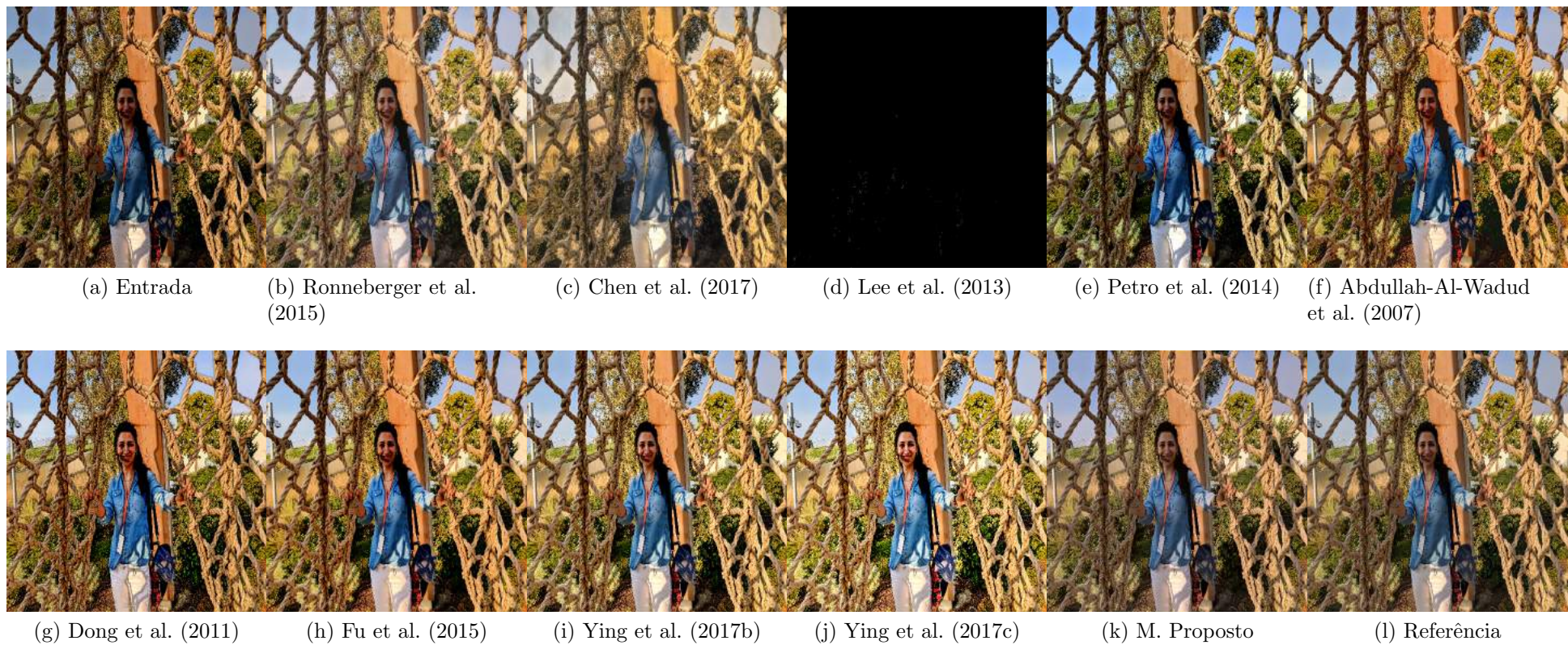


Figura 5.3: Resultados para restauração de imagem subexposta sintética gerada a partir do *dataset* HDR+*burst*. A imagem de entrada apresenta variações bruscas de intensidade, alternando entre regiões muito escuras e muito claras. O modelo proposto apresenta resultados mais semelhantes à imagem referência

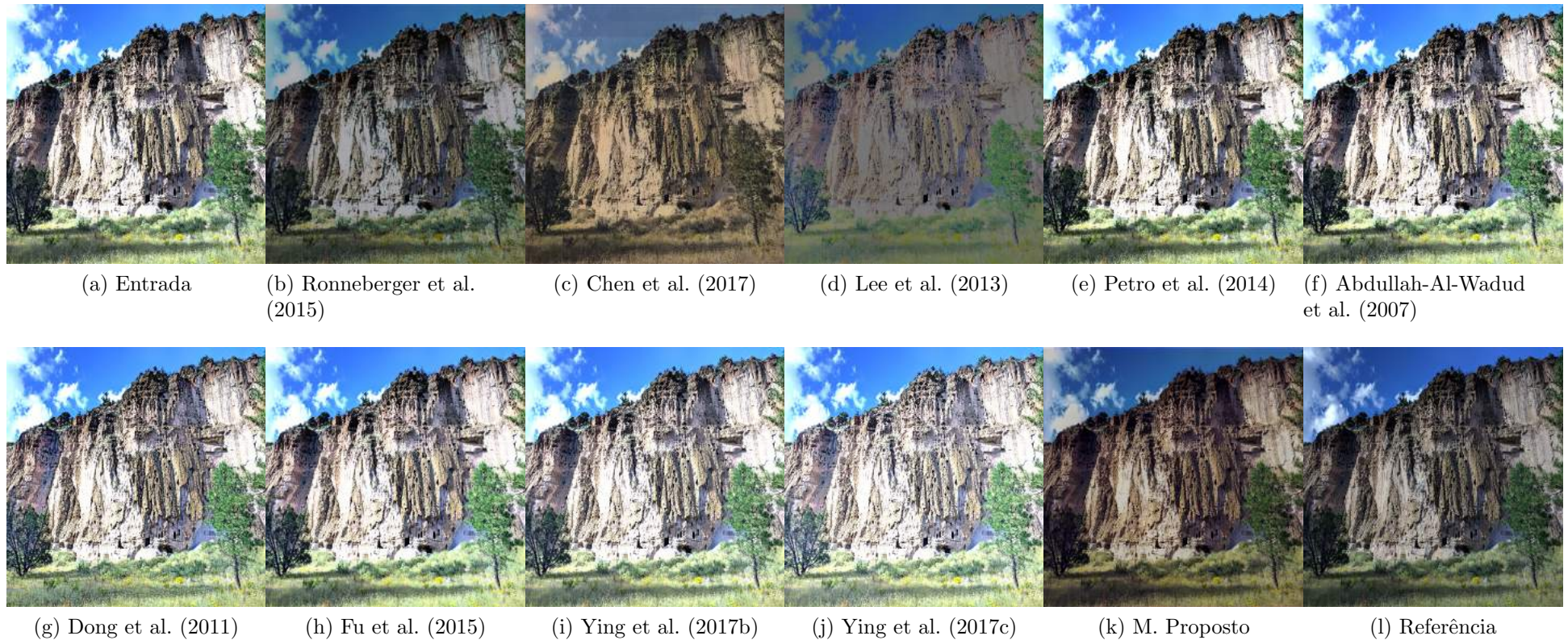


Figura 5.4: Resultados da restauração de imagem severamente saturada gerada a partir do *dataset* HDR+*burst*. Na região central da imagem de entrada, observam-se vários blocos com saturação total dos valores de intensidade. Os modelos de restauração clássicos mostram-se insuficientes para o tratamento da saturação. Dentre os modelos baseados em aprendizagem, o modelo proposto se sobressai aos demais por apresentar maior correção de cor, menor incidência de artefatos de bloco e melhor definição em regiões de borda

5.2.3 A6300 *Multi-Exposure Dataset* (real)

Este *dataset* se difere dos anteriores por conter imagens com exposição inadequada reais. O *dataset* é composto de conjuntos de 4 imagens para cada cena: uma imagem apropriadamente exposta utilizando uma única fotografia, uma imagem subexposta, uma imagem sobre-exposta e uma composição das anteriores utilizando o método de *Tone Mapping* de Mertens et al. (2007). As imagens sub e sobre-expostas são obtidas pela câmera a partir de compensação de exposição com prioridade de abertura, com valores de exposição (EV) no intervalo EV -0.7 – EV +0.7. As cenas retratadas incluem ambientes internos e externos. Todas as imagens são arquivadas utilizando o compressão JPEG com perdas, de acordo com o algoritmo implementado pelo fabricante da câmera. Os dados apresentados a consideram 50 amostras aleatórias simples extraídas do *dataset* sem repetição.

Inicialmente, apresenta-se um resumo estatístico para os resultados atingidos pelo método proposto considerando distintas medidas de qualidade. A Tabela 22 apresenta os resultados da aplicação do modelo proposto para correção de imagens subexpostas neste dataset. Por razões de exequibilidade, os dados são calculados sobre uma amostra de 50 imagens. Verifica-se, de imediato, que os resultados apresentados pelo método proposto em um dataset real são inferiores aos observados na restauração de danos simulados. Tal condição é esperada, visto que a interação dos elementos ópticos da lente, do sensor de imagem e da eletrônica embarcada são complexos, difíceis de reproduzir em simulação. Um comparativo entre as Tabelas 22 e 23 mostra resultados equilibrados para restauração de imagens sub e sobre-expostas. Em ambos os casos, os valores de média e mediana são bastante próximos, indicando pouca assimetria nos resultados observados (não existem pontos fora da curva que gerem distorção as medidas de tendência central).

Para verificar se as diferenças entre os resultados produzidos pelos diversos modelos são estatisticamente significativas utilizou-se das estatísticas Qui-Quadrado de Friedman (Friedman, 1937, 1939) e Teste de Postos Sinalizados de Wilcoxon (Wilcoxon, 1992). Cabe reforçar que a utilização de estatísticas não paramétricas se dá em função da distribuição dos valores não segue uma distribuição normal. Seguiu-se o mesmo procedimento reportado na Seção 5.2.1. No teste de Friedman identificou-se que existe diferença significativa entre os tratamentos para todas as métricas, considerando um nível de significância $p < 0.05$. Realizou-se então os teste de Wilcoxon para identificar como o modelo proposto se compara aos demais.

Considerando imagens subexpostas, os dados apresentados nas Tabelas 26 e 24 permitem observar que, para a maior parte das métricas utilizadas, existe diferença estatisticamente significativa entre os resultados atingidos pelos diversos modelos de restauração. É interessante notar que neste *dataset*, ao contrário do que ocorre nos *datasets* discutidos anteriormente, os métodos baseados em processamento clássico de imagens apresentam os melhores resultados. Neste cenário, a diferença entre estes modelos e o modelo proposto é significativa e o método proposto se equipara em termos de resultados ao modelo U-Net para as métricas PSNR, MAE, diferença de histogramas e CIEDE 2000. Em todas as métricas, a imagem restaurada através

Tabela 22: Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens subexpostas do *dataset* A6300

	PSNR↑	MSE↓	MAE↓	SSIM↑	Sobel IoU ↑	Canny IoU ↑	Hist. Diff. ↓	GMSD↓	VIFP↑	FSIM↑	FSIMc↑	RECO	CIEDE 2000 ↓
Média	17,492	0,023	0,131	0,860	0,662	0,437	5,703	2,731	0,910	0,940	0,936	1,010	13,009
D. Padrão	3,666	0,013	0,046	0,072	0,111	0,107	2,536	1,304	0,212	0,022	0,024	0,197	4,041
Mínimo	12,600	0,001	0,026	0,612	0,235	0,107	1,493	0,162	0,574	0,848	0,841	0,578	2,826
25%	15,440	0,015	0,111	0,826	0,618	0,394	4,090	2,123	0,772	0,934	0,929	0,909	11,145
Mediana	16,888	0,020	0,135	0,867	0,683	0,423	5,042	2,643	0,899	0,941	0,937	0,993	12,864
75%	18,208	0,029	0,153	0,899	0,724	0,494	7,118	3,329	1,002	0,952	0,950	1,078	15,040
Máximo	30,536	0,055	0,225	0,985	0,812	0,670	11,709	6,583	1,803	0,980	0,978	1,724	21,114

Tabela 23: Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens sobre-expostas do *dataset* A6300

	PSNR↑	MSE↓	MAE↓	SSIM↑	Sobel IoU ↑	Canny IoU ↑	Hist. Diff. ↓	GMSD↓	VIFP↑	FSIM↑	FSIMc↑	RECO	CIEDE 2000 ↓
Média	18,647	0,021	0,113	0,864	0,580	0,347	5,059	4,352	0,744	0,900	0,895	0,760	11,021
D. Padrão	4,628	0,021	0,059	0,083	0,125	0,142	2,533	2,214	0,206	0,038	0,039	0,175	4,861
Mínimo	8,809	0,001	0,032	0,630	0,287	0,013	1,681	0,607	0,215	0,834	0,828	0,276	4,523
25%	16,050	0,012	0,084	0,829	0,509	0,265	2,996	3,107	0,641	0,873	0,866	0,638	9,251
Mediana	17,235	0,019	0,111	0,875	0,577	0,327	4,420	4,408	0,754	0,893	0,887	0,786	11,054
75%	19,238	0,025	0,139	0,923	0,628	0,435	6,023	5,423	0,879	0,918	0,913	0,830	12,640
Máximo	28,440	0,132	0,360	0,973	0,837	0,629	11,650	9,940	1,347	0,975	0,973	1,272	35,091

Tabela 24: Mediana para restauração de imagens subexpostas do dataset A6300 (Steffens et al., 2018a) (real). Os três melhores resultados para cada medida de qualidade são sublinhados

Método	PSNR↑	MSE↓	MAE↓	SSIM↑	Sobel IoU↑	Canny IoU↑	Hist. Diff.↓	GMSD↓	VIFP↑	FSIM↑	FSIMc↑	RECO	CIEDE 2000↓
M. Proposto	16,888	0,020	0,135	<u>0,867</u>	<u>0,683</u>	<u>0,423</u>	5,042	<u>2,643</u>	<u>0,899</u>	<u>0,941</u>	<u>0,937</u>	<u>0,993</u>	12,864
Ronneberger et al. (2015)	17,177	0,019	0,125	<u>0,853</u>	0,579	0,402	<u>4,910</u>	<u>2,954</u>	<u>0,985</u>	0,921	0,917	<u>1,182</u>	12,475
Chen et al. (2017)	15,089	0,031	0,168	0,818	0,537	0,362	6,630	3,645	1,235	0,915	0,904	1,619	16,447
Abdullah-Al-Wadud et al. (2007)	18,154	<u>0,015</u>	<u>0,106</u>	0,819	0,592	0,331	5,472	3,892	0,383	0,903	0,896	0,625	<u>11,721</u>
Ying et al. (2017b)	<u>15,757</u>	<u>0,027</u>	<u>0,155</u>	0,840	<u>0,637</u>	<u>0,410</u>	5,891	3,306	0,942	<u>0,934</u>	<u>0,931</u>	1,222	15,134
Fu et al. (2015)	12,787	0,053	0,224	0,722	0,609	<u>0,412</u>	7,488	5,075	0,979	<u>0,921</u>	<u>0,917</u>	1,290	20,441
Lee et al. (2013)	12,548	0,056	0,205	0,764	0,576	0,421	8,902	5,712	1,005	0,928	0,924	1,693	19,711
Petro et al. (2014)	11,973	0,063	0,227	0,617	0,499	0,294	6,524	8,641	0,575	0,864	0,858	0,659	21,747
Dong et al. (2011)	<u>22,333</u>	<u>0,015</u>	<u>0,100</u>	0,806	0,540	0,209	<u>4,354</u>	4,687	0,442	0,875	0,870	0,484	<u>11,050</u>
Ying et al. (2017c)	<u>21,303</u>	<u>0,007</u>	<u>0,069</u>	<u>0,914</u>	<u>0,693</u>	<u>0,470</u>	<u>3,651</u>	<u>1,768</u>	0,834	<u>0,952</u>	<u>0,949</u>	<u>0,994</u>	<u>7,379</u>
Não Tratada	8,151	0,153	0,373	0,326	0,307	0,082	10,298	13,021	4,274	0,801	0,796	5,127	34,742

Tabela 25: Mediana para restauração de imagens sobre-expostas do dataset A6300 (Steffens et al., 2018a) (real). Os três melhores resultados para cada medida de qualidade são sublinhados

Método	PSNR↑	MSE↓	MAE↓	SSIM↑	Sobel IoU↑	Canny IoU↑	Hist. Diff.↓	GMSD↓	VIFP↑	FSIM↑	FSIMc↑	RECO	CIEDE 2000↓
M. Proposto	17,235	<u>0,019</u>	<u>0,111</u>	0,875	0,577	<u>0,327</u>	4,420	4,408	0,754	0,893	0,887	0,786	11,054
Ronneberger et al. (2015)	<u>15,983</u>	<u>0,025</u>	<u>0,141</u>	<u>0,855</u>	0,571	<u>0,327</u>	<u>6,681</u>	<u>3,821</u>	<u>1,053</u>	<u>0,910</u>	<u>0,905</u>	1,265	<u>13,217</u>
Chen et al. (2017)	11,686	0,068	0,230	0,814	0,551	0,291	<u>6,504</u>	<u>5,395</u>	0,734	0,885	0,869	1,306	20,006
Abdullah-Al-Wadud et al. (2007)	11,727	0,067	0,225	0,788	0,539	0,228	7,706	10,429	0,450	0,840	0,837	0,455	17,774
Ying et al. (2017b)	8,097	0,155	0,379	0,744	<u>0,601</u>	<u>0,339</u>	9,664	12,253	0,796	0,873	0,871	<u>1,124</u>	28,936
Fu et al. (2015)	7,702	0,170	0,396	0,737	0,535	0,281	9,786	12,718	<u>0,849</u>	0,861	0,859	1,297	29,910
Lee et al. (2013)	<u>14,707</u>	<u>0,034</u>	<u>0,145</u>	<u>0,818</u>	0,487	0,216	8,926	6,671	<u>1,728</u>	0,838	0,835	1,878	<u>14,132</u>
Petro et al. (2014)	10,087	0,098	0,293	<u>0,792</u>	0,580	0,314	8,307	10,146	0,461	0,871	0,865	0,527	22,182
Dong et al. (2011)	12,530	0,172	0,398	0,718	0,509	0,258	10,134	14,028	0,741	0,826	0,823	0,868	30,476
Ying et al. (2017c)	7,682	0,171	0,399	<u>0,727</u>	<u>0,596</u>	0,313	10,235	12,459	0,845	0,868	0,864	1,283	30,276
Não Tratada	8,445	0,143	0,363	0,763	<u>0,636</u>	0,374	9,473	10,756	<u>0,897</u>	<u>0,898</u>	<u>0,896</u>	<u>1,155</u>	27,727

Tabela 26: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do *dataset* A6300 *Multi-Exposure Dataset* (valor-p)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	0,053	0,044	0,073	0,000	0,000	0,000	0,075	0,000	0,000	0,000	0,000	0,000	0,088
Chen et al. (2017)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Abdullah-Al-Wadud et al. (2007)	0,046	0,037	0,006	0,000	0,000	0,000	0,843	0,000	0,000	0,000	0,000	0,000	0,011
Ying et al. (2017b)	0,025	0,010	0,005	0,002	0,000	0,001	0,004	0,001	0,055	0,006	0,026	0,003	0,006
Fu et al. (2015)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Lee et al. (2013)	0,000	0,000	0,000	0,000	0,000	0,098	0,000	0,000	0,057	0,000	0,000	0,000	0,000
Petro et al. (2014)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Dong et al. (2011)	0,000	0,022	0,002	0,000	0,000	0,000	0,038	0,000	0,000	0,000	0,000	0,000	0,014
Ying et al. (2017c)	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,000	0,008	0,000	0,000	0,032	0,000
Não Tratada	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Tabela 27: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do *dataset* A6300 *Multi-Exposure Dataset* (valor-p)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	0,001	0,000	0,000	0,024	0,261	0,222	0,000	0,208	0,000	0,746	0,579	0,000	0,003
Chen et al. (2017)	0,000	0,000	0,000	0,002	0,033	0,034	0,000	0,001	0,080	0,004	0,000	0,000	0,000
Abdullah-Al-Wadud et al. (2007)	0,000	0,000	0,000	0,000	0,022	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017b)	0,000	0,000	0,000	0,000	0,357	0,172	0,000	0,000	0,046	0,004	0,011	0,000	0,000
Fu et al. (2015)	0,000	0,000	0,000	0,000	0,002	0,000	0,000	0,000	0,005	0,000	0,000	0,000	0,000
Lee et al. (2013)	0,000	0,000	0,000	0,005	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Petro et al. (2014)	0,000	0,000	0,000	0,000	0,599	0,070	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Dong et al. (2011)	0,000	0,000	0,000	0,000	0,003	0,000	0,000	0,000	0,431	0,000	0,000	0,047	0,000
Ying et al. (2017c)	0,000	0,000	0,000	0,000	0,472	0,005	0,000	0,000	0,002	0,000	0,000	0,000	0,000
Não Tratada	0,000	0,000	0,000	0,000	0,017	0,019	0,000	0,000	0,000	0,776	0,553	0,000	0,000

do modelo é significativamente melhor que a imagem não tratada (rejeita-se a hipótese nula).

Considerando as imagens em condição de sobre-exposição, cujo teste Teste dos Postos Significados é apresentado na Tabela 27 e as medianas são apresentadas na Tabela 25 verifica-se que existe diferença significativa entre os dados restaurados pelo modelo proposto e a maioria dos métodos avaliados. Nesta condição, o teste de Wilcoxon não aponta diferença significativa entre o modelo proposto, Ronneberger et al. (2015), Ying et al. (2017b), Petro et al. (2014), e Ying et al. (2017c) para a avaliação de intersecção de bordas utilizando Sobel. Quando avaliada a intersecção de bordas utilizando Canny o modelo proposto se compara a Ronneberger et al. (2015) e Ying et al. (2017b). Em termos de similaridade de características mensuradas pelas métricas FSIM e FSIMc verifica-se que a diferença não é significativa para as imagens não tratadas e para as imagens tratadas utilizando Ronneberger et al. (2015).

Ao comparar-se os valores apresentados na Tabela 24 com os valores correspondentes dos *datasets* simulados (Tabelas 12 e 18) identifica-se um aumento expressivo nas medidas de erro entre a imagem danificada não processada e a sua referência. Merece especial atenção a métrica SSIM, cujos valores variam de 0,853 (HDR+*burst*) e 0,859 (Fivek), para 0,307 (A6300). Nota-se, portanto, uma degradação bastante expressiva. Uma vez que a imagem de entrada utilizada pelos modelos para restauração é notavelmente mais danificada, é natural que os resultados da restauração sejam inferiores.

Através do comparativo apresentado na Tabela 24 verifica-se, em contraposição aos resultados discutidos nas Seções 5.2.1 e 5.2.2, que os métodos baseados em processamento de imagem clássicos se sobressaem aos métodos baseados em redes neurais convolucionais. Neste sentido destacam-se os métodos de Ying et al. (2017c), Dong et al. (2011), Ying et al. (2017b) e Abdullah-Al-Wadud et al. (2007). Dentre todos os métodos testados, Ying et al. (2017c) apresentou resultados superiores para as medidas MSE, MAE, SSIM, Sobel IoU, Canny IoU, diferença de histogramas, GMSD, FSIM, FSIMc, RECO e CIEDE 2000.

Dentre os métodos baseados em redes neurais, o método proposto apresenta melhores resultados para as medidas PSNR, SSIM, Sobel IoU, Canny IoU, GMSD, FSIM e FSIMc. Para esta parte do *dataset*, a rede U-Net, de Ronneberger et al. (2015), produziu melhores resultados em termos de ajuste de histogramas e informação em *features* VIFP. Por fim, cabe salientar que todos os métodos de restauração aplicados obtiveram contribuição significativa na melhoria de imagem.

A Tabela 24 apresenta um comparativo entre os métodos para restauração de imagens sobre-expostas. Nesta condição, os modelos baseados em redes neurais voltam a apresentar resultados melhores que os modelos baseados em processamento de imagens clássico. O modelo proposto supera os demais nas medidas PSNR, MSE, MAE, SSIM, Canny IoU, diferença de histogramas e CIEDE 2000. Nas medidas FSIM e FSIMc, o modelo proposto apresenta ligeira piora quando comparado à imagem sobre-exposta de entrada.

A Figura 5.5 apresenta um comparativo visual entre os métodos de restauração quando aplicados à restauração de imagem subexposta com EV-0.7 do *dataset* A6300. Trata-se de uma cena em ambiente interno, apresentando alto contraste entre regiões bem iluminadas e

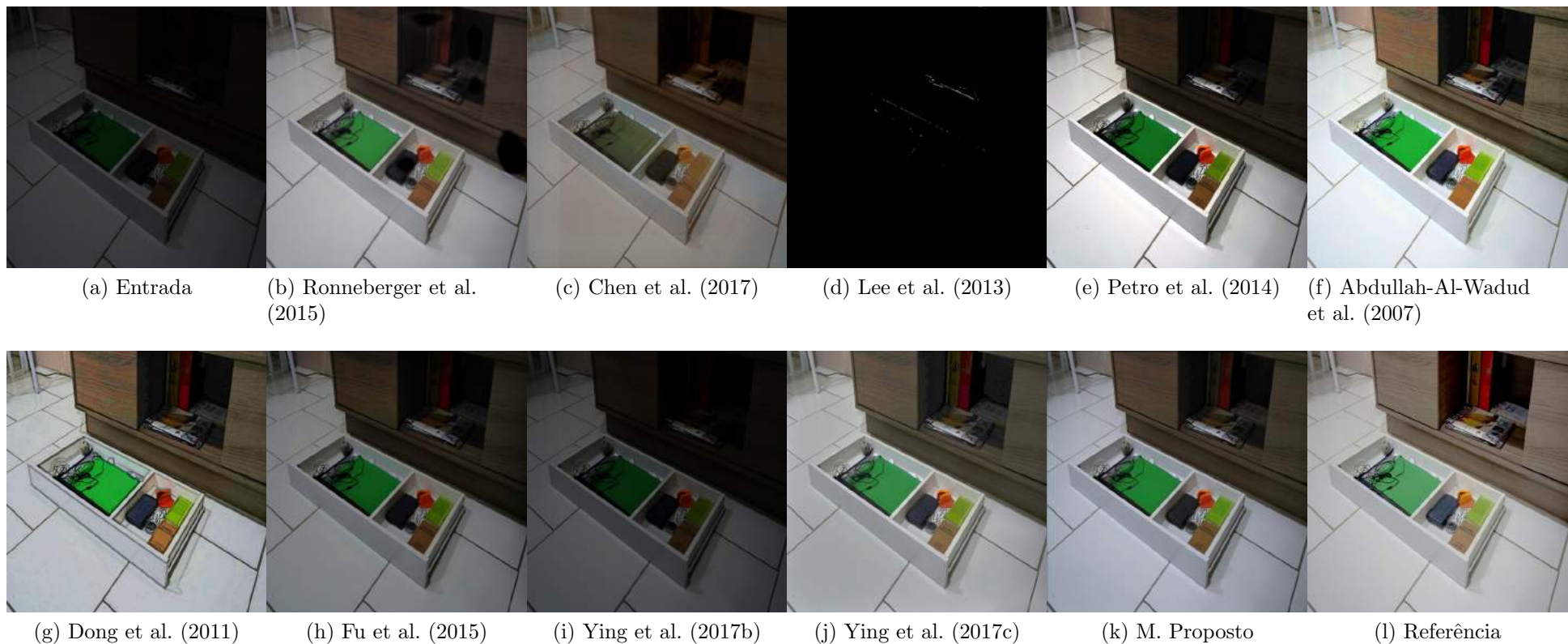


Figura 5.5: Resultados obtidos por diferentes métodos na restauração de imagem subexposta real. Observa-se que o modelo proposto apresenta equilíbrio entre restauração dos detalhes e constância de cor. Especialmente na região central da imagem, pode-se observar que os demais métodos comparados são incapazes de restaurar a visibilidade e amplificam os efeitos adversos de ruídos e artefatos de compressão

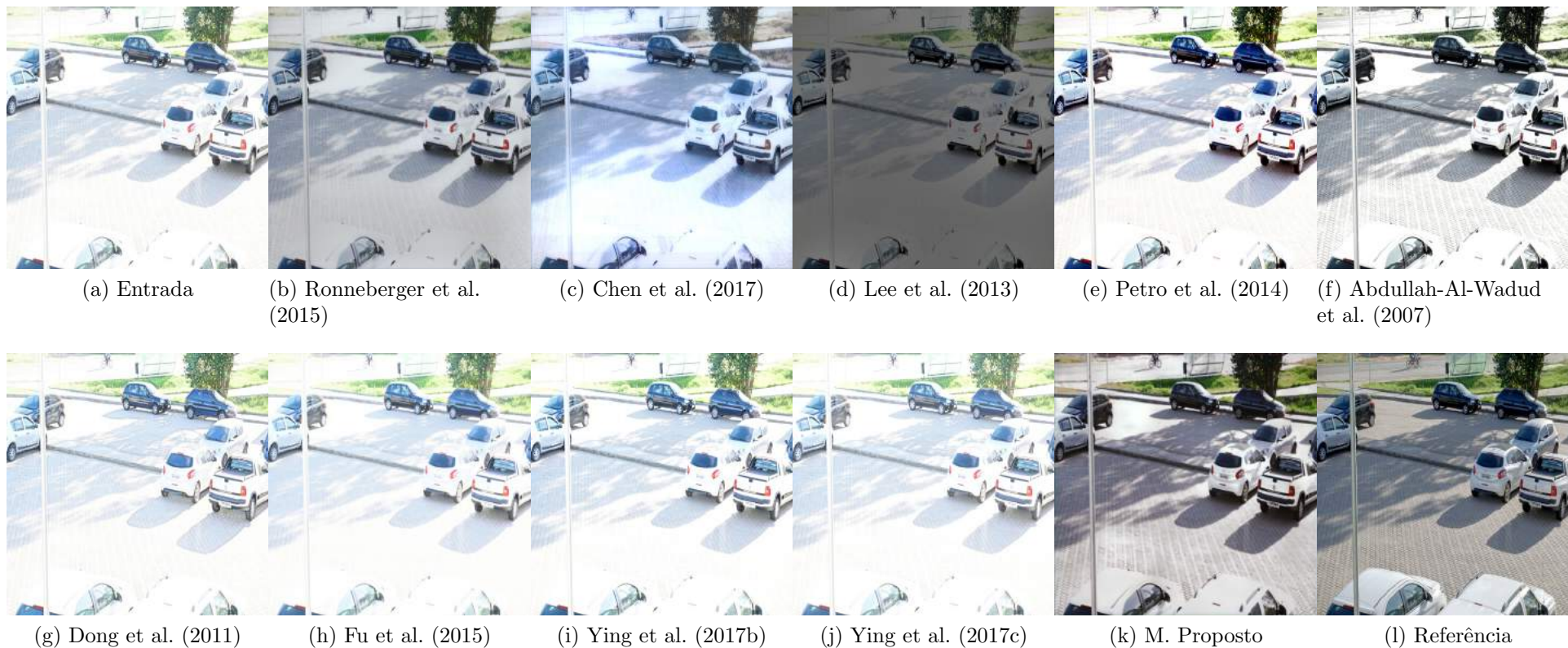


Figura 5.6: Resultados obtidos por diferentes métodos na restauração de imagem sobre-exposta real em ambiente externo. Destaca-se que a imagem de entrada apresenta saturação severa em diversas partes. A cena contém pouca informação de cor, sendo a maior parte dos elementos em tons de cinza. Nenhum dos métodos testados produz resultados semelhantes à imagem referência. Dentre todos, o método proposto é o que apresenta maior equilíbrio entre correção de cor e restauração de texturas

regiões de sombra. Observa-se que o modelo proposto apresenta equilíbrio entre restauração dos detalhes e constância de cor. Especialmente na região central da imagem, observa-se que os demais métodos comparados são incapazes de restaurar a visibilidade e amplificam os efeitos adversos de ruídos e artefatos de compressão. Com relação aos modelos Ronneberger et al. (2015) e Chen et al. (2017), verifica-se que o modelo proposto apresenta resultados superiores ao preservar informação de textura e regiões de borda. Com relação aos demais métodos, vê-se que a amplificação do sinal teve como consequência a amplificação do ruído indesejável.

A Figura 5.6 apresenta os resultados de restauração para uma imagem sobre-exposta. Destaca-se que a imagem de entrada apresenta saturação severa em diversas partes. A cena contém pouca informação de cor, sendo a maior parte dos elementos em tons de cinza. Nenhum dos métodos testados produz resultados semelhantes à imagem referência. Esta condição pode ser atribuída à insuficiência de dados nas regiões vizinhas, fazendo com que os modelos não encontrem subsídio suficiente para o preenchimento das regiões saturadas. Dentre todos, o método proposto é o que apresenta maior equilíbrio entre correção de cor e restauração de texturas, especialmente perceptível nos blocos de pavimentação.

O Apêndice 7.4.2 apresenta estatísticas adicionais sobre os resultados neste conjunto de dados. Resultados visuais mais abrangentes podem ser visualizados no Apêndice 7.9, onde apresenta-se um comparativo utilizando uma variedade maior de cenas.

5.2.4 Cai2018 *Multi-Exposure Dataset* (real)

O mais desafiador dentre os *datasets* utilizados para avaliar a arquitetura de rede proposta, este dataset é composto de imagens reais obtidas através de distintos equipamentos fotográficos. Cada cena é fotografada utilizando a técnica de *bracketing*, em que múltiplas fotografias são obtidas utilizando distintos tempos de exposição do sensor. Os valores de *exposure compensation* EV, variam no intervalo [EV -3; EV +3], resultando em imagens danificadas de forma severa por sub e sobre-exposição. Os dados apresentados a seguir foram obtidos considerando 50 amostras aleatórias simples sem repetição.

As tabelas 28 e 29 apresentam respectivamente os resultados da aplicação do método proposto em imagens sub e sobre-expostas. Verifica-se um equilíbrio nos resultados obtidos para ambas as condições. O intervalo interquartil, por outro lado, mostra grande dispersão dos valores. O valor de SSIM e PSNR obtido é baixo se comparado aos outros *datasets*, refletindo a condição de dano das imagens de entrada.

Fez-se a verificação dos níveis de significância utilizando as estatísticas não paramétricas Qui-Quadrado de Friedman (Friedman, 1937, 1939) e Teste de Postos Sinalizados de Wilcoxon (Wilcoxon, 1992). Seguiu-se o mesmo procedimento reportado na Seção 5.2.1. Novamente o teste de Friedman indica que as diferenças entre os tratamentos são estatisticamente significativas. O teste de Wilcoxon em imagens subexpostas, apresentado na Tabela 32 permitem observar que na métrica PSNR o modelo U-Net (Ronneberger et al., 2015) atingiu resultados significativamente superiores ao modelo proposto. Considerando PSNR, não existem diferenças

Tabela 28: Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens subexpostas do *dataset* Cai et al. (2018)

	PSNR $\uparrow$	MSE $\downarrow$	MAE $\downarrow$	SSIM $\uparrow$	Sobel IoU $\uparrow$	Canny IoU $\uparrow$	Hist. Diff. $\downarrow$	GMSD $\downarrow$	VIFP $\uparrow$	FSIM $\uparrow$	FSIMc $\uparrow$	RECO	CIEDE 2000 $\downarrow$
Média	16,295	0,035	0,141	0,725	0,506	0,389	7,261	5,317	1,116	0,880	0,867	1,772	15,973
D. Padrão	4,136	0,028	0,068	0,178	0,187	0,156	1,777	4,633	0,726	0,077	0,080	2,226	6,558
Mínimo	9,751	0,004	0,048	0,318	0,024	0,065	3,970	0,714	0,349	0,680	0,660	0,676	6,490
25%	12,438	0,011	0,086	0,583	0,393	0,280	5,635	1,721	0,737	0,844	0,832	0,931	10,740
Mediana	16,349	0,023	0,130	0,769	0,541	0,394	7,132	3,429	0,988	0,910	0,901	1,228	14,707
75%	19,438	0,057	0,192	0,875	0,649	0,509	8,970	8,086	1,223	0,937	0,924	1,569	21,059
Máximo	24,073	0,106	0,279	0,940	0,781	0,711	10,343	20,240	4,985	0,969	0,964	13,492	32,340

Tabela 29: Resumo estatístico para os resultados obtidos pelo modelo proposto quando aplicado em imagens sobre-expostas do *dataset* Cai et al. (2018)

	PSNR $\uparrow$	MSE $\downarrow$	MAE $\downarrow$	SSIM $\uparrow$	Sobel IoU $\uparrow$	Canny IoU $\uparrow$	Hist. Diff. $\downarrow$	GMSD $\downarrow$	VIFP $\uparrow$	FSIM $\uparrow$	FSIMc $\uparrow$	RECO	CIEDE 2000 $\downarrow$
Média	16,994	0,028	0,123	0,779	0,563	0,352	5,598	6,128	0,709	0,854	0,838	1,746	16,323
D. Padrão	3,606	0,026	0,059	0,119	0,182	0,187	2,085	4,837	0,272	0,090	0,095	1,380	6,678
Mínimo	8,333	0,004	0,046	0,446	0,134	0,018	2,462	1,051	0,069	0,591	0,565	0,804	7,455
25%	14,837	0,012	0,080	0,728	0,438	0,218	3,581	2,336	0,579	0,801	0,783	0,992	11,518
Mediana	16,816	0,021	0,114	0,789	0,570	0,363	5,397	4,547	0,785	0,882	0,865	1,195	14,239
75%	19,294	0,033	0,155	0,868	0,705	0,481	7,393	8,230	0,873	0,922	0,911	1,911	19,357
Máximo	23,926	0,147	0,332	0,954	0,857	0,695	9,542	20,099	1,364	0,975	0,970	7,931	40,163

Tabela 30: Mediana para restauração de imagens subexpostas do *dataset* Cai et al. (2018) (real). Os três melhores resultados para cada medida de qualidade são sublinhados

Método	PSNR $\uparrow$	MSE $\downarrow$	MAE $\downarrow$	SSIM $\uparrow$	Sobel IoU $\uparrow$	Canny IoU $\uparrow$	Hist. Diff. $\downarrow$	GMSD $\downarrow$	VIFP $\uparrow$	FSIM $\uparrow$	FSIMc $\uparrow$	RECO	CIEDE 2000 $\downarrow$
M. Proposto	16,349	<u>0,023</u>	<u>0,130</u>	<u>0,769</u>	<u>0,541</u>	<u>0,394</u>	7,132	3,429	0,988	<u>0,910</u>	<u>0,901</u>	<u>1,228</u>	<u>14,707</u>
Ronneberger et al. (2015)	<u>17,421</u>	<u>0,018</u>	<u>0,110</u>	<u>0,739</u>	0,365	0,296	<u>6,851</u>	<u>4,057</u>	<u>1,139</u>	0,853	0,850	2,575	<u>12,052</u>
Chen et al. (2017)	<u>16,500</u>	<u>0,022</u>	<u>0,127</u>	<u>0,746</u>	0,429	0,333	7,650	4,457	1,262	<u>0,874</u>	0,864	1,828	<u>14,624</u>
Abdullah-Al-Wadud et al. (2007)	14,382	0,036	0,153	0,656	<u>0,536</u>	0,315	8,036	5,577	0,518	<u>0,867</u>	0,850	<u>1,123</u>	17,085
Ying et al. (2017b)	12,564	0,055	0,203	0,665	<u>0,483</u>	0,333	7,929	6,141	1,284	<u>0,874</u>	<u>0,867</u>	1,700	22,027
Fu et al. (2015)	10,365	0,092	0,275	0,532	0,415	<u>0,354</u>	8,867	8,497	1,368	<u>0,845</u>	<u>0,835</u>	2,170	29,573
Lee et al. (2013)	10,545	0,088	0,260	0,593	0,381	<u>0,261</u>	11,363	10,589	1,120	0,824	0,814	1,748	26,766
Petro et al. (2014)	10,382	0,092	0,271	0,447	0,417	0,243	8,658	10,424	0,829	0,796	0,786	1,264	28,522
Dong et al. (2011)	14,380	0,051	0,192	0,630	0,496	0,144	<u>7,525</u>	7,371	0,669	0,827	0,819	<u>0,878</u>	19,412
Ying et al. (2017c)	<u>16,413</u>	0,024	0,135	<u>0,736</u>	<u>0,518</u>	<u>0,392</u>	<u>7,674</u>	3,194	<u>1,111</u>	0,893	<u>0,887</u>	1,557	15,570
Não Tratada	7,173	0,192	0,401	0,215	0,123	0,087	10,807	17,188	4,794	0,703	0,690	5,739	38,071

Tabela 31: Mediana para restauração de imagens sobre-expostas do *dataset* Cai et al. (2018) (real). Os três melhores resultados para cada medida de qualidade são sublinhados

Método	PSNR $\uparrow$	MSE $\downarrow$	MAE $\downarrow$	SSIM $\uparrow$	Sobel IoU $\uparrow$	Canny IoU $\uparrow$	Hist. Diff. $\downarrow$	GMSD $\downarrow$	VIFP $\uparrow$	FSIM $\uparrow$	FSIMc $\uparrow$	RECO	CIEDE 2000 $\downarrow$
M. Proposto	16,816	<u>0,021</u>	0,114	<u>0,789</u>	<u>0,570</u>	<u>0,363</u>	5,397	4,547	0,785	0,882	0,865	1,195	14,239
Ronneberger et al. (2015)	<u>15,819</u>	<u>0,026</u>	<u>0,124</u>	<u>0,719</u>	0,430	0,228	<u>5,391</u>	<u>6,307</u>	1,424	<u>0,809</u>	0,795	2,064	<u>14,628</u>
Chen et al. (2017)	<u>16,247</u>	<u>0,024</u>	<u>0,117</u>	<u>0,787</u>	<u>0,555</u>	<u>0,286</u>	<u>4,328</u>	<u>4,382</u>	0,701	<u>0,855</u>	<u>0,844</u>	<u>1,257</u>	<u>14,078</u>
Abdullah-Al-Wadud et al. (2007)	10,318	0,093	0,263	0,678	0,525	0,240	6,405	14,058	0,430	0,784	0,769	0,629	20,176
Ying et al. (2017b)	8,309	0,148	0,355	0,646	0,458	0,211	8,381	14,328	0,854	0,793	0,784	1,645	26,801
Fu et al. (2015)	8,184	0,152	0,368	0,633	0,399	0,193	8,642	14,674	1,099	0,783	0,774	2,287	27,691
Lee et al. (2013)	12,481	0,056	0,189	0,696	0,353	0,178	8,099	10,441	2,044	0,768	0,753	2,696	20,377
Petro et al. (2014)	9,844	0,104	0,282	0,674	<u>0,499</u>	<u>0,242</u>	6,896	13,532	0,463	0,785	0,771	<u>0,702</u>	22,452
Dong et al. (2011)	9,381	0,147	0,359	0,631	<u>0,449</u>	<u>0,229</u>	8,474	14,825	<u>0,963</u>	0,773	0,760	1,500	28,031
Ying et al. (2017c)	8,203	0,151	0,367	0,639	0,446	0,222	8,752	14,679	<u>1,063</u>	0,784	0,778	1,949	29,058
Não Tratada	8,628	0,137	0,344	0,664	0,462	0,255	8,451	13,734	<u>1,064</u>	0,805	<u>0,797</u>	1,814	25,642

Tabela 32: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do *dataset* Cai2018 *Multi-Exposure* (valor-p)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	0,000	0,000	0,000	0,012	0,000	0,000	0,261	0,001	0,000	0,000	0,000	0,000	0,000
Chen et al. (2017)	0,163	0,003	0,057	0,081	0,000	0,000	0,004	0,000	0,000	0,000	0,000	0,000	0,896
Abdullah-Al-Wadud et al. (2007)	0,057	0,042	0,110	0,000	0,043	0,000	0,000	0,046	0,000	0,000	0,000	0,054	0,055
Ying et al. (2017b)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,004	0,000	0,000	0,000	0,000
Fu et al. (2015)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Lee et al. (2013)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,146	0,000	0,000	0,996	0,000
Petro et al. (2014)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,114	0,000	0,000	0,912	0,000
Dong et al. (2011)	0,014	0,000	0,000	0,000	0,000	0,000	0,028	0,000	0,000	0,000	0,000	0,040	0,000
Ying et al. (2017c)	0,454	0,000	0,001	0,000	0,640	0,000	0,167	0,000	0,057	0,000	0,020	0,000	0,010
Não Tratada	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Tabela 33: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do *dataset* Cai2018 *Multi-Exposure* (valor-p)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	0,014	0,169	0,067	0,000	0,000	0,000	0,927	0,000	0,000	0,000	0,000	0,000	0,083
Chen et al. (2017)	0,002	0,006	0,021	0,000	0,000	0,000	0,000	0,005	0,791	0,000	0,001	0,308	0,028
Abdullah-Al-Wadud et al. (2007)	0,000	0,000	0,000	0,000	0,003	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Ying et al. (2017b)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,005	0,000
Fu et al. (2015)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Lee et al. (2013)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Petro et al. (2014)	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Dong et al. (2011)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,006	0,000
Ying et al. (2017c)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Não Tratada	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

estatisticamente significantes entre os resultados obtidos com o método proposto, Chen et al. (2017), Abdullah-Al-Wadud et al. (2007) e Ying et al. (2017c).

A Tabela 32 permite verificar ainda que, considerando a métrica SSIM, o teste de Wilcoxon mostra que o modelo proposto é superior ao modelo U-Net ($p < 0.05$, diferença estatisticamente significativa entre os grupos). Quando considera-se a diferença de histogramas, a diferença entre os resultados do modelo proposto e o modelo U-Net (Ronneberger et al., 2015) e o método de Ying et al. (2017c) podem ser atribuídos ao acaso, não apresentando, portanto, incremento ou decremento significativo na restauração. Quando considera-se a métrica RECO, que avalia a integridade de bordas, percebe-se que os resultados do modelo proposto não são estatisticamente diferentes dos resultados obtidos por Abdullah-Al-Wadud et al. (2007); Lee et al. (2013) e Petro et al. (2014). Por fim, na avaliação de cor mensurada pela métrica CIEDE 2000 verifica-se que o modelo proposto não tem diferença estatisticamente significativa para os modelos de Chen et al. (2017) e Abdullah-Al-Wadud et al. (2007), tendo performance superior ou equivalente aos demais modelos testados.

Nas imagens sobre-expostas, combinando as Tabelas 31 e 33 também observam-se poucas situações em que não se rejeita a hipótese nula. o modelo proposto atingiu a melhor performance nas métricas PSNR, MSE, MAE, SSIM, intersecção de bordas com Sobel e Canny, FSIM, FSIMc, RECO. A diferença é estatisticamente significativa nas métricas supracitadas exceto MSE, MAE, Diferença de histogramas e CIEDE 2000 para o modelo U-Net (Ronneberger et al., 2015).

A Tabela 30 apresenta o comparativo entre métodos para restauração de imagens subexpostas. Observa-se, neste caso, uma preponderância dos métodos baseados em redes neurais convolucionais. O modelo proposto é superior ao modelo de Ronneberger et al. (2015) nas medidas que avaliam primordialmente gradiente e bordas (Sobel IoU, Canny IoU, GMSD e RECO). Nas medidas que ponderam a cor da imagem com maior peso (MAE, MSE, diferença de histogramas e CIEDE 2000), o modelo de Ronneberger et al. (2015) apresenta performance levemente superior. Merece destaque o método de Ying et al. (2017c) com performance muito próxima aos modelos treinados.

A Tabela 31 apresenta os resultados para restauração de imagens sobre-expostas. Novamente verifica-se que os modelos baseados em redes convolucionais são capazes de produzir resultados superiores em termos de cor e preservação de bordas. Todos os modelos baseados em CNN apresentaram ganho expressivo com relação as imagens não processadas. Dentre estes, o modelo proposto atinge os melhores resultados para PSNR, MSE, MAE, SSIM, Sobel IoU, Canny IoU, FSIM, FSIMc, e RECO.

Fora do quadro comparativo, destaca-se que o trabalho original de Cai et al. (2018) reporta para imagens subexpostas um valor de 19,770 para PSNR e 0,934 para FSIM. Em imagens sobre-expostas, Cai et al. (2018) reporta um valor 20,21 para PSNR e 0,93 para FSIM. Destaca-se, no entanto, que estes resultados são obtidos considerando somente imagens subexpostas e sobre-expostas com compensação de exposição $EV \pm 1$. Devido à falta de detalhamento do experimento, não é possível executar o comparativo no mesmo conjunto de imagens utilizadas

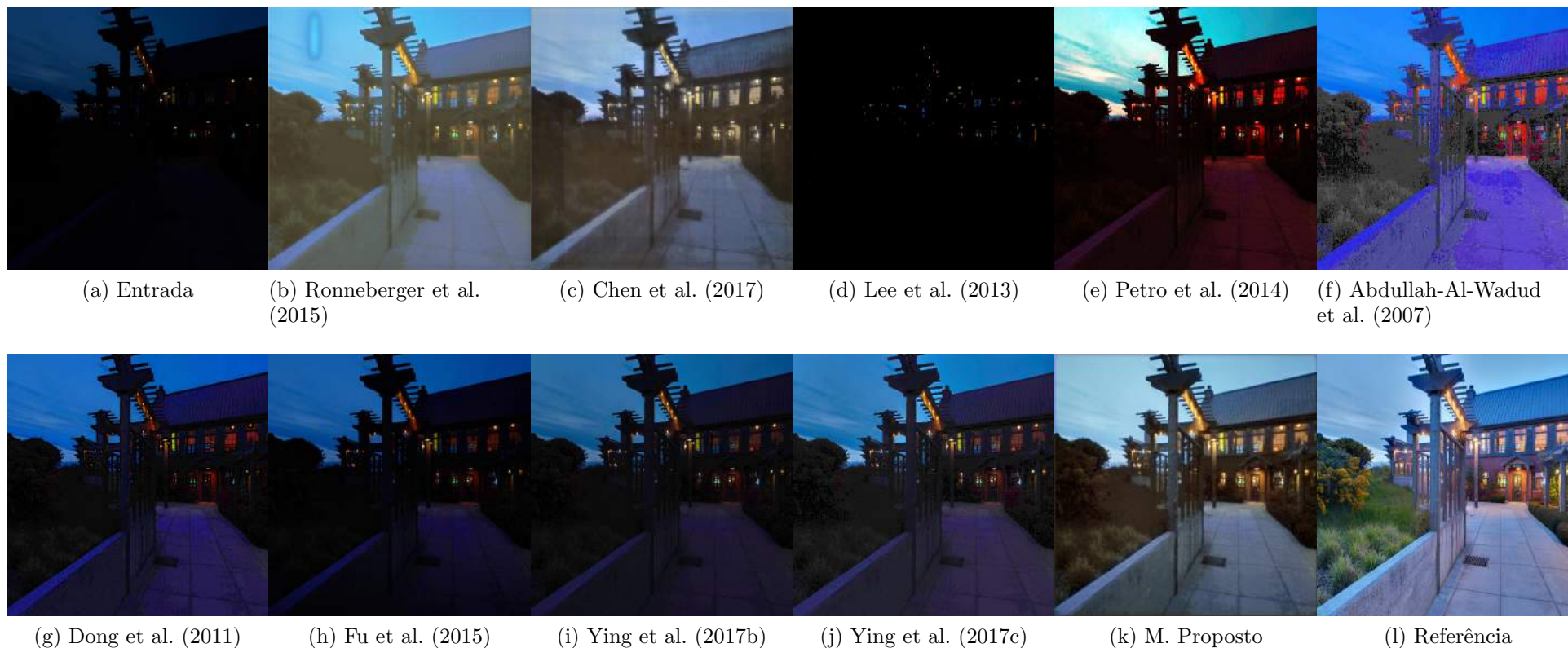


Figura 5.7: Resultado de restauração para imagem noturna extremamente subexposta do *dataset* Cai et al. (2018). Observa-se alto contraste com regiões muito claras próximas aos pontos de iluminação e muito escuras no restante. Dentre os modelos avaliados, o modelo proposto apresenta os melhores resultados no sentido de permitir a visualização de alguns detalhes nas regiões muito escuras. Observa-se, no entanto, que todos os modelos testados são significativamente afetados pelos artefatos da compressão JPEG, resultando na presença de blocos com transição abrupta de intensidade na imagem de saída

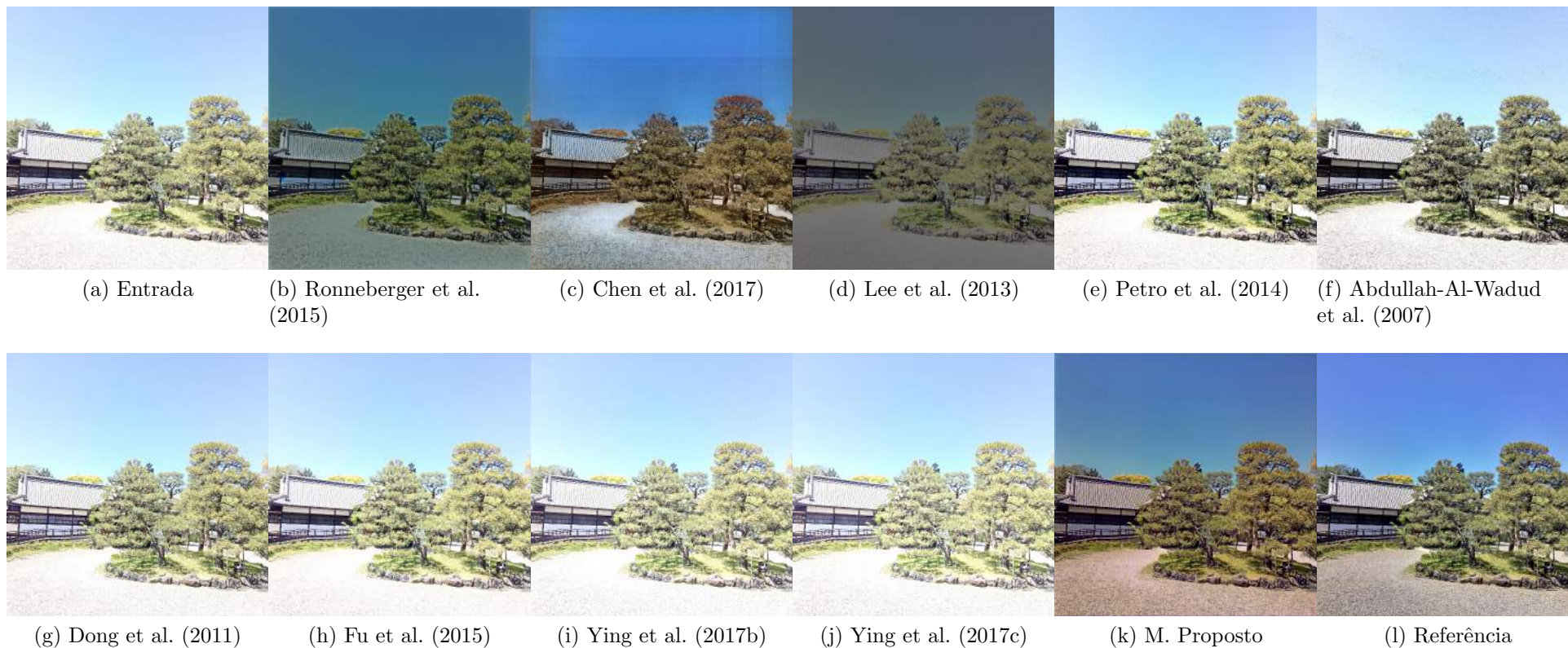


Figura 5.8: Resultados de restauração de imagens diurna sobre-exposta do *dataset* Cai et al. (2018). Observa-se que, dentre os modelos baseados em aprendizagem, o modelo proposto é o que produz resultados mais próximos da referência, preservando nitidez nas regiões borda e correção de cor. Nas regiões saturadas observa-se a preservação de texturas com aparência natural

pelo autor.

A Figura 5.7 apresenta o resultado de restauração para uma imagem noturna subexposta do *dataset* Cai et al. (2018). Observa-se alto contraste, com regiões muito claras próximas ao pontos de iluminação e muito escuras no restante. Dentre os modelos avaliados, o modelo proposto apresenta os melhores resultados no sentido de permitir a visualização de alguns detalhes nas regiões muito escuras (telhado, vegetação). Observa-se que todos os modelos testados são significativamente afetados pelos artefatos da compressão JPEG, resultando na presença de blocos com transição abrupta de intensidade na imagem de saída.

A Figura 5.8 apresenta os resultados de restauração de imagens diurna sobre-exposta do *dataset* Cai et al. (2018). Observa-se que, dentre os modelos baseados em aprendizagem, o modelo proposto é o que produz resultados mais próximos da referência, preservando nitidez nas regiões borda e correção de cor. Nas regiões saturadas observa-se a preservação de texturas com aparência natural. Observa-se ainda que os modelos de restauração baseados no processamento de imagens clássico resultam em pouca transformação da imagem de saída, apresentando, desta forma, pouca contribuição para correção dos efeitos adversos da saturação.

O Apêndice 7.10 apresenta resultados visuais mais abrangentes da aplicação do modelo proposto e dos demais métodos utilizados no comparativo.

5.3 Outros Comparativos

Esta Seção apresenta os demais métodos estado-da-técnica para os quais não foi factível uma avaliação detalhada dos resultados. Os quadros comparativos a seguir levam em consideração somente os dados publicados pelos respectivos autores, sendo limitados aos dados publicamente disponíveis. Estes modelos não foram implementados. Os dados disponíveis podem, portanto, ser incompletos e apresentar discrepâncias com outros dados na literatura. Ressalta-se que a maioria dos modelos foi desenhada ou treinada para trabalhar apenas com imagens subexpostas, limitando desta forma a análise comparativa.

A Tabela 34 apresenta um comparativo entre os métodos de restauração para imagens subexpostas, utilizando como critério o melhor valor reportado pelos autores em cada métrica. As setas indicam a interpretação de cada uma das medidas de qualidade empregadas: $\uparrow$ indica que um valor mais alto é representa uma melhor performance enquanto $\downarrow$ indica que um valor mais baixo indica melhor performance. Destaca-se que estas medidas levam em consideração distintos *datasets*, não fornecendo, portanto, uma comparação exata entre os métodos, mas sim um indicativo sobre a performance. No entanto, nos casos em que os autores não disponibilizam publicamente os dados utilizados ou os procedimentos são sub-especificados esta é a única forma plausível de fazer a avaliação.

Cabe salientar ainda que alguns dos modelos de restauração consideram apenas imagens geradas de forma sintética, ao passo que outros utilizam para validação apenas dados reais. Conforme discutido nas Seções 5.2.1, 5.2.2, 5.2.3 e 5.2.4 o modelo proposto apresenta uma degradação expressiva no desempenho quando comparamos os diferentes conjuntos de dados.

Tabela 34: Comparativo entre métodos de restauração para imagens subexpostas. Os dados apresentados representam os melhores resultados reportados pelos autores, independente do *dataset* utilizado. Exceções são marcadas com ^(a), que indica que a fonte do dado é Liang et al. (2020), ou ^(b) que indica que a fonte do dado é Zhang et al. (2019b). ^(c) indica que os resultados foram obtidos em dados brutos (RAW). ^(d) indica modelos leves, desenhados para funcionar equipamentos pouco poderosos

Nome	PSNR↑	SSIM↑	VIFP↑	FSIM↑	FSIMc↑	CIEDE2000↓	NIQE↓	LOE↓	BRISQUE↓
Modelo Proposto (simulado) ^d	29,010	0,970	0,940	0,990	0,983	3,838	-	-	-
Modelo Proposto (real) ^d	16,888	0,867	0,899	0,941	0,937	12,864	-	-	-
DeepUPE (Wang et al., 2019c)	30,970	0,856	-	-	-	-	7,864 ^a	284,920 ^a	7,864 ^a
SID ^c (Chen et al., 2018)	28,880	0,787	-	-	-	-	-	-	-
CWAN (Atoum et al., 2020)	28,560	0,909	-	-	-	-	-	-	-
MBLLEN (Lv et al., 2018) ^d	25,970	0,870	0,490	-	-	-	-	-	-
AgLLNet (Lv et al., 2021)	25,240	0,940	0,670	-	-	-	-	495,480	-
LLIE-Net (Guo et al., 2019)	23,680	0,910	-	-	-	-	-	-	-
RDGAN (Wang et al., 2019a)	22,340	-	-	-	0,958	-	-	-	-
(Liang et al., 2020)	22,216	0,786	-	-	-	-	3,635	289,050	21,778
(Xu et al., 2020)	22,130	0,717	-	-	-	-	-	-	-
VP (Li et al., 2020b)	21,119	0,840	-	-	0,952	-	-	-	-
NLHD (Hao et al., 2021)	21,112	0,810	-	-	-	13,610	3,580	262,990	-
KinD (Zhang et al., 2019b)	20,866	0,802	-	-	-	-	5,146	2012,200	26,644
DRBN (Yang et al., 2020)	20,130	0,849	-	-	-	-	-	-	-
(Xiong et al., 2020)	20,040	0,820	-	-	-	-	-	-	-
SICE (Cai et al., 2018) ^d	19,770	-	-	0,934	-	-	-	-	-
GLADNet (Wang et al., 2018b)	19,720 ^b	0,703 ^b	-	-	-	-	19,720 ^b	902,500 ^b	-
(Afifi et al., 2020)	19,685	0,742	-	-	-	-	-	-	-
ZeroDCE (Guo et al., 2020) ^d	19,570	0,590	-	-	-	-	-	-	-
(Zhang et al., 2020a)	19,150	0,710	-	-	-	-	4,793	1384,100	-
Retinex-Net (Wei et al., 2018) ^d	16,774 ^a	0,559 ^a	-	-	-	-	9,730 ^b	993,290 ^a	8,879 ^a
ZeroDCE++ (Li et al., 2021) ^d	16,420	0,580	-	-	-	-	-	-	-
BIMEF (Ying et al., 2017a)	13,875	0,577 ^a	-	-	-	-	7,515	287,000	27,651 ^a
Ying 2017 (Ying et al., 2017b)	-	-	-	-	-	-	-	287,000	-
DeclipNet (Honig & Werman, 2018)	-	-	-	-	-	-	-	-	-
DALE (Kwon et al., 2020)	-	-	-	-	-	-	3,610	714,600	22,200
EnlightenGAN (Jiang et al., 2019) ^d	-	-	-	-	-	-	3,385	-	-

É, portanto, plausível que o mesmo ocorra também para os outros métodos de restauração.

Exceto pelo modelo SID (Chen et al., 2018), todos os valores apresentados foram obtidos no processamento de imagens no formato sRGB. No caso do Modelo Proposto (simulado), leva-se em consideração os resultados obtidos utilizando o *dataset* MIT-Adobe FiveK. No caso do Modelo Proposto (real) leva-se em consideração os resultados obtidos utilizando o *dataset* Multi-exposição A6300. A tabela está ordenada seguindo a métrica PSNR. Uma vez que os outros trabalhos não apresentam seus resultados considerando as métricas MSE, RECO e GMSD, estas não foram incluídas na listagem.

Observa-se na Tabela 34 que os resultados obtidos utilizando o modelo de rede neural proposto para a restauração de imagens estão consistentes com os demais dados da literatura. Admite-se, portanto, que o modelo atinge resultados equivalentes ao estado da arte na área de restauração de imagens impactadas por exposição inadequada. Considerando dados simulados e imagens sRGB, o modelo proposto tem o melhor índice de similaridade estrutural (SSIM) e a segunda melhor relação sinal-ruído de pico (PSNR). quando levam-se em conta somente os modelos leves, o modelo proposto é o que apresenta a maior PSNR.

No entanto, se considerados apenas os resultados da restauração de imagem subexposta em dados reais, verifica-se que o modelo proposto assume uma posição intermediária no *ranking* por PSNR e SSIM. Conforme já discutido anteriormente na Seção 5.1, a maior parte dos estimadores de qualidade de imagem baseados em referência apresenta uma correlação forte entre si, fazendo com que uma avaliação das métricas SSIM e PSNR possa ser generalizada também para as outras métricas.

Por fim, ressalta-se que para o método proposto não foram calculadas as métricas cegas NIQE (Wang et al., 2013), LOE (Wei et al., 2018) e BRISQUE (Mittal et al., 2012). Conforme já evidenciado por Zhang et al. (2019b) estas medidas são suscetíveis à condições como resolução da imagem e, portanto, menos robustas que as métricas que consideram dados pareados. Em casos onde existe uma imagem referência conhecida, as métricas e medidas que levam em consideração um par de imagem apresentam um poder de avaliação maior que as métricas cegas.

Tabela 35: Comparativo entre métodos de restauração para imagens sobre-expostas. Os dados apresentados representam os melhores resultados reportados pelos autores, independente do *dataset* utilizado. ^(a) indica modelos leves, desenhados para funcionar equipamentos pouco poderosos

Nome	PSNR↑	SSIM↑	VIFP↑	FSIM↑	FSIMc↑	CIEDE 2000 ↓
Modelo Proposto (simulado) ^a	23,116	0,945	1,157	0,966	0,959	8,114
Modelo Proposto (real) ^a	17,235	0,875	0,744	0,893	0,887	11,054
SICE (Cai et al., 2018) ^a	20,210	-	-	0,935	-	-
(Afifi et al., 2020)	19,349	0,737	-	-	-	-
DeclipNet(Honig & Werman, 2018)	-	-	-	-	-	-

A Tabela 35 apresenta os resultados conforme reportados pelos autores para a restauração de imagens sobre-expostas. Destaca-se que a restauração de imagens sobre-expostas tem uma quantidade menor de modelos disponíveis, uma vez que a maior parcela dos trabalhos relacionados concentra-se na restauração de imagens obtidas em condições de pouca luz. Neste cenário, observa-se que o modelo proposto e o modelo SICE (Cai et al., 2018) foram desenhados para operar em equipamentos pouco poderosos. Cai et al. (2018), embora não forneçam dados sobre o tamanho da rede, alegam que SICE pode ser executado em CPU, com tempos comparáveis aos obtidos por métodos de equalização de histograma.

Quanto aos números apresentados na Tabela 35, verifica-se novamente que o modelo proposto apresenta resultados similares aos apresentados pelos seus pares. Cabe salientar que, uma vez que os *datasets* utilizados são distintos e sub-especificados, a comparação direta entre os métodos também é prejudicada. Em dados reais, verifica-se que o modelo proposto apresenta média de PSNR e SSIM inferior ao modelo SICE. Com relação ao modelo de Afifi et al. (2020) (avaliado em dados sintéticos) o modelo atinge performance superior em ambas as métricas. Não existem dados disponíveis para o modelo DeclipNet (Honig & Werman, 2018) e as informações publicadas são insuficientes para a reprodução dos experimentos.

De forma geral, o comparativo com os dados reportados na literatura mostra que o modelo proposto atinge resultados satisfatórios tanto com imagens subexpostas quanto com imagens sobre-expostas. Com relação ao tamanho do modelo, mensurado em parâmetros da rede neural, o modelo proposto tem 380.899 parâmetros, sendo o terceiro menor entre os avaliados com dados disponíveis. Por ordem, do menor para o maior, tem-se KinD++ (Li et al., 2021) com 10.000 parâmetros, KinD (Zhang et al., 2019b) com 79.000 parâmetros, MBLLEN (Lv et al., 2018) com 450.000, Retinex-Net (Wei et al., 2018) com 555.000, SICE (Cai et al., 2018) e o modelo de Afifi et al. (2020) com 7.000.000 de parâmetros.

5.4 Validação a Nível de Aplicação

Para além da restauração de imagens para que tenham uma melhor qualidade percebida, mensurada a partir de métricas e medidas de similaridade entre imagens a nível de cor, similaridade estrutural, contraste, gradiente e outros, tem-se nesta pesquisa o objetivo de restaurar a informação presente na imagem. A importância desta restauração se dá na condição de que o modelo pode ser utilizado para mitigar os efeitos indesejados da exposição inapropriada em diversas aplicações que se utilizam da computação visual para obter as informações para tomada de decisão. Nesta Secção, faz-se uma investigação do impacto da subexposição e da sobre-exposição nestas aplicações e demonstra-se como o modelo de restauração proposto pode contribuir na restauração do conteúdo da cena representado pela imagem.

A fim de avaliar a resiliência de diversos modelos de reconhecimento de imagens diante de situações de sub e sobre-exposição faz-se o uso de imagens geradas sinteticamente. Os modelos de classificação foram usados como originalmente propostos, isto é, com conjuntos idênticos de pesos, formato de entrada e camadas de ativação interna utilizados na referência. Utilizam-se

os modelos preparados para o desafio ImageNet ILSRVC Challenge (Russakovsky et al., 2015).

5.4.1 Reconhecimento de Imagens

Fez-se uma avaliação abrangente dos impactos gerados pela subexposição e sobre-exposição em distintos modelos de reconhecimento de imagens baseados em redes neurais convolucionais. O Apêndice 7.6 apresenta um quadro resumo dos resultados obtidos em dados simulando diferentes condições de exposição. De maneira geral, observa-se que os modelos de reconhecimento de imagem que obtiveram a maior acurácia no conjunto original de imagens, também obtêm os melhores resultados quando aplicados nas imagens que foram distorcidas e manipuladas. Em uma análise por distorção, nota-se que a classificação dos modelos de melhor desempenho raramente muda de posição. As exceções a esta condição são limitadas à falta de exposição extrema obtida por meio da transformação Gama com $\gamma = [\frac{1}{8}; 8]$. Nestas condições, a precisão obtida por alguns dos modelos torna-os inúteis para aplicações práticas.

O impacto das distorções está, em geral, associado ao número de pesos treináveis na rede neural de classificação. NASNetLarge tem um desempenho melhor do que todos os outros modelos avaliados neste estudo, independentemente da distorção aplicada à imagem de entrada. Inception-ResNet-v2 e Xception também se mostram robustos, apesar de apresentarem uma queda expressiva na acurácia. Esses são os três modelos que apresentaram o melhor desempenho entre os modelos de reconhecimento considerados. Identifica-se também que modelos maiores (em termos de parâmetros treináveis) lidam melhor com as distorções. A importância do número de pesos é evidenciada quando se leva em consideração a condição observada com NASNetMobile e NASNetLarge, que possuem células base idênticas (arquitetura), mas uma grande diferença quando se trata do número de pesos. O modelo VGG-16, mais amplo levado em consideração, é um contra-exemplo. Lançado em 2014, o VGG-16 apresenta a menor robustez de todos os modelos considerados neste estudo.

Uma vez que se verifica que os modelos de reconhecimento são, de fato, afetados pela qualidade das imagens de entrada, propõe-se um *pipeline* de restauração de imagens baseado no modelo proposto. Optou-se por explorar como o *pipeline* de restauração impacta o modelo VGG-16 (Simonyan & Zisserman, 2014). Conforme resultados da avaliação inicial, VGG-16 (Simonyan & Zisserman, 2014) é altamente suscetível a distorção de imagem e todas as distorções de imagem resultaram em uma queda de acurácia maior que 10 % para este modelo sendo, desta forma, um estudo de caso adequado para a avaliação de um *pipeline* que considere o pré-processamento com restauração da imagem. Levou-se em conta ainda o fato de o modelo VGG servir como base para uma para um número de técnicas e práticas amplamente difundidas na área (Johnson et al., 2016; Long et al., 2015; Pravitasari et al., 2020; Lee et al., 2019a).

A Tabela 36 compara o impacto das imagens danificadas e os efeitos da restauração no modelo VGG-16 (Simonyan & Zisserman, 2014). A partir dos resultados obtidos com uma abordagem *pipeline*, pode-se verificar que o problema de exposição incorreta pode ser minimizado pelo uso de métodos de aprimoramento de imagem. O modelo de restauração fornece um

Tabela 36: Acurácia Top-1 para o modelo VGG-16 (Simonyan & Zisserman, 2014) considerando imagens com diferentes níveis de subexposição e sobre-exposição simulada e restauradas pelo modelo de restauração proposto

Distorção simulada	Sem restauração	Com restauração
Sem distorção	0,612	-
Gama $\frac{1}{2}$	0,584	0,605
Gama $\frac{1}{4}$	0,455	0,612
Gama $\frac{1}{8}$	0,236	0,618
Gama 2	0,566	0,610
Gama 4	0,401	0,575
Gama 8	0,175	0,429
Truncado Q_1 (subexposta)	0,541	0,601
Truncado Q_3 (sobre-exposta)	0,548	0,608

ganho expressivo em condições em que os pixels são truncados por saturação e subexposição.

O *pipeline* de restauração oferece um ganho expressivo mesmo em condições extremas. Para subexposição, simulada por transformação Gama com $\gamma = 8$, a abordagem de *pipeline* do modelo é capaz de melhorar a Precisão Top-1 de 0,175 para 0,429. Para subexposição extrema, simulada pelo operador de potência Gama com $\gamma = \frac{1}{8}$, a abordagem de *pipeline* foi capaz de restaurar a acurácia de reconhecimento de objeto de 0,236 a 0,618. É interessante notar que esta precisão de 0,618 Top-1 é maior do que a precisão de 0,612 obtida no conjunto original de imagens.

5.4.2 Segmentação Semântica de Imagens

Outras aplicações de percepção para robótica, sistemas autônomos e percepção de máquina podem apresentar queda de desempenho semelhante quando submetidas à degradação da imagem por ruído ou exposição inadequada. Seleção de objetos, localização e mapeamento, navegação, fechamento de loop, prevenção de obstáculos, robôs de coleta e sistemas de interação humano-robô costumam compartilhar os mesmos blocos de construção básicos com classificadores de imagem baseados em aprendizagem profunda (Zhang et al., 2019a; Ha et al., 2018; Teso-Fz-Betoño et al., 2020; Qiu et al., 2018; Liu et al., 2018; Jia et al., 2020; Ma et al., 2019; Du et al., 2018). Para entender melhor como essas distorções de imagem comuns podem afetar os sistemas autônomos e robóticos, explora-se como elas afetam a segmentação de instâncias.

Na tarefa de segmentação de instância, o objetivo é distinguir cada instância de cada objeto dentro da imagem no nível do pixel. Avalia-se como Mask-RCNN (He et al., 2017), um modelo projetado para detectar objetos em uma imagem de forma eficiente e, ao mesmo tempo, gerar uma máscara de segmentação para cada instância, atua sob condições não ideais.

Os resultados visuais para segmentação de instância de uma cena urbana usando o modelo Mask-RCNN são mostrados na Figura 5.9. Esta imagem mostra uma cena de alto contraste que representa uma situação prática e plausível na navegação externa autônoma. No geral, nota-se um impacto significativo nos resultados, especialmente sob condições severas de falta

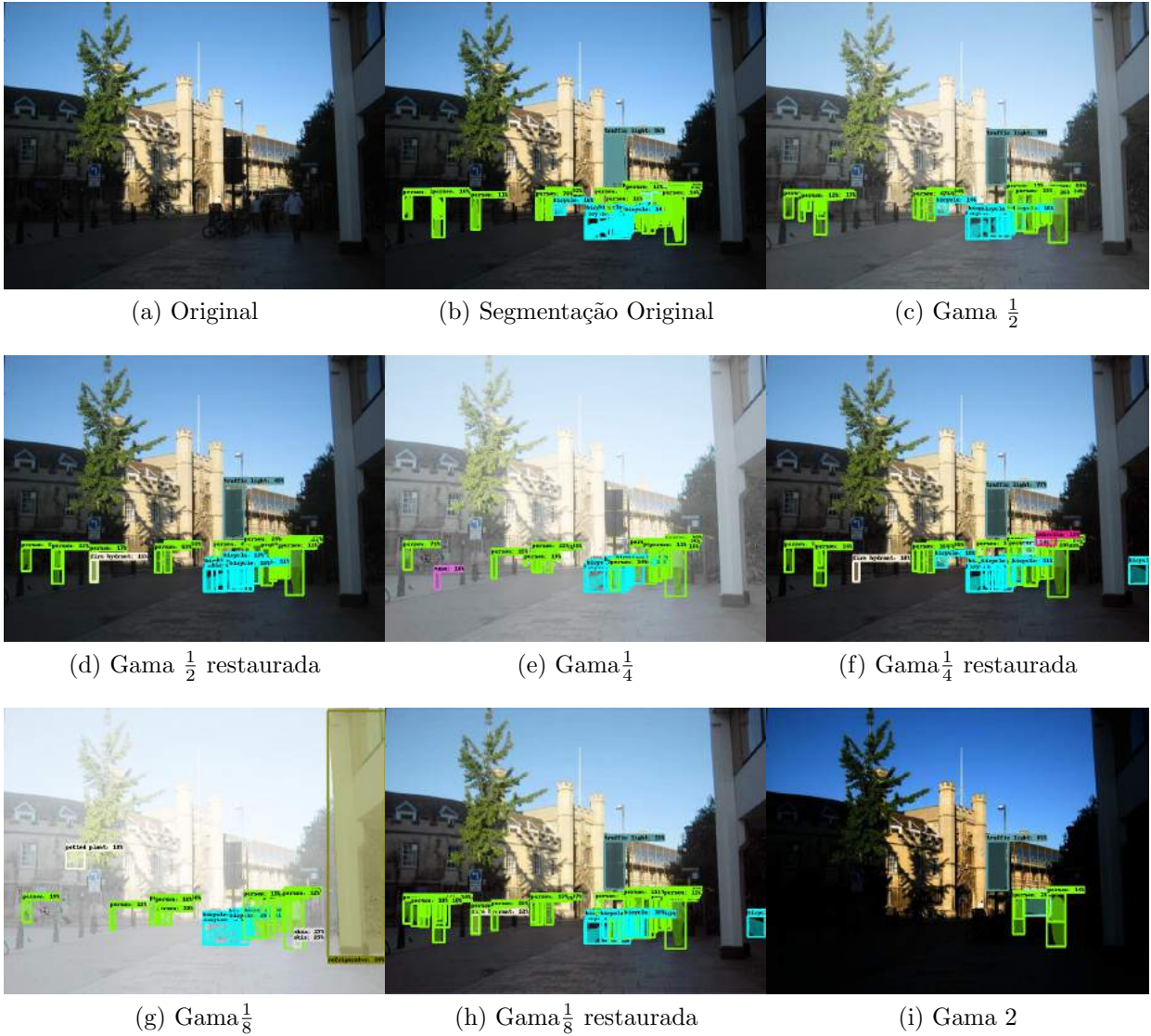


Figura 5.9: Impactos da distorção na tarefa de segmentação de instância: (a) Imagem original, (b) Segmentação na imagem original, (c) Gama $\frac{1}{2}$, (d) Gama $\frac{1}{2}$ restaurada, (e) Gama $\frac{1}{4}$, (f) Gama $\frac{1}{4}$ restaurada, (g) Gama $\frac{1}{8}$, (h) Gama $\frac{1}{8}$ restaurada, (i) Gama 2. Rótulos de classe seguem: limão - pessoa; azul claro - bicicleta; cinza - cadeira; amarelo - vaso de planta; rosa - vaso; rosa pink - guarda-chuva; roxo - roupas; verde-mar - carro; e branco - esquis

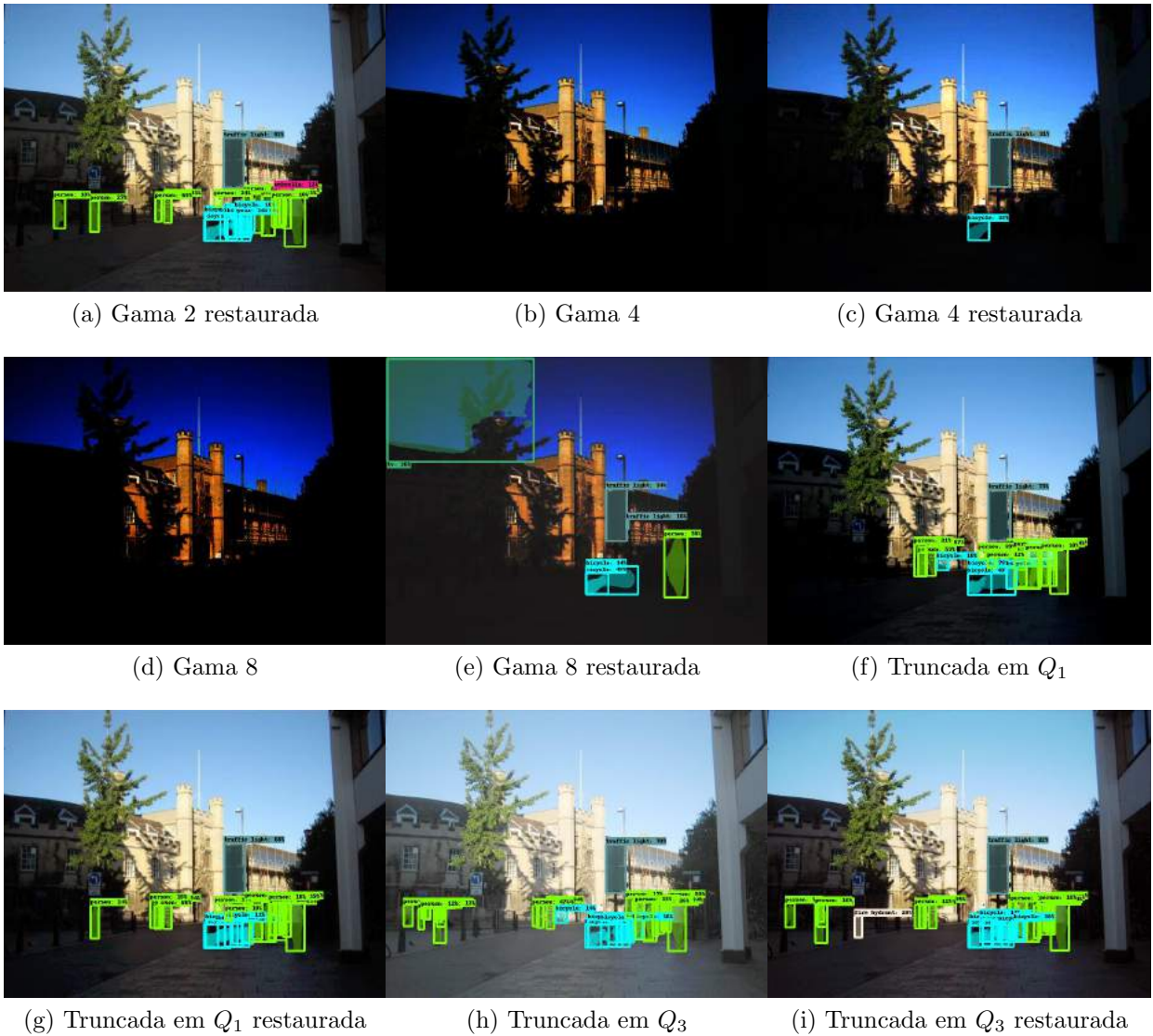


Figura 5.10: Impactos da distorção na tarefa de segmentação de instância (continuação): (a) Gama 2 restaurada, (b) Gama 4, (c) Gama 4 restaurada, (d) Gama 8, (e) Gama 8 restaurada, (f) Truncada em Q_1 , (g) Truncada em Q_1 restaurada, (h) Imagem truncada em Q_3 , e (i) Imagem truncada em Q_3 restaurada. Rótulos de classe seguem: limão - pessoa; azul claro - bicicleta; cinza - cadeira; amarelo - vaso de planta; rosa - vaso; rosa pink - guarda-chuva; roxo - roupas; verde-mar - carro; e creme - hidrante

de exposição e ruído. Pode-se identificar a ocorrência de falsos positivos e falsos negativos. Considerando um hipotético sistema de direção autônoma, PFs e NFs, como os apresentados nesta amostra, podem resultar em mau funcionamento, dados insuficientes para a tomada de ações, ou mesmo em decisões autônomas que colocam vidas em risco.

Na ordem de apresentação, a Figura 5.9 mostra uma imagem devidamente exposta da cena (5.9a); os resultados da segmentação da instância na imagem original (5.9b), imagem afetada por Gama $\frac{1}{2}$ (5.9c), imagem afetada por Gama $\frac{1}{4}$ (5.9e), imagem afetada por Gama $\frac{1}{8}$ (5.9g), imagem afetada por Gama 2 (5.9i), imagem afetada por Gama 4 (5.10b), imagem afetada pela distorção Gama 8 (5.10d), imagem afetada pelo Truncamento Q_1 (5.10f), imagem afetada por e Truncamento Q_3 (o). Cada cor representa um rótulo de classe, definido da seguinte forma: limão representa uma pessoa; o azul claro representa uma bicicleta; cinza representa uma cadeira; amarelo representa um vaso de planta; rosa representa um vaso; rosa pink representa um guarda-chuva; roxo representa roupas; o verde-mar representa um carro; e o branco representa esquis.

A Figura 5.9b mostra que, na imagem original, o modelo de segmentação Mask-RCNN é capaz de identificar corretamente pessoas, bicicletas e sinais de trânsito. Nessa condição, o sistema autônomo poderia contar com os resultados da segmentação para realizar a localização, o mapeamento e a prevenção de obstáculos. As Figuras 5.10b e 5.10d mostram que a subexposição gerada pela transformação Gama com $\gamma = [4; 8]$ inviabilizam o uso do modelo de segmentação de instâncias para qualquer aplicação prática. Na Figura 5.9i, que mostra a subexposição gerada por $\gamma = 2$, vê-se que a quantidade de objetos detectados é significativamente reduzida.

A maioria das distorções de imagem resulta em um aumento expressivo de falsos negativos. Indo além, nas Figuras 5.9e e 5.9g, observa-se que Mask-RCNN resulta em falsos positivos, incluindo instâncias de objetos como cadeira, mochila, vaso, vasos de plantas e esquis. A gravidade dos impactos sobre robôs e aplicativos autônomos que dependem desses sistemas certamente está em discussão.

A visualização lado a lado dos resultados de segmentação de instância evidencia um impacto na quantidade de instâncias que o segmentador é capaz de encontrar. As Figuras 5.9d, 5.9f, 5.9h, 5.10a, 5.10c, 5.10e, 5.10g, 5.10i, restauradas antes da segmentação, apresentam mais itens segmentados. Cabe ressaltar, no entanto, que mesmo em condições ideais os modelos de segmentação estado da técnica como Fang et al. (2021), Wu et al. (2020), Liu et al. (2021), Hu et al. (2017) apresentam uma precisão média entre 0,5 e 0,65 a depender do *dataset* em que são utilizados. Logo, mesmo um incremento marginal na precisão destes modelos é uma contribuição para que a sua saída possa ser utilizada com maior confiabilidade.

Muitas aplicações recentes na robótica dependem da percepção visual do ambiente. Robótica e automação, interação homem-robô, interfaces e interação homem-máquina, robótica social e de serviço, robótica médica, sistemas não tripulados, sistemas autônomos, sistemas ciber-físicos e outros campos relacionados se beneficiaram dos avanços na visão de máquina fornecidos por redes neurais profundas. Os resultados deste experimento mostram que distorções comuns po-

dem fazer com que esses sistemas se tornem não confiáveis ou mesmo perigosos. Acredita-se que as técnicas de pré-processamento de imagem e melhores sensores de imagem também desempenham um papel significativo para a percepção baseada na imagem e podem ser aplicadas para tornar esses sistemas mais confiáveis.

5.4.3 Análise de Desempenho

Considerando aplicações em robótica e sistemas automatizados, o tempo que o modelo de rede neural leva para fazer a restauração da imagem é uma característica primordial. Tempos de processamento elevados inviabilizam a aplicação em tarefas que exigem tomada de decisão em tempo-real. Nesta Seção apresentam-se os resultados de uma avaliação do tempo de inferência do modelo proposto em distintas resoluções de imagens e equipamentos.

Sabe-se que, em geral, a análise de desempenho de algoritmos computacionais é realizada considerando-se a complexidade. No entanto, para os modelos baseados em redes neurais profundas esta é uma tarefa complexa e sujeita a nuances. É notório que a maior parte dos modelos de restauração comparados não disponibilizam qualquer informação sobre a complexidade assintótica ou mesmo sobre a quantidade de operações aritméticas necessárias. Desta forma, restringe-se a avaliação ao tempo necessário para que o modelo proposto faça a restauração de imagens em várias resoluções de entrada em diferentes configurações de hardware.

Utilizaram-se quatro diferentes configurações de equipamentos:

- E1: Equipamento com uma GPU Tesla V100-SXM2 (640 NVIDIA Tensor *Cores*, 5120 CUDA *Cores*, 16Gib de memória VRAM), CPU Intel Xeon 2.00GHz e 12GB de memória RAM. Via Google Colaboratory.
- E2: Equipamento com uma TPU v2-8 (8 TPU cores, 64 GiB), CPU Intel Xeon CPU 2.00GHz e 14GB de memória RAM. Via Google Colaboratory.
- E3: Equipamento sem GPU, equipado com uma CPU AMD EPYC 7B12 2.25 GHz (64 *Cores* / 128 *Threads*) e 32GB de memória RAM. Via Google Colaboratory.
- E4: Notebook pessoal sem GPU, equipado com CPU Intel Core i7-7500U e 8GB de memória RAM.

O procedimento de medição foi realizado considerando um conjunto de 50 imagens e 100 iterações. No total, para cada equipamento foram feitas 5000 inferências. Os valores apresentados consideram o tempo total incluindo o carregamento da imagem, o redimensionamento para as dimensões de entrada suportadas pela rede neural, a restauração utilizando o modelo, o redimensionamento para a resolução original (quando aplicável) e o salvamento da imagem. Desta forma, tem-se uma avaliação honesta para aplicações no mundo real. Para fazer a inferência, utilizou-se das bibliotecas computacionais compiladas com a melhor configuração para o equipamento.

A Tabela 37 apresenta o tempo médio (em milissegundos) de inferência por imagem nas resoluções 512×512 , 1024×1024 , 2048×2048 e 4096×4096 píxeis. Verifica-se que os equipamentos utilizados foram capazes de realizar a inferência em todas as resoluções testadas. Na resolução 512×512 todos os equipamentos testados atingem tempos médios de processamento inferiores a 61 milissegundos sendo, portanto, capazes de processar aproximadamente 16 quadros por segundo. Observa-se ainda que na resolução 512×512 têm desempenhos semelhantes, embora exista uma diferença expressiva de poder computacional.

Tabela 37: Tempo médio de inferência em milissegundos para restauração de imagens utilizando o modelo proposto em milissegundos

Equipamento / Resolução	512×512 px.	1024×1024 px.	2048×2048 px.	4096×4096 px.
E1 - Tesla V100-SXM2	59,25	227,50	360,00	472,50
E2 - TPU	61,00	262,50	425,00	497,00
E3 - AMD EPYC 7B12	45,75	181,75	300,00	460,00
E4 - Intel Core i7-7500U	60,50	395,00	787,50	1207,50

Ainda na Tabela 37 identifica-se que ao dobrar a resolução de 512×512 para 1024×1024 (quadruplicando a quantidade de píxeis na imagem) o tempo de inferência por imagem aumenta em valor proporcional para os equipamentos E1, E2 e E3. Para o equipamento E4 o tempo de inferência aumenta em aproximadamente sete vezes. Em resoluções mais altas este aumento proporcional no tempo de processamento é atenuado. Para exemplificar, a quantidade de píxeis nas imagens 4096×4096 é 64 vezes maior do que a quantidade de píxeis nas imagens 512×512 , mas isto não se reflete nos tempos de processamento.

Nos equipamentos mais poderosos o modelo de restauração atinge, no mínimo, dois quadros por segundo na resolução 4096×4096 . Apenas no equipamento E4 o modelo de restauração leva mais de um segundo para fazer a inferência. Destaca-se que a utilização de imagens nesta resolução em aplicações de visão computacional é incomum. Modelos estabelecidos de classificação de imagens, por exemplo, tendem a utilizar imagens com resolução inferior a 512×512 , como é o caso de Simonyan & Zisserman (2014), He et al. (2016a), Szegedy et al. (2017), Huang et al. (2017), Chollet (2017), Sandler et al. (2018) e Zoph et al. (2018). Modelos para direção autônoma de veículos baseados em datasets como Kitty Geiger et al. (2013) e Xu et al. (2017) também trabalham com resoluções que permitiriam ao modelo processar vários quadros por segundo.

Verifica-se pelos resultados apresentados que o modelo de restauração de imagens proposto é uma alternativa factível para a utilização em sistemas perceptivos baseados em visão aplicados em sistemas robóticos e autônomos. O modelo pode ser utilizado mesmo em equipamentos modestos, sendo integrado a um *pipeline* de processamento.

5.5 Teste de Ablação

De forma a validar as principais contribuições de cada uma das partes do modelo proposto fez-se um teste quantitativo de ablação, isto é, substituíram-se algumas partes da arquitetura

e da função objetivo como forma de identificar a real contribuição. Para cada desmembramento, fez-se um retreino completo do modelo de restauração aplicando procedimento idêntico ao aplicado para o modelo principal. Foram avaliadas as seguintes variações sobre o modelo original:

1. Sem bloco de convoluções dilatadas: Todas as camadas que empregavam blocos de convoluções dilatadas foram substituídas por convoluções 3×3 simples, mantendo a quantidade de filtros e por conseguinte, a quantidade de pesos treináveis.
2. Utilizando erro médio quadrático como função objetivo: substitui-se a função objetivo desenhada especificamente para o problema de restauração de imagens subexpostas e sobre-expostas por uma função de erro genérica, sem discriminação específica relacionada ao problema em questão.
3. Utilizando somente DSSIM como função objetivo: Substitui-se a função objetivo original por uma função objetivo que considera somente a dissimilaridade estrutural.
4. Utilizando somente a combinação de MSE e DSSIM como função objetivo: Remove-se da função objetivo original o mapa de pesos.

Para este comparativo, empregou-se apenas a métrica SSIM, que, conforme discutido na Seção 5.1, apresenta uma correlação significativa com as demais métricas e medidas pareadas de qualidade de imagens. A Tabela 38 apresenta os resultados para imagens subexpostas. Os dados apresentados consideram a mediana para 50 amostras do dataset de Cai et al. (2018). Observa-se que para esta condição, os resultados evidenciam a efetividade das decisões de arquitetura adotadas. Apesar de todos os modelos apresentarem melhora expressiva com relação a imagem não tratada, todas ao remover-se alguma das partes do modelo ou da função objetivo, verifica-se que as métricas PSNR e SSIM apresentam piora. Considerando SSIM, o modelo treinado sem o mapa de pesos na função objetivo é o que apresenta menor capacidade de restauração, seguido pelo modelo ajustado considerando somente DSSIM, o modelo ajustado considerando somente MSE e o modelo baseado em uma arquitetura sem o bloco de convoluções dilatadas.

Tabela 38: Mediana para restauração de imagens subexpostas do *dataset* Cai et al. (2018) (real) no teste de ablação

Modelo	PSNR	SSIM
Modelo base	16.349	0.769
1 - Sem bloco de convoluções dilatadas	15.291	0.707
2 - Somente MSE	15.128	0.700
3 - Somente DSSIM	14.637	0.677
4 - Sem ponderação da função objetivo	14.560	0.661
Não Tratada	7.173	0.215

A Tabela 39 apresenta as medianas para o teste de ablação em imagens sobre-expostas. Novamente, nessa condição observa-se que as decisões de arquitetura e de função objetivo

Tabela 39: Mediana para restauração de imagens sobre-expostas do *dataset* Cai et al. (2018) (real) no teste de ablação

Modelo	PSNR	SSIM
Modelo base	16.816	0.789
1 - Sem bloco de convoluções dilatadas	13.621	0.639
2 - Somente MSE	15.407	0.721
3 - Somente DSSIM	13.631	0.639
4 - Sem ponderação da função objetivo	12.494	0.584
Não tratada	8.628	0.664

repercutem na qualidade da restauração obtida. Nesta situação, o modelo menos efetivo é o modelo sem ponderação na função objetivo, seguido pelo modelo treinado apenas utilizando DSSIM e pelo modelo baseado em uma arquitetura sem blocos convolucionais. Em todos os cenários, verifica-se que os modelos apresentam uma melhora do PSNR. Já quando considera-se a métrica SSIM verifica-se que algumas das combinações impactam negativamente esta métrica com relação ao conjunto de imagens não tratadas. Ressalta-se, no entanto, que os valores de SSIM mensurados nas imagens sobre-expostas não tratadas são bastante distintos dos valores de SSIM mensurados em imagens sub-expostas não tratadas.

5.6 Visualização da Ativação Interna do Modelo

Com atenção especial no bloco de convoluções dilatadas introduzido na Seção 4.2, a presente Seção discute os principais benefícios e limitações do modelo proposto, levando em consideração os mapas de atenção para cada camada convolucional do modelo. A visualização da ativação interna do modelo auxilia na identificação de quais regiões tem mais relevância e permite saber se a rede está utilizando as partes corretas para realizar a restauração.

5.6.1 Aplicação em Imagem Sub-Exposta

O remanescente desta seção leva em consideração as imagens apresentadas na Figura 5.12. Trata-se de uma imagem severamente danificada em função do uso de tempo de exposição inadequado. Apesar de bastante escura, a imagem de entrada apresenta pouco ruído. Para melhor visualização, recomenda-se a utilização do formato digital, de forma que seja possível observar os detalhes com maior fidelidade. A escala de cores utilizada em todos os mapas de ativação é apresentada na Figura 5.11.



Figura 5.11: Escala de cores utilizada para representação dos mapas de ativação. Azul indica atenção mínima. Vermelho indica atenção máxima.



Figura 5.12: Imagem de entrada subexposta, saída do modelo e referência.

A Figura 5.13 apresenta a atuação de cada tipo de convolução dentro do primeiro bloco convolucional na rede proposta. Neste bloco as imagens estão na resolução total, não tendo passado por qualquer redimensionamento. As quatro camadas apresentadas encontram-se em um arranjo paralelo, permitindo que todas tenham acesso direto à imagem de entrada. Pode-se observar que, nesta primeira etapa, a rede utiliza as convoluções 3×3 sem dilatação para preservar a informação nas regiões mais claras da imagem de entrada. A preservação destas regiões contribui para que o modelo seja capaz de preservar informação local, como texturas e bordas. As convoluções com dilatação 2 e 4 atribuem maior relevância para as regiões escuras da imagem. Já as convoluções com dilatação 8 destacam a ocorrência de transições de intensidade.

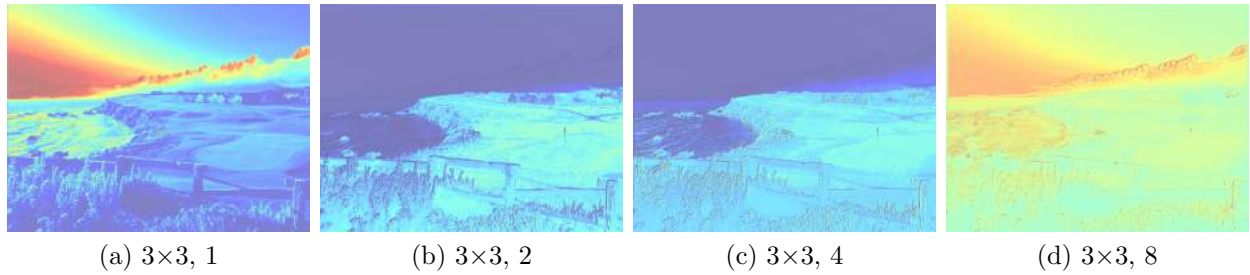


Figura 5.13: Primeiro bloco convolucional, atuando sobre a *imagem* na resolução original na restauração de uma imagem subexposta. (a),(b),(c) e (d) são paralelas.

A Figura 5.14 apresenta o fluxo dos dados dentro da arquitetura *encoder*. (a),(b),(c) e (d) apresentam o primeiro bloco convolucional dentro do *encoder*, atuando sobre a imagem com $\frac{1}{2}$ da resolução original. Identifica-se que a convolução sem dilatação, juntamente com convoluções dilatadas 2 e 4, apresentam maior impacto na restauração da imagem. Observa-se também uma especialização por padrão de textura na imagem. Este padrão é relevante pois permite que o modelo preserve *features* locais e agregue correção de cor e luminância da imagem como um todo. Destaca-se ainda que, neste ponto, a rede já identificou os pontos mais críticos para a restauração.

Ainda dentro do fluxo *encoder*, o próximo bloco convolucional atua sobre a imagem com $\frac{1}{4}$ da resolução original. (e),(f),(g) e (h) na Figura 5.14 são bastante heterogêneas. Identifica-se aqui que a maior parte da atenção é concentrada nas convoluções 3×3 com dilatação 2. As convoluções sem dilatação agregam informação somente em regiões da imagem ricas em textura. A convolução de dilatação 8 contribui para o preenchimento das regiões mais críticas da imagem.

O último bloco convolucional dentro do *encoder*, representada pelas imagens (i),(j),(k) e (l) na Figura 5.14, atua sobre a imagem com $\frac{1}{8}$ da resolução original. Observa-se que, nestas condições, o fluxo de dados é definido majoritariamente pelas convoluções sem dilatação. As convoluções com dilatação 2 e 4 aparentam não destacar elementos específicos da imagem, apesar de contribuírem significativamente. Já as convoluções com dilatação 8 continuam dando atenção à região mais crítica da imagem. Este tipo de comportamento é esperado, uma vez que as etapas iniciais do *encoder* já transformaram significativamente a imagem.

A visualização das ativações em cada uma das camadas do *encoder* permite observar o funcionamento do bloco convolucional em diferentes níveis de profundidade na rede neural. Observa-se, de forma explícita, a atuação das convoluções dilatadas na agregação de contexto para a restauração das partes mais impactadas pela sub-exposição e para a preservação das texturas presentes na imagem inapropriadamente exposta.

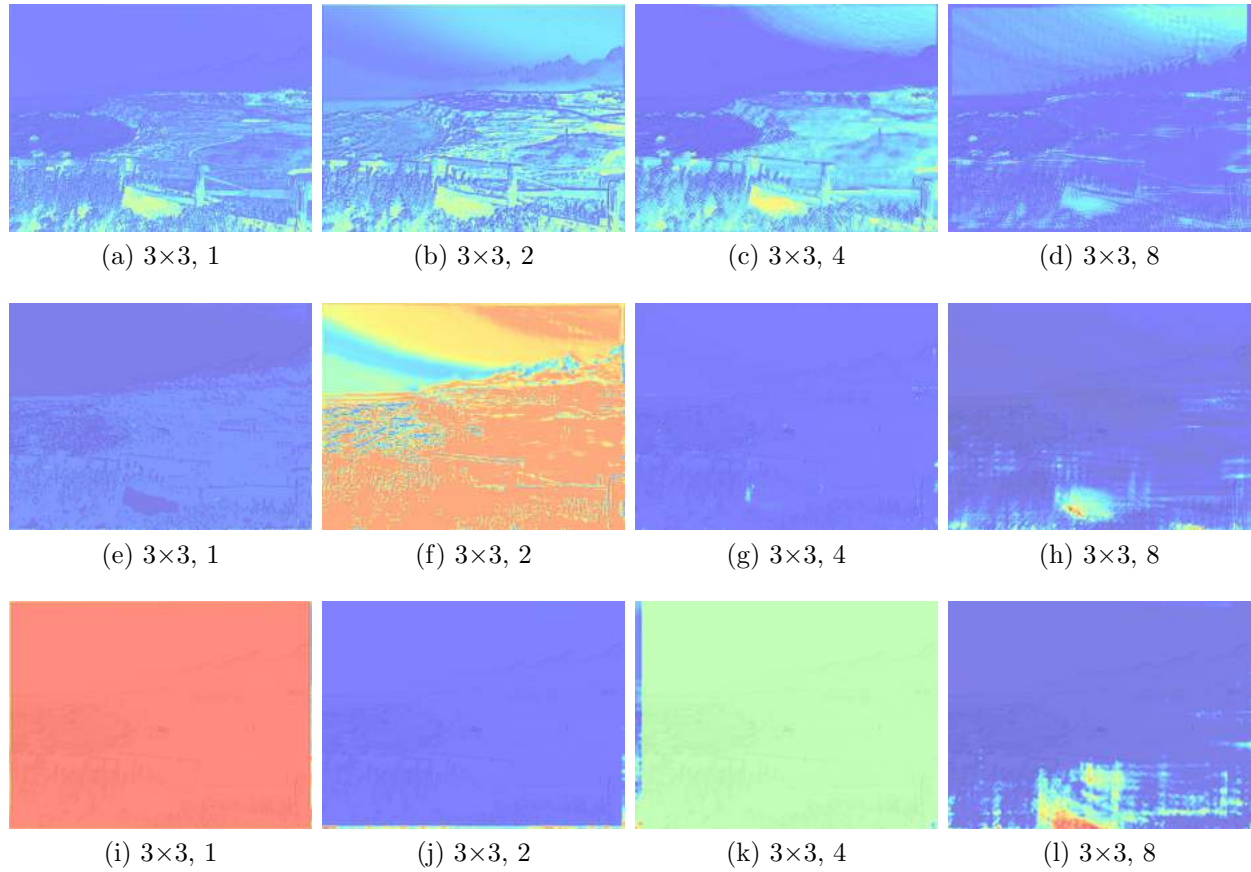


Figura 5.14: Blocos convolucionais dentro do *encoder* na restauração de uma imagem subexposta. Cada linha apresenta as convoluções paralelas dentro do bloco convolucional.

A Figura 5.15 apresenta a atenção dentro das camadas do *decoder*. Observa-se, na primeira camada do *decoder*, que as convoluções sem dilatação (a) e com dilatação 2 (b) carregam informações de toda a imagem, de forma quase homogênea. Já as convoluções com dilatação maior (c) e (d), e por consequência, com acesso à uma região maior de vizinhança, continuam apresentando contribuição significativa para o preenchimento de regiões mais críticas para a restauração.

Na segunda camada do *decoder*, representada pelas imagens (e), (f), (g) e (h), da Figura 5.15, observa-se maior contribuição por parte das convoluções sem dilatação. Estas contribuem ainda para a transformação da imagem, ajustando os valores na região mais escura. Nas convoluções dilatadas não se observa o mesmo destaque, ou seja, elas contribuem de forma menos expressiva para a formação da imagem final. Verifica-se no entanto que as convoluções dilatadas neste nível da rede são utilizadas pelo modelo para a preservação de zonas de transição de intensidade, contribuindo para a preservação da similaridade estrutural entre a imagem de entrada e a imagem restaurada pelo modelo.

Na terceira e última camada do *decoder*, representada pelas imagens (i), (j), (k) e (l), da Figura 5.15, observa-se que todas as convoluções apresentam pouca contribuição na transformação da imagem. Por um lado, para a imagem avaliada, tanto a convolução regular quanto as convoluções dilatadas não parecem oferecer contribuição para a correção de pontos muito escuros ou muito claros. Pelo outro, destaca-se que as regiões que apresentam detalhes ricos em textura, como a textura do oceano, contornos de nuvens e vegetação recebem maior atenção. Desta forma, verifica-se que mesmo na última camada do *decoder* a arquitetura utilizando um bloco de convoluções paralelas favorece a performance do modelo.

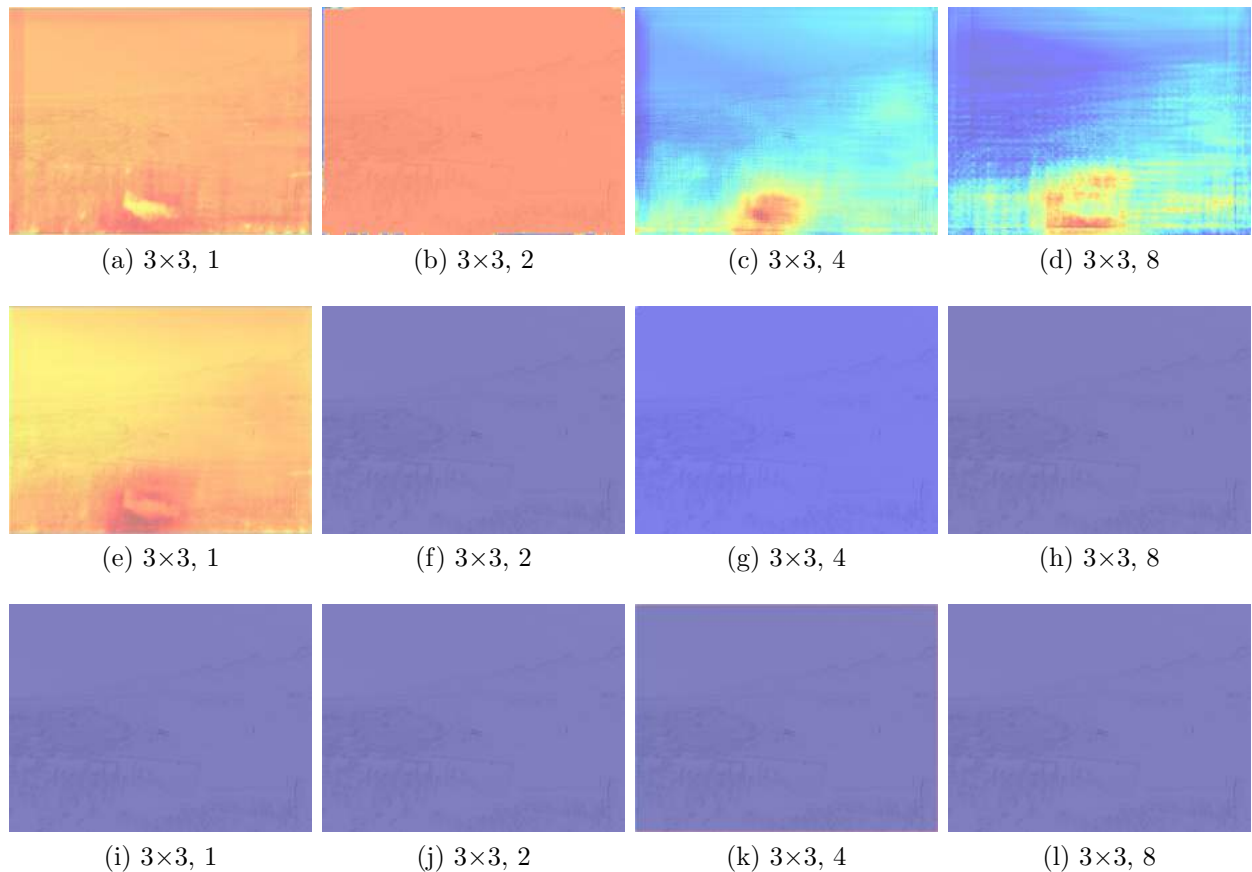


Figura 5.15: Blocos convolucionais dentro do *decoder* na restauração de uma imagem subexposta. Cada linha apresenta as convoluções paralelas dentro do bloco convolucional.

A Figura 5.16 apresenta o fluxo paralelo à estrutura *encoder-decoder*. Este fluxo é desenhado com o intuito de preservar *features* locais na imagem de entrada. Para tanto, é composto de

duas camadas convolucionais 3×3 sem dilatação em sequência. Observa-se que para a imagem avaliada, somente a primeira camada (a) oferece contribuição significativa na transformação da imagem, tendo sua atenção concentrada nas regiões mais escuras e com muitos detalhes. A segunda camada (b) atua apenas como passagem para o fluxo de dados, não sendo identificada contribuição para restauração de qualquer parte específica da imagem.

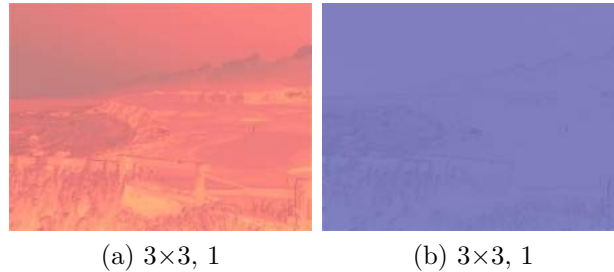


Figura 5.16: Fluxo paralelo ao *encoder-decoder* composto por convoluções 3×3 sem dilatação, atuando sobre a imagem na resolução original na restauração de uma imagem subexposta. (a) e (b) são sequenciais

Por fim, a Figura 5.17 mostra a atenção da rede nas camadas finais. Estas camadas são precedidas pela concatenação dos fluxos *encoder-decoder* com o fluxo da imagem na resolução original e uma etapa de *Instance Normalization*. Todas utilizam convoluções 3×3 sem dilatação. Nota-se que nestas camadas a atenção da rede é concentrada em regiões com transição brusca de intensidade, bem como regiões que receberam menos evidência nas camadas anteriores.

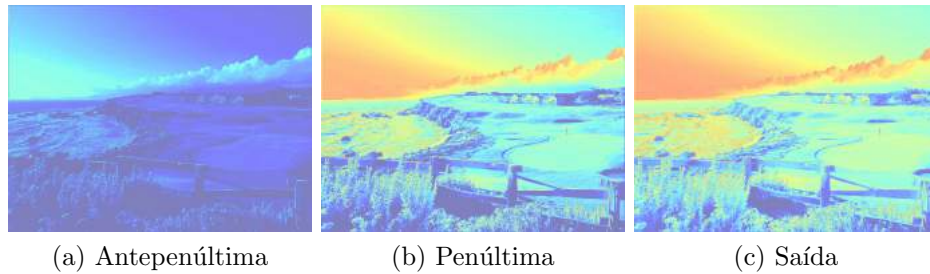


Figura 5.17: Três últimas camadas sequenciais da rede atuando na restauração de uma imagem subexposta. (a),(b) e (c) são sequenciais

5.6.2 Aplicação em Imagem Sobre-Exposta

A Figura 5.18 apresenta uma imagem de ambiente interno afetada pelos efeitos de sobre-exposição. Destaca-se a região central da imagem, onde diversos objetos que compõe a cena não são visíveis na imagem sobre-exposta (a). Nota-se ainda a existência de reflexo no plano do vidro. Na imagem restaurada (b) pelo modelo proposto, objetos e texturas desta região ficam visíveis, sendo facilmente identificados pelo observador humano. O mesmo ocorre nas mudanças repentina de intensidade, como nos marcos da janela e regiões com texto. Com relação à imagem referência (c), observa-se uma leve distorção da cor.

A Figura 5.19 apresenta a atuação de cada tipo de convolução dentro do primeiro bloco convolucional na rede proposta. Esse bloco opera sobre as imagens sem redimensionamento.



Figura 5.18: Imagem de entrada sobre-exposta, saída do modelo e referência.

Todo o fluxo da rede passa pelo bloco. As quatro camadas apresentadas encontram-se em um arranjo paralelo, permitindo que todas tenham acesso direto à imagem de entrada. Pode-se observar que, nesta primeira etapa, a rede utiliza as convoluções 3×3 sem dilatação para preservar a informação nas regiões mais claras da imagem de entrada. Estas regiões contribuem para que o modelo seja capaz de preservar informação local, como texturas e bordas. Para a imagem em questão, as convoluções com dilatação não destacam nenhuma região em particular.

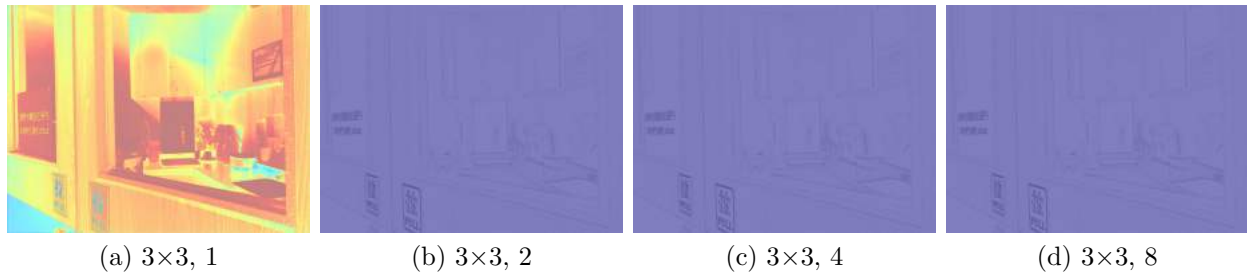


Figura 5.19: Primeiro bloco convolucional, atuando sobre a *imagem* na resolução original na restauração de uma imagem sobre-exposta. (a),(b),(c) e (d) são paralelas

A Figura 5.20 apresenta o fluxo dos dados dentro da arquitetura *encoder*. (a),(b),(c) e (d) apresentam o primeiro bloco convolucional dentro do *encoder*, atuando sobre a imagem com $\frac{1}{2}$ da resolução original. Identifica-se que a convolução sem dilatação (a), juntamente com as convoluções de dilatação 8 (b), concentram a maior parte da atenção na rede. Para este caso, observa-se um comportamento distinto do observado na imagem subexposta (Figura 5.14), em que convoluções dilatadas em 2 e 4 apresentam maior impacto na restauração da imagem. Observa-se também uma especialização por padrão de textura na imagem. Este padrão é relevante pois permite que o modelo preserve *features* locais e agregue correção de cor e luminância da imagem como um todo. Destaca-se ainda que, neste ponto, as convoluções com dilatação 8 (d) parecem atuar na identificação de regiões com iluminação não homogênea (algo similar ao que é feito em processamento de imagens clássico utilizando esquemas de cor que separam tonalidade de cor e saturação).

O bloco convolucional seguinte atua sobre a imagem com $\frac{1}{4}$ da resolução original, ainda dentro do fluxo *encoder*. (e), (f), (g) são bastante homogêneas, não sendo possível destacar regiões específicas da imagem nas quais contribuem mais. Observa-se que a maior parte da atenção é concentrada nas convoluções 3×3 com dilatação 8 (h), que contribui para o preenchimento das regiões mais críticas da imagem.

Por fim, o último bloco convolucional dentro do *encoder*, representado pelas imagens (i),(j),(k) e (l), atua sobre a imagem com $\frac{1}{8}$ da resolução original. Observa-se que, nestas condições, o fluxo de dados se dá majoritariamente pelas convoluções com dilatação 8 (l) e a (k). As convoluções sem dilatação (i), com dilatação 2 (j) e 4 (k) aparentam não destacar elementos específicos da imagem.

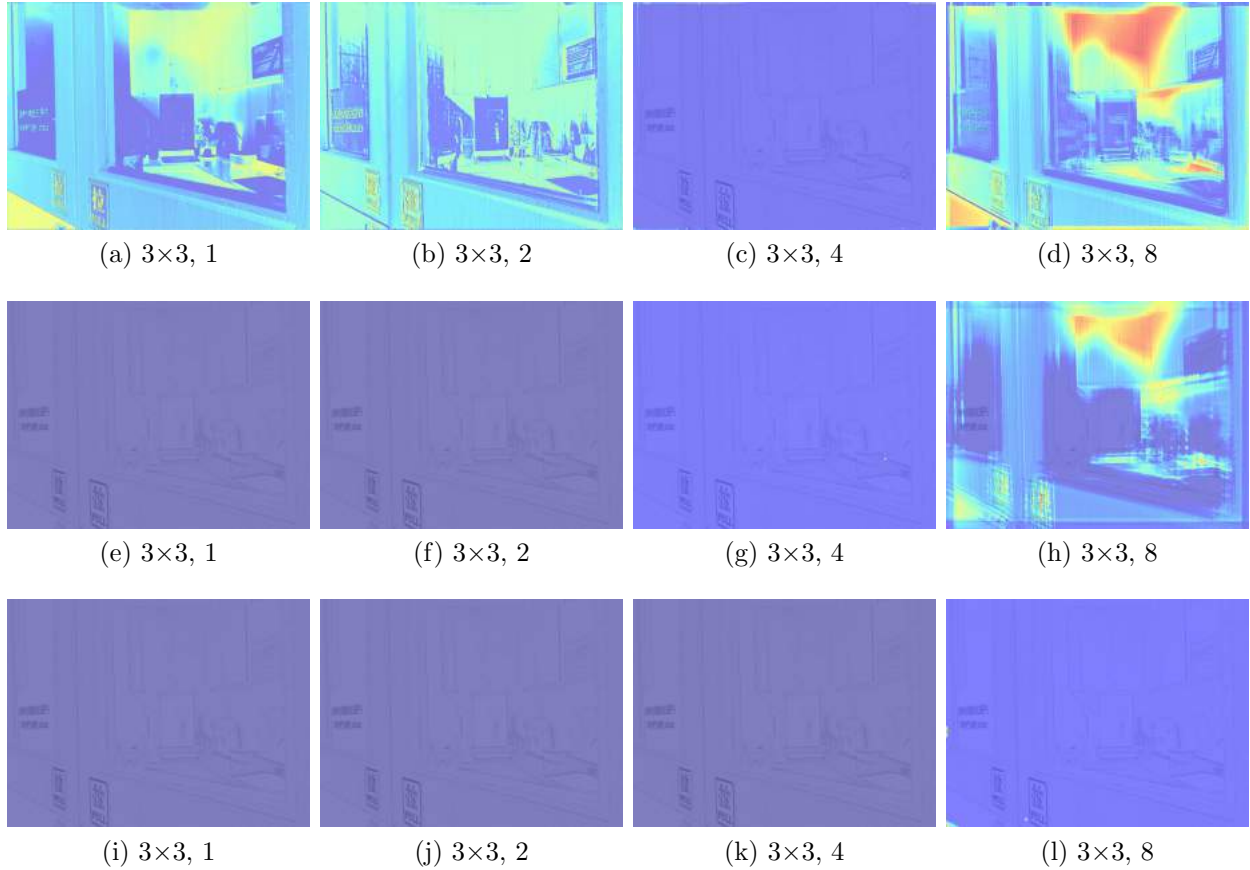


Figura 5.20: Blocos convolucionais dentro do *encoder*, atuando na restauração de uma imagem sobre-exposta. Cada linha apresenta as convoluções paralelas dentro do bloco convolucional.

A Figura 5.21 apresenta a atenção dentro das camadas do *decoder*. Observa-se na primeira camada do *decoder* que convolução sem dilatação (a) carrega informações de forma homogênea, não sendo possível visualizar algum ponto específico de maior relevância. Já as convoluções com dilatação 2 (b) concentram atenção na restauração do sinal atenuado por reflexo. As convoluções com dilatação 8 (b) carregam informações de toda a imagem, ponderando com maior relevância as variações de iluminação. Neste bloco do *decoder*, convoluções com dilatação 2 e 8 são as que mais oferecem contribuição para o resultado final da restauração.

Na segunda camada do *decoder*, representada pelas imagens (e), (f), (g) e (h) da Figura 5.21, observa-se que todas as convoluções contribuem de maneira homogênea, não prendendo atenção em qualquer região específica da imagem. Por fim, na terceira e última camada do *decoder*, observa-se que as convoluções de dilatação maior (k) e (l) contribuem de forma mais acentuada para a formação da imagem restaurada.

Observa-se, comparada à atuação do *decoder* para restaurações de imagens sub e sobre-expostas (Figuras 5.21 e 5.15), que existe uma mudança expressiva nos mapas de ativação para cada bloco convolucional. Para imagens subexpostas, na terceira camada do *decoder*, tanto a convolução regular quanto as convoluções dilatadas não parecem oferecer contribuição para a correção de pontos específicos. Já na restauração de imagens sobre-expostas, verifica-se que a segunda camada atua de forma mais homogênea, ao passo que a terceira camada apresenta um maior fluxo de informação a partir das convoluções com maior campo receptivo.

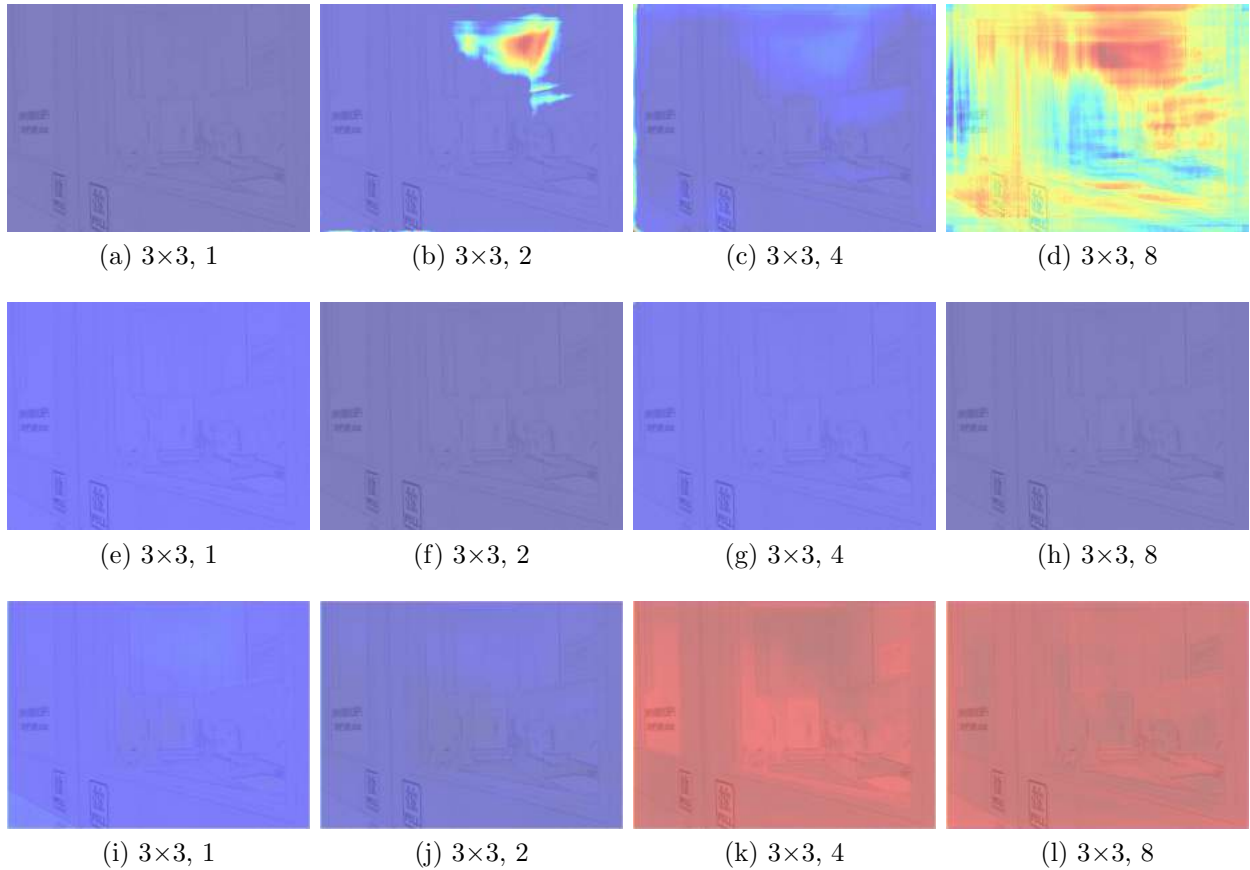


Figura 5.21: Blocos convolucionais dentro do *decoder*, atuando na restauração de uma imagem sobre-exposta. Cada linha apresenta as convoluções paralelas dentro do bloco convolucional.

A Figura 5.22 apresenta o fluxo paralelo à estrutura *encoder-decoder*. Este fluxo é desenhado com o intuito de preservar *features* locais na imagem de entrada e opera sobre a imagem na resolução original. Para tanto, é composto de duas camadas convolucionais 3×3 sem dilatação em sequência. De forma similar ao que ocorre na imagem subexposta (Figura 5.15), observa-se que, para a imagem avaliada, somente a primeira camada (a) oferece contribuição significativa na transformação da imagem. No entanto, percebe-se aqui que a atenção desta camada é concentrada nas regiões mais claras, e, portanto, mais suscetíveis à saturação. Esta camada destaca ainda as regiões com mudança brusca de gradiente e ricas em textura. A camada seguinte (b) atua apenas como passagem para o fluxo de dados, não sendo identificada contribuição para restauração de qualquer parte da imagem em específico.

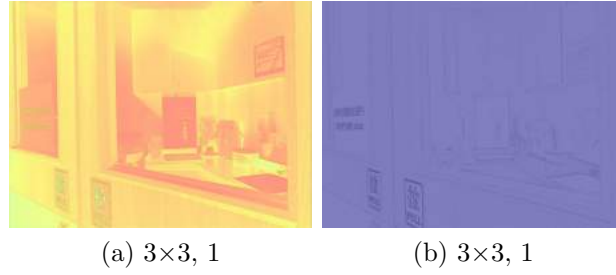


Figura 5.22: Fluxo paralelo ao *encoder-decoder* composto por convoluções 3×3 sem dilatação, atuando sobre a imagem na resolução original na restauração de uma imagem sobre-exposta. (a) e (b) são sequenciais

Por fim, a Figura 5.23 mostra a atenção da rede nas camadas finais. Estas camadas são precedidas pela concatenação dos fluxos *encoder-decoder* com o fluxo da imagem na resolução original e uma etapa de *Instance Normalization*. Todas utilizam convoluções 3×3 sem dilatação. Observa-se nessas camadas finais da rede que a atenção é concentrada em regiões que demandam maior transformação com relação à imagem de entrada. As três camadas da rede privilegiam a restauração de grandes regiões claras e a manutenção de detalhes de textura.

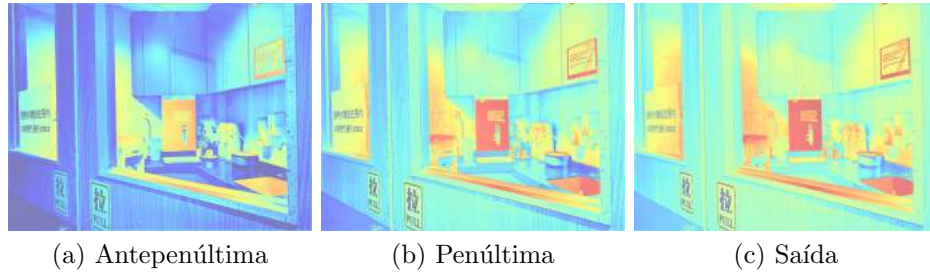


Figura 5.23: Três últimas camadas sequenciais da rede. (a),(b) e (c) são sequenciais

5.6.3 Considerações sobre a Visualização

A aplicação do método de visualização da atenção nas camadas internas do modelo proposto permite uma melhor compreensão sobre a extração e utilização de *features* pela rede. Naturalmente, o problema de predição imagem-imagem apresenta distinções expressivas em relação aos métodos de classificação, campo onde as técnicas de visualização já se encontram em um estágio mais maduro. Desta forma, a avaliação visual, por si só, não oferece subsídio suficiente para avaliar a robustez e confiabilidade do modelo (Protas et al., 2018).

Apesar das limitações, a visualização permite percorrer as diversas camadas dentro da rede neural convolucional proposta, identificando quais as características da imagem de entrada são levadas em consideração pelo modelo para fazer a restauração da imagem. Esse tipo de estudo, ainda incomum na área de pesquisa, pode ajudar no desenvolvimento de arquiteturas melhores e mais eficientes.

Com relação ao bloco de convoluções dilatadas, desenhado com o intuito de expandir exponencialmente o campo receptivo e permitir a utilização de modelos menos profundos, a

visualização permite identificar a contribuição de cada nível de dilatação para a formação da imagem final. Observa-se, que, como era desejado, a rede se beneficia desse arranjo, utilizando as convoluções sem dilatação para preservação de aspectos locais e as convoluções dilatadas para agregação de contexto, a partir da vizinhança de cada píxel.

Observa-se ainda que em algumas etapas a rede utiliza as convoluções muito dilatadas para destacar regiões de tonalidade semelhante, mas afetadas por iluminação não homogênea (posição de iluminação e sombra). Pode-se conjecturar conseqüentemente, que a rede aprende a decompor a imagem de forma similar aos algoritmos de restauração de imagens clássicos baseados em constância de cor. Algoritmos de constância de cor buscam, geralmente, isolar as componentes de iluminação da cena, de forma a minimizar seu impacto em sua representação.

Nos fluxos externos à estrutura *encoder-decoder*, compostos por convoluções 3×3 regulares, observa-se grande magnitude na atenção, identificando-se claramente as regiões que receberam maior importância na restauração. Já nas camadas finais, com acesso tanto à saída da rede *encoder-decoder* com *skip connections* quanto ao fluxo paralelo, evidencia-se que a atenção maior está sempre concentrada nas regiões mais claras da imagem de entrada e regiões com grande variabilidade de intensidade.

Empiricamente, mostra-se que a arquitetura proposta, apesar de pouco profunda, utiliza os recursos disponíveis de maneira efetiva, sendo capaz de restaurar o sinal afetado por baixo contraste decorrente de exposição equivocada. Evidencia-se ainda que nas regiões onde ocorre saturação total do sinal, o modelo é capaz de utilizar informações existentes em regiões próximas e interpolar cor e textura.

5.7 Limitações conhecidas

O modelo de rede neural convolucional proposto aprende uma função não linear para mapear imagens sub ou sobre-expostas em imagens com bom contraste, restaurando características de cor, textura e semântica. Além disso, se beneficia de conhecimento agregado durante a etapa de treinamento para interpolar valores de píxel em regiões nas quais uma parcela significativa da informação foi truncada por extrapolar os limites do sistema de aquisição. O potencial da aplicação do modelo para restauração de imagens afetadas por exposição inadequada fica evidenciado qualitativamente e quantitativamente através das métricas aplicadas na avaliação.

Existem, no entanto, limitações e problemas conhecidos que podem ser trabalhados de forma a aprimorar os resultados obtidos. Dentre estes, destaca-se:

1. Limitações no preenchimento e recuperação de detalhes de textura em grandes blocos saturados: Esta limitação está associada diretamente ao problema de *inpainting*, no qual partes faltantes de uma imagem são preenchidas de modo a gerar uma imagem completa. Este é um problema de tratamento complexo, de forma que existe um ramo de pesquisa específico para este problema. Quando toda informação de textura e cor é perdida como consequência da exposição equivocada da cena, o modelo precisa atuar de forma a interpolar a partir de dados preservados na vizinhança. Em geral, observa-se que o modelo

proposto preenche a região com uma média das cores próximas sem gerar texturas ou artefatos inconsistentes com a cena.

2. Artefatos de blocos oriundos da compressão JPEG, especialmente perceptível em regiões da imagem que apresentam baixo contraste: Ao comprimir a imagem já danificada pela exposição inapropriada, a compressão JPEG introduz perdas adicionais. Uma vez que a compressão se dá por blocos, a imagem comprimida passa a apresentar zonas de transição abrupta inconsistentes com a visualização da cena. Sem conhecimento destes artefatos, o modelo de restauração acaba ampliando os efeitos adversos. Além disso, pequenas estruturas na imagem e gradientes suaves não podem ser recuperados pelo modelo proposto.

6 CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

A restauração de imagens é uma das tarefas-chave de processamento de imagem. O trabalho apresentado tem particular interesse na restauração de imagens capturadas em condições não ideais de exposição, englobando subexposição e sobre-exposição. Imagens capturadas com exposição inapropriada apresentam, com frequência, baixo contraste e perda de informação nas regiões muito claras ou muito escuras da cena. Esses efeitos são oriundos de limitações do sensor de imagem e do arranjo óptico utilizado para captura. Além dos efeitos de amostragem e quantização, as imagens também, frequentemente, sofrem com artefatos e perda de dados decorrentes de compressão.

Saturação, subexposição e artefatos de compressão impedem que os detalhes da cena sejam representados adequadamente na imagem capturada. Neste trabalho propõe-se e avalia-se uma alternativa baseada em redes neurais convolucionais ponta-a-ponta para a restauração de imagens danificadas pelas condições supracitadas. Esta abordagem é apropriada para o problema em questão, uma vez que a aquisição é afetada por muitas variáveis não controladas ou sequer mensuráveis.

Inicialmente, faz-se uma revisão do estado da técnica relacionada com a restauração de imagens afetadas pela exposição inadequada, bem como de abordagens empregadas no problema de transformação imagem imagem de forma mais abrangente. Identificam-se, a partir deste levantamento inicial, algumas lacunas na área de conhecimento ainda em estágio inicial. Estas lacunas estão relacionadas à limitação dos conjuntos de dados adequados para a investigação de técnicas na área, à ausência de avaliações experimentais sobre métricas de qualidade de imagens quando empregadas para condições de subexposição e sobre-exposição – uma vez que seu emprego é geralmente ligado à estimação de ruído ou perdas decorrentes de compressão –, e ao baixo número de técnicas e modelos dedicados ao tratamento de sobre-exposição.

Como primeira contribuição da pesquisa, tem-se a criação e disponibilização para uso público de um *dataset* de imagens pareadas contemplando condições de subexposição e sobre-exposição. Juntamente com o *dataset* SICE (Cai et al., 2018), disponibilizado publicamente no mesmo período de realização da pesquisa, este é um dos poucos *datasets* públicos que contempla a também a condição de sobre-exposição. Destaca-se que estes *datasets* de cenas cotidianas e criados com o propósito específico de permitir o desenvolvimento e avaliação de modelos de restauração de imagens inapropriadamente expostas são catalisadores para o desenvolvimento da área de pesquisa.

Como segunda contribuição da pesquisa tem-se a avaliação de um conjunto amplo de métricas e medidas de qualidade de imagem considerando a perspectiva da exposição inadequada. Embora muitas destas métricas já venham sendo aplicadas empiricamente na avaliação de métodos de restauração pouco se sabe sobre o comportamento das mesmas quando aplicadas ao problema. Nesta pesquisa apresenta-se uma avaliação detalhada, evidenciando que na maior

parte apresentam correlação entre si, mas mantêm características individuais que justificam o seu emprego para um propósito específico. Além disso, evidencia-se que geralmente os resultados fornecidos pelas métricas de qualidade não apresentam distribuição normal, não sendo, portanto, recomendada a utilização da média aritmética simples como medida de tendência central, prática que é amplamente difundida entre pesquisadores da área.

Como contribuição principal, desenvolve-se um modelo de rede convolucional com a finalidade de restaurar imagens sub e sobre-expostas. O modelo, apresentado no Capítulo 4, combina em sua arquitetura aspectos de modelos U-Net Ronneberger et al. (2015), CAN (*Context Aggregation Network*)(Chen et al., 2017) e NIN (*Network In Network*)(Lin et al., 2013) levando em consideração as particularidades do problema tratado. O modelo traz também elementos destinados a acelerar o tempo de treinamento, reduzir os requisitos de memória e aprimorar o resultado da restauração. Verifica-se que a rede proposta é capaz de convergir em um modelo de transformação de imagens adequado ao problema tratado. Os resultados obtidos são avaliados usando vários índices de qualidade de imagem, indicando que a rede proposta é apta a mitigar efeitos de imagens danificadas pela exposição heterogênea. Verifica-se que o método proposto oferece ganho significativo em relação aos demais métodos comparados, tanto em dados simulados quanto em dados reais.

Os resultados qualitativos e quantitativos indicam que o modelo de restauração proposto, baseado em rede neural convolucional, é capaz de restaurar e reconstruir imagens impactadas por exposição inapropriada, preservando as características mais representativas para a visão computacional como definição, contraste, nitidez e correção de cor. Uma avaliação através de 14 medidas de qualidade de imagem indica que o modelo supera os métodos utilizados no comparativo na maior parte dos casos. A avaliação utilizando testes de hipótese mostra que o ganho obtido é significativo. Nas situações em que os trabalhos relacionados não disponibilizaram todos os detalhes de arquitetura e implementação da rede, os dados necessários não estavam publicamente acessíveis, fez-se uma avaliação subjetiva considerando os melhores resultados apresentados na literatura disponível.

Os resultados a nível de aplicação, discutidos nas Seção 5.4, mostram que o problema tratado deteriora expressivamente o desempenho de distintos modelos de reconhecimento de imagens e segmentação de instâncias. Mostra-se ainda que a utilização do modelo proposto em *pipeline* permite mitigar os efeitos indesejados causados pela exposição inadequada das imagens sem necessidade de retreinamento. Por fim, faz-se uma avaliação do tempo de inferência do modelo em diferentes condições, verificando-se a viabilidade de uso em aplicações práticas. Evidencia-se, desta forma a contribuição do modelo dentro dos objetivos propostos.

Por meio da visualização das camadas internas do modelo, verifica-se que a arquitetura de rede proposta é utilizada de forma efetiva, agregando informação de diversas partes da imagem para realizar o preenchimento de regiões saturadas ou subexpostas. A estrutura permite que a rede preserve informação nas regiões com forte transição de intensidade, conservando texturas e bordas.

Dentre as principais limitações do modelo proposto, identifica-se uma dificuldade no pre-

enchimento em locais com saturação severa (grandes áreas) e realce dos artefatos de blocos oriundos da compressão JPEG, especialmente perceptível em regiões da imagem que apresentam baixo contraste. A primeira limitação está relacionada diretamente com o problema de interpolação de dados. A segunda tem relação com ruído de alta frequência nas bordas de cada bloco comprimido e com a perda de informação inerente à compressão com perdas.

Como continuação desta proposta, propõe-se a investigação dos impactos de compressão de imagens em condições de subexposição e sobre-exposição para as aplicações de visão computacional e sobre os modelos de restauração. Observado-que a compressão é uma condição inerente a estas aplicações e que provoca a perda de informações sobre os níveis de intensidade da imagem, bem como a ocorrência de artefatos de blocos e perda de textura, verifica-se um ponto que demanda investigação mais aprofundada.

Da mesma forma, considerando aplicações que exigem o processamento de uma sequência de imagens, acredita-se que a pesquisa envolvendo o processamento de vídeo como forma de obter informações sobre cor, textura e mesmo a semântica da cena possa viabilizar o desenvolvimento de modelos mais robustos. Ainda como trabalho futuro, considera-se a utilização de modelos de perda perceptual (Blau & Michaeli, 2018; Egiazarian et al., 2018) como função objetivo, de forma a maximizar a capacidade de restauração de características visualmente relevantes da imagem e mitigar os efeitos de bloco oriundos da restauração de imagens comprimidas.

Ainda como trabalho futuro, cabe aprofundar a discussão sobre métricas de qualidade de imagem voltadas para a avaliação de condições de exposição inapropriada. Conforme o levantamento apresentado as métricas atualmente empregadas foram concebidas considerando, em geral, a avaliação de perdas causadas por compressão ou ruído. Embora estas tenham se mostrado úteis na avaliação dos resultados atingidos pelo modelo proposto, especialmente quando considera-se a aplicação em dados pareados, verifica-se que o estudo de métricas desenhadas especificamente para o problema da exposição inapropriada é ainda um problema em aberto.

Resultados parciais e contribuições do trabalho foram publicados em Hüttner et al. (2017), onde são explorados os aspectos que compõe a arquitetura e otimizadores utilizados para treinamento de redes neurais; Vaz-Jr et al. (2017), Vaz-Jr et al. (2018a) e Vaz-Jr et al. (2018b), onde são explorados algoritmos baseados em fusão de imagens para restauração de imagens cuja visibilidade é prejudicada pela presença de barreiras físicas entre a câmera e a cena; Steffens et al. (2017) onde são avaliadas distintas métricas cegas de qualidade de imagens com foco em superfícies reflexivas ou polidas; Steffens et al. (2018b) onde apresenta-se uma análise exploratória de dados da distribuição de intensidade em fotografias noturnas; Steffens et al. (2018a) Steffens et al. (2019) e Steffens et al. (2020a) onde se apresentam diferentes versões do modelo de restauração proposto; e Steffens et al. (2020b) e Steffens et al. (2021) onde se discutem os aspectos de robustez dos modelos de reconhecimento de imagens e se demonstra a viabilidade de aplicação do modelo de restauração proposto em aplicações autônomas e robóticas.

Referências

- Abdullah-Al-Wadud, M., Kabir, M. H., Dewan, M. A. A., & Chae, O. (2007). A dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(2):593–600.
- Affi, M., Derpanis, K. G., Ommer, B., & Brown, M. S. (2020). Learning to correct overexposed and underexposed photos. *arXiv preprint arXiv:2003.11596*.
- Ai, S. & Kwon, J. (2020). Extreme low-light image enhancement for surveillance cameras using attention u-net. *Sensors*, 20(2):495.
- Amudha, J., Pradeepa, N., & Sudhakar, R. (2012). A survey on digital image restoration. *Procedia engineering*, 38:2378–2382.
- Ancuti, C. & Ancuti, C. O. (2013). Single image dehazing by multi-scale fusion. *IEEE Transactions on Image Processing*, 22(8):3271–3282.
- Ancuti, C., Ancuti, C. O., Haber, T., & Bekaert, P. (2012). Enhancing underwater images and videos by fusion. *IEEE Conference of Computer Vision and Pattern Recognition (CVPR)*, pages 81–88.
- Ancuti, C. O., Ancuti, C., Vleeschouwer, C. D., & Bovik, A. C. (2017). Single-scale fusion: An effective approach to merging images. *IEEE Transactions on Image Processing*, 26(1):65–78.
- Appuhamy, E. & Madhusanka, B. (2018). Development of a gpu-based human emotion recognition robot eye for service robot by using convolutional neural network. In *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, pages 433–438. IEEE.
- Aravena, N. C., Hermosilla, G., Vera, E., & Fariás, G. (2018). Cj: An intelligent robotic head based on deep learning for hri. In *2018 IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA)*, pages 1–6. IEEE.
- Arvind, C., Mishra, R., Vishal, K., & Gundimeda, V. (2018). Vision based speed breaker detection for autonomous vehicle. In *Tenth International Conference on Machine Vision (ICMV 2017)*, volume 10696, page 106960E. International Society for Optics and Photonics.
- Atoum, Y., Ye, M., Ren, L., Tai, Y., & Liu, X. (2020). Color-wise attention network for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 506–507.
- Awodele, O. & Jegede, O. (2009). Neural networks and its application in engineering. *Science & IT*.

- Baroncini, V., Capodiferro, L., Di Claudio, E. D., & Jacovitti, G. (2009). The polar edge coherence: a quasi blind metric for video quality assessment. In *Signal Processing Conference, 2009 17th European*, pages 564–568. IEEE.
- Benenson, R., Popov, S., & Ferrari, V. (2019). Large-scale interactive object segmentation with human annotators. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11700–11709.
- Bengio, I. G. Y. & Courville, A. (2016). Deep learning. Book in preparation for MIT Press.
- Blau, Y., Mechrez, R., Timofte, R., Michaeli, T., & Zelnik-Manor, L. (2018). The 2018 pirm challenge on perceptual image super-resolution. In *European Conference on Computer Vision*, pages 334–355. Springer.
- Blau, Y. & Michaeli, T. (2018). The perception-distortion tradeoff. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6228–6237.
- Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26.
- Bychkovsky, V., Paris, S., Chan, E., & Durand, F. (2011). Learning photographic global tonal adjustment with a database of input / output image pairs. In *The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Cai, B., Xu, X., Jia, K., Qing, C., & Tao, D. (2016). Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198.
- Cai, J., Gu, S., & Zhang, L. (2018). Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Transactions on Image Processing*, 27(4):2049–2062.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, 8(6):679–698.
- Cepeda-Negrete, J. & Sanchez-Yanez, R. E. (2013). Gray-world assumption on perceptual color spaces. In *Pacific-Rim Symposium on Image and Video Technology*, pages 493–504. Springer.
- Chen, C., Chen, Q., Xu, J., & Koltun, V. (2018). Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Chen, C., Seff, A., Kornhauser, A., & Xiao, J. (2015a). Deepdriving: Learning affordance for direct perception in autonomous driving. In *The IEEE International Conference on Computer Vision (ICCV)*.
- Chen, J., Adams, A., Wadhwa, N., & Hasinoff, S. W. (2016). Bilateral guided upsampling. *ACM Transactions on Graphics (TOG)*, 35(6):203.

- Chen, Q., Xu, J., & Koltun, V. (2017). Fast image processing with fully-convolutional networks. In *IEEE International Conference on Computer Vision*, volume 9, pages 2516–2525.
- Chen, Y., Xiao, X., Liu, H.-l., & Feng, P. (2015b). Dynamic color image resolution compensation under low light. *Optik*, 126(6):603–608.
- Chi, W., Wang, J., & Meng, M. Q.-H. (2018). A gait recognition method for human following in service robots. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 48(9):1429–1440.
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258.
- Clevert, D.-A., Unterthiner, T., & Hochreiter, S. (2015). Fast and accurate deep network learning by exponential linear units (elus). *arXiv preprint arXiv:1511.07289*.
- Croux, C. & Dehon, C. (2010). Influence functions of the spearman and kendall correlation measures. *Statistical methods & applications*, 19(4):497–515.
- Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2007). Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095.
- Dawson-Howe, K. (2014). *A Practical Introduction to Computer Vision with OpenCV*. Wiley.
- de Stoutz, E., Ignatov, A., Kobyshev, N., Timofte, R., & Van Gool, L. (2018). Fast perceptual image enhancement. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 0–0.
- Debevec, P. E. & Malik, J. (2008). Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes*, page 31. ACM.
- Diane, S. A., Lesiv, E. A., Pesheva, I. A., & Neschetnaya, A. Y. (2019). Multi-aspect environment mapping with a group of mobile robots. In *2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, pages 478–482. IEEE.
- Dominguez-Molina, J. A., González-Farías, G., Rodríguez-Dagnino, R. M., & Monterrey, I. C. (2003). A practical procedure to estimate the shape parameter in the generalized gaussian distribution. *technique report I-01-18_eng. pdf*, available through http://www.cimat.mx/reportes/enlinea/I-01-18_eng.pdf, 1.
- Dong, X., Wang, G., Pang, Y., Li, W., Wen, J., Meng, W., & Lu, Y. (2011). Fast efficient algorithm for enhancement of low lighting video. In *2011 IEEE International Conference on Multimedia and Expo*, pages 1–6. IEEE.

- Drews-Jr, P., Hernández, E., Elfes, A., Nascimento, E. R., & Campos, M. (2016). Real-time monocular obstacle avoidance using underwater dark channel prior. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4672–4677.
- Drozdal, M., Vorontsov, E., Chartrand, G., Kadoury, S., & Pal, C. (2016). The importance of skip connections in biomedical image segmentation. In *Deep Learning and Data Labeling for Medical Applications*, pages 179–187. Springer.
- Du, D., Qi, Y., Yu, H., Yang, Y., Duan, K., Li, G., Zhang, W., Huang, Q., & Tian, Q. (2018). The unmanned aerial vehicle benchmark: Object detection and tracking. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 370–386.
- Egiazarian, K., Ponomarenko, M., Lukin, V., & Ieremeiev, O. (2018). Statistical evaluation of visual quality metrics for image denoising. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6752–6756. IEEE.
- Fairfield, N., Urmson, C. P., & Montemerlo, M. S. (2016). Camera based localization. US Patent 9,476,970.
- Fang, Y., Yang, S., Wang, X., Li, Y., Fang, C., Shan, Y., Feng, B., & Liu, W. (2021). Instances as queries.
- Forsyth, D. & Ponce, J. (2015). *Computer Vision: A Modern Approach*. Pearson Education Limited.
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association*, 32(200):675–701.
- Friedman, M. (1939). A correction: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association. American Statistical Association*, 34(205):109.
- Fu, X., Liao, Y., Zeng, D., Huang, Y., Zhang, X.-P., & Ding, X. (2015). A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation. *IEEE Transactions on Image Processing*, 24(12):4965–4977.
- Gao, F., Wang, C., Li, L., & Zhang, D. (2019). Altitude information acquisition of uav based on monocular vision and mems. *Journal of Intelligent & Robotic Systems*, pages 1–12.
- Gaya, J. O., Gonçalves, L. T., Duarte, A. C., Zanchetta, B., Drews-Jr, P., & Botelho, S. S. C. (2016). Vision-based obstacle avoidance using deep learning. In *2016 XIII Latin American Robotics Symposium and IV Brazilian Robotics Symposium (LARS/SBR)*, pages 7–12.
- Geiger, A., Lenz, P., Stiller, C., & Urtasun, R. (2013). Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, 32(11):1231–1237.

- Gharbi, M., Chen, J., Barron, J. T., Hasinoff, S. W., & Durand, F. (2017). Deep bilateral learning for real-time image enhancement. *ACM Transactions on Graphics (TOG)*, 36(4):118.
- Gijssenij, A., Gevers, T., & Van De Weijer, J. (2011). Computational color constancy: Survey and experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489.
- Glorot, X. & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256.
- Gonçalves, L. T., de Oliveira Gaya, J. F., Junior, P. J. L. D., & da Costa Botelho, S. S. (2018). Guidednet: Single image dehazing using an end-to-end convolutional neural network. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 79–86. IEEE.
- Gonzalez, R. & Woods, R. (2009). *Processamento digital de imagens (3a. ed.)*. Pearson Educación.
- Guo, C., Li, C., Guo, J., Loy, C. C., Hou, J., Kwong, S., & Cong, R. (2020). Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1780–1789.
- Guo, Y., Ke, X., Ma, J., & Zhang, J. (2019). A pipeline neural network for low-light image enhancement. *IEEE Access*, 7:13737–13744.
- Gupta, A. & Rush, A. M. (2017). Dilated convolutions for modeling long-distance genomic dependencies. *arXiv preprint arXiv:1710.01278*.
- Ha, I., Kim, H., Park, S., & Kim, H. (2018). Image retrieval using bim and features from pretrained vgg network for indoor localization. *Building and Environment*, 140:23–31.
- Hao, H., Yingkun, H., Yuxuan, S., Benzhen, W., & Jun, X. (2021). Nlhd: A pixel-level non-local retinex model for low-light image enhancement. *arXiv preprint arXiv:2106.06971*.
- Hasinoff, S. W. (2014). Saturation (imaging). In *Computer Vision*, pages 699–701. Springer.
- Hasinoff, S. W., Sharlet, D., Geiss, R., Adams, A., Barron, J. T., Kainz, F., Chen, J., & Levoy, M. (2016). Burst photography for high dynamic range and low-light imaging on mobile cameras. *ACM Transactions on Graphics (TOG)*, 35(6):192.
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969.
- He, K., Sun, J., & Tang, X. (2013). Guided image filtering. *IEEE transactions on pattern analysis & machine intelligence*, 35(6):1397–1409.

- He, K., Zhang, X., Ren, S., & Sun, J. (2016a). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016b). Identity mappings in deep residual networks. In *European conference on computer vision*, pages 630–645. Springer.
- Hendrycks, D. & Dietterich, T. (2019). Benchmarking neural network robustness to common corruptions and perturbations. *arXiv preprint arXiv:1903.12261*.
- Hendrycks, D. & Gimpel, K. (2016). Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*.
- Hochreiter, S., Bengio, Y., Frasconi, P., Schmidhuber, J., et al. (2001). *Gradient flow in recurrent nets: the difficulty of learning long-term dependencies*. A field guide to dynamical recurrent neural networks. IEEE Press.
- Honig, S. & Werman, M. (2018). Image declipping with deep networks. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3923–3927. IEEE.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
- Hu, Y., He, H., Xu, C., Wang, B., & Lin, S. (2018). Exposure: A white-box photo post-processing framework. *ACM Transactions on Graphics (TOG)*, 37(2):1–17.
- Hu, Y.-T., Huang, J.-B., & Schwing, A. (2017). Maskrcnn: Instance level video object segmentation. In *Advances in Neural Information Processing Systems*, pages 325–334.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2261–2269.
- Huang, J., Xiong, Z., Fu, X., Liu, D., & Zha, Z.-J. (2019). Hybrid image enhancement with progressive laplacian enhancing unit. In *Proceedings of the 27th ACM International Conference on Multimedia*, pages 1614–1622.
- Huang, Y., Sun, Z., Jin, L., & Luo, C. (2020). Epan: Effective parts attention network for scene text recognition. *Neurocomputing*, 376:202–213.
- Hüttner, V., Steffens, C. R., & da Costa Botelho, S. S. (2017). First response fire combat: Deep leaning based visible fire detection. In *Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR), 2017 Latin American*, pages 1–6. IEEE.

- Huynh-Thu, Q. & Ghanbari, M. (2008). Scope of validity of psnr in image/video quality assessment. *Electronics letters*, 44(13):800–801.
- Ibrahim, H. & Kong, N. S. P. (2007). Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(4):1752–1758.
- Ignatov, A., Kobyshev, N., Timofte, R., Vanhoey, K., & Van Gool, L. (2017). Dslr-quality photos on mobile devices with deep convolutional networks. In *Proceedings of the IEEE international conference on computer vision*.
- Ignatov, A., Kobyshev, N., Timofte, R., Vanhoey, K., & Van Gool, L. (2018). Wespe: Weakly supervised photo enhancer for digital cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 691–700.
- Ignatov, A. & Timofte, R. (2019). Ntire 2019 challenge on image enhancement: Methods and results. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 0–0.
- Iocchi, L., Holz, D., Ruiz-del Solar, J., Sugiura, K., & Van Der Zant, T. (2015). Robocup@home: Analysis and results of evolving competitions for domestic and service robots. *Artificial Intelligence*, 229:258–281.
- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1125–1134.
- Ito, K., Okano, T., & Aoki, T. (2017). Recent advances in biometric security: A case study of liveness detection in face recognition. In *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 220–227. IEEE.
- Jain, D. & Raman, S. (2021). Deep over and under exposed region detection. In Singh, S. K., Roy, P., Raman, B., & Nagabhushan, P., editors, *Computer Vision and Image Processing*, pages 34–45, Singapore. Springer Singapore.
- Jarque, C. M. & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3):255–259.
- Jia, W., Tian, Y., Luo, R., Zhang, Z., Lian, J., & Zheng, Y. (2020). Detection and segmentation of overlapped fruits based on optimized mask r-cnn application in apple harvesting robot. *Computers and Electronics in Agriculture*, 172:105380.
- Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., & Wang, Z. (2019). Enlightengan: Deep light enhancement without paired supervision. *arXiv preprint arXiv:1906.06972*.

- Jobson, D. J., Rahman, Z.-u., & Woodell, G. A. (1997). A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image processing*, 6(7):965–976.
- Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*, pages 694–711. Springer.
- Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1/2):81–93.
- Kendall, M. G. (1945). The treatment of ties in ranking problems. *Biometrika*, 33(3):239–251.
- Kingma, D. & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kohli, P. & Chadha, A. (2019). Enabling pedestrian safety using computer vision techniques: A case study of the 2018 uber inc. self-driving car crash. In *Future of Information and Communication Conference*, pages 261–279. Springer.
- Kwon, D., Kim, G., & Kwon, J. (2020). Dale: Dark region-aware low-light image enhancement. *arXiv preprint arXiv:2008.12493*.
- Lai, W., Huang, J., Ahuja, N., & Yang, M. (2017). Deep laplacian pyramid networks for fast and accurate super-resolution. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5835–5843.
- Land, E. H. (1977). The retinex theory of color vision. *Scientific american*, 237(6):108–129.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4681–4690.
- Lee, C.-H., Shih, J.-L., Lien, C.-C., & Han, C.-C. (2013). Adaptive multiscale retinex for image contrast enhancement. In *Signal-Image Technology & Internet-Based Systems (SITIS), 2013 International Conference on*, pages 43–50. IEEE.
- Lee, J., Kim, E., Lee, S., Lee, J., & Yoon, S. (2019a). Ficklenet: Weakly and semi-supervised semantic image segmentation using stochastic inference. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5267–5276.
- Lee, T.-j., Kim, C.-h., & Cho, D.-i. D. (2019b). A monocular vision sensor-based efficient slam method for indoor service robots. *IEEE Transactions on Industrial Electronics*, 66(1):318–328.

- Li, C., Guo, C., & Loy, C. C. (2021). Learning to enhance low-light image via zero-reference deep curve estimation. *arXiv preprint arXiv:2103.00860*.
- Li, M., Zhang, T., Chen, Y., & Smola, A. J. (2014). Efficient mini-batch training for stochastic optimization. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 661–670.
- Li, Q., Wu, H., Xu, L., Wang, L., Lv, Y., & Kang, X. (2020a). Low-light image enhancement based on deep symmetric encoder–decoder convolutional networks. *Symmetry*, 12(3):446.
- Li, X., Guo, X., Mei, L., Shang, M., Gao, J., Shu, M., & Wang, X. (2020b). Visual perception model for rapid and adaptive low-light image enhancement. *arXiv preprint arXiv:2005.07343*.
- Liang, J., Xu, Y., Quan, Y., Wang, J., Ling, H., & Ji, H. (2020). Deep bilateral retinex for low-light image enhancement. *arXiv preprint arXiv:2007.02018*.
- Lin, M., Chen, Q., & Yan, S. (2013). Network in network. *arXiv preprint arXiv:1312.4400*.
- Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer.
- Liu, W., Hu, J., & Wang, W. (2020). A novel camera fusion method based on switching scheme and occlusion-aware object detection for real-time robotic grasping. *Journal of Intelligent & Robotic Systems*, pages 1–18.
- Liu, Y.-P., Yang, C.-H., Ling, H., Mabu, S., & Kuremoto, T. (2018). A visual system of citrus picking robot using convolutional neural networks. In *2018 5th international conference on systems and informatics (ICSAI)*, pages 344–349. IEEE.
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. *arXiv preprint arXiv:2103.14030*.
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440.
- Luo, C., Jin, L., & Sun, Z. (2019). Moran: A multi-object rectified attention network for scene text recognition. *Pattern Recognition*, 90:109–118.
- Lv, F., Li, Y., & Lu, F. (2021). Attention guided low-light image enhancement with a large scale low-light simulation dataset. *International Journal of Computer Vision*, 129(7):2175–2193.

- Lv, F., Lu, F., Wu, J., & Lim, C. (2018). Mblen: Low-light image/video enhancement using cnns. In *British Machine Vision Conference (BMVC)*.
- Ma, C., Yang, C.-Y., Yang, X., & Yang, M.-H. (2017). Learning a no-reference quality metric for single-image super-resolution. *Computer Vision and Image Understanding*, 158:1–16.
- Ma, L. Y., Xie, W., & Huang, H. B. (2019). Convolutional neural network based obstacle detection for unmanned surface vehicle. *Mathematical biosciences and engineering: MBE*, 17(1):845–861.
- Malla, A. M., Davidson, P. R., Bones, P. J., Green, R., & Jones, R. D. (2010). Automated video-based measurement of eye closure for detecting behavioral microsleep. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 6741–6744. IEEE.
- Mangal, R., Nori, A. V., & Orso, A. (2019). Robustness of neural networks: a probabilistic and practical approach. In *2019 IEEE/ACM 41st International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)*, pages 93–96. IEEE.
- Mao, X.-J., Shen, C., & Yang, Y.-B. (2016). Image restoration using convolutional auto-encoders with symmetric skip connections. *arXiv preprint arXiv:1606.08921*.
- Mertens, T., Kautz, J., & Van Reeth, F. (2007). Exposure fusion. In *Computer Graphics and Applications, 2007. PG'07. 15th Pacific Conference on*, pages 382–390. IEEE.
- Messias, L. R., Steffens, C. R., Drews-Jr, P. L., & Botelho, S. S. (2020). Ucan: A learning-based model to enhance poorly exposed images. In *Anais Estendidos do XXXIII Conference on Graphics, Patterns and Images*, pages 171–174. SBC.
- Milletari, F., Navab, N., & Ahmadi, S.-A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *3D Vision (3DV), 2016 Fourth International Conference on*, pages 565–571. IEEE.
- Mittal, A., Moorthy, A. K., & Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 21(12):4695–4708.
- Mohapatra, B. R., Mishra, A., & Rout, S. K. (2014). A comprehensive review on image restoration techniques. *International Journal of Research in Advent Technology*, 2(3):101–105.
- Molina, M., Frau, P., & Maravall, D. (2018). A collaborative approach for surface inspection using aerial robots and computer vision. *Sensors*, 18(3):893.
- Moran, S., Marza, P., McDonagh, S., Parisot, S., & Slabaugh, G. (2020). Deeplpf: Deep local parametric filters for image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12826–12835.

- Narmadha, J., Ranjithapriya, S., & Kannaambaal, T. (2017). Survey on image processing under image restoration. In *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*, pages 1–5.
- Newson, R. (2002). Parameters behind “nonparametric” statistics: Kendall’s tau, somers’ d and median differences. *The Stata Journal*, 2(1):45–64.
- Odena, A., Dumoulin, V., & Olah, C. (2016). Deconvolution and checkerboard artifacts. *Distill.*
- Pan, X., Luo, P., Shi, J., & Tang, X. (2018). Two at once: Enhancing learning and generalization capacities via ibn-net. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 464–479.
- Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., & Efros, A. (2016). Context encoders: Feature learning by inpainting. In *Computer Vision and Pattern Recognition (CVPR)*.
- Perrone, M. P., Khan, H., Kim, C., Kyrillidis, A., Quinn, J., & Salapura, V. (2019). Optimal mini-batch size selection for fast gradient descent. *arXiv preprint arXiv:1911.06459*.
- Petro, A. B., Sbert, C., & Morel, J.-M. (2014). Multiscale retinex. *Image Processing On Line*, pages 71–88.
- Piyathilaka, L. & Kodagoda, S. (2015). Human activity recognition for domestic robots. In *Field and Service Robotics*, pages 395–408. Springer.
- Prashnani, E., Cai, H., Mostofi, Y., & Sen, P. (2018). Pieapp: Perceptual image-error assessment through pairwise preference. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Pravitasari, A. A., Iriawan, N., Almuahyar, M., Azmi, T., Fithriasari, K., Purnami, S. W., Ferriastuti, W., et al. (2020). Unet-vgg16 with transfer learning for mri-based brain tumor segmentation. *Telkomnika*, 18(3):1310–1318.
- Protas, É., Bratti, J. D., Gaya, J. F., Drews, P., & Botelho, S. S. (2018). Visualization methods for image transformation convolutional neural networks. *IEEE transactions on neural networks and learning systems*.
- Pyatykh, S., Hesser, J., & Zheng, L. (2013). Image noise level estimation by principal component analysis. *IEEE transactions on image processing*, 22(2):687–699.
- Qi, B., Kun, G., Tian, Y.-x., & Zhu, Z.-y. (2013). A novel false color mapping model-based fusion method of visual and infrared images. In *2013 International Conference on Optical Instruments and Technology: Optoelectronic Imaging and Processing Technology*, volume 9045, page 904519. International Society for Optics and Photonics.

- Qian, K., Jing, X., Duan, Y., Zhou, B., Fang, F., Xia, J., & Ma, X. (2020). Grasp pose detection with affordance-based task constraint learning in single-view point clouds. *JOURNAL OF INTELLIGENT & ROBOTIC SYSTEMS*.
- Qian, X. & Klabjan, D. (2020). The impact of the mini-batch size on the variance of gradients in stochastic gradient descent. *arXiv preprint arXiv:2004.13146*.
- Qiu, K., Ai, Y., Tian, B., Wang, B., & Cao, D. (2018). Siamese-resnet: implementing loop closure detection based on siamese network. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 716–721. IEEE.
- Radford, A., Metz, L., & Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks. *ICLR*.
- Rao, Y., Lin, W. Y., & Chen, L. (2010). Image-based fusion for video enhancement of night-time surveillance. *Optical Engineering*, 49(12):120501.
- Raskar, R., Ilie, A., & Yu, J. (2005). Image fusion for context enhancement and video surrealism. In *ACM SIGGRAPH 2005 Courses*, pages 4–es. Association for Computing Machinery.
- Redmon, J. & Farhadi, A. (2017). Yolo9000: better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7263–7271.
- Ren, W., Liu, S., Ma, L., Xu, Q., Xu, X., Cao, X., Du, J., & Yang, M. (2019). Low-light image enhancement via a deep hybrid network. *IEEE Transactions on Image Processing*, 28(9):4364–4375.
- Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., & Yang, M.-H. (2016). Single image dehazing via multi-scale convolutional neural networks. In *European conference on computer vision*, pages 154–169. Springer.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252.
- Saini, D. K., Ahir, D., & Ganatra, A. (2016). Techniques and challenges in building intelligent systems: anomaly detection in camera surveillance. In *Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems: Volume 2*, pages 11–21. Springer.

- Sánchez-Ramírez, E. E., Rosales-Silva, A. J., & Alfaro-Flores, R. A. (2020). High-precision visual-tracking using the imm algorithm and discrete gpi observers (imm-dgpi). *Journal of Intelligent & Robotic Systems*, 99(3):815–835.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4510–4520.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 618–626.
- Sezan, M. I. & Tekalp, A. M. (1990). Survey of recent developments in digital image restoration. *Optical Engineering*, 29(5):393–405.
- Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611.
- Sharma, G., Wu, W., & Dalal, E. N. (2005). The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur*, 30(1):21–30.
- Sheikh, H. R. & Bovik, A. C. (2004). Image information and visual quality. In *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*, volume 3, pages iii–709. IEEE.
- Shi, W., Alawieh, M. B., Li, X., & Yu, H. (2017). Algorithm and hardware implementation for visual perception system in autonomous vehicle: a survey. *Integration, the VLSI Journal*, 59:148–156.
- Shirmohammadi, S. & Ferrero, A. (2014). Camera as the instrument: the rising trend of vision based measurement. *IEEE Instrumentation & Measurement Magazine*, 17(3):41–47.
- Simonelli, M. & Quaglio, A. (2015). Surveillance camera. US Patent App. 29/507,172.
- Simonyan, K. & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Singh, K. & Parihar, A. S. (2020). A comparative analysis of illumination estimation based image enhancement techniques. In *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, pages 1–5.

- Soares, L. B., Weis, Á. A., Rodrigues, R. N., Drews, P. L., Guterres, B., Botelho, S. S., & Nelson Filho, D. (2017). Seam tracking and welding bead geometry analysis for autonomous welding robot. In *2017 Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR)*, pages 1–6. IEEE.
- Steffens, C., Drews-Jr, P., & Botelho, S. (2018a). Deep learning based exposure correction for image exposure correction with application in computer vision for robotics. In *Latin American Robotic Symposium and Brazilian Symposium on Robotics (LARS/SBR)*, pages 194–200. IEEE.
- Steffens, C. R., Drews Jr, P. L. J., Botelho, s. S. C., & Mattos, V. L. D. (2018b). Análise exploratória de dados de imagens digitais noturnas. *Conferência Sul em Modelagem Computacional*, 1(1):1–16.
- Steffens, C. R., Hüttner, V., & da Costa Botelho, S. S. (2017). Blind iqa for pictures in extreme conditions: Experimental evaluation on metallic surfaces. In *Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR), 2017 Latin American*, pages 1–6. IEEE.
- Steffens, C. R., Messias, L. R., Drews-Jr, P. J., & Botelho, S. S. d. C. (2020a). Cnn based image restoration. *Journal of Intelligent & Robotic Systems*, pages 1–19.
- Steffens, C. R., Messias, L. R. V., Drews, P. L. J., & d. C. Botelho, S. S. (2019). Can exposure, noise and compression affect image recognition? an assessment of the impacts on state-of-the-art convnets. In *2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE)*, pages 61–66.
- Steffens, C. R., Messias, L. R. V., Drews-Jr, P. J. L., & da Costa Botelho, S. S. (2021). On robustness of robotic and autonomous systems perception. *Journal of Intelligent & Robotic Systems*, 101(3):1–17.
- Steffens, C. R., Messias, L. R. V., Drews-Jr, P. L. J., & da Costa Botelho, S. S. (2020b). A pipelined approach to deal with image distortion in computer vision. In Cerri, R. & Prati, R. C., editors, *Intelligent Systems*, pages 212–225, Cham. Springer International Publishing.
- Stephens, M. A. (1974). Edf statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*, 69(347):730–737.
- Stephens, M. A. (1976). Asymptotic results for goodness-of-fit statistics with unknown parameters. *The Annals of Statistics*, pages 357–369.
- Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In *AAAI*, volume 4, page 12.

- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826.
- Teso-Fz-Betoño, D., Zulueta, E., Sánchez-Chica, A., Fernandez-Gamiz, U., & Saenz-Aguirre, A. (2020). Semantic segmentation to develop an indoor navigation system for an autonomous mobile robot. *Mathematics*, 8(5):855.
- Toet, A. (1992). Multiscale contrast enhancement with applications to image fusion. *Optical Engineering*, 31(5):1026–1032.
- Toet, A. (2005). Colorizing single band intensified nightvision images. *Displays*, 26(1):15–21.
- Ulyanov, D., Vedaldi, A., & Lempitsky, V. (2016). Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*.
- Van Den Oord, A., Kalchbrenner, N., & Kavukcuoglu, K. (2016). Pixel recurrent neural networks. In *International Conference on Machine Learning*, pages 1747–1756.
- van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., Yu, T., & the scikit-image contributors (2014). Scikit-image: image processing in Python. *PeerJ*, 2:e453.
- Vaz-Jr, E. S., Drews-Jr, P. J. L., & Steffens, C. R. (2018a). Restoration of images affected by welding fume. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 72–78. IEEE.
- Vaz-Jr, E. S., Drews-Jr, P. L. J., & Steffens, C. R. (2017). Restauração de imagens afetadas por fuligem de solda. *Revista Junior De Iniciação Científica Em Ciências Exatas E Engenharia*, 1(1):1–8.
- Vaz-Jr, E. S. V., Drews-Jr, P. L., Weis, Á. A., Steffens, C. R., & da Costa Botelho, S. S. (2018b). Image processing for automated welding robot: Reducing interference due to fume in camera lenses. In *2018 Latin American Robotic Symposium, 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education (WRE)*, pages 207–212. IEEE.
- Voigtlaender, P., Krause, M., Osep, A., Luiten, J., Sekar, B. B. G., Geiger, A., & Leibe, B. (2019). Mots: Multi-object tracking and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7942–7951.
- Wang, J., Tan, W., Niu, X., & Yan, B. (2019a). Rdgan: Retinex decomposition based adversarial learning for low-light enhancement. In *2019 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1186–1191. IEEE.

- Wang, P., Huang, X., Cheng, X., Zhou, D., Geng, Q., & Yang, R. (2019b). The apolloscape open dataset for autonomous driving and its application. *IEEE transactions on pattern analysis and machine intelligence*.
- Wang, R., Zhang, Q., Fu, C.-W., Shen, X., Zheng, W.-S., & Jia, J. (2019c). Underexposed photo enhancement using deep illumination estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6849–6857.
- Wang, S., Zheng, J., Hu, H.-M., & Li, B. (2013). Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Transactions on Image Processing*, 22(9):3538–3548.
- Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J., & Catanzaro, B. (2018a). High-resolution image synthesis and semantic manipulation with conditional gans. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8798–8807.
- Wang, W., Wei, C., Yang, W., & Liu, J. (2018b). Gladnet: Low-light enhancement network with global awareness. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pages 751–755. IEEE.
- Wang, W., Wu, X., Yuan, X., & Gao, Z. (2020). An experiment-based review of low-light image enhancement methods. *IEEE Access*, 8:87884–87917.
- Wang, W., Wu, X., Yuan, X., & Gao, Z. (2020). An experiment-based review of low-light image enhancement methods. *IEEE Access*, 8:87884–87917.
- Wang, Y.-M., Sun, Z.-L., & Han, F.-Q. (2018c). An effective low-light image enhancement algorithm via fusion model. In *International Conference on Intelligent Computing*, pages 388–396. Springer.
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612.
- Wanner, S. & Goldluecke, B. (2013). Reconstructing reflective and transparent surfaces from epipolar plane images. In *German Conference on Pattern Recognition*, pages 1–10. Springer.
- Weber, F., Rosa, G., Terra, F., Oldoni, A., & Drew-Jr, P. (2018). A low cost system to optimize pesticide application based on mobile technologies and computer vision. In *2018 Latin American Robotic Symposium, 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education (WRE)*, pages 345–350.
- Wei, C., Wang, W., Yang, W., & Liu, J. (2018). Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*.

- Weis, Á. A., Mor, J. L., Soares, L. B., Steffens, C. R., Drews-Jr, P. L., de Faria, M. F., Evald, P. J., Azzolin, R. Z., Nelson Filho, D., & Botelho, S. S. d. C. (2017). Automated seam tracking system based on passive monocular vision for automated linear robotic welding process. In *2017 IEEE 15th International Conference on Industrial Informatics (INDIN)*, pages 305–310. IEEE.
- Wilcoxon, F. (1992). Individual comparisons by ranking methods. In *Breakthroughs in statistics*, pages 196–202. Springer.
- Wu, C.-Y., Hu, X., Happold, M., Xu, Q., & Neumann, U. (2020). Geometry-aware instance segmentation with disparity maps. *arXiv preprint arXiv:2006.07802*.
- Xie, H., Fang, S., Zha, Z.-J., Yang, Y., Li, Y., & Zhang, Y. (2019). Convolutional attention networks for scene text recognition. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 15(1s):1–17.
- Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017). Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1492–1500.
- Xiong, W., Liu, D., Shen, X., Fang, C., & Luo, J. (2020). Unsupervised real-world low-light image enhancement with decoupled networks. *arXiv preprint arXiv:2005.02818*.
- Xu, H., Gao, Y., Yu, F., & Darrell, T. (2017). End-to-end learning of driving models from large-scale video datasets. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2174–2182.
- Xu, K., Yang, X., Yin, B., & Lau, R. W. (2020). Learning to restore low-light images via decomposition-and-enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2281–2290.
- Xue, W., Zhang, L., & Mou, X. (2013). Learning without human scores for blind image quality assessment. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 995–1002.
- Xue, W., Zhang, L., Mou, X., & Bovik, A. C. (2014). Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Transactions on Image Processing*, 23(2):684–695.
- Yang, S., Scherer, S. A., Yi, X., & Zell, A. (2017). Multi-camera visual slam for autonomous navigation of micro aerial vehicles. *Robotics and Autonomous Systems*, 93:116–134.
- Yang, W., Wang, S., Fang, Y., Wang, Y., & Liu, J. (2020). From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3063–3072.

- Yeganeh, H. & Wang, Z. (2012). Objective quality assessment of tone-mapped images. *IEEE Transactions on Image Processing*, 22(2):657–667.
- Yi, Z., Zhang, H. R., Tan, P., & Gong, M. (2017). Dualgan: Unsupervised dual learning for image-to-image translation. In *ICCV*, pages 2868–2876.
- Ying, Z., Li, G., & Gao, W. (2017a). A bio-inspired multi-exposure fusion framework for low-light image enhancement. arxiv 2017. *arXiv preprint arXiv:1711.00591*.
- Ying, Z., Li, G., Ren, Y., Wang, R., & Wang, W. (2017b). A new image contrast enhancement algorithm using exposure fusion framework. In *International Conference on Computer Analysis of Images and Patterns*, pages 36–46. Springer.
- Ying, Z., Li, G., Ren, Y., Wang, R., & Wang, W. (2017c). A new low-light image enhancement algorithm using camera response model. In *Computer Vision Workshop (ICCVW), 2017 IEEE International Conference on*, pages 3015–3022. IEEE.
- Young, K.-y., Cheng, S.-L., Ko, C.-H., & Tsou, H.-W. (2020). Development of a comfort-based motion guidance system for a robot walking helper. *Journal of Intelligent & Robotic Systems*, pages 1–10.
- Yu, F. & Koltun, V. (2015). Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*.
- Zhang, J., Li, C., Kosov, S., Grzegorzec, M., Shirahama, K., Jiang, T., Sun, C., Li, Z., & Li, H. (2021). Lcu-net: A novel low-cost u-net for environmental microorganism image segmentation. *Pattern Recognition*, 115:107885.
- Zhang, L., Zhang, L., & Bovik, A. C. (2015). A feature-enriched completely blind image quality evaluator. *IEEE Transactions on Image Processing*, 24(8):2579–2591.
- Zhang, L., Zhang, L., Mou, X., Zhang, D., et al. (2011). Fsim: a feature similarity index for image quality assessment. *IEEE transactions on Image Processing*, 20(8):2378–2386.
- Zhang, W., Zhou, H., Sun, S., Wang, Z., Shi, J., & Loy, C. C. (2019a). Robust multi-modality multi-object tracking. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2365–2374.
- Zhang, Y., Di, X., Zhang, B., & Wang, C. (2020a). Self-supervised image enhancement network: Training with low light images only. *arXiv*, pages arXiv–2002.
- Zhang, Y., Zhang, J., & Guo, X. (2019b). Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the 27th ACM International Conference on Multimedia*, pages 1632–1640.

- Zhang, Z., Wu, C., Coleman, S., & Kerr, D. (2020b). Dense-inception u-net for medical image segmentation. *Computer methods and programs in biomedicine*, 192:105395.
- Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2881–2890.
- Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2242–2251. IEEE.
- Zhu, X. & Milanfar, P. (2010). Automatic parameter selection for denoising algorithms using a no-reference measure of image content. *IEEE Transactions on Image Processing*, 19(12):3116–3132.
- Zoph, B., Vasudevan, V., Shlens, J., & Le, Q. V. (2018). Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8697–8710.
- Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. *Graphic Gems IV San Diego: Academic Press Professional*, page 474–485.

7 APÊNDICE

7.1 Teste de Correlação entre Indicadores de Qualidade de Imagem

Tabela G.40: Teste de correlação de Pearson (PCC) para medidas de qualidade de imagem no Dataset A6300 incluindo condições de sub e sobre-exposição. Os resultados indicam que a maior parcela das métricas avaliadas apresenta entre si correlação forte ou muito forte. Métricas populares como PSNR, MSE, MAE, SSIM e FSIM só não apresentam correlação forte com as medidas VIFP e RECO

	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Dif. Hist.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
PSNR	1.00	-0.84	-0.93	0.76	0.76	0.85	-0.88	-0.82	-0.32	0.79	0.79	-0.39	-0.92
MSE	-0.84	1.00	0.97	-0.64	-0.66	-0.70	0.74	0.95	0.28	-0.75	-0.75	0.37	0.96
MAE	-0.93	0.97	1.00	-0.71	-0.72	-0.78	0.84	0.92	0.29	-0.78	-0.78	0.38	0.98
SSIM	0.76	-0.64	-0.71	1.00	0.88	0.79	-0.72	-0.67	-0.63	0.83	0.83	-0.67	-0.79
Sobel IoU	0.76	-0.66	-0.72	0.88	1.00	0.83	-0.69	-0.70	-0.54	0.87	0.87	-0.59	-0.75
Canny IoU	0.85	-0.70	-0.78	0.79	0.83	1.00	-0.78	-0.72	-0.45	0.79	0.79	-0.52	-0.81
Dif. Hist.	-0.88	0.74	0.84	-0.72	-0.69	-0.78	1.00	0.71	0.35	-0.68	-0.68	0.43	0.83
GMSD	-0.82	0.95	0.92	-0.67	-0.70	-0.72	0.71	1.00	0.30	-0.81	-0.81	0.39	0.92
VIFP	-0.32	0.28	0.29	-0.63	-0.54	-0.45	0.35	0.30	1.00	-0.40	-0.40	0.90	0.39
FSIM	0.79	-0.75	-0.78	0.83	0.87	0.79	-0.68	-0.81	-0.40	1.00	1.00	-0.44	-0.81
FSIMc	0.79	-0.75	-0.78	0.83	0.87	0.79	-0.68	-0.81	-0.40	1.00	1.00	-0.43	-0.81
RECO	-0.39	0.37	0.38	-0.67	-0.59	-0.52	0.43	0.39	0.90	-0.44	-0.43	1.00	0.49
CIEDE 2000	-0.92	0.96	0.98	-0.79	-0.75	-0.81	0.83	0.92	0.39	-0.81	-0.81	0.49	1.00

7.2 Resultados para Dataset Multi-Exposição baseado em FiveK (sintético)

7.2.1 Testes de Normalidade

Tabela G.41: Resultados para o teste de normalidade para a saída dos modelos utilizando o *dataset* FiveK subexposto. Se o valor-p é menor que o nível de significância $\alpha = 0,05$ então rejeita-se a hipótese nula de que os dados tem distribuição normal. Para o teste de Anderson-Darling apresenta-se o valor da estatística e o valor crítico. Resultados considerando 50 amostras extraídas por amostragem aleatória simples sem reposição

Métodos	Métrica	χ^2	Shapiro-Wilk	Jarque-Bera	Anderson-Darling
M. Proposto	PSNR	0.004*	0.07	0.001*	0.395/0.736*
	MSE	0.000*	0.000*	0.000*	6.390/0.736
	MAE	0.000*	0.000*	0.000*	1.772/0.736
	SSIM	0.000*	0.000*	0.000*	3.983/0.736
	Sobel IoU	0.000*	0.000*	0.000*	1.844/0.736
	Canny IoU	0.000*	0.000*	0.000*	1.091/0.736
	Dif. Hist.	0.000*	0.000*	0.000*	1.525/0.736
	GMSD	0.000*	0.000*	0.000*	3.394/0.736
	VIFP	0.004*	0.003*	0.005*	1.201/0.736
	FSIM	0.001*	0.000*	0.001*	3.282/0.736
	FSIMc	0.001*	0.000*	0.001*	3.274/0.736
	RECO	0.002*	0.003*	0.002*	1.198/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	3.005/0.736
Ronneberger et al. (2015)	PSNR	0.165	0.042*	0.197	0.798/0.736
	MSE	0.000*	0.000*	0.000*	5.830/0.736
	MAE	0.000*	0.000*	0.000*	2.982/0.736
	SSIM	0.000*	0.000*	0.000*	2.007/0.736
	Sobel IoU	0.072	0.009*	0.099	1.328/0.736
	Canny IoU	0.000*	0.000*	0.000*	1.064/0.736
	Dif. Hist.	0.065	0.001*	0.067	1.623/0.736
	GMSD	0.000*	0.000*	0.000*	4.819/0.736
	VIFP	0.219	0.065	0.281	0.711/0.736*
	FSIM	0.001*	0.000*	0.000*	3.716/0.736
	FSIMc	0.001*	0.000*	0.001*	3.567/0.736
	RECO	0.513	0.879	0.733	0.181/0.736*
	CIEDE 2000	0.000*	0.000*	0.000*	2.018/0.736
Chen et al. (2017)	PSNR	0.191	0.058	0.214	0.747/0.736
	MSE	0.000*	0.000*	0.000*	5.175/0.736
	MAE	0.000*	0.000*	0.000*	2.455/0.736
	SSIM	0.000*	0.000*	0.000*	4.929/0.736
	Sobel IoU	0.000*	0.000*	0.000*	2.369/0.736
	Canny IoU	0.000*	0.000*	0.000*	1.417/0.736
	Dif. Hist.	0.164	0.034*	0.232	0.992/0.736
	GMSD	0.000*	0.000*	0.000*	2.176/0.736
	VIFP	0.218	0.112	0.25	0.458/0.736*

	FSIM	0.000*	0.000*	0.000*	3.605/0.736
	FSIMc	0.006*	0.000*	0.004*	2.811/0.736
	RECO	0.000*	0.000*	0.000*	0.857/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	1.894/0.736
Abdullah-Al-Wadud et al. (2007)	PSNR	0.617	0.464	0.6	0.394/0.736*
	MSE	0.000*	0.000*	0.000*	3.381/0.736
	MAE	0.000*	0.000*	0.000*	1.822/0.736
	SSIM	0.002*	0.001*	0.001*	1.478/0.736
	Sobel IoU	0.325	0.034*	0.424	1.111/0.736
	Canny IoU	0.71	0.906	0.816	0.231/0.736*
	Dif. Hist.	0.819	0.197	0.919	0.677/0.736*
	GMSD	0.000*	0.001*	0.000*	0.798/0.736
	VIFP	0.000*	0.000*	0.000*	1.087/0.736
	FSIM	0.141	0.011*	0.162	0.930/0.736
	FSIMc	0.136	0.015*	0.178	0.941/0.736
	RECO	0.1	0.121	0.311	0.717/0.736*
	CIEDE 2000	0.000*	0.001*	0.000*	1.460/0.736
Ying et al. (2017b)	PSNR	0.987	0.995	0.956	0.175/0.736*
	MSE	0.000*	0.000*	0.000*	2.002/0.736
	MAE	0.026*	0.053	0.043*	0.665/0.736*
	SSIM	0.000*	0.000*	0.000*	2.440/0.736
	Sobel IoU	0.128	0.068	0.19	0.874/0.736
	Canny IoU	0.048*	0.023*	0.08	0.871/0.736
	Dif. Hist.	0.786	0.523	0.734	0.258/0.736*
	GMSD	0.886	0.974	0.804	0.184/0.736*
	VIFP	0.33	0.605	0.448	0.348/0.736*
	FSIM	0.179	0.469	0.431	0.288/0.736*
	FSIMc	0.121	0.371	0.394	0.339/0.736*
	RECO	0.000*	0.000*	0.000*	1.816/0.736
	CIEDE 2000	0.004*	0.006*	0.005*	1.014/0.736
Fu et al. (2015)	PSNR	0.020*	0.040*	0.028*	0.999/0.736
	MSE	0.000*	0.000*	0.000*	4.618/0.736
	MAE	0.000*	0.000*	0.000*	1.515/0.736
	SSIM	0.061	0.114	0.093	0.548/0.736*
	Sobel IoU	0.013*	0.018*	0.020*	0.741/0.736
	Canny IoU	0.594	0.871	0.831	0.261/0.736*
	Dif. Hist.	0.004*	0.001*	0.005*	1.136/0.736
	GMSD	0.438	0.534	0.459	0.330/0.736*
	VIFP	0.053	0.151	0.094	0.525/0.736*
	FSIM	0.335	0.38	0.459	0.322/0.736*
	FSIMc	0.271	0.329	0.424	0.390/0.736*
	RECO	0.000*	0.003*	0.000*	0.723/0.736*
	CIEDE 2000	0.000*	0.005*	0.000*	0.879/0.736
Lee et al. (2013)	PSNR	0.003*	0.006*	0.003*	0.824/0.736
	MSE	0.174	0.034*	0.193	0.720/0.736*
	MAE	0.396	0.445	0.54	0.260/0.736*
	SSIM	0.000*	0.000*	0.000*	11.494/0.736
	Sobel IoU	0.000*	0.000*	0.000*	11.454/0.736

	Canny IoU	0.000*	0.000*	0.000*	10.252/0.736
	Dif. Hist.	0.000*	0.000*	0.000*	14.587/0.736
	GMSD	0.33	0.14	0.43	0.391/0.736*
	VIFP	0.001*	0.000*	0.000*	1.723/0.736
	FSIM	0.193	0.015*	0.214	0.782/0.736
	FSIMc	0.148	0.008*	0.161	0.917/0.736
	RECO	0.000*	0.000*	0.000*	11.932/0.736
	CIEDE 2000	0.371	0.346	0.478	0.294/0.736*
Petro et al. (2014)	PSNR	0.82	0.893	0.756	0.249/0.736*
	MSE	0.000*	0.000*	0.000*	3.462/0.736
	MAE	0.020*	0.003*	0.033*	0.841/0.736
	SSIM	0.096	0.086	0.15	0.538/0.736*
	Sobel IoU	0.641	0.272	0.621	0.312/0.736*
	Canny IoU	0.681	0.134	0.65	0.547/0.736*
	Dif. Hist.	0.875	0.913	0.988	0.217/0.736*
	GMSD	0.004*	0.002*	0.005*	0.916/0.736
	VIFP	0.000*	0.000*	0.000*	1.597/0.736
	FSIM	0.004*	0.000*	0.004*	1.414/0.736
	FSIMc	0.006*	0.001*	0.007*	1.027/0.736
	RECO	0.185	0.071	0.358	0.601/0.736*
	CIEDE 2000	0.089	0.104	0.14	0.469/0.736*
Dong et al. (2011)	PSNR	0.89	0.782	0.941	0.253/0.736*
	MSE	0.004*	0.001*	0.005*	1.476/0.736
	MAE	0.118	0.018*	0.141	0.994/0.736
	SSIM	0.021*	0.001*	0.027*	1.782/0.736
	Sobel IoU	0.063	0.008*	0.076	1.124/0.736
	Canny IoU	0.505	0.315	0.697	0.630/0.736*
	Dif. Hist.	0.667	0.744	0.731	0.224/0.736*
	GMSD	0.321	0.487	0.485	0.284/0.736*
	VIFP	0.020*	0.064	0.028*	0.655/0.736*
	FSIM	0.168	0.074	0.247	0.642/0.736*
	FSIMc	0.154	0.09	0.219	0.577/0.736*
	RECO	0.000*	0.001*	0.000*	1.392/0.736
	CIEDE 2000	0.002*	0.001*	0.002*	1.571/0.736
Ying et al. (2017c)	PSNR	0.114	0.234	0.374	0.405/0.736*
	MSE	0.092	0.004*	0.107	1.214/0.736
	MAE	0.274	0.238	0.356	0.477/0.736*
	SSIM	0.000*	0.000*	0.000*	2.107/0.736
	Sobel IoU	0.24	0.402	0.374	0.485/0.736*
	Canny IoU	0.193	0.221	0.29	0.690/0.736*
	Dif. Hist.	0.52	0.622	0.537	0.243/0.736*
	GMSD	0.678	0.185	0.66	0.667/0.736*
	VIFP	0.000*	0.001*	0.000*	1.416/0.736
	FSIM	0.372	0.28	0.44	0.503/0.736*
	FSIMc	0.392	0.323	0.461	0.500/0.736*
	RECO	0.000*	0.000*	0.000*	2.470/0.736
	CIEDE 2000	0.193	0.057	0.259	0.559/0.736*
	PSNR	0.000*	0.000*	0.000*	1.150/0.736

MSE	0.000*	0.000*	0.000*	3.057/0.736
MAE	0.132	0.034*	0.164	0.753/0.736
SSIM	0.15	0.191	0.251	0.553/0.736*
Sobel IoU	0.199	0.108	0.243	0.467/0.736*
Canny IoU	0.068	0.028*	0.116	0.552/0.736*
Dif. Hist.	0.866	0.514	0.928	0.400/0.736*
GMSD	0.000*	0.000*	0.000*	3.356/0.736
VIFP	0.141	0.028*	0.152	0.870/0.736
FSIM	0.004*	0.000*	0.003*	2.624/0.736
FSIMc	0.004*	0.000*	0.003*	2.526/0.736
RECO	0.010*	0.005*	0.015*	1.081/0.736
CIEDE 2000	0.105	0.048*	0.147	0.644/0.736*

Tabela G.42: Resultados para o teste de normalidade para a saída dos modelos utilizando o *dataset* FiveK sobre-exposto. Se o valor-p é menor que o nível de significância $\alpha = 0,05$ então rejeita-se a hipótese nula de que os dados tem distribuição normal. Para o teste de Anderson-Darling apresenta-se o valor da estatística e o valor crítico. Resultados considerando 50 amostras extraídas por amostragem aleatória simples sem reposição

Métodos	Métrica	χ^2	Shapiro-Wilk	Jarque-Bera	Anderson-Darling
M. Proposto	PSNR	0.718	0.318	0.677	0.664/0.736*
	MSE	0.000*	0.000*	0.000*	4.260/0.736
	MAE	0.000*	0.000*	0.000*	2.451/0.736
	SSIM	0.000*	0.000*	0.000*	3.913/0.736
	Sobel IoU	0.000*	0.000*	0.000*	1.442/0.736
	Canny IoU	0.515	0.561	0.525	0.379/0.736*
	Dif. Hist.	0.069	0.002*	0.073	1.223/0.736
	GMSD	0.000*	0.000*	0.000*	2.979/0.736
	VIFP	0.55	0.744	0.551	0.209/0.736*
	FSIM	0.000*	0.000*	0.000*	2.936/0.736
	FSIMc	0.000*	0.000*	0.000*	2.789/0.736
	RECO	0.666	0.819	0.668	0.167/0.736*
	CIEDE 2000	0.001*	0.000*	0.000*	1.676/0.736
Ronneberger et al. (2015)	PSNR	0.095	0.116	0.162	0.382/0.736*
	MSE	0.000*	0.000*	0.000*	7.209/0.736
	MAE	0.000*	0.000*	0.000*	2.514/0.736
	SSIM	0.000*	0.000*	0.000*	3.748/0.736
	Sobel IoU	0.000*	0.000*	0.000*	2.358/0.736
	Canny IoU	0.126	0.010*	0.155	0.947/0.736
	Dif. Hist.	0.148	0.029*	0.189	0.867/0.736
	GMSD	0.000*	0.000*	0.000*	3.991/0.736
	VIFP	0.22	0.273	0.299	0.354/0.736*
	FSIM	0.000*	0.000*	0.000*	3.565/0.736
	FSIMc	0.000*	0.000*	0.000*	3.354/0.736
	RECO	0.907	0.872	0.974	0.247/0.736*
	CIEDE 2000	0.000*	0.000*	0.000*	1.707/0.736
	PSNR	0.264	0.545	0.479	0.295/0.736*

	MSE	0.001*	0.000*	0.000*	1.518/0.736
	MAE	0.036*	0.028*	0.058	0.649/0.736*
	SSIM	0.000*	0.000*	0.000*	4.337/0.736
	Sobel IoU	0.000*	0.000*	0.000*	1.942/0.736
	Canny IoU	0.249	0.158	0.284	0.647/0.736*
	Dif. Hist.	0.092	0.011*	0.101	1.270/0.736
	GMSD	0.000*	0.000*	0.000*	3.262/0.736
	VIFP	0.362	0.435	0.414	0.386/0.736*
	FSIM	0.000*	0.000*	0.000*	4.553/0.736
	FSIMc	0.000*	0.000*	0.000*	3.968/0.736
	RECO	0.371	0.557	0.532	0.362/0.736*
	CIEDE 2000	0.226	0.177	0.284	0.438/0.736*
Abdullah-Al-Wadud et al. (2007)	PSNR	0.176	0.207	0.237	0.488/0.736*
	MSE	0.000*	0.000*	0.000*	1.222/0.736
	MAE	0.213	0.318	0.335	0.282/0.736*
	SSIM	0.000*	0.000*	0.000*	1.576/0.736
	Sobel IoU	0.42	0.49	0.635	0.446/0.736*
	Canny IoU	0.102	0.016*	0.124	0.881/0.736
	Dif. Hist.	0.754	0.216	0.705	0.547/0.736*
	GMSD	0.26	0.103	0.286	0.476/0.736*
	VIFP	0.329	0.144	0.347	0.697/0.736*
	FSIM	0.000*	0.001*	0.000*	0.677/0.736*
	FSIMc	0.000*	0.001*	0.000*	0.689/0.736*
	RECO	0.000*	0.000*	0.000*	1.776/0.736
	CIEDE 2000	0.119	0.204	0.187	0.397/0.736*
Ying et al. (2017b)	PSNR	0.006*	0.026*	0.008*	0.759/0.736
	MSE	0.000*	0.001*	0.000*	0.709/0.736*
	MAE	0.385	0.716	0.591	0.219/0.736*
	SSIM	0.016*	0.015*	0.024*	0.811/0.736
	Sobel IoU	0.279	0.448	0.405	0.397/0.736*
	Canny IoU	0.144	0.37	0.232	0.275/0.736*
	Dif. Hist.	0.027*	0.031*	0.046*	0.475/0.736*
	GMSD	0.003*	0.002*	0.003*	0.981/0.736
	VIFP	0.427	0.476	0.458	0.406/0.736*
	FSIM	0.000*	0.000*	0.000*	1.017/0.736
	FSIMc	0.000*	0.000*	0.000*	0.989/0.736
	RECO	0.091	0.309	0.159	0.414/0.736*
	CIEDE 2000	0.418	0.652	0.634	0.227/0.736*
Fu et al. (2015)	PSNR	0.002*	0.007*	0.001*	0.968/0.736
	MSE	0.000*	0.004*	0.000*	0.701/0.736*
	MAE	0.663	0.96	0.844	0.185/0.736*
	SSIM	0.001*	0.002*	0.000*	1.023/0.736
	Sobel IoU	0.039*	0.075	0.068	0.695/0.736*
	Canny IoU	0.048*	0.086	0.084	0.420/0.736*
	Dif. Hist.	0.089	0.158	0.325	0.440/0.736*
	GMSD	0.029*	0.003*	0.041*	1.018/0.736
	VIFP	0.617	0.695	0.599	0.287/0.736*
	FSIM	0.000*	0.000*	0.000*	1.011/0.736

	FSIMc	0.000*	0.000*	0.000*	0.984/0.736
	RECO	0.657	0.946	0.744	0.198/0.736*
	CIEDE 2000	0.79	0.955	0.967	0.162/0.736*
Lee et al. (2013)	PSNR	0.52	0.397	0.581	0.512/0.736*
	MSE	0.000*	0.000*	0.000*	3.951/0.736
	MAE	0.001*	0.000*	0.000*	2.558/0.736
	SSIM	0.000*	0.000*	0.000*	3.173/0.736
	Sobel IoU	0.418	0.145	0.57	0.346/0.736*
	Canny IoU	0.201	0.050*	0.27	0.727/0.736*
	Dif. Hist.	0.000*	0.004*	0.148	1.220/0.736
	GMSD	0.107	0.009*	0.113	0.980/0.736
	VIFP	0.182	0.278	0.408	0.496/0.736*
	FSIM	0.000*	0.005*	0.000*	0.657/0.736*
	FSIMc	0.000*	0.004*	0.000*	0.712/0.736*
	RECO	0.000*	0.005*	0.000*	1.004/0.736
	CIEDE 2000	0.059	0.005*	0.066	1.298/0.736
Petro et al. (2014)	PSNR	0.000*	0.000*	0.000*	1.374/0.736
	MSE	0.000*	0.000*	0.000*	1.327/0.736
	MAE	0.099	0.284	0.175	0.422/0.736*
	SSIM	0.000*	0.000*	0.000*	1.340/0.736
	Sobel IoU	0.515	0.475	0.751	0.519/0.736*
	Canny IoU	0.002*	0.001*	0.002*	1.500/0.736
	Dif. Hist.	0.868	0.601	0.79	0.266/0.736*
	GMSD	0.168	0.019*	0.184	0.815/0.736
	VIFP	0.317	0.17	0.35	0.732/0.736*
	FSIM	0.000*	0.002*	0.000*	0.631/0.736*
	FSIMc	0.000*	0.003*	0.000*	0.628/0.736*
	RECO	0.028*	0.043*	0.047*	0.690/0.736*
	CIEDE 2000	0.079	0.247	0.138	0.385/0.736*
Dong et al. (2011)	PSNR	0.523	0.482	0.517	0.367/0.736*
	MSE	0.012*	0.005*	0.017*	0.950/0.736
	MAE	0.554	0.725	0.547	0.196/0.736*
	SSIM	0.009*	0.007*	0.012*	0.933/0.736
	Sobel IoU	0.26	0.274	0.338	0.398/0.736*
	Canny IoU	0.182	0.29	0.306	0.276/0.736*
	Dif. Hist.	0.09	0.098	0.157	0.350/0.736*
	GMSD	0.047*	0.003*	0.059	1.290/0.736
	VIFP	0.84	0.881	0.77	0.273/0.736*
	FSIM	0.000*	0.001*	0.000*	0.978/0.736
	FSIMc	0.000*	0.001*	0.000*	1.005/0.736
	RECO	0.065	0.296	0.115	0.399/0.736*
	CIEDE 2000	0.716	0.727	0.677	0.229/0.736*
Ying et al. (2017c)	PSNR	0.326	0.384	0.358	0.429/0.736*
	MSE	0.003*	0.013*	0.003*	0.598/0.736*
	MAE	0.904	0.95	0.857	0.131/0.736*
	SSIM	0.037*	0.029*	0.054	0.689/0.736*
	Sobel IoU	0.332	0.419	0.428	0.394/0.736*
	Canny IoU	0.263	0.501	0.386	0.234/0.736*

	Dif. Hist.	0.016 [*]	0.016 [*]	0.026 [*]	0.562/0.736 [*]
	GMSD	0.011 [*]	0.002 [*]	0.016 [*]	1.111/0.736
	VIFP	0.398	0.349	0.43	0.433/0.736 [*]
	FSIM	0.000 [*]	0.000 [*]	0.000 [*]	1.032/0.736
	FSIMc	0.000 [*]	0.001 [*]	0.000 [*]	0.994/0.736
	RECO	0.003 [*]	0.012 [*]	0.002 [*]	0.808/0.736
	CIEDE 2000	0.945	0.806	0.921	0.274/0.736 [*]
Não Tratada	PSNR	0.000 [*]	0.000 [*]	0.000 [*]	3.747/0.736
	MSE	0.000 [*]	0.000 [*]	0.000 [*]	1.082/0.736
	MAE	0.501	0.72	0.681	0.271/0.736 [*]
	SSIM	0.000 [*]	0.000 [*]	0.000 [*]	1.209/0.736
	Sobel IoU	0.362	0.686	0.543	0.305/0.736 [*]
	Canny IoU	0.013 [*]	0.014 [*]	0.020 [*]	0.891/0.736
	Dif. Hist.	0.973	0.693	0.966	0.252/0.736 [*]
	GMSD	0.010 [*]	0.001 [*]	0.013 [*]	1.198/0.736
	VIFP	0.615	0.74	0.601	0.357/0.736 [*]
	FSIM	0.000 [*]	0.000 [*]	0.000 [*]	1.201/0.736
	FSIMc	0.000 [*]	0.000 [*]	0.000 [*]	1.221/0.736
	RECO	0.895	0.884	0.836	0.256/0.736 [*]
	CIEDE 2000	0.726	0.924	0.859	0.149/0.736 [*]

7.2.2 Média

Tabela G.43: Valores de media para restauração de imagens subexpostas do *dataset* FiveK

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
M. Proposto	24,675	0,004	0,053	0,938	0,851	0,711	5,585	0,711	0,783	0,980	0,978	0,939	6,599
Ronneberger et al. (2015)	24,985	0,006	0,055	0,892	0,752	0,641	5,954	0,927	0,957	0,972	0,970	0,922	6,266
Chen et al. (2017)	23,448	0,007	0,064	0,907	0,794	0,610	6,623	1,260	0,893	0,967	0,959	1,009	9,849
Abdullah-Al-Wadud et al. (2007)	17,417	0,024	0,120	0,761	0,598	0,346	6,604	4,244	0,429	0,894	0,887	0,608	13,079
Ying et al. (2017b)	19,537	0,013	0,090	0,800	0,648	0,439	5,795	3,616	0,479	0,922	0,917	0,523	9,833
Fu et al. (2015)	19,537	0,013	0,088	0,797	0,667	0,403	6,230	3,202	0,452	0,925	0,920	0,547	9,474
Lee et al. (2013)	8,305	0,190	0,353	0,127	0,097	0,045	11,128	24,822	0,838	0,585	0,569	-162,893	33,257
Petro et al. (2014)	19,281	0,019	0,099	0,772	0,729	0,578	5,699	3,215	0,681	0,938	0,926	0,569	11,329
Dong et al. (2011)	16,358	0,026	0,123	0,711	0,514	0,300	6,453	6,388	0,341	0,846	0,839	0,389	13,591
Ying et al. (2017c)	15,628	0,032	0,141	0,727	0,576	0,327	6,552	9,590	0,347	0,863	0,855	0,419	15,195
Não Tratada	21,456	0,016	0,102	0,772	0,836	0,720	6,124	1,488	0,765	0,968	0,966	0,798	8,808

Tabela G.44: Valores de media para restauração de imagens sobre-expostas do *dataset* FiveK

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
M. Proposto	20,164	0,015	0,082	0,891	0,724	0,487	5,757	3,624	0,762	0,926	0,919	0,758	9,445
Ronneberger et al. (2015)	18,977	0,021	0,097	0,808	0,600	0,424	6,162	5,547	0,940	0,895	0,888	1,139	10,848
Chen et al. (2017)	20,374	0,011	0,079	0,898	0,690	0,455	5,823	3,156	1,067	0,936	0,928	1,254	9,859
Abdullah-Al-Wadud et al. (2007)	13,324	0,062	0,190	0,720	0,525	0,264	6,487	12,263	0,401	0,807	0,800	0,423	18,275
Ying et al. (2017b)	10,791	0,104	0,284	0,669	0,580	0,337	8,187	9,746	0,483	0,855	0,846	0,677	26,891
Fu et al. (2015)	11,755	0,087	0,249	0,720	0,599	0,313	7,617	9,188	0,446	0,858	0,849	0,636	24,260
Lee et al. (2013)	15,047	0,041	0,143	0,806	0,563	0,246	7,455	9,608	1,552	0,854	0,848	2,114	14,594
Petro et al. (2014)	14,174	0,058	0,181	0,753	0,576	0,360	6,421	9,751	0,438	0,844	0,834	0,443	17,961
Dong et al. (2011)	16,773	0,096	0,266	0,658	0,504	0,278	7,903	10,681	0,390	0,805	0,795	0,461	26,187
Ying et al. (2017c)	10,070	0,116	0,303	0,649	0,560	0,313	8,315	10,567	0,461	0,844	0,834	0,692	28,888
Não Tratada	14,187	0,066	0,203	0,773	0,627	0,409	6,734	7,877	0,489	0,874	0,866	0,583	19,754

7.2.3 Teste dos Postos Sinalizados de Wilcoxon

Tabela G.45: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do *dataset* FiveK (estatística do teste)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	584	507	585	144	0	1	575	390	51	78	91	500	502
Chen et al. (2017)	400	342	407	123	46	29	212	12	163	8	6	389	40
Abdullah-Al-Wadud et al. (2007)	3	3	10	3	0	1	271	3	12	0	0	16	5
Ying et al. (2017b)	16	7	20	0	0	0	579	1	0	1	1	2	56
Fu et al. (2015)	6	6	22	13	1	0	295	3	0	0	0	0	24
Lee et al. (2013)	0	0	0	0	0	1	21	0	636	0	0	394	0
Petro et al. (2014)	58	40	101	18	44	144	560	61	243	44	27	0	103
Dong et al. (2011)	0	0	0	0	0	0	358	0	0	0	0	0	0
Ying et al. (2017c)	0	0	1	0	0	0	362	0	0	0	0	3	4
Não Tratada	286	179	165	59	443	587	428	290	436	221	272	16	338

Tabela G.46: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do *dataset* FiveK (estatística do teste)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	484	558	503	42	17	155	564	371	239	247	255	11	478
Chen et al. (2017)	583	484	582	609	338	391	540,5	516	34	377	389	0	542
Abdullah-Al-Wadud et al. (2007)	7	6	7	0	0	0	341,5	1	0	0	0	36	42
Ying et al. (2017b)	2	2	2	3	13	17	57	3	43	35	41	293	8
Fu et al. (2015)	7	6	5	12	19	5	95	11	11	31	45	185	19
Lee et al. (2013)	161	200	200	88	52	0	298	57	38	30	34	9	175
Petro et al. (2014)	72	14	14	10	54	103	347	16	0	21	26	0	60
Dong et al. (2011)	64	0	0	0	0	3	66	0	0	0	0	25	3
Ying et al. (2017c)	0	0	0	0	6	5	48	0	26	16	18	374	0
Não Tratada	105	15	19	54	109	208	280	50	19	116	133	84	53

7.3 Resultados para HDR+ *Burst Photography Dataset* (sintético)

7.3.1 Testes de Normalidade

Tabela G.47: Resultados para o teste de normalidade para a saída dos modelos utilizando o *dataset* HDR+*burst* subexposto. Se o valor-p é menor que o nível de significância $\alpha = 0,05$ então rejeita-se a hipótese nula de que os dados tem distribuição normal. Para o teste de Anderson-Darling apresenta-se o valor da estatística e o valor crítico. Resultados considerando 50 amostras extraídas por amostragem aleatória simples sem reposição

Métodos	Métrica	χ^2	Shapiro-Wilk	Jarque-Bera	Anderson-Darling
M. Proposto	PSNR	0.000*	0.000*	0.000*	1.757/0.736
	MSE	0.000*	0.000*	0.000*	13.297/0.736
	MAE	0.000*	0.000*	0.000*	6.740/0.736
	SSIM	0.000*	0.000*	0.000*	5.361/0.736
	Sobel IoU	0.000*	0.000*	0.000*	3.436/0.736
	Canny IoU	0.000*	0.000*	0.000*	2.122/0.736
	Dif. Hist.	0.019*	0.001*	0.024*	1.391/0.736
	GMSD	0.000*	0.000*	0.000*	10.691/0.736
	VIFP	0.000*	0.000*	0.000*	2.025/0.736
	FSIM	0.000*	0.000*	0.000*	7.260/0.736
	FSIMc	0.000*	0.000*	0.000*	7.153/0.736
	RECO	0.000*	0.000*	0.000*	1.369/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	6.646/0.736
Ronneberger et al. (2015)	PSNR	0.000*	0.002*	0.000*	0.615/0.736*
	MSE	0.000*	0.000*	0.000*	9.060/0.736
	MAE	0.000*	0.000*	0.000*	3.931/0.736
	SSIM	0.001*	0.006*	0.000*	0.771/0.736
	Sobel IoU	0.000*	0.003*	0.000*	0.967/0.736
	Canny IoU	0.000*	0.000*	0.000*	1.400/0.736
	Dif. Hist.	0.005*	0.003*	0.006*	1.202/0.736
	GMSD	0.000*	0.000*	0.000*	7.637/0.736
	VIFP	0.000*	0.000*	0.000*	0.998/0.736
	FSIM	0.000*	0.000*	0.000*	4.466/0.736
	FSIMc	0.000*	0.000*	0.000*	4.047/0.736
	RECO	0.000*	0.000*	0.000*	0.997/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	3.572/0.736
Chen et al. (2017)	PSNR	0.000*	0.008*	0.000*	0.490/0.736*
	MSE	0.000*	0.000*	0.000*	5.968/0.736
	MAE	0.000*	0.000*	0.000*	1.807/0.736
	SSIM	0.000*	0.000*	0.000*	3.930/0.736
	Sobel IoU	0.000*	0.000*	0.000*	1.939/0.736
	Canny IoU	0.079	0.12	0.141	0.751/0.736
	Dif. Hist.	0.003*	0.006*	0.002*	0.817/0.736
	GMSD	0.000*	0.000*	0.000*	6.935/0.736
	VIFP	0.286	0.606	0.426	0.235/0.736*
	FSIM	0.000*	0.000*	0.000*	3.939/0.736

	FSIMc	0.000*	0.000*	0.000*	3.340/0.736
	RECO	0.000*	0.000*	0.000*	2.399/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	1.762/0.736
Abdullah-Al-Wadud et al. (2007)	PSNR	0.996	0.779	0.928	0.277/0.736*
	MSE	0.000*	0.000*	0.000*	4.257/0.736
	MAE	0.003*	0.000*	0.003*	1.969/0.736
	SSIM	0.008*	0.000*	0.007*	2.228/0.736
	Sobel IoU	0.472	0.45	0.668	0.343/0.736*
	Canny IoU	0.756	0.485	0.711	0.271/0.736*
	Dif. Hist.	0.085	0.001*	0.098	1.262/0.736
	GMSD	0.000*	0.000*	0.000*	2.962/0.736
	VIFP	0.939	0.985	0.847	0.148/0.736*
	FSIM	0.002*	0.000*	0.002*	2.055/0.736
	FSIMc	0.004*	0.000*	0.004*	2.068/0.736
	RECO	0.591	0.616	0.628	0.274/0.736*
	CIEDE 2000	0.040*	0.002*	0.046*	1.460/0.736
Ying et al. (2017b)	PSNR	0.000*	0.002*	0.000*	0.773/0.736
	MSE	0.000*	0.000*	0.000*	7.372/0.736
	MAE	0.000*	0.000*	0.000*	2.699/0.736
	SSIM	0.000*	0.000*	0.000*	1.735/0.736
	Sobel IoU	0.000*	0.010*	0.000*	0.574/0.736*
	Canny IoU	0.445	0.32	0.457	0.427/0.736*
	Dif. Hist.	0.001*	0.004*	0.001*	1.022/0.736
	GMSD	0.000*	0.000*	0.000*	3.870/0.736
	VIFP	0.09	0.083	0.13	0.517/0.736*
	FSIM	0.000*	0.000*	0.000*	2.285/0.736
	FSIMc	0.000*	0.000*	0.000*	2.260/0.736
	RECO	0.393	0.234	0.402	0.397/0.736*
	CIEDE 2000	0.000*	0.000*	0.000*	2.205/0.736
Fu et al. (2015)	PSNR	0.000*	0.000*	0.000*	1.747/0.736
	MSE	0.000*	0.000*	0.000*	7.926/0.736
	MAE	0.000*	0.000*	0.000*	3.978/0.736
	SSIM	0.000*	0.000*	0.000*	1.332/0.736
	Sobel IoU	0.001*	0.014*	0.000*	0.518/0.736*
	Canny IoU	0.407	0.32	0.62	0.559/0.736*
	Dif. Hist.	0.018*	0.001*	0.020*	1.572/0.736
	GMSD	0.000*	0.000*	0.000*	6.336/0.736
	VIFP	0.366	0.081	0.541	0.521/0.736*
	FSIM	0.000*	0.000*	0.000*	2.620/0.736
	FSIMc	0.000*	0.000*	0.000*	2.571/0.736
	RECO	0.377	0.345	0.392	0.328/0.736*
	CIEDE 2000	0.000*	0.000*	0.000*	2.395/0.736
Lee et al. (2013)	PSNR	0.000*	0.000*	0.000*	1.593/0.736
	MSE	0.002*	0.002*	0.002*	1.712/0.736
	MAE	0.074	0.056	0.105	0.989/0.736
	SSIM	0.000*	0.000*	0.000*	13.085/0.736
	Sobel IoU	0.000*	0.000*	0.000*	10.715/0.736
	Canny IoU	0.000*	0.000*	0.000*	15.923/0.736

	Dif. Hist.	0.000*	0.000*	0.000*	15.330/0.736
	GMSD	0.009*	0.055	0.003*	0.520/0.736*
	VIFP	0.218	0.417	0.429	0.340/0.736*
	FSIM	0.371	0.183	0.483	0.333/0.736*
	FSIMc	0.238	0.106	0.354	0.384/0.736*
	RECO	0.000*	0.000*	0.000*	13.255/0.736
	CIEDE 2000	0.024*	0.045*	0.038*	0.995/0.736
Petro et al. (2014)	PSNR	0.945	0.614	0.942	0.405/0.736*
	MSE	0.000*	0.000*	0.000*	6.091/0.736
	MAE	0.000*	0.000*	0.000*	2.113/0.736
	SSIM	0.000*	0.000*	0.000*	0.984/0.736
	Sobel IoU	0.001*	0.011*	0.000*	0.455/0.736*
	Canny IoU	0.445	0.269	0.545	0.418/0.736*
	Dif. Hist.	0.017*	0.002*	0.145	1.491/0.736
	GMSD	0.000*	0.000*	0.000*	2.983/0.736
	VIFP	0.721	0.582	0.714	0.311/0.736*
	FSIM	0.000*	0.000*	0.000*	1.920/0.736
	FSIMc	0.000*	0.000*	0.000*	1.222/0.736
	RECO	0.962	0.79	0.871	0.239/0.736*
	CIEDE 2000	0.000*	0.001*	0.000*	0.902/0.736
Dong et al. (2011)	PSNR	0.1	0.041*	0.152	0.673/0.736*
	MSE	0.000*	0.000*	0.000*	3.361/0.736
	MAE	0.000*	0.000*	0.000*	1.464/0.736
	SSIM	0.001*	0.005*	0.000*	0.607/0.736*
	Sobel IoU	0.215	0.643	0.359	0.286/0.736*
	Canny IoU	0.362	0.27	0.54	0.554/0.736*
	Dif. Hist.	0.048*	0.001*	0.049*	1.418/0.736
	GMSD	0.000*	0.000*	0.000*	1.495/0.736
	VIFP	0.724	0.804	0.702	0.173/0.736*
	FSIM	0.139	0.036*	0.157	0.824/0.736
	FSIMc	0.15	0.043*	0.169	0.784/0.736
	RECO	0.135	0.164	0.376	0.678/0.736*
	CIEDE 2000	0.001*	0.001*	0.000*	1.498/0.736
Ying et al. (2017c)	PSNR	0.535	0.454	0.532	0.318/0.736*
	MSE	0.097	0.040*	0.127	0.675/0.736*
	MAE	0.735	0.789	0.69	0.177/0.736*
	SSIM	0.003*	0.001*	0.002*	1.343/0.736
	Sobel IoU	0.028*	0.019*	0.032*	1.005/0.736
	Canny IoU	0.886	0.803	0.803	0.214/0.736*
	Dif. Hist.	0.318	0.22	0.359	0.476/0.736*
	GMSD	0.085	0.232	0.15	0.508/0.736*
	VIFP	0.888	0.796	0.805	0.224/0.736*
	FSIM	0.023*	0.006*	0.033*	1.032/0.736
	FSIMc	0.025*	0.006*	0.035*	1.065/0.736
	RECO	0.07	0.067	0.285	0.563/0.736*
	CIEDE 2000	0.521	0.32	0.525	0.357/0.736*
	PSNR	0.001*	0.025*	0.000*	0.477/0.736*
	MSE	0.000*	0.000*	0.000*	7.919/0.736

MAE	0.000*	0.000*	0.000*	2.058/0.736
SSIM	0.000*	0.004*	0.000*	0.560/0.736*
Sobel IoU	0.000*	0.000*	0.000*	1.443/0.736
Canny IoU	0.109	0.017*	0.124	0.958/0.736
Dif. Hist.	0.201	0.054	0.243	0.709/0.736*
GMSD	0.000*	0.000*	0.000*	7.611/0.736
VIFP	0.000*	0.000*	0.000*	2.374/0.736
FSIM	0.000*	0.000*	0.000*	3.617/0.736
FSIMc	0.000*	0.000*	0.000*	3.624/0.736
RECO	0.000*	0.000*	0.000*	1.419/0.736
CIEDE 2000	0.000*	0.000*	0.000*	2.055/0.736

Tabela G.48: Resultados para o teste de normalidade para a saída dos modelos utilizando o *dataset* HDR+burst sobre-exposto. Se o valor-p é menor que o nível de significância $\alpha = 0,05$ então rejeita-se a hipótese nula de que os dados tem distribuição normal. Para o teste de Anderson-Darling apresenta-se o valor da estatística e o valor crítico. Resultados considerando 50 amostras extraídas por amostragem aleatória simples sem reposição

Métodos	Métrica	χ^2	Shapiro-Wilk	Jarque-Bera	Anderson-Darling
M. Proposto	PSNR	0.027*	0.037*	0.045*	0.718/0.736*
	MSE	0.000*	0.000*	0.000*	6.734/0.736
	MAE	0.000*	0.000*	0.000*	4.058/0.736
	SSIM	0.035*	0.013*	0.048*	0.811/0.736
	Sobel IoU	0.000*	0.000*	0.000*	1.525/0.736
	Canny IoU	0.000*	0.000*	0.000*	1.978/0.736
	Dif. Hist.	0.262	0.068	0.277	0.587/0.736*
	GMSD	0.000*	0.000*	0.000*	2.315/0.736
	VIFP	0.174	0.045*	0.204	0.888/0.736
	FSIM	0.000*	0.000*	0.000*	1.790/0.736
	FSIMc	0.000*	0.000*	0.000*	1.609/0.736
	RECO	0.857	0.94	0.782	0.145/0.736*
	CIEDE 2000	0.000*	0.000*	0.000*	3.686/0.736
Ronneberger et al. (2015)	PSNR	0.316	0.821	0.48	0.263/0.736*
	MSE	0.000*	0.000*	0.000*	4.599/0.736
	MAE	0.000*	0.000*	0.000*	2.611/0.736
	SSIM	0.035*	0.08	0.06	0.567/0.736*
	Sobel IoU	0.000*	0.000*	0.000*	2.243/0.736
	Canny IoU	0.003*	0.000*	0.000*	1.753/0.736
	Dif. Hist.	0.084	0.004*	0.131	1.168/0.736
	GMSD	0.001*	0.000*	0.001*	1.915/0.736
	VIFP	0.294	0.614	0.427	0.320/0.736*
	FSIM	0.000*	0.000*	0.000*	2.445/0.736
	FSIMc	0.000*	0.000*	0.000*	2.570/0.736
	RECO	0.93	0.945	0.91	0.129/0.736*
	CIEDE 2000	0.000*	0.000*	0.000*	1.950/0.736
	PSNR	0.092	0.17	0.152	0.321/0.736*
	MSE	0.000*	0.000*	0.000*	3.456/0.736

	MAE	0.000*	0.000*	0.000*	1.672/0.736
	SSIM	0.000*	0.000*	0.000*	5.440/0.736
	Sobel IoU	0.000*	0.000*	0.000*	3.997/0.736
	Canny IoU	0.000*	0.000*	0.000*	2.390/0.736
	Dif. Hist.	0.034*	0.008*	0.159	1.042/0.736
	GMSD	0.000*	0.000*	0.000*	3.094/0.736
	VIFP	0.000*	0.000*	0.000*	2.077/0.736
	FSIM	0.000*	0.000*	0.000*	2.151/0.736
	FSIMc	0.000*	0.000*	0.000*	1.513/0.736
	RECO	0.792	0.884	0.927	0.260/0.736*
	CIEDE 2000	0.000*	0.001*	0.000*	0.884/0.736
Abdullah-Al-Wadud et al. (2007)	PSNR	0.003*	0.022*	0.003*	0.816/0.736
	MSE	0.010*	0.051	0.014*	0.570/0.736*
	MAE	0.885	0.619	0.962	0.349/0.736*
	SSIM	0.000*	0.001*	0.000*	0.954/0.736
	Sobel IoU	0.001*	0.006*	0.000*	0.790/0.736
	Canny IoU	0.926	0.761	0.954	0.235/0.736*
	Dif. Hist.	0.030*	0.000*	0.108	1.639/0.736
	GMSD	0.000*	0.000*	0.000*	1.309/0.736
	VIFP	0.483	0.753	0.531	0.234/0.736*
	FSIM	0.002*	0.003*	0.001*	1.037/0.736
	FSIMc	0.001*	0.004*	0.001*	0.974/0.736
	RECO	0.198	0.071	0.421	0.612/0.736*
	CIEDE 2000	0.625	0.926	0.852	0.148/0.736*
Ying et al. (2017b)	PSNR	0.000*	0.000*	0.000*	2.920/0.736
	MSE	0.093	0.492	0.15	0.218/0.736*
	MAE	0.010*	0.153	0.009*	0.350/0.736*
	SSIM	0.000*	0.001*	0.000*	1.061/0.736
	Sobel IoU	0.002*	0.012*	0.001*	0.870/0.736
	Canny IoU	0.074	0.112	0.093	0.514/0.736*
	Dif. Hist.	0.001*	0.002*	0.157	1.247/0.736
	GMSD	0.000*	0.000*	0.000*	1.563/0.736
	VIFP	0.547	0.297	0.727	0.638/0.736*
	FSIM	0.000*	0.000*	0.000*	1.679/0.736
	FSIMc	0.000*	0.000*	0.000*	1.648/0.736
	RECO	0.606	0.439	0.708	0.402/0.736*
	CIEDE 2000	0.016*	0.145	0.006*	0.478/0.736*
Fu et al. (2015)	PSNR	0.000*	0.000*	0.000*	2.059/0.736
	MSE	0.104	0.349	0.183	0.304/0.736*
	MAE	0.324	0.557	0.477	0.291/0.736*
	SSIM	0.000*	0.000*	0.000*	1.179/0.736
	Sobel IoU	0.006*	0.016*	0.007*	0.824/0.736
	Canny IoU	0.943	0.976	0.99	0.241/0.736*
	Dif. Hist.	0.039*	0.004*	0.208	1.066/0.736
	GMSD	0.000*	0.000*	0.000*	1.318/0.736
	VIFP	0.841	0.515	0.91	0.487/0.736*
	FSIM	0.000*	0.000*	0.000*	1.272/0.736
	FSIMc	0.000*	0.000*	0.000*	1.321/0.736

	RECO	0.831	0.657	0.883	0.302/0.736*
	CIEDE 2000	0.16	0.499	0.234	0.374/0.736*
Lee et al. (2013)	PSNR	0.035*	0.218	0.052	0.338/0.736*
	MSE	0.000*	0.000*	0.000*	4.028/0.736
	MAE	0.000*	0.000*	0.000*	1.381/0.736
	SSIM	0.001*	0.005*	0.000*	0.769/0.736
	Sobel IoU	0.877	0.595	0.935	0.308/0.736*
	Canny IoU	0.049*	0.040*	0.082	0.803/0.736
	Dif. Hist.	0.000*	0.001*	0.129	1.306/0.736
	GMSD	0.003*	0.009*	0.002*	0.983/0.736
	VIFP	0.000*	0.002*	0.000*	0.883/0.736
	FSIM	0.54	0.712	0.649	0.219/0.736*
	FSIMc	0.585	0.756	0.68	0.191/0.736*
	RECO	0.000*	0.000*	0.000*	1.231/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	1.580/0.736
Petro et al. (2014)	PSNR	0.005*	0.012*	0.006*	0.962/0.736
	MSE	0.001*	0.008*	0.000*	0.598/0.736*
	MAE	0.888	0.923	0.845	0.214/0.736*
	SSIM	0.000*	0.000*	0.000*	1.922/0.736
	Sobel IoU	0.000*	0.000*	0.000*	1.562/0.736
	Canny IoU	0.704	0.285	0.671	0.402/0.736*
	Dif. Hist.	0.022*	0.003*	0.15	1.215/0.736
	GMSD	0.000*	0.000*	0.000*	1.779/0.736
	VIFP	0.000*	0.000*	0.000*	1.480/0.736
	FSIM	0.000*	0.000*	0.000*	1.786/0.736
	FSIMc	0.000*	0.000*	0.000*	1.740/0.736
	RECO	0.731	0.664	0.838	0.226/0.736*
	CIEDE 2000	0.315	0.804	0.474	0.194/0.736*
Dong et al. (2011)	PSNR	0.002*	0.004*	0.002*	0.882/0.736
	MSE	0.367	0.376	0.448	0.449/0.736*
	MAE	0.433	0.252	0.623	0.389/0.736*
	SSIM	0.001*	0.006*	0.000*	0.826/0.736
	Sobel IoU	0.06	0.011*	0.099	1.166/0.736
	Canny IoU	0.793	0.792	0.763	0.245/0.736*
	Dif. Hist.	0.004*	0.004*	0.197	0.977/0.736
	GMSD	0.000*	0.000*	0.000*	1.113/0.736
	VIFP	0.935	0.47	0.911	0.438/0.736*
	FSIM	0.000*	0.002*	0.000*	0.951/0.736
	FSIMc	0.000*	0.002*	0.000*	0.905/0.736
	RECO	0.388	0.256	0.485	0.521/0.736*
	CIEDE 2000	0.139	0.304	0.219	0.341/0.736*
Ying et al. (2017c)	PSNR	0.000*	0.000*	0.000*	5.252/0.736
	MSE	0.203	0.488	0.252	0.360/0.736*
	MAE	0.000*	0.002*	0.000*	0.911/0.736
	SSIM	0.004*	0.007*	0.001*	0.969/0.736
	Sobel IoU	0.007*	0.041*	0.009*	0.659/0.736*
	Canny IoU	0.696	0.844	0.673	0.257/0.736*
	Dif. Hist.	0.028*	0.005*	0.258	1.179/0.736

	GMSD	0.000 [*]	0.000 [*]	0.000 [*]	1.309/0.736
	VIFP	0.645	0.211	0.754	0.609/0.736 [*]
	FSIM	0.000 [*]	0.000 [*]	0.000 [*]	1.296/0.736
	FSIMc	0.000 [*]	0.000 [*]	0.000 [*]	1.295/0.736
	RECO	0.081	0.12	0.132	0.653/0.736 [*]
	CIEDE 2000	0.000 [*]	0.000 [*]	0.000 [*]	1.491/0.736
Não Tratada	PSNR	0.000 [*]	0.000 [*]	0.000 [*]	1.546/0.736
	MSE	0.008 [*]	0.037 [*]	0.011 [*]	0.530/0.736 [*]
	MAE	0.986	0.943	0.901	0.203/0.736 [*]
	SSIM	0.000 [*]	0.000 [*]	0.000 [*]	1.592/0.736
	Sobel IoU	0.023 [*]	0.099	0.038 [*]	0.538/0.736 [*]
	Canny IoU	0.848	0.288	0.777	0.479/0.736 [*]
	Dif. Hist.	0.012 [*]	0.002 [*]	0.141	1.222/0.736
	GMSD	0.000 [*]	0.000 [*]	0.000 [*]	1.795/0.736
	VIFP	0.51	0.513	0.719	0.501/0.736 [*]
	FSIM	0.000 [*]	0.000 [*]	0.000 [*]	1.806/0.736
	FSIMc	0.000 [*]	0.000 [*]	0.000 [*]	1.762/0.736
	RECO	0.981	0.453	0.977	0.343/0.736 [*]
	CIEDE 2000	0.609	0.907	0.787	0.258/0.736 [*]

7.3.2 Média

Tabela G.49: Valores de media para restauração de imagens subexpostas do *dataset* HDR+*burst*

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
M. Proposto	28,090	0,004	0,038	0,951	0,873	0,757	4,836	0,788	0,915	0,980	0,978	0,919	4,738
Ronneberger et al. (2015)	22,662	0,007	0,061	0,858	0,712	0,639	4,570	1,084	0,999	0,962	0,959	1,038	6,560
Chen et al. (2017)	21,203	0,009	0,075	0,899	0,755	0,578	6,015	1,394	1,004	0,952	0,944	1,055	9,373
Abdullah-Al-Wadud et al. (2007)	18,615	0,022	0,110	0,824	0,688	0,444	6,690	3,997	0,516	0,910	0,905	0,695	11,257
Ying et al. (2017b)	20,591	0,011	0,081	0,864	0,720	0,513	5,857	3,758	0,534	0,931	0,929	0,563	8,160
Fu et al. (2015)	20,330	0,011	0,081	0,863	0,719	0,469	6,148	3,561	0,528	0,932	0,929	0,587	8,341
Lee et al. (2013)	6,895	0,228	0,394	0,040	0,029	0,006	11,486	30,707	0,890	0,445	0,434	16,384	35,446
Petro et al. (2014)	21,538	0,013	0,079	0,853	0,789	0,647	5,402	2,540	0,636	0,955	0,948	0,653	8,679
Dong et al. (2011)	17,939	0,019	0,102	0,791	0,626	0,376	6,719	6,011	0,445	0,874	0,871	0,478	10,252
Ying et al. (2017c)	14,999	0,035	0,147	0,760	0,639	0,367	6,872	11,995	0,375	0,854	0,849	0,424	14,404
Não Tratada	23,454	0,011	0,077	0,850	0,855	0,740	5,283	1,164	0,829	0,969	0,969	0,869	6,655

Tabela G.50: Valores de media para restauração de imagens sobre-expostas do *dataset* HDR+*burst*

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
M. Proposto	22,424	0,008	0,063	0,939	0,804	0,644	5,795	2,005	1,208	0,964	0,956	1,298	8,621
Ronneberger et al. (2015)	20,444	0,012	0,075	0,845	0,627	0,535	5,844	3,133	1,328	0,944	0,939	1,274	8,128
Chen et al. (2017)	20,501	0,011	0,074	0,902	0,740	0,555	6,200	2,587	1,374	0,947	0,933	1,336	10,300
Abdullah-Al-Wadud et al. (2007)	13,392	0,053	0,191	0,785	0,645	0,433	7,189	8,655	0,480	0,880	0,875	0,489	17,003
Ying et al. (2017b)	12,039	0,073	0,243	0,731	0,651	0,458	7,899	8,099	0,517	0,886	0,880	0,551	21,466
Fu et al. (2015)	12,753	0,063	0,218	0,780	0,676	0,444	7,489	7,954	0,507	0,893	0,888	0,534	19,574
Lee et al. (2013)	14,800	0,039	0,152	0,747	0,450	0,202	8,174	10,444	1,709	0,804	0,798	3,251	14,627
Petro et al. (2014)	14,626	0,041	0,161	0,855	0,699	0,561	7,000	6,741	0,547	0,916	0,910	0,505	14,487
Dong et al. (2011)	13,869	0,070	0,229	0,716	0,583	0,377	7,682	9,656	0,451	0,838	0,832	0,425	20,733
Ying et al. (2017c)	10,481	0,102	0,290	0,673	0,611	0,378	8,081	10,799	0,476	0,853	0,846	0,564	25,803
Não Tratada	14,791	0,042	0,166	0,854	0,722	0,588	6,986	6,148	0,553	0,924	0,919	0,547	14,735

7.3.3 Teste dos Postos Sinalizados de Wilcoxon

Tabela G.51: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do *dataset* HDR+*burst* (estatística do teste)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	10	89	70	1	1	1	515,5	208	136	0	0	4	63
Chen et al. (2017)	17	95	64	41	1	0	218	79	321	0	0	110	33
Abdullah-Al-Wadud et al. (2007)	8	5	17	10	0	0	132	0	0	0	0	92	15
Ying et al. (2017b)	0	0	16	0	0	0	303	0	0	0	0	0	39
Fu et al. (2015)	1	32	32	15	0	0	177	0	0	0	0	0	27
Lee et al. (2013)	0	0	0	1	0	0	1	0	618	0	0	140	0
Petro et al. (2014)	34	21	23	8	23	117	376	24	1	34	34	0	56
Dong et al. (2011)	0	0	1	0	0	0	102	0	0	0	0	0	2
Ying et al. (2017c)	2	21	17	0	0	0	148	0	0	0	0	0	17
Não Tratada	142	80	78	48	336	530	256	369	10	208	244	168	214

Tabela G.52: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do *dataset* HDR+*burst* (estatística do teste)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	105	77	174	0	0	8	534	3	284	0	0	525	426
Chen et al. (2017)	226	328	370	105	29	14	361,5	79	364	7	1	481	212
Abdullah-Al-Wadud et al. (2007)	29	48	48	38	37	11	195	10	0	13	16	0	72
Ying et al. (2017b)	51	29	37	13	62	49	109	20	0	13	17	0	48
Fu et al. (2015)	55	41	47	27	63	20	160	22	0	14	17	0	56
Lee et al. (2013)	0	0	0	0	0	0	21	0	127	0	0	10	1
Petro et al. (2014)	47	52	63	110	144	213	261	23	0	55	67	0	106
Dong et al. (2011)	56	31	36	7	7	0	125	10	0	1	1	0	50
Ying et al. (2017c)	49	11	22	6	28	0	90	8	0	2	3	0	39
Não Tratada	85	81	84	136	230	346	293	59	0	106	132	0	119

7.4 Resultados para A6300 *Multi-Exposure Dataset* (real)

7.4.1 Testes de Normalidade

Tabela G.53: Resultados para o teste de normalidade para a saída dos modelos utilizando o *dataset* A6300 subexposto. Se o valor-p é menor que o nível de significância $\alpha = 0,05$ então rejeita-se a hipótese nula de que os dados tem distribuição normal. Para o teste de Anderson-Darling apresenta-se o valor da estatística e o valor crítico. Resultados considerando 50 amostras extraídas por amostragem aleatória simples sem reposição

Métodos	Métrica	χ^2	Shapiro-Wilk	Jarque-Bera	Anderson-Darling
M. Proposto	PSNR	0.000*	0.000*	0.000*	2.194/0.736
	MSE	0.112	0.022*	0.153	0.836/0.736
	MAE	0.701	0.329	0.783	0.466/0.736*
	SSIM	0.007*	0.028*	0.006*	0.697/0.736*
	Sobel IoU	0.000*	0.000*	0.000*	1.763/0.736
	Canny IoU	0.095	0.104	0.155	0.747/0.736
	Dif. Hist.	0.035*	0.002*	0.042*	1.589/0.736
	GMSD	0.122	0.184	0.214	0.547/0.736*
	VIFP	0.000*	0.000*	0.000*	0.830/0.736
	FSIM	0.000*	0.000*	0.000*	1.051/0.736
	FSIMc	0.000*	0.001*	0.000*	1.223/0.736
	RECO	0.000*	0.001*	0.000*	1.114/0.736
	CIEDE 2000	0.789	0.444	0.874	0.450/0.736*
Ronneberger et al. (2015)	PSNR	0.022*	0.015*	0.033*	0.987/0.736
	MSE	0.005*	0.010*	0.006*	0.682/0.736*
	MAE	0.925	0.667	0.837	0.258/0.736*
	SSIM	0.000*	0.004*	0.000*	0.665/0.736*
	Sobel IoU	0.000*	0.000*	0.000*	1.683/0.736
	Canny IoU	0.154	0.568	0.265	0.276/0.736*
	Dif. Hist.	0.083	0.011*	0.099	0.873/0.736
	GMSD	0.07	0.183	0.124	0.561/0.736*
	VIFP	0.000*	0.000*	0.000*	1.208/0.736
	FSIM	0.000*	0.000*	0.000*	1.213/0.736
	FSIMc	0.000*	0.000*	0.000*	1.168/0.736
	RECO	0.000*	0.000*	0.000*	1.432/0.736
	CIEDE 2000	0.736	0.612	0.694	0.250/0.736*
Chen et al. (2017)	PSNR	0.000*	0.000*	0.000*	2.541/0.736
	MSE	0.066	0.084	0.116	0.572/0.736*
	MAE	0.182	0.081	0.294	0.898/0.736
	SSIM	0.000*	0.003*	0.000*	0.810/0.736
	Sobel IoU	0.000*	0.000*	0.000*	1.178/0.736
	Canny IoU	0.000*	0.004*	0.000*	0.787/0.736
	Dif. Hist.	0.036*	0.006*	0.157	1.070/0.736
	GMSD	0.128	0.426	0.217	0.329/0.736*
	VIFP	0.000*	0.005*	0.000*	0.612/0.736*
	FSIM	0.000*	0.000*	0.000*	1.541/0.736

	FSIMc	0.000*	0.000*	0.000*	1.427/0.736
	RECO	0.043*	0.132	0.07	0.574/0.736*
	CIEDE 2000	0.345	0.52	0.457	0.319/0.736*
Abdullah-Al-Wadud et al. (2007)	PSNR	0.000*	0.001*	0.000*	1.240/0.736
	MSE	0.000*	0.000*	0.000*	7.460/0.736
	MAE	0.000*	0.000*	0.000*	3.714/0.736
	SSIM	0.000*	0.000*	0.000*	1.830/0.736
	Sobel IoU	0.032*	0.000*	0.031*	2.635/0.736
	Canny IoU	0.755	0.831	0.883	0.257/0.736*
	Dif. Hist.	0.034*	0.006*	0.048*	0.844/0.736
	GMSD	0.000*	0.000*	0.000*	3.521/0.736
	VIFP	0.559	0.928	0.779	0.173/0.736*
	FSIM	0.000*	0.000*	0.000*	3.352/0.736
	FSIMc	0.000*	0.000*	0.000*	3.381/0.736
	RECO	0.029*	0.016*	0.045*	0.793/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	3.575/0.736
Ying et al. (2017b)	PSNR	0.018*	0.024*	0.026*	0.764/0.736
	MSE	0.000*	0.000*	0.000*	3.020/0.736
	MAE	0.132	0.06	0.178	0.784/0.736
	SSIM	0.000*	0.000*	0.000*	1.758/0.736
	Sobel IoU	0.000*	0.000*	0.000*	0.974/0.736
	Canny IoU	0.627	0.635	0.67	0.453/0.736*
	Dif. Hist.	0.054	0.005*	0.147	1.156/0.736
	GMSD	0.000*	0.000*	0.000*	2.911/0.736
	VIFP	0.000*	0.000*	0.000*	3.479/0.736
	FSIM	0.000*	0.000*	0.000*	2.535/0.736
	FSIMc	0.000*	0.000*	0.000*	2.490/0.736
	RECO	0.000*	0.000*	0.000*	1.452/0.736
	CIEDE 2000	0.065	0.013*	0.085	1.077/0.736
Fu et al. (2015)	PSNR	0.000*	0.000*	0.000*	4.297/0.736
	MSE	0.686	0.758	0.698	0.277/0.736*
	MAE	0.000*	0.002*	0.000*	1.094/0.736
	SSIM	0.964	0.885	0.921	0.236/0.736*
	Sobel IoU	0.000*	0.000*	0.000*	1.359/0.736
	Canny IoU	0.076	0.016*	0.035*	1.114/0.736
	Dif. Hist.	0.428	0.025*	0.643	0.662/0.736*
	GMSD	0.15	0.028*	0.213	1.014/0.736
	VIFP	0.000*	0.000*	0.000*	1.412/0.736
	FSIM	0.007*	0.059	0.003*	0.405/0.736*
	FSIMc	0.007*	0.057	0.002*	0.427/0.736*
	RECO	0.000*	0.004*	0.000*	1.096/0.736
	CIEDE 2000	0.005*	0.012*	0.006*	0.732/0.736*
Lee et al. (2013)	PSNR	0.000*	0.000*	0.000*	2.614/0.736
	MSE	0.000*	0.000*	0.000*	3.892/0.736
	MAE	0.000*	0.000*	0.000*	1.262/0.736
	SSIM	0.000*	0.000*	0.000*	4.300/0.736
	Sobel IoU	0.000*	0.000*	0.000*	1.404/0.736
	Canny IoU	0.098	0.127	0.173	0.673/0.736*

	Dif. Hist.	0.116	0.002*	0.125	1.171/0.736
	GMSD	0.000*	0.000*	0.000*	1.230/0.736
	VIFP	0.006*	0.001*	0.000*	1.464/0.736
	FSIM	0.000*	0.000*	0.000*	7.162/0.736
	FSIMc	0.000*	0.000*	0.000*	7.098/0.736
	RECO	0.000*	0.000*	0.000*	17.908/0.736
	CIEDE 2000	0.001*	0.001*	0.000*	1.260/0.736
Petro et al. (2014)	PSNR	0.11	0.102	0.153	0.618/0.736*
	MSE	0.09	0.05	0.124	0.573/0.736*
	MAE	0.845	0.862	0.776	0.183/0.736*
	SSIM	0.15	0.043*	0.169	0.623/0.736*
	Sobel IoU	0.019*	0.074	0.029*	0.594/0.736*
	Canny IoU	0.337	0.147	0.525	0.407/0.736*
	Dif. Hist.	0.381	0.027*	0.428	0.997/0.736
	GMSD	0.000*	0.000*	0.000*	1.808/0.736
	VIFP	0.438	0.612	0.5	0.228/0.736*
	FSIM	0.486	0.628	0.571	0.299/0.736*
	FSIMc	0.537	0.756	0.625	0.268/0.736*
	RECO	0.012*	0.009*	0.018*	0.854/0.736
	CIEDE 2000	0.591	0.758	0.637	0.215/0.736*
Dong et al. (2011)	PSNR	0.603	0.541	0.598	0.377/0.736*
	MSE	0.000*	0.000*	0.000*	1.646/0.736
	MAE	0.001*	0.008*	0.000*	0.694/0.736*
	SSIM	0.001*	0.011*	0.001*	0.852/0.736
	Sobel IoU	0.003*	0.000*	0.002*	1.849/0.736
	Canny IoU	0.000*	0.000*	0.000*	2.487/0.736
	Dif. Hist.	0.000*	0.000*	0.000*	3.038/0.736
	GMSD	0.469	0.325	0.611	0.518/0.736*
	VIFP	0.000*	0.000*	0.000*	1.141/0.736
	FSIM	0.459	0.413	0.507	0.507/0.736*
	FSIMc	0.475	0.418	0.487	0.506/0.736*
	RECO	0.001*	0.009*	0.000*	0.982/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	1.185/0.736
Ying et al. (2017c)	PSNR	0.535	0.664	0.619	0.284/0.736*
	MSE	0.000*	0.000*	0.000*	2.199/0.736
	MAE	0.007*	0.003*	0.009*	1.089/0.736
	SSIM	0.000*	0.001*	0.000*	0.843/0.736
	Sobel IoU	0.000*	0.000*	0.000*	1.517/0.736
	Canny IoU	0.049*	0.12	0.029*	0.626/0.736*
	Dif. Hist.	0.000*	0.000*	0.000*	3.252/0.736
	GMSD	0.06	0.081	0.106	0.527/0.736*
	VIFP	0.000*	0.000*	0.000*	1.784/0.736
	FSIM	0.000*	0.001*	0.000*	0.614/0.736*
	FSIMc	0.000*	0.002*	0.000*	0.614/0.736*
	RECO	0.000*	0.010*	0.000*	0.714/0.736*
	CIEDE 2000	0.003*	0.001*	0.004*	1.196/0.736
	PSNR	0.000*	0.000*	0.000*	1.213/0.736
	MSE	0.366	0.073	0.421	0.763/0.736

MAE	0.015*	0.012*	0.024*	0.894/0.736
SSIM	0.328	0.229	0.342	0.579/0.736*
Sobel IoU	0.132	0.326	0.225	0.533/0.736*
Canny IoU	0.028*	0.000*	0.026*	2.076/0.736
Dif. Hist.	0.049*	0.000*	0.164	1.917/0.736
GMSD	0.693	0.353	0.775	0.418/0.736*
VIFP	0.000*	0.000*	0.000*	1.990/0.736
FSIM	0.909	0.207	0.991	0.515/0.736*
FSIMc	0.856	0.161	0.989	0.552/0.736*
RECO	0.000*	0.000*	0.000*	1.036/0.736
CIEDE 2000	0.152	0.065	0.228	0.548/0.736*

Tabela G.54: Resultados para o teste de normalidade para a saída dos modelos utilizando o *dataset* A6300 sobre-exposto. Se o valor-p é menor que o nível de significância $\alpha = 0,05$ então rejeita-se a hipótese nula de que os dados tem distribuição normal. Para o teste de Anderson-Darling apresenta-se o valor da estatística e o valor crítico. Resultados considerando 50 amostras extraídas por amostragem aleatória simples sem reposição

Métodos	Métrica	χ^2	Shapiro-Wilk	Jarque-Bera	Anderson-Darling
M. Proposto	PSNR	0.065	0.000*	0.076	2.976/0.736
	MSE	0.000*	0.000*	0.000*	3.379/0.736
	MAE	0.000*	0.000*	0.000*	1.150/0.736
	SSIM	0.010*	0.004*	0.015*	0.847/0.736
	Sobel IoU	0.925	0.278	0.885	0.575/0.736*
	Canny IoU	0.674	0.064	0.678	0.934/0.736
	Dif. Hist.	0.008*	0.001*	0.009*	1.391/0.736
	GMSD	0.409	0.143	0.533	0.387/0.736*
	VIFP	0.234	0.316	0.321	0.455/0.736*
	FSIM	0.219	0.012*	0.257	1.103/0.736
	FSIMc	0.159	0.005*	0.192	1.294/0.736
	RECO	0.158	0.34	0.205	0.494/0.736*
	CIEDE 2000	0.000*	0.000*	0.000*	1.959/0.736
Ronneberger et al. (2015)	PSNR	0.84	0.827	0.984	0.248/0.736*
	MSE	0.000*	0.000*	0.000*	3.846/0.736
	MAE	0.000*	0.000*	0.000*	0.954/0.736
	SSIM	0.061	0.167	0.105	0.544/0.736*
	Sobel IoU	0.58	0.359	0.567	0.379/0.736*
	Canny IoU	0.312	0.227	0.329	0.446/0.736*
	Dif. Hist.	0.776	0.337	0.722	0.403/0.736*
	GMSD	0.001*	0.001*	0.000*	1.127/0.736
	VIFP	0.031*	0.012*	0.048*	1.163/0.736
	FSIM	0.412	0.187	0.427	0.652/0.736*
	FSIMc	0.467	0.287	0.473	0.570/0.736*
	RECO	0.000*	0.000*	0.000*	5.658/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	1.391/0.736
	PSNR	0.040*	0.002*	0.168	1.821/0.736
	MSE	0.000*	0.000*	0.000*	1.918/0.736

	MAE	0.026*	0.002*	0.044*	0.974/0.736
	SSIM	0.039*	0.018*	0.067	0.679/0.736*
	Sobel IoU	0.558	0.684	0.597	0.305/0.736*
	Canny IoU	0.184	0.145	0.331	0.575/0.736*
	Dif. Hist.	0.018*	0.035*	0.253	0.629/0.736*
	GMSD	0.001*	0.001*	0.000*	0.749/0.736
	VIFP	0.117	0.008*	0.126	1.290/0.736
	FSIM	0.455	0.554	0.568	0.279/0.736*
	FSIMc	0.352	0.602	0.515	0.245/0.736*
	RECO	0.000*	0.000*	0.000*	2.660/0.736
	CIEDE 2000	0.000*	0.001*	0.000*	1.009/0.736
Abdullah-Al-Wadud et al. (2007)	PSNR	0.313	0.258	0.457	0.582/0.736*
	MSE	0.000*	0.000*	0.000*	0.930/0.736
	MAE	0.053	0.12	0.094	0.340/0.736*
	SSIM	0.081	0.020*	0.098	0.835/0.736
	Sobel IoU	0.020*	0.021*	0.029*	0.762/0.736
	Canny IoU	0.001*	0.026*	0.19	0.871/0.736
	Dif. Hist.	0.012*	0.002*	0.137	1.216/0.736
	GMSD	0.037*	0.071	0.064	0.480/0.736*
	VIFP	0.000*	0.000*	0.000*	0.867/0.736
	FSIM	0.276	0.187	0.368	0.400/0.736*
	FSIMc	0.23	0.248	0.378	0.346/0.736*
	RECO	0.054	0.109	0.297	0.583/0.736*
	CIEDE 2000	0.000*	0.001*	0.000*	0.711/0.736*
Ying et al. (2017b)	PSNR	0.000*	0.000*	0.000*	5.350/0.736
	MSE	0.498	0.003*	0.731	1.427/0.736
	MAE	0.059	0.000*	0.073	2.840/0.736
	SSIM	0.362	0.018*	0.389	1.197/0.736
	Sobel IoU	0.051	0.061	0.307	0.517/0.736*
	Canny IoU	0.972	0.239	0.973	0.610/0.736*
	Dif. Hist.	0.040*	0.000*	0.042*	2.180/0.736
	GMSD	0.599	0.009*	0.583	1.246/0.736
	VIFP	0.000*	0.000*	0.000*	1.887/0.736
	FSIM	0.331	0.021*	0.391	0.971/0.736
	FSIMc	0.364	0.019*	0.412	1.015/0.736
	RECO	0.000*	0.000*	0.000*	4.804/0.736
	CIEDE 2000	0.171	0.001*	0.292	2.103/0.736
Fu et al. (2015)	PSNR	0.108	0.001*	0.125	2.277/0.736
	MSE	0.000*	0.000*	0.000*	1.116/0.736
	MAE	0.292	0.018*	0.444	1.095/0.736
	SSIM	0.378	0.264	0.464	0.543/0.736*
	Sobel IoU	0.197	0.218	0.38	0.398/0.736*
	Canny IoU	0.79	0.075	0.956	0.875/0.736
	Dif. Hist.	0.044*	0.000*	0.047*	1.483/0.736
	GMSD	0.022*	0.030*	0.025*	0.531/0.736*
	VIFP	0.000*	0.000*	0.000*	2.615/0.736
	FSIM	0.574	0.272	0.618	0.421/0.736*
	FSIMc	0.617	0.263	0.639	0.437/0.736*

	RECO	0.000*	0.000*	0.000*	9.073/0.736
	CIEDE 2000	0.000*	0.003*	0.000*	0.721/0.736*
Lee et al. (2013)	PSNR	0.032*	0.001*	0.035*	1.942/0.736
	MSE	0.004*	0.011*	0.002*	0.486/0.736*
	MAE	0.064	0.029*	0.089	0.730/0.736*
	SSIM	0.131	0.008*	0.212	1.024/0.736
	Sobel IoU	0.211	0.068	0.287	0.561/0.736*
	Canny IoU	0.305	0.474	0.438	0.364/0.736*
	Dif. Hist.	0.103	0.002*	0.131	1.212/0.736
	GMSD	0.028*	0.026*	0.3	0.645/0.736*
	VIFP	0.000*	0.000*	0.000*	0.973/0.736
	FSIM	0.000*	0.002*	0.111	1.403/0.736
	FSIMc	0.000*	0.004*	0.123	1.287/0.736
	RECO	0.021*	0.089	0.277	0.648/0.736*
	CIEDE 2000	0.51	0.074	0.724	0.728/0.736*
Petro et al. (2014)	PSNR	0.052	0.000*	0.052	2.329/0.736
	MSE	0.118	0.013*	0.205	0.755/0.736
	MAE	0.621	0.014*	0.601	1.250/0.736
	SSIM	0.086	0.065	0.129	0.533/0.736*
	Sobel IoU	0.006*	0.047*	0.003*	0.512/0.736*
	Canny IoU	0.027*	0.034*	0.205	0.786/0.736
	Dif. Hist.	0.082	0.021*	0.277	0.771/0.736
	GMSD	0.000*	0.000*	0.000*	1.673/0.736
	VIFP	0.558	0.089	0.64	0.752/0.736
	FSIM	0.203	0.135	0.317	0.698/0.736*
	FSIMc	0.25	0.083	0.284	0.692/0.736*
	RECO	0.622	0.307	0.619	0.583/0.736*
	CIEDE 2000	0.034*	0.012*	0.020*	0.792/0.736
Dong et al. (2011)	PSNR	0.000*	0.001*	0.127	1.984/0.736
	MSE	0.000*	0.000*	0.000*	1.035/0.736
	MAE	0.005*	0.024*	0.001*	0.680/0.736*
	SSIM	0.008*	0.002*	0.009*	1.331/0.736
	Sobel IoU	0.964	0.374	0.873	0.390/0.736*
	Canny IoU	0.601	0.072	0.583	1.088/0.736
	Dif. Hist.	0.048*	0.000*	0.051	1.424/0.736
	GMSD	0.000*	0.002*	0.000*	0.972/0.736
	VIFP	0.000*	0.000*	0.000*	2.099/0.736
	FSIM	0.963	0.98	0.932	0.138/0.736*
	FSIMc	0.923	0.865	0.929	0.242/0.736*
	RECO	0.000*	0.000*	0.000*	4.759/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	1.312/0.736
Ying et al. (2017c)	PSNR	0.4	0.116	0.607	0.664/0.736*
	MSE	0.000*	0.000*	0.000*	1.204/0.736
	MAE	0.000*	0.002*	0.000*	0.746/0.736
	SSIM	0.006*	0.006*	0.007*	1.142/0.736
	Sobel IoU	0.142	0.021*	0.166	0.920/0.736
	Canny IoU	0.447	0.024*	0.546	1.174/0.736
	Dif. Hist.	0.062	0.000*	0.073	1.597/0.736

Não Tratada	GMSD	0.000*	0.005*	0.000*	0.579/0.736*
	VIFP	0.000*	0.000*	0.000*	2.111/0.736
	FSIM	0.599	0.495	0.641	0.376/0.736*
	FSIMc	0.66	0.5	0.671	0.385/0.736*
	RECO	0.000*	0.000*	0.000*	8.849/0.736
	CIEDE 2000	0.000*	0.000*	0.000*	1.920/0.736
	PSNR	0.000*	0.000*	0.000*	7.528/0.736
	MSE	0.003*	0.000*	0.000*	1.379/0.736
	MAE	0.122	0.000*	0.143	3.154/0.736
	SSIM	0.258	0.001*	0.274	1.910/0.736
	Sobel IoU	0.050*	0.066	0.299	0.495/0.736*
	Canny IoU	0.501	0.089	0.508	0.754/0.736
	Dif. Hist.	0.020*	0.000*	0.017*	3.216/0.736
	GMSD	0.117	0.003*	0.185	0.998/0.736
	VIFP	0.000*	0.000*	0.000*	3.508/0.736
	FSIM	0.327	0.009*	0.42	1.087/0.736
	FSIMc	0.33	0.008*	0.417	1.117/0.736
	RECO	0.000*	0.000*	0.000*	7.829/0.736
	CIEDE 2000	0.208	0.000*	0.296	3.097/0.736

7.4.2 Média

Tabela G.55: Valores de média para restauração de imagens subexpostas do *dataset* A6300 *Multi-Exposure Dataset*

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
M. Proposto	17,492	0,023	0,131	0,860	0,662	0,437	5,703	2,731	0,910	0,940	0,936	1,010	13,009
Ronneberger et al. (2015)	17,882	0,020	0,124	0,844	0,578	0,397	5,457	3,184	1,045	0,922	0,918	1,215	12,427
Chen et al. (2017)	16,013	0,031	0,158	0,812	0,529	0,349	7,188	3,590	1,239	0,912	0,903	1,705	15,963
Abdullah-Al-Wadud et al. (2007)	17,421	0,024	0,121	0,791	0,520	0,342	5,672	5,364	0,376	0,887	0,881	0,659	12,454
Ying et al. (2017b)	15,857	0,038	0,168	0,807	0,619	0,397	6,642	4,007	1,154	0,928	0,925	1,332	16,334
Fu et al. (2015)	13,593	0,052	0,212	0,727	0,613	0,407	7,784	4,798	1,079	0,920	0,917	1,362	19,772
Lee et al. (2013)	13,273	0,059	0,201	0,745	0,560	0,410	8,599	6,230	0,975	0,907	0,903	-3,021	19,630
Petro et al. (2014)	12,510	0,067	0,227	0,603	0,493	0,294	7,081	9,292	0,586	0,863	0,856	0,739	21,196
Dong et al. (2011)	22,331	0,019	0,107	0,798	0,508	0,227	4,954	4,713	0,433	0,878	0,873	0,494	11,869
Ying et al. (2017c)	21,742	0,009	0,076	0,904	0,685	0,468	4,223	1,822	0,874	0,951	0,949	1,040	8,063
Não Tratada	8,354	0,151	0,369	0,343	0,332	0,108	10,044	12,988	5,272	0,805	0,800	6,032	33,631

Tabela G.56: Valores de média para restauração de imagens sobre-exposta do *dataset* A6300 *Multi-Exposure Dataset*

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
M. Proposto	18,647	0,021	0,113	0,864	0,580	0,347	5,059	4,352	0,744	0,900	0,895	0,760	11,021
Ronneberger et al. (2015)	16,443	0,031	0,149	0,845	0,567	0,308	6,984	4,679	1,141	0,903	0,899	1,626	13,344
Chen et al. (2017)	13,622	0,071	0,218	0,823	0,544	0,277	6,848	5,767	0,877	0,889	0,874	1,535	18,469
Abdullah-Al-Wadud et al. (2007)	12,073	0,072	0,228	0,774	0,529	0,244	8,028	10,899	0,434	0,838	0,835	0,431	17,808
Ying et al. (2017b)	9,961	0,138	0,341	0,758	0,614	0,329	8,885	11,336	0,859	0,884	0,880	1,250	26,406
Fu et al. (2015)	8,617	0,157	0,372	0,728	0,536	0,260	9,386	12,642	1,029	0,864	0,861	1,941	28,976
Lee et al. (2013)	15,620	0,034	0,140	0,835	0,488	0,220	8,255	6,374	1,909	0,851	0,847	1,931	13,900
Petro et al. (2014)	11,375	0,090	0,264	0,789	0,580	0,295	8,186	10,924	0,444	0,864	0,857	0,519	21,415
Dong et al. (2011)	11,174	0,169	0,388	0,701	0,514	0,220	9,813	15,036	0,827	0,826	0,821	1,066	30,653
Ying et al. (2017c)	7,989	0,169	0,393	0,717	0,586	0,288	9,818	13,444	1,000	0,873	0,870	1,700	30,879
Não Tratada	12,701	0,122	0,304	0,787	0,639	0,387	8,133	9,914	1,046	0,900	0,897	1,486	23,424

7.4.3 Teste dos Postos Sinalizados de Wilcoxon

Tabela G.57: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do *dataset* A6300 *Multi-Exposure Dataset* (estatística do teste)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	437	429	452	148	36	250	453	157	16	3	1	20	461
Chen et al. (2017)	101	110	75	33	0	48	202	64	30	0	0	1	25
Abdullah-Al-Wadud et al. (2007)	431	421	354	160	133	140	617	270	0	120	113	125	375
Ying et al. (2017b)	406	369	344	315	273	295	338	278	439	354	407	334	354
Fu et al. (2015)	4	1	1	0	87	170	78	1	160	33	36	7	0
Lee et al. (2013)	2	3	16	2	79	466	147	72	440	92	101	94	5
Petro et al. (2014)	36	21	46	0	13	5	205	0	83	0	0	239	42
Dong et al. (2011)	112	400	311	28	35	7	423	46	0	1	1	0	383
Ying et al. (2017c)	117	76	93	96	77	220	280	113	364	26	16	415	81
Não Tratada	0	0	0	0	3	0	0	0	0	0	0	0	0

Tabela G.58: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do *dataset* A6300 *Multi-Exposure Dataset* (estatística do teste)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	282	268	267	403	521	511	104	507	5	604	580	0	332
Chen et al. (2017)	29	66	72	310	416	418	141	287	456	339	251	0	69
Abdullah-Al-Wadud et al. (2007)	6	6	6	207	401	324	12	14	38	128	137	0	30
Ying et al. (2017b)	0	0	0	64	542	496	10	33	431	341	375	104	0
Fu et al. (2015)	0	0	0	33	322	182	3	2	349	100	106	128	0
Lee et al. (2013)	146	159	226	345	134	106	89	192	9	42	41	27	196
Petro et al. (2014)	0	0	0	201	583	450	5	12	0	206	185	42	1
Dong et al. (2011)	10	0	0	18	332	129	0	0	556	13	13	432	0
Ying et al. (2017c)	0	0	0	19	563	346	0	0	324	155	166	72	0
Não Tratada	84	40	38	219	391	394	115	101	173	608	576	16	55

7.5 Resultados para Cai2018 *Multi-Exposure Dataset* (real)

7.5.1 Testes de Normalidade

Tabela G.59: Resultados para o teste de normalidade para a saída dos modelos utilizando o *dataset* Cai et al. (2018) subexposto. Se o valor-p é menor que o nível de significância $\alpha = 0,05$ então rejeita-se a hipótese nula de que os dados tem distribuição normal. Para o teste de Anderson-Darling apresenta-se o valor da estatística e o valor crítico. Resultados considerando 50 amostras extraídas por amostragem aleatória simples sem reposição

Métodos	Métrica	χ^2	Shapiro-Wilk	Jarque-Bera	Anderson-Darling
M. Proposto	PSNR	0.010*	0.031*	0.228	0.761/0.736
	MSE	0.051	0.000*	0.054	2.087/0.736
	MAE	0.037*	0.008*	0.158	1.031/0.736
	SSIM	0.062	0.001*	0.076	1.549/0.736
	Sobel IoU	0.124	0.016*	0.133	0.913/0.736
	Canny IoU	0.462	0.417	0.499	0.342/0.736*
	Dif. Hist.	0.000*	0.014*	0.184	1.177/0.736
	GMSD	0.001*	0.000*	0.001*	2.513/0.736
	VIFP	0.000*	0.000*	0.000*	4.877/0.736
	FSIM	0.014*	0.000*	0.014*	1.968/0.736
	FSIMc	0.012*	0.000*	0.012*	2.051/0.736
	RECO	0.000*	0.000*	0.000*	10.344/0.736
	CIEDE 2000	0.069	0.002*	0.072	1.405/0.736
Ronneberger et al. (2015)	PSNR	0.093	0.191	0.348	0.555/0.736*
	MSE	0.017*	0.001*	0.022*	1.351/0.736
	MAE	0.231	0.047*	0.293	0.651/0.736*
	SSIM	0.065	0.005*	0.149	1.228/0.736
	Sobel IoU	0.000*	0.010*	0.146	1.065/0.736
	Canny IoU	0.124	0.162	0.397	0.379/0.736*
	Dif. Hist.	0.143	0.112	0.192	0.535/0.736*
	GMSD	0.000*	0.000*	0.000*	2.959/0.736
	VIFP	0.000*	0.000*	0.000*	6.090/0.736
	FSIM	0.051	0.000*	0.051	1.626/0.736
	FSIMc	0.059	0.000*	0.059	1.645/0.736
	RECO	0.000*	0.000*	0.000*	8.147/0.736
	CIEDE 2000	0.196	0.045*	0.212	0.713/0.736*
Chen et al. (2017)	PSNR	0.012*	0.06	0.233	0.682/0.736*
	MSE	0.071	0.001*	0.087	1.577/0.736
	MAE	0.014*	0.032*	0.202	0.739/0.736
	SSIM	0.076	0.001*	0.089	1.438/0.736
	Sobel IoU	0.032*	0.027*	0.21	0.787/0.736
	Canny IoU	0.301	0.117	0.356	0.698/0.736*
	Dif. Hist.	0.102	0.203	0.378	0.480/0.736*
	GMSD	0.001*	0.000*	0.000*	2.950/0.736
	VIFP	0.000*	0.000*	0.000*	3.527/0.736
	FSIM	0.011*	0.000*	0.012*	2.023/0.736

	FSIMc	0.014 [*]	0.000 [*]	0.014 [*]	1.961/0.736
	RECO	0.000 [*]	0.000 [*]	0.000 [*]	10.787/0.736
	CIEDE 2000	0.152	0.030 [*]	0.164	0.815/0.736
Abdullah-Al-Wadud et al. (2007)	PSNR	0.118	0.016 [*]	0.173	0.985/0.736
	MSE	0.123	0.006 [*]	0.174	1.085/0.736
	MAE	0.038 [*]	0.095	0.295	0.565/0.736 [*]
	SSIM	0.227	0.027 [*]	0.246	0.614/0.736 [*]
	Sobel IoU	0.494	0.488	0.562	0.412/0.736 [*]
	Canny IoU	0.216	0.42	0.454	0.334/0.736 [*]
	Dif. Hist.	0.026 [*]	0.032 [*]	0.295	0.874/0.736
	GMSD	0.006 [*]	0.008 [*]	0.007 [*]	0.777/0.736
	VIFP	0.794	0.65	0.735	0.268/0.736 [*]
	FSIM	0.217	0.069	0.276	0.713/0.736 [*]
	FSIMc	0.286	0.131	0.312	0.521/0.736 [*]
	RECO	0.000 [*]	0.000 [*]	0.000 [*]	4.029/0.736
	CIEDE 2000	0.014 [*]	0.047 [*]	0.253	0.854/0.736
Ying et al. (2017b)	PSNR	0.054	0.004 [*]	0.064	0.994/0.736
	MSE	0.040 [*]	0.001 [*]	0.103	1.489/0.736
	MAE	0.020 [*]	0.09	0.274	0.593/0.736 [*]
	SSIM	0.001 [*]	0.005 [*]	0.132	1.175/0.736
	Sobel IoU	0.002 [*]	0.008 [*]	0.181	1.179/0.736
	Canny IoU	0.159	0.23	0.383	0.392/0.736 [*]
	Dif. Hist.	0.121	0.043 [*]	0.356	0.740/0.736
	GMSD	0.007 [*]	0.000 [*]	0.008 [*]	1.749/0.736
	VIFP	0.006 [*]	0.022 [*]	0.008 [*]	0.660/0.736 [*]
	FSIM	0.087	0.002 [*]	0.095	1.303/0.736
	FSIMc	0.09	0.002 [*]	0.097	1.283/0.736
	RECO	0.000 [*]	0.000 [*]	0.000 [*]	11.466/0.736
	CIEDE 2000	0.036 [*]	0.188	0.296	0.488/0.736 [*]
Fu et al. (2015)	PSNR	0.008 [*]	0.000 [*]	0.008 [*]	1.809/0.736
	MSE	0.001 [*]	0.016 [*]	0.184	0.918/0.736
	MAE	0.020 [*]	0.028 [*]	0.201	0.839/0.736
	SSIM	0.004 [*]	0.047 [*]	0.234	0.640/0.736 [*]
	Sobel IoU	0.002 [*]	0.032 [*]	0.202	0.980/0.736
	Canny IoU	0.015 [*]	0.041 [*]	0.274	0.838/0.736
	Dif. Hist.	0.373	0.006 [*]	0.421	0.928/0.736
	GMSD	0.109	0.006 [*]	0.12	1.007/0.736
	VIFP	0.11	0.008 [*]	0.118	1.036/0.736
	FSIM	0.001 [*]	0.000 [*]	0.097	1.817/0.736
	FSIMc	0.002 [*]	0.000 [*]	0.096	1.807/0.736
	RECO	0.000 [*]	0.000 [*]	0.000 [*]	10.637/0.736
	CIEDE 2000	0.031 [*]	0.062	0.234	0.681/0.736 [*]
Lee et al. (2013)	PSNR	0.862	0.003 [*]	0.786	1.554/0.736
	MSE	0.010 [*]	0.000 [*]	0.014 [*]	4.742/0.736
	MAE	0.025 [*]	0.000 [*]	0.058	2.791/0.736
	SSIM	0.000 [*]	0.000 [*]	0.049 [*]	3.456/0.736
	Sobel IoU	0.000 [*]	0.000 [*]	0.097	1.901/0.736
	Canny IoU	0.013 [*]	0.000 [*]	0.195	1.966/0.736

	Dif. Hist.	0.004*	0.000*	0.003*	4.114/0.736
	GMSD	0.012*	0.000*	0.010*	3.253/0.736
	VIFP	0.002*	0.003*	0.002*	1.039/0.736
	FSIM	0.018*	0.000*	0.014*	3.379/0.736
	FSIMc	0.021*	0.000*	0.018*	3.297/0.736
	RECO	0.000*	0.000*	0.000*	18.581/0.736
	CIEDE 2000	0.036*	0.004*	0.125	1.317/0.736
Petro et al. (2014)	PSNR	0.128	0.010*	0.208	1.046/0.736
	MSE	0.148	0.004*	0.159	0.928/0.736
	MAE	0.327	0.078	0.478	0.459/0.736*
	SSIM	0.328	0.327	0.511	0.302/0.736*
	Sobel IoU	0.11	0.062	0.288	0.747/0.736
	Canny IoU	0.346	0.147	0.447	0.463/0.736*
	Dif. Hist.	0.221	0.151	0.446	0.386/0.736*
	GMSD	0.001*	0.000*	0.000*	1.665/0.736
	VIFP	0.011*	0.007*	0.015*	1.026/0.736
	FSIM	0.089	0.005*	0.094	1.230/0.736
	FSIMc	0.126	0.013*	0.144	1.027/0.736
	RECO	0.000*	0.000*	0.000*	3.722/0.736
	CIEDE 2000	0.331	0.235	0.513	0.387/0.736*
Dong et al. (2011)	PSNR	0.001*	0.002*	0.133	1.423/0.736
	MSE	0.030*	0.000*	0.072	2.290/0.736
	MAE	0.000*	0.002*	0.12	1.472/0.736
	SSIM	0.004*	0.002*	0.12	1.351/0.736
	Sobel IoU	0.178	0.038*	0.192	0.950/0.736
	Canny IoU	0.384	0.284	0.399	0.377/0.736*
	Dif. Hist.	0.469	0.143	0.562	0.433/0.736*
	GMSD	0.000*	0.000*	0.000*	1.852/0.736
	VIFP	0.000*	0.000*	0.000*	1.577/0.736
	FSIM	0.013*	0.001*	0.016*	1.508/0.736
	FSIMc	0.012*	0.000*	0.013*	1.648/0.736
	RECO	0.000*	0.000*	0.000*	12.056/0.736
	CIEDE 2000	0.006*	0.008*	0.155	1.143/0.736
Ying et al. (2017c)	PSNR	0.24	0.018*	0.378	0.951/0.736
	MSE	0.037*	0.000*	0.07	2.972/0.736
	MAE	0.000*	0.000*	0.093	2.011/0.736
	SSIM	0.024*	0.000*	0.091	1.676/0.736
	Sobel IoU	0.173	0.11	0.309	0.714/0.736*
	Canny IoU	0.000*	0.003*	0.129	1.328/0.736
	Dif. Hist.	0.004*	0.05	0.229	0.678/0.736*
	GMSD	0.006*	0.000*	0.006*	2.973/0.736
	VIFP	0.31	0.381	0.401	0.336/0.736*
	FSIM	0.037*	0.000*	0.035*	2.294/0.736
	FSIMc	0.038*	0.000*	0.036*	2.312/0.736
	RECO	0.000*	0.000*	0.000*	13.995/0.736
	CIEDE 2000	0.001*	0.000*	0.107	1.891/0.736
	PSNR	0.000*	0.000*	0.000*	2.639/0.736
	MSE	0.043*	0.012*	0.174	0.959/0.736

MAE	0.081	0.002*	0.085	1.218/0.736
SSIM	0.005*	0.000*	0.006*	1.540/0.736
Sobel IoU	0.041*	0.000*	0.046*	1.659/0.736
Canny IoU	0.000*	0.000*	0.000*	2.772/0.736
Dif. Hist.	0.004*	0.000*	0.004*	2.994/0.736
GMSD	0.491	0.69	0.575	0.203/0.736*
VIFP	0.000*	0.000*	0.000*	7.877/0.736
FSIM	0.083	0.081	0.272	0.729/0.736*
FSIMc	0.086	0.084	0.282	0.708/0.736*
RECO	0.000*	0.000*	0.000*	6.087/0.736
CIEDE 2000	0.107	0.005*	0.114	1.130/0.736

Tabela G.60: Resultados para o teste de normalidade para a saída dos modelos utilizando o *dataset* Cai et al. (2018) sobre-exposto. Se o valor-p é menor que o nível de significância $\alpha = 0,05$ então rejeita-se a hipótese nula de que os dados tem distribuição normal. Para o teste de Anderson-Darling apresenta-se o valor da estatística e o valor crítico. Resultados considerando 50 amostras extraídas por amostragem aleatória simples sem reposição

Métodos	Métrica	χ^2	Shapiro-Wilk	Jarque-Bera	Anderson-Darling
M. Proposto	PSNR	0.521	0.475	0.604	0.381/0.736*
	MSE	0.000*	0.000*	0.000*	2.946/0.736
	MAE	0.000*	0.001*	0.000*	1.075/0.736
	SSIM	0.217	0.022*	0.284	0.716/0.736*
	Sobel IoU	0.012*	0.029*	0.203	0.773/0.736
	Canny IoU	0.128	0.16	0.374	0.377/0.736*
	Dif. Hist.	0.001*	0.013*	0.188	0.888/0.736
	GMSD	0.000*	0.000*	0.000*	3.577/0.736
	VIFP	0.062	0.003*	0.111	1.820/0.736
	FSIM	0.041*	0.002*	0.052	1.195/0.736
	FSIMc	0.059	0.003*	0.076	1.146/0.736
	RECO	0.000*	0.000*	0.000*	3.865/0.736
	CIEDE 2000	0.000*	0.001*	0.000*	0.939/0.736
Ronneberger et al. (2015)	PSNR	0.403	0.711	0.613	0.320/0.736*
	MSE	0.000*	0.000*	0.000*	2.412/0.736
	MAE	0.000*	0.000*	0.000*	2.217/0.736
	SSIM	0.119	0.174	0.31	0.550/0.736*
	Sobel IoU	0.027*	0.027*	0.206	0.834/0.736
	Canny IoU	0.001*	0.007*	0.19	0.856/0.736
	Dif. Hist.	0.010*	0.008*	0.237	0.949/0.736
	GMSD	0.000*	0.000*	0.000*	2.020/0.736
	VIFP	0.000*	0.000*	0.000*	2.142/0.736
	FSIM	0.125	0.004*	0.153	1.317/0.736
	FSIMc	0.121	0.004*	0.16	1.263/0.736
	RECO	0.000*	0.000*	0.000*	4.422/0.736
	CIEDE 2000	0.000*	0.002*	0.000*	0.760/0.736
	PSNR	0.107	0.179	0.359	0.643/0.736*
	MSE	0.000*	0.000*	0.000*	2.520/0.736

	MAE	0.037*	0.002*	0.045*	1.345/0.736
	SSIM	0.153	0.011*	0.171	0.815/0.736
	Sobel IoU	0.025*	0.07	0.237	0.744/0.736
	Canny IoU	0.020*	0.069	0.273	0.664/0.736*
	Dif. Hist.	0.049*	0.149	0.322	0.465/0.736*
	GMSD	0.000*	0.000*	0.000*	3.886/0.736
	VIFP	0.96	0.672	0.991	0.260/0.736*
	FSIM	0.095	0.003*	0.109	1.212/0.736
	FSIMc	0.122	0.003*	0.136	1.221/0.736
	RECO	0.000*	0.000*	0.000*	1.950/0.736
	CIEDE 2000	0.020*	0.001*	0.024*	1.371/0.736
Abdullah-Al-Wadud et al. (2007)	PSNR	0.016*	0.034*	0.025*	0.575/0.736*
	MSE	0.002*	0.001*	0.001*	1.360/0.736
	MAE	0.465	0.255	0.47	0.517/0.736*
	SSIM	0.605	0.658	0.63	0.233/0.736*
	Sobel IoU	0.26	0.301	0.477	0.388/0.736*
	Canny IoU	0.217	0.047*	0.262	0.513/0.736*
	Dif. Hist.	0.000*	0.000*	0.116	1.788/0.736
	GMSD	0.000*	0.000*	0.000*	1.594/0.736
	VIFP	0.007*	0.022*	0.010*	0.704/0.736*
	FSIM	0.647	0.477	0.623	0.248/0.736*
	FSIMc	0.645	0.504	0.638	0.251/0.736*
	RECO	0.000*	0.000*	0.000*	10.260/0.736
	CIEDE 2000	0.116	0.274	0.197	0.381/0.736*
Ying et al. (2017b)	PSNR	0.049*	0.032*	0.08	0.689/0.736*
	MSE	0.314	0.182	0.369	0.510/0.736*
	MAE	0.823	0.24	0.825	0.545/0.736*
	SSIM	0.569	0.511	0.569	0.232/0.736*
	Sobel IoU	0.119	0.289	0.38	0.388/0.736*
	Canny IoU	0.309	0.094	0.35	0.457/0.736*
	Dif. Hist.	0.047*	0.000*	0.1	2.329/0.736
	GMSD	0.001*	0.002*	0.000*	1.144/0.736
	VIFP	0.000*	0.000*	0.000*	2.090/0.736
	FSIM	0.109	0.129	0.286	0.610/0.736*
	FSIMc	0.068	0.098	0.271	0.670/0.736*
	RECO	0.000*	0.000*	0.000*	17.697/0.736
	CIEDE 2000	0.402	0.432	0.52	0.386/0.736*
Fu et al. (2015)	PSNR	0.001*	0.006*	0.001*	0.783/0.736
	MSE	0.31	0.247	0.337	0.475/0.736*
	MAE	0.47	0.411	0.583	0.415/0.736*
	SSIM	0.763	0.459	0.712	0.242/0.736*
	Sobel IoU	0.099	0.246	0.377	0.353/0.736*
	Canny IoU	0.272	0.019*	0.302	0.639/0.736*
	Dif. Hist.	0.011*	0.000*	0.085	2.736/0.736
	GMSD	0.001*	0.002*	0.001*	1.178/0.736
	VIFP	0.000*	0.000*	0.000*	4.704/0.736
	FSIM	0.301	0.306	0.371	0.448/0.736*
	FSIMc	0.269	0.284	0.378	0.451/0.736*

	RECO	0.000*	0.000*	0.000*	15.903/0.736
	CIEDE 2000	0.795	0.881	0.872	0.222/0.736*
Lee et al. (2013)	PSNR	0.481	0.244	0.709	0.642/0.736*
	MSE	0.000*	0.000*	0.000*	2.323/0.736
	MAE	0.007*	0.048*	0.004*	0.610/0.736*
	SSIM	0.000*	0.000*	0.000*	0.941/0.736
	Sobel IoU	0.672	0.617	0.641	0.401/0.736*
	Canny IoU	0.295	0.23	0.489	0.417/0.736*
	Dif. Hist.	0.09	0.001*	0.166	1.328/0.736
	GMSD	0.000*	0.000*	0.000*	1.597/0.736
	VIFP	0.003*	0.012*	0.003*	0.684/0.736*
	FSIM	0.19	0.302	0.282	0.371/0.736*
	FSIMc	0.248	0.361	0.347	0.341/0.736*
	RECO	0.000*	0.000*	0.000*	16.553/0.736
	CIEDE 2000	0.779	0.853	0.884	0.167/0.736*
Petro et al. (2014)	PSNR	0.000*	0.000*	0.000*	1.727/0.736
	MSE	0.045*	0.038*	0.073	0.631/0.736*
	MAE	0.777	0.695	0.746	0.202/0.736*
	SSIM	0.217	0.497	0.453	0.284/0.736*
	Sobel IoU	0.112	0.09	0.306	0.576/0.736*
	Canny IoU	0.073	0.006*	0.089	0.893/0.736
	Dif. Hist.	0.000*	0.000*	0.118	1.833/0.736
	GMSD	0.000*	0.000*	0.000*	1.500/0.736
	VIFP	0.69	0.684	0.657	0.219/0.736*
	FSIM	0.399	0.165	0.498	0.536/0.736*
	FSIMc	0.332	0.223	0.488	0.454/0.736*
	RECO	0.000*	0.000*	0.000*	12.352/0.736
	CIEDE 2000	0.634	0.776	0.87	0.227/0.736*
Dong et al. (2011)	PSNR	0.000*	0.002*	0.143	1.719/0.736
	MSE	0.081	0.039*	0.106	0.746/0.736
	MAE	0.976	0.754	0.981	0.339/0.736*
	SSIM	0.522	0.353	0.553	0.265/0.736*
	Sobel IoU	0.046*	0.087	0.259	0.584/0.736*
	Canny IoU	0.18	0.064	0.409	0.579/0.736*
	Dif. Hist.	0.048*	0.000*	0.06	2.571/0.736
	GMSD	0.000*	0.000*	0.000*	1.789/0.736
	VIFP	0.000*	0.000*	0.000*	4.267/0.736
	FSIM	0.248	0.134	0.296	0.604/0.736*
	FSIMc	0.233	0.137	0.319	0.600/0.736*
	RECO	0.000*	0.000*	0.000*	16.652/0.736
	CIEDE 2000	0.306	0.286	0.382	0.329/0.736*
Ying et al. (2017c)	PSNR	0.705	0.762	0.835	0.265/0.736*
	MSE	0.052	0.020*	0.072	0.987/0.736
	MAE	0.719	0.427	0.898	0.550/0.736*
	SSIM	0.721	0.63	0.682	0.252/0.736*
	Sobel IoU	0.023*	0.14	0.264	0.528/0.736*
	Canny IoU	0.209	0.034*	0.388	0.643/0.736*
	Dif. Hist.	0.048*	0.000*	0.078	2.802/0.736

Não Tratada	GMSD	0.002 [*]	0.000 [*]	0.001 [*]	1.764/0.736
	VIFP	0.000 [*]	0.000 [*]	0.000 [*]	4.311/0.736
	FSIM	0.242	0.174	0.308	0.532/0.736 [*]
	FSIMc	0.214	0.159	0.315	0.548/0.736 [*]
	RECO	0.000 [*]	0.000 [*]	0.000 [*]	16.503/0.736
	CIEDE 2000	0.053	0.042 [*]	0.086	0.471/0.736 [*]
	PSNR	0.000 [*]	0.000 [*]	0.000 [*]	2.027/0.736
	MSE	0.222	0.06	0.238	0.712/0.736 [*]
	MAE	0.451	0.104	0.501	0.582/0.736 [*]
	SSIM	0.364	0.658	0.53	0.212/0.736 [*]
	Sobel IoU	0.026 [*]	0.131	0.295	0.494/0.736 [*]
	Canny IoU	0.236	0.016 [*]	0.256	0.701/0.736 [*]
	Dif. Hist.	0.040 [*]	0.000 [*]	0.075	2.443/0.736
	GMSD	0.001 [*]	0.002 [*]	0.001 [*]	1.022/0.736
	VIFP	0.000 [*]	0.000 [*]	0.000 [*]	4.887/0.736
	FSIM	0.154	0.046 [*]	0.263	0.756/0.736
	FSIMc	0.124	0.047 [*]	0.268	0.747/0.736
	RECO	0.000 [*]	0.000 [*]	0.000 [*]	16.611/0.736
	CIEDE 2000	0.691	0.34	0.82	0.461/0.736 [*]

7.5.2 Média

Tabela G.61: Valores de média para restauração de imagens subexpostas do *dataset* Cai2018 *Multi-Exposure Dataset*

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
M. Proposto	16,295	0,035	0,141	0,725	0,506	0,389	7,261	5,317	1,116	0,880	0,867	1,772	15,973
Ronneberger et al. (2015)	17,673	0,021	0,112	0,695	0,358	0,289	7,032	6,376	1,464	0,825	0,817	2,958	12,980
Chen et al. (2017)	16,432	0,030	0,135	0,716	0,421	0,314	7,579	6,153	1,425	0,839	0,827	2,829	16,009
Abdullah-Al-Wadud et al. (2007)	15,244	0,041	0,157	0,649	0,530	0,301	8,196	6,192	0,522	0,852	0,838	1,537	17,428
Ying et al. (2017b)	13,384	0,072	0,217	0,617	0,450	0,319	8,268	8,118	1,401	0,837	0,828	4,143	22,788
Fu et al. (2015)	11,369	0,103	0,273	0,506	0,396	0,295	8,911	10,642	1,534	0,805	0,796	4,274	27,433
Lee et al. (2013)	9,909	0,134	0,301	0,456	0,345	0,234	10,131	15,133	1,258	0,726	0,716	404,338	29,904
Petro et al. (2014)	10,831	0,103	0,271	0,445	0,384	0,242	8,740	11,904	0,911	0,772	0,755	1,675	27,743
Dong et al. (2011)	15,210	0,074	0,212	0,578	0,456	0,153	7,704	9,702	0,778	0,797	0,787	0,589	22,256
Ying et al. (2017c)	15,994	0,052	0,171	0,677	0,504	0,343	7,507	6,601	1,121	0,858	0,849	0,596	18,386
Não Tratada	8,083	0,178	0,380	0,259	0,179	0,116	10,323	17,532	7,815	0,667	0,656	13,723	36,819

Tabela G.62: Valores de média para restauração de imagens sobre-expostas do *dataset* Cai2018 *Multi-Exposure Dataset*

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
M. Proposto	16,994	0,028	0,123	0,779	0,563	0,352	5,598	6,128	0,709	0,854	0,838	1,746	16,323
Ronneberger et al. (2015)	15,956	0,029	0,131	0,697	0,402	0,229	5,727	7,961	1,652	0,801	0,789	3,461	14,753
Chen et al. (2017)	16,356	0,031	0,131	0,767	0,533	0,285	4,609	5,900	0,712	0,843	0,829	1,490	14,895
Abdullah-Al-Wadud et al. (2007)	10,980	0,100	0,257	0,680	0,514	0,250	7,278	14,628	0,479	0,773	0,763	0,817	20,587
Ying et al. (2017b)	8,493	0,156	0,360	0,635	0,463	0,240	8,583	14,763	1,035	0,782	0,774	-7,297	28,081
Fu et al. (2015)	8,501	0,162	0,362	0,631	0,392	0,199	8,659	15,503	1,658	0,768	0,760	105,486	28,130
Lee et al. (2013)	12,930	0,056	0,185	0,681	0,375	0,167	7,808	11,178	2,336	0,754	0,742	2,168	20,191
Petro et al. (2014)	10,675	0,109	0,275	0,690	0,523	0,268	7,431	14,181	0,538	0,783	0,771	1,082	22,250
Dong et al. (2011)	10,730	0,160	0,360	0,631	0,420	0,203	8,641	15,657	1,404	0,763	0,754	12,382	28,200
Ying et al. (2017c)	8,053	0,169	0,377	0,621	0,433	0,220	8,831	15,855	1,478	0,774	0,765	15,223	29,505
Não Tratada	9,489	0,143	0,329	0,674	0,471	0,270	8,221	14,011	1,503	0,796	0,788	12,823	25,454

7.5.3 Teste dos Postos Sinalizados de Wilcoxon

Tabela G.63: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens subexpostas do *dataset* Cai2018 *Multi-Exposure Dataset* (estatística do teste)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	254	182	216	376	3	167	521	282	205	24	38	54	103
Chen et al. (2017)	493	329	440	457	0	51	341	255	65	3	2	51	624
Abdullah-Al-Wadud et al. (2007)	440	427	472	145	428	150	147	431	18	170	158	438	439
Ying et al. (2017b)	219	111	122	74	246	266	155,5	123	339	167	204	140	138
Fu et al. (2015)	110	41	48	9	22	164	73	32	211	47	68	106	63
Lee et al. (2013)	7	8	10	0	114	89	23	8	487	36	48	637	8
Petro et al. (2014)	18	19	18	0	57	92	105	15	474	11	8	626	19
Dong et al. (2011)	384	86	104	7	146	0	410	5	125	1	1	425	141
Ying et al. (2017c)	560	224	278	207	589	260	494,5	187	440	266	397	218	370
Não Tratada	3	2	2	0	0	34	5	1	9	2	3	172	3

Tabela G.64: Teste dos Postos Sinalizados de Wilcoxon para restauração de imagens sobre-expostas do *dataset* Cai2018 *Multi-Exposure Dataset* (estatística do teste)

Método	PSNR	MSE	MAE	SSIM	Sobel IoU	Canny IoU	Hist. Diff.	GMSD	VIFP	FSIM	FSIMc	RECO	CIEDE 2000
Ronneberger et al. (2015)	383	495	448	68	0	7	628	78	0	0	1	0	458
Chen et al. (2017)	319	354	399	254	177	60	198	347	610	190	297	532	410
Abdullah-Al-Wadud et al. (2007)	0	0	0	48	334	33	274	0	74	3	7	54	42
Ying et al. (2017b)	0	0	0	0	11	26	93	0	161	0	9	344	0
Fu et al. (2015)	0	0	0	17	1	2	79	0	127	0	0	141	0
Lee et al. (2013)	59	85	80	45	9	1	182	18	7	0	3	198	178
Petro et al. (2014)	1	1	5	102	287	99	219	0	96	10	11	101	60
Dong et al. (2011)	5	0	0	0	0	6	87	0	257	0	0	353	0
Ying et al. (2017c)	0	0	0	0	6	9	70	0	160	0	1	127	0
Não Tratada	1	1	5	105	126	139	137	1	56	24	51	157	25

7.6 Avaliação do Impacto da Subexposição e Sobre-exposição em Aplicações de Reconhecimento de Imagens

Tabela G.65: Avaliação do impacto de distorções simuladas no desempenho de modelos de reconhecimento de imagens

	VGG-16			Resnet			Inception-v3			Inception Resnet-v2			DenseNet		
Metrica	Top-1	Top-5	F1-Score	Top-1	Top-5	F1-Score	Top-1	Top-5	F1-Score	Top-1	Top-5	F1-Score	Top-1	Top-5	F1-Score
Original	0,612	0,838	0,609	0,668	0,870	0,666	0,747	0,920	0,744	0,773	0,936	0,770	0,663	0,870	0,664
Gama 1/2	0,584	0,817	0,582	0,634	0,848	0,633	0,727	0,908	0,724	0,755	0,926	0,752	0,602	0,830	0,608
Gama 1/4	0,455	0,707	0,464	0,501	0,745	0,509	0,645	0,854	0,647	0,683	0,879	0,685	0,445	0,696	0,465
Gama 1/8	0,236	0,452	0,252	0,280	0,503	0,302	0,469	0,703	0,487	0,516	0,743	0,533	0,222	0,429	0,243
Gama 2	0,566	0,800	0,564	0,623	0,838	0,621	0,719	0,905	0,716	0,746	0,920	0,742	0,625	0,841	0,626
Gama 4	0,401	0,635	0,408	0,459	0,693	0,467	0,591	0,809	0,593	0,626	0,836	0,626	0,458	0,685	0,468
Gama 8	0,175	0,334	0,192	0,217	0,390	0,235	0,342	0,552	0,361	0,385	0,595	0,402	0,224	0,398	0,242
Truncado Q_1	0,541	0,780	0,539	0,593	0,814	0,592	0,713	0,900	0,709	0,737	0,916	0,734	0,642	0,854	0,641
Truncado Q_3	0,548	0,780	0,546	0,603	0,816	0,602	0,721	0,903	0,718	0,750	0,922	0,747	0,642	0,854	0,642
	Xception			Mobilenet			Mobilenet V2			NASNetLarge			NASNetMobile		
Metrica	Top-1	Top-5	F1-Score	Top-1	Top-5	F1-Score	Top-1	Top-5	F1-Score	Top-1	Top-5	F1-Score	Top-1	Top-5	F1-Score
Original	0,763	0,929	0,760	0,657	0,863	0,655	0,600	0,827	0,603	0,806	0,951	0,803	0,693	0,888	0,689
Gama 1/2	0,746	0,920	0,743	0,623	0,840	0,624	0,497	0,751	0,511	0,794	0,946	0,791	0,663	0,868	0,661
Gama 1/4	0,662	0,869	0,665	0,504	0,746	0,515	0,295	0,535	0,316	0,745	0,919	0,743	0,551	0,780	0,558
Gama 1/8	0,452	0,452	0,252	0,285	0,513	0,313	0,130	0,285	0,139	0,612	0,828	0,620	0,324	0,542	0,349
Gama 2	0,734	0,915	0,731	0,614	0,832	0,612	0,564	0,795	0,566	0,787	0,941	0,784	0,652	0,858	0,649
Gama 4	0,612	0,828	0,612	0,452	0,686	0,461	0,376	0,605	0,389	0,691	0,882	0,690	0,486	0,716	0,491
Gama 8	0,370	0,586	0,386	0,218	0,396	0,238	0,157	0,305	0,174	0,467	0,681	0,480	0,239	0,416	0,257
Truncado Q_1	0,729	0,910	0,726	0,612	0,831	0,610	0,577	0,806	0,577	0,777	0,936	0,774	0,647	0,857	0,643
Truncado Q_3	0,734	0,914	0,731	0,621	0,839	0,619	0,587	0,816	0,588	0,785	0,940	0,782	0,659	0,863	0,655

7.7 Dataset Fivek



Figura 7.1: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem sobre-exposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a), reduzindo o número de pixels truncados e tornando-se a imagem com uma cor mais fiel à original

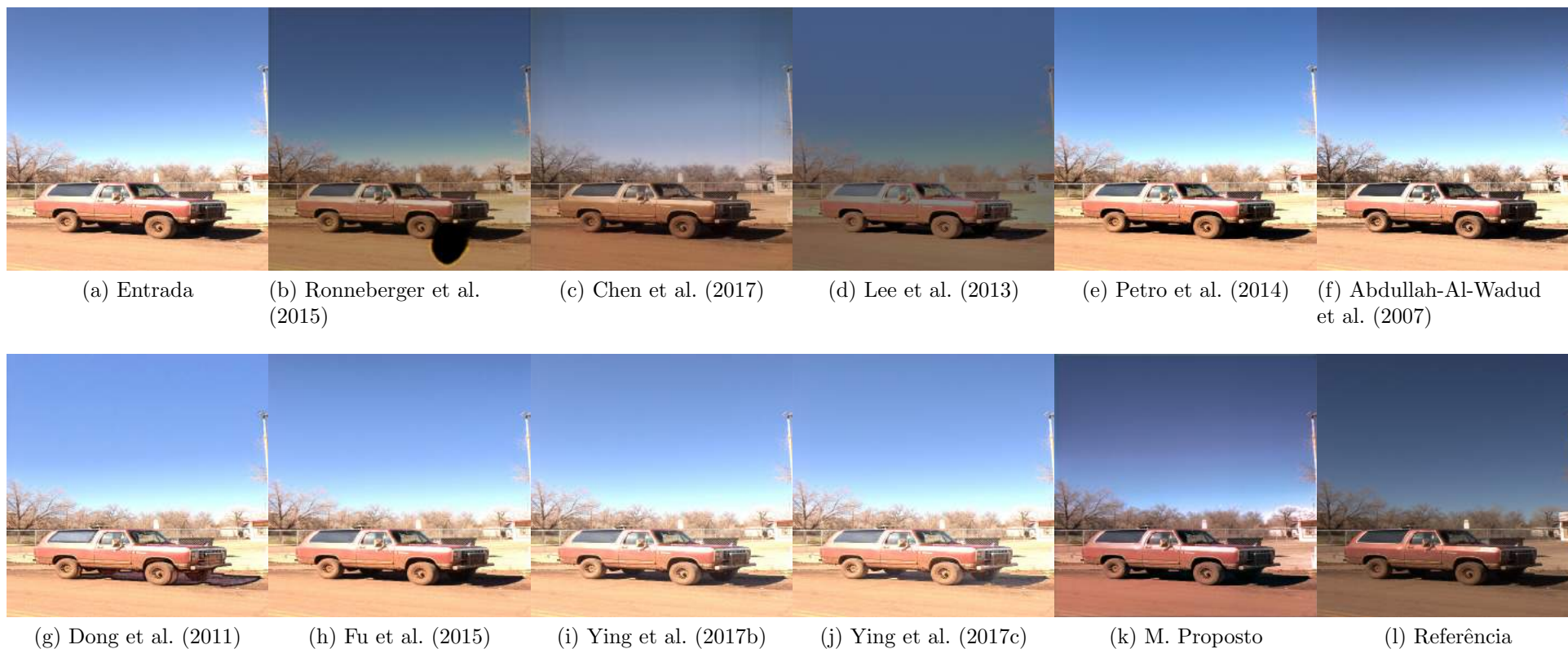


Figura 7.2: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem sobre-exposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a), reduzindo o número de pixels truncados no fundo da imagem, obtendo-se um melhor detalhamento dos objetos



(a) Entrada

(b) Ronneberger et al. (2015)

(c) Chen et al. (2017)

(d) Lee et al. (2013)

(e) Petro et al. (2014)

(f) Abdullah-Al-Wadud et al. (2007)



(g) Dong et al. (2011)

(h) Fu et al. (2015)

(i) Ying et al. (2017b)

(j) Ying et al. (2017c)

(k) M. Proposto

(l) Referência

Figura 7.3: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem subexposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a), reduzindo o número de pixels truncados e tornando a cor mais coerente com a referência

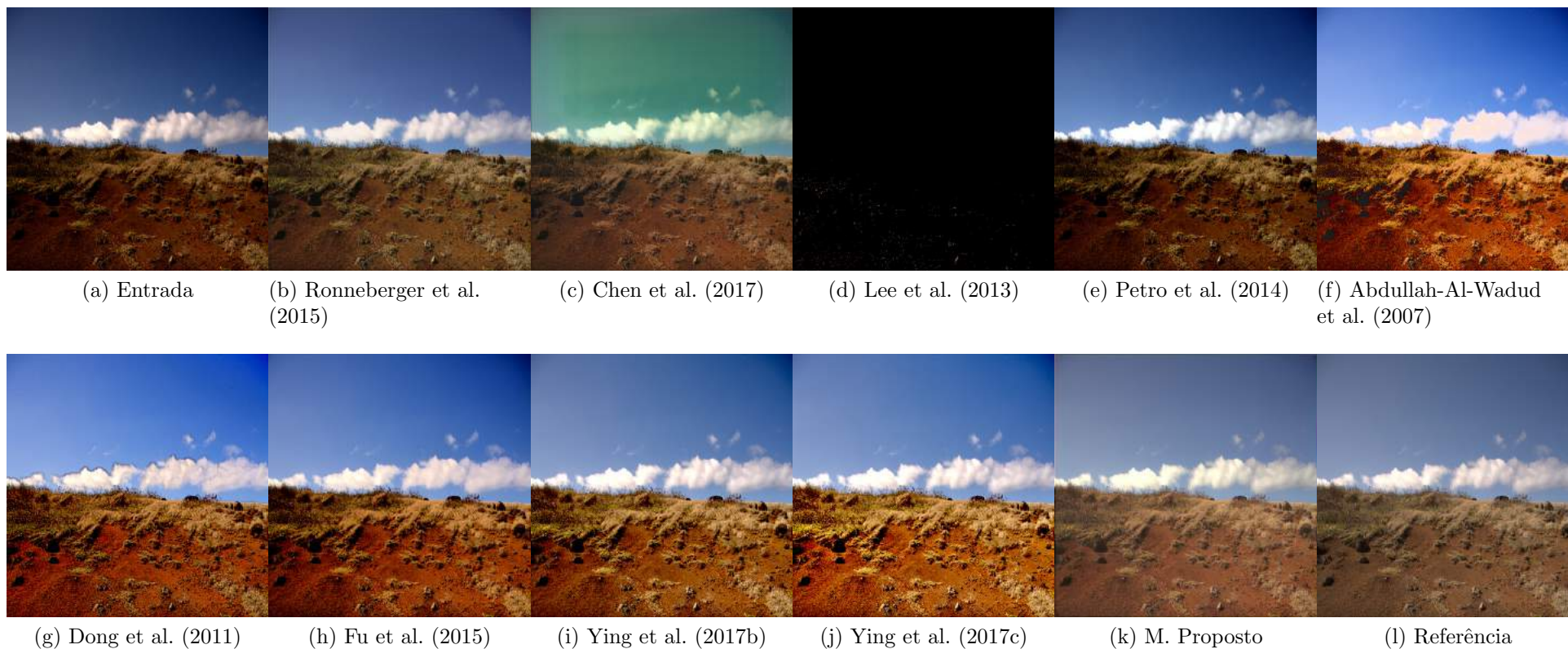


Figura 7.4: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem subexposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a), tornando a cor e detalhes mais coerente com a referência

7.8 Dataset HDR+Burst

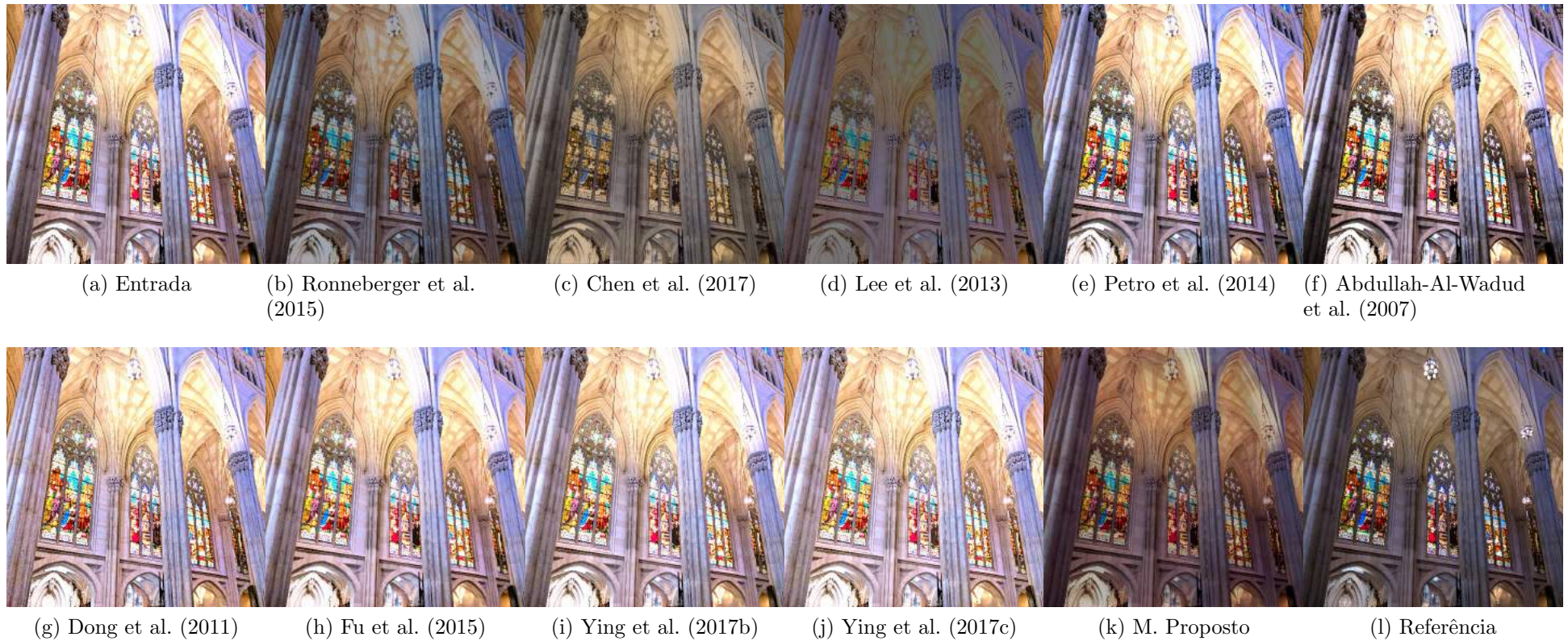


Figura 7.5: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem sobre-exposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a), principalmente na parte superior, tornando a cor e detalhes mais coerente com a referência

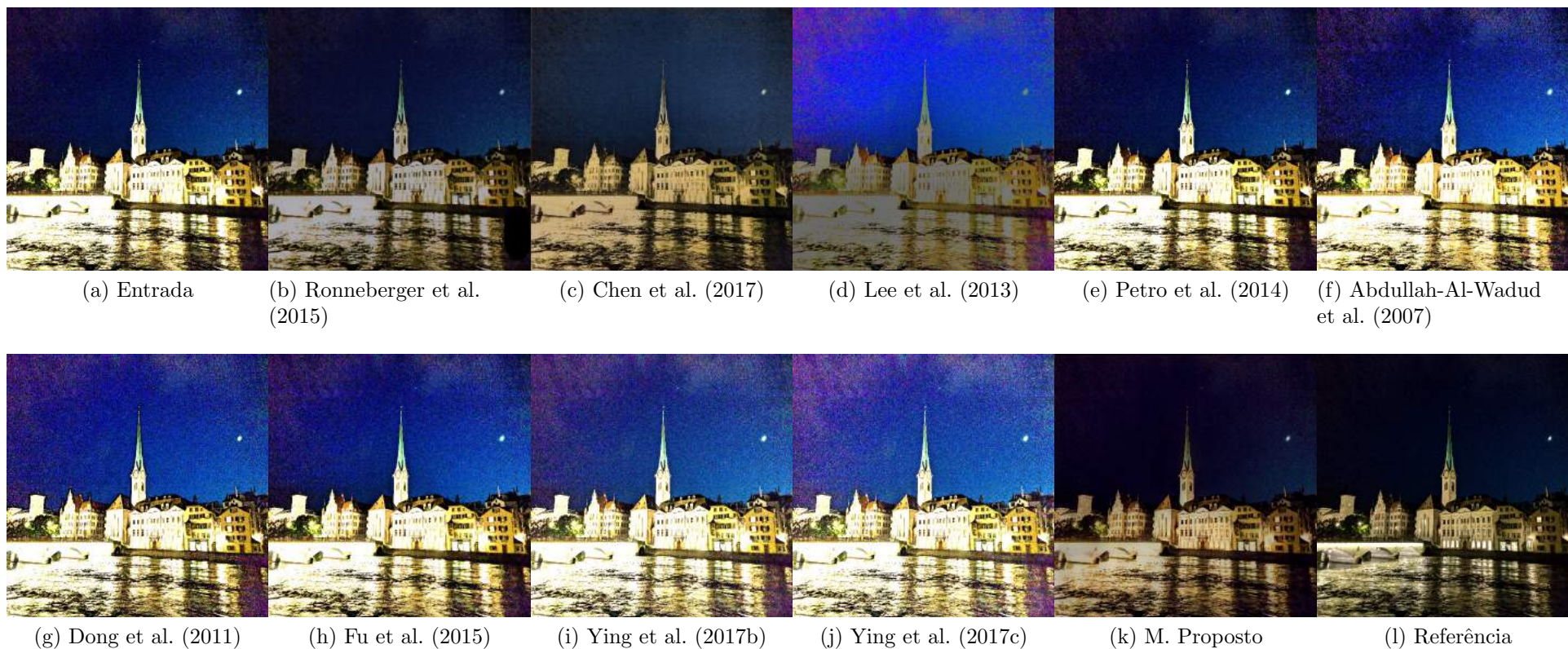
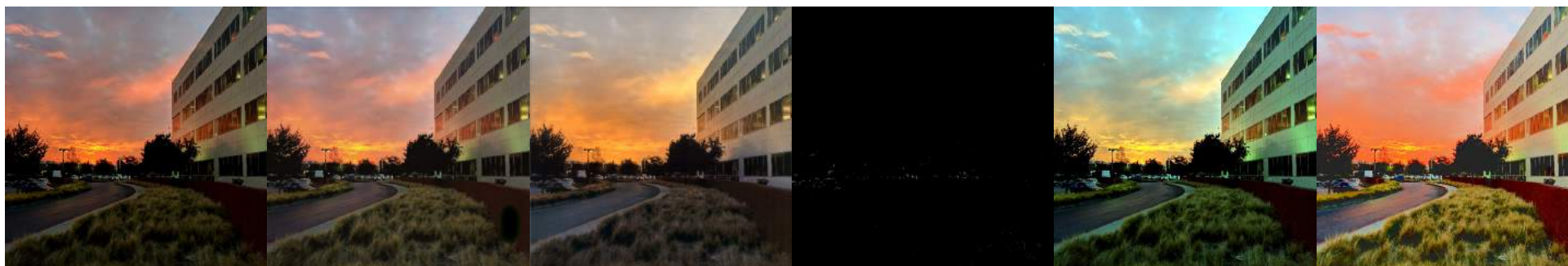


Figura 7.6: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem sobre-exposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a), filtrando grande parte dos ruídos, diferente dos outros métodos que enalteceram-os



Figura 7.7: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem subexposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a), reduzindo o número de pixels truncados e tornando a cor mais homogenia, como na imagem de referência



(a) Entrada

(b) Ronneberger et al. (2015)

(c) Chen et al. (2017)

(d) Lee et al. (2013)

(e) Petro et al. (2014)

(f) Abdullah-Al-Wadud et al. (2007)



(g) Dong et al. (2011)

(h) Fu et al. (2015)

(i) Ying et al. (2017b)

(j) Ying et al. (2017c)

(k) M. Proposto

(l) Referência

Figura 7.8: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem subexposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a), reduzindo o número de pixels truncados e tornando a cor mais homogênea, como na imagem de referência

7.9 Dataset A6300

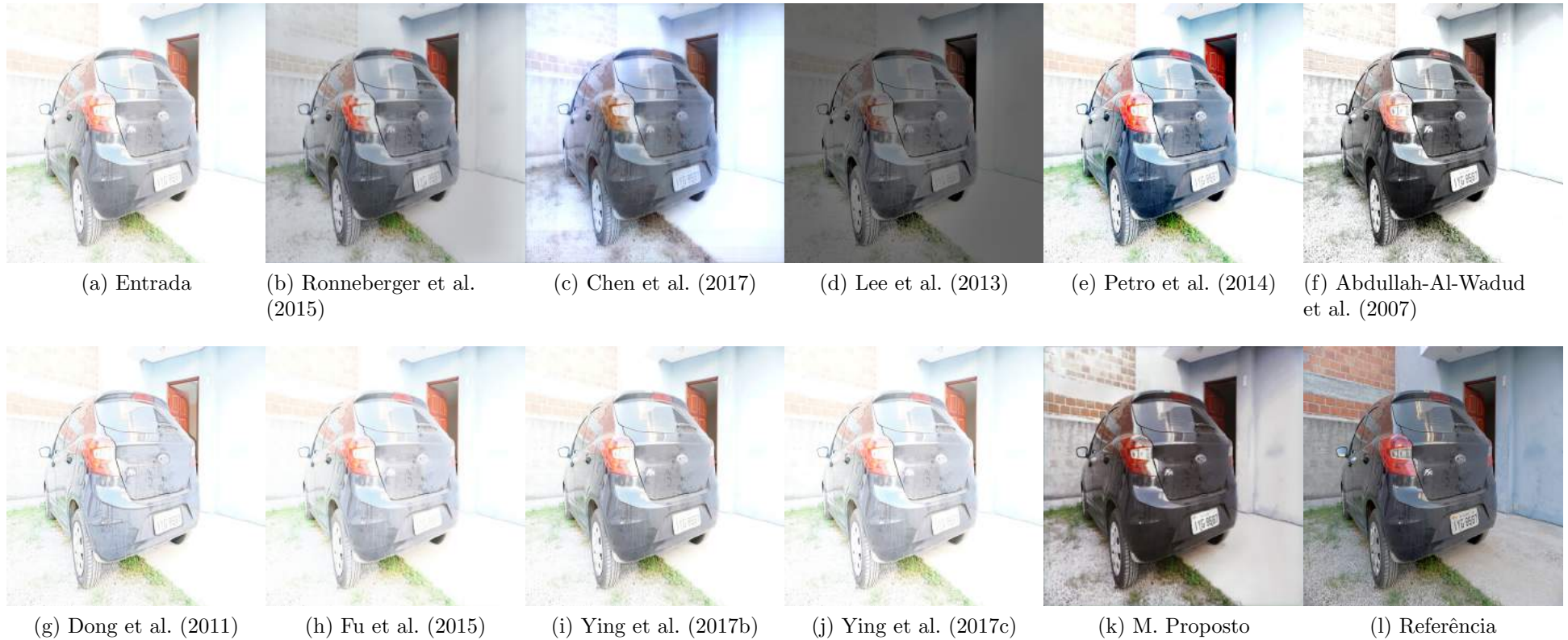


Figura 7.9: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem sobre-exposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a) reduzindo o número de pixels truncados, tornando-se mais fiel à imagem de referência.

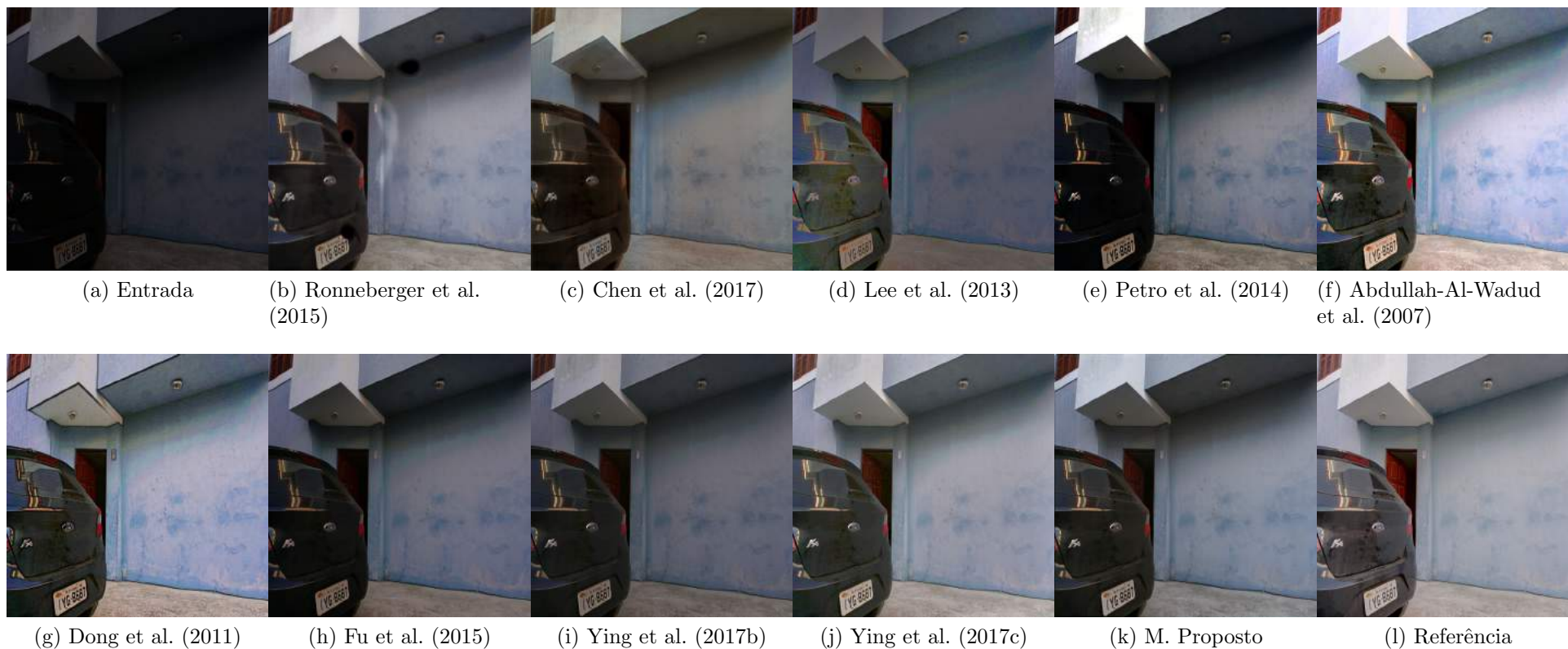
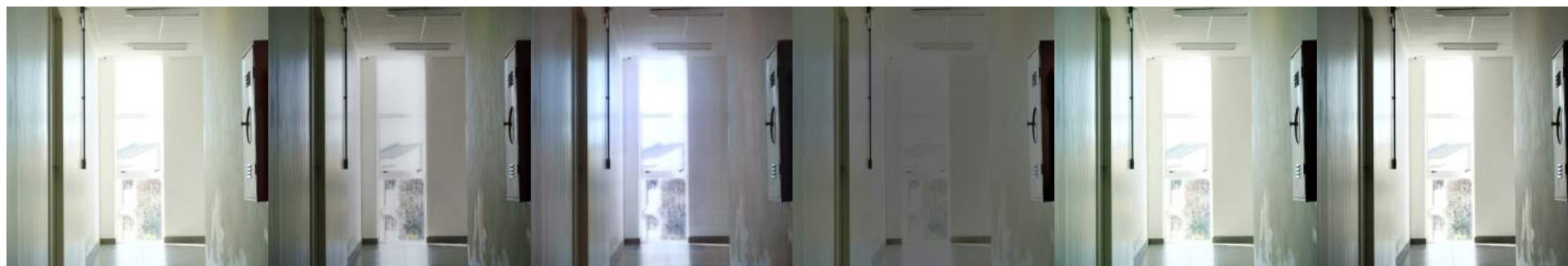


Figura 7.10: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem subexposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a) reduzindo o número de pixels truncados, tornando-se mais fiel à imagem de referência.



(a) Entrada

(b) Ronneberger et al. (2015)

(c) Chen et al. (2017)

(d) Lee et al. (2013)

(e) Petro et al. (2014)

(f) Abdullah-Al-Wadud et al. (2007)



(g) Dong et al. (2011)

(h) Fu et al. (2015)

(i) Ying et al. (2017b)

(j) Ying et al. (2017c)

(k) M. Proposto

(l) Referência

Figura 7.11: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem sobre-exposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a), reduzindo o número de pixels truncados e não denegrindo o entorno da região saturada, tornando-se mais fiel à imagem de referência.



(a) Entrada

(b) Ronneberger et al. (2015)

(c) Chen et al. (2017)

(d) Lee et al. (2013)

(e) Petro et al. (2014)

(f) Abdullah-Al-Wadud et al. (2007)



(g) Dong et al. (2011)

(h) Fu et al. (2015)

(i) Ying et al. (2017b)

(j) Ying et al. (2017c)

(k) M. Proposto

(l) Referência

Figura 7.12: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem subexposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a) reduzindo o número de pixels truncados, tornando-se a imagem com cores mais uniformes e com um maior contraste.

7.10 Dataset Cai2018

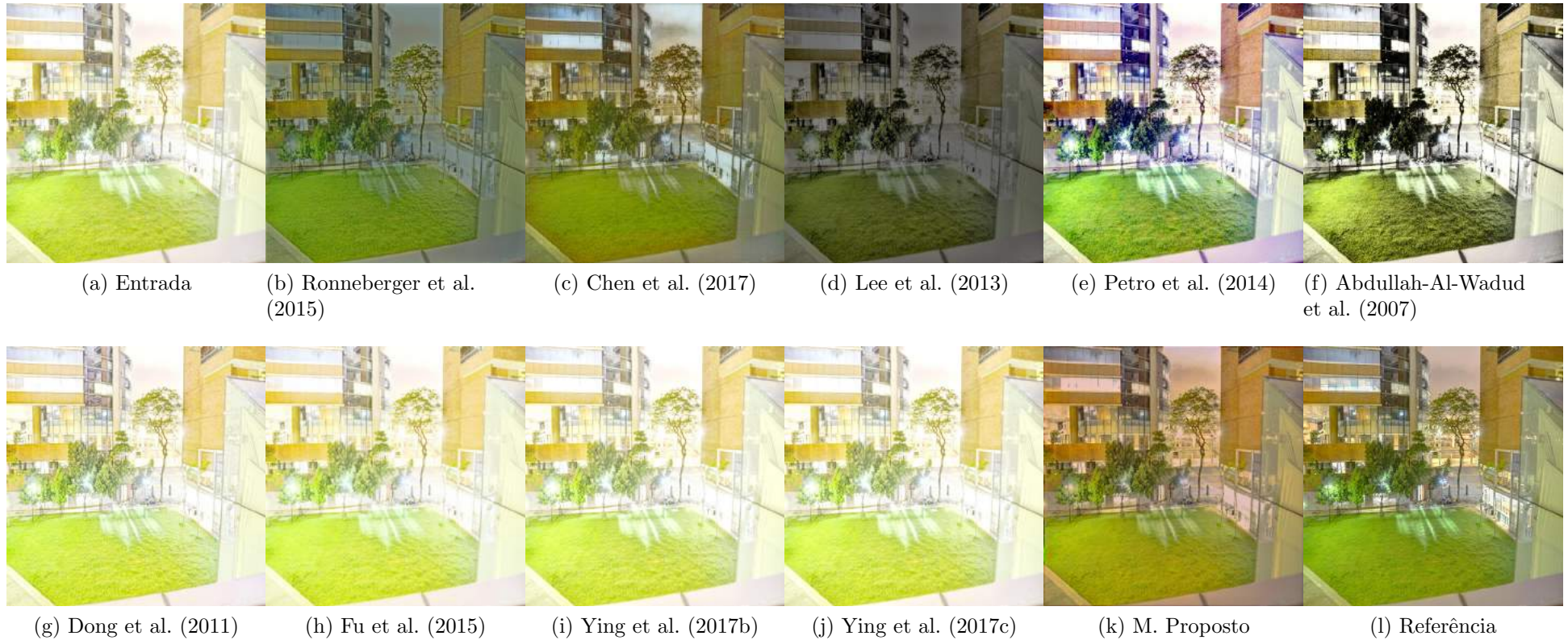


Figura 7.13: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem sobre-exposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a) reduzindo o número de pixels truncados, tornando-se mais fiel à imagem de referência.

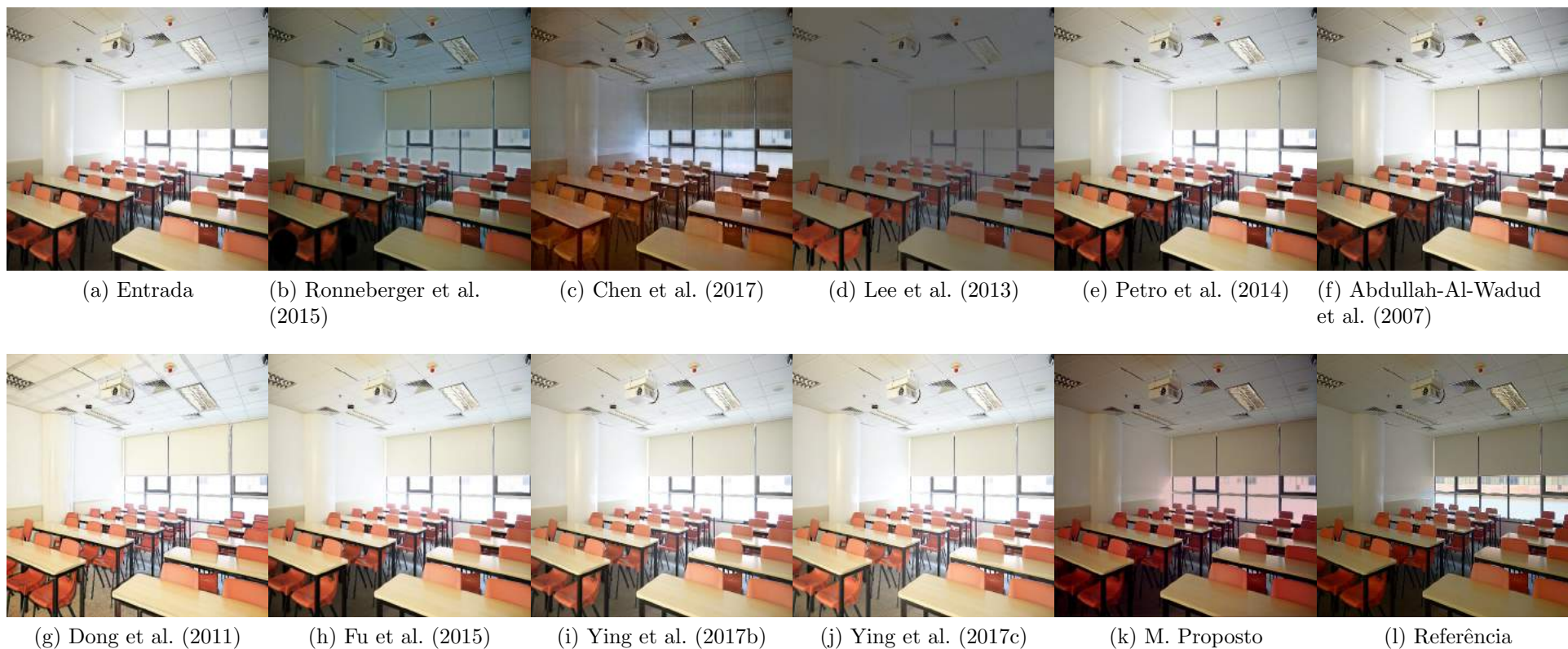


Figura 7.14: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem sobre-exposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração da imagem de entrada (a) reduzindo o número de pixels truncados, tornando-se mais fiel à imagem de referência.

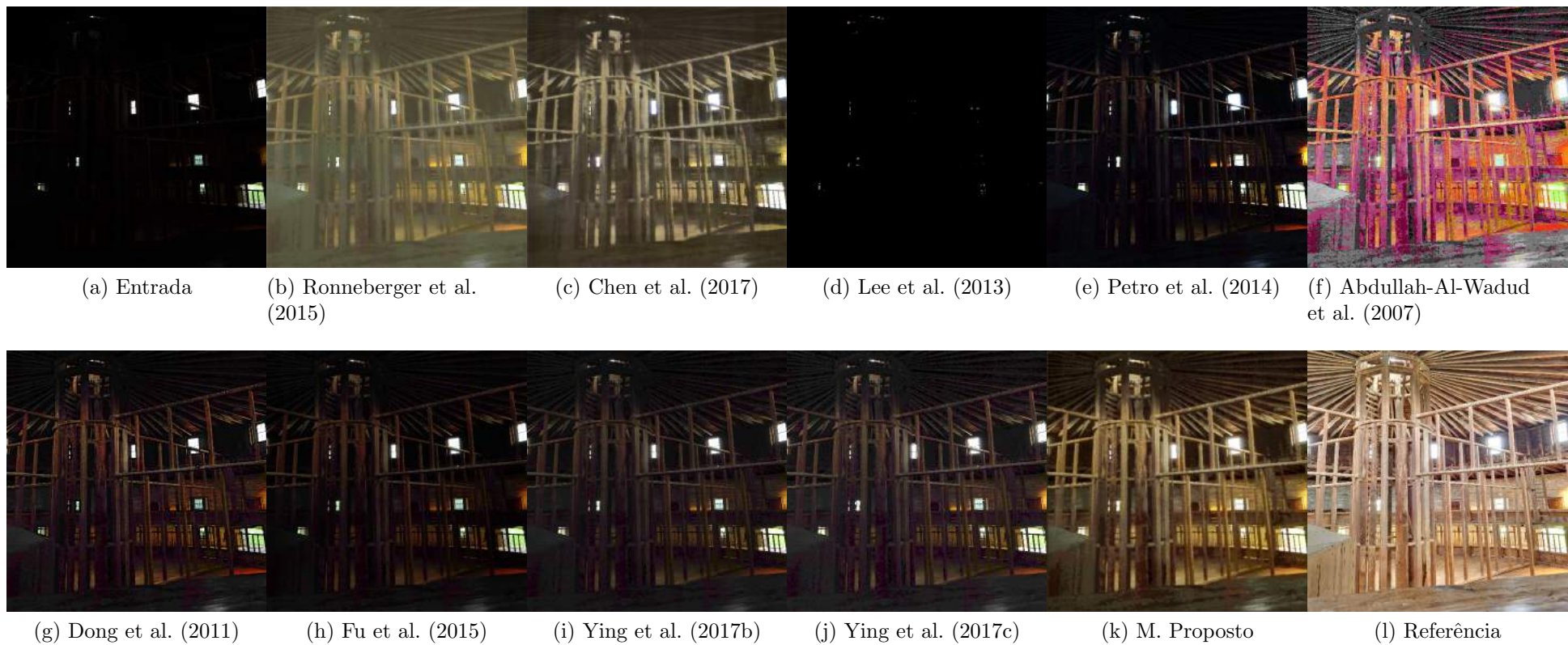


Figura 7.15: Resultados qualitativos dos métodos relacionados (b - j) em uma imagem subexposta. Nela, percebe-se que o Método Proposto (k) obteve um melhor restauração e detalhamento da imagem de entrada (a), reduzindo o número de pixels truncados e tornando-se a imagem com uma cor mais fiel à original.