



UNIVERSIDADE FEDERAL DO RIO GRANDE - FURG
CENTRO DE CIÊNCIAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM COMPUTACIONAL

Tese de Doutorado

T-Net: um modelo de rede neural de aprendizado profundo para predição de perfis de sedimentos no interior de tubulação a partir de imagens térmicas

Jovania Menezes Dias

Tese de Doutorado apresentada ao PPPGMC da Universidade Federal do Rio Grande, como requisito parcial para a obtenção do grau de Doutor em Modelagem Computacional.

Orientador: Prof. Dra. Silvia Silva da Costa Botelho

Rio Grande, 2025

Ficha Catalográfica

D541t Dias, Jovania Menezes.

T-Net: um modelo de rede neural de aprendizado profundo para predição de perfis de sedimentos no interior de tubulação a partir de imagens térmicas / Jovania Menezes Dias. – 2025.

120 f.

Tese (doutorado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-Graduação em Modelagem Computacional, Rio Grande/RS, 2025.

Orientadora: Dra. Silvia Silva da Costa Botelho.

1. Inspeção não invasiva 2. Imagem térmica 3. *Deep learning*
I. Botelho, Silvia Silva da Costa II. Título.

CDU 004

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

Jovania Menezes Dias

“T-Net: um modelo de rede neural de aprendizado profundo para predição de perfis de sedimentos no interior de tubulação a partir de imagens térmicas”

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 Doutora. Área de concentração: Modelagem Computacional.

Aprovado em: 23/05/2025

BANCA EXAMINADORA

Profa. Dra. Silvia Silva da Costa Botelho (Orientadora - FURG)



Documento assinado digitalmente
RODRIGO ANDRADE DE BEM
Data: 09/07/2025 11:25:57-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Rodrigo Andrade de Bem (FURG)



Documento assinado digitalmente
PAULO JEFFERSON DIAS DE OLIVEIRA EVALD
Data: 09/07/2025 12:49:09-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Paulo Jefferson Dias de Oliveira Evald (UFPEL)



Documento assinado digitalmente
DIANA FRANCISCA ADAMATTI
Data: 09/07/2025 10:50:00-0300
Verifique em <https://validar.iti.gov.br>

Profa Dra. Diana Francisca Adamatti (FURG)

Rio Grande - RS
2025



Documento assinado digitalmente

SILVIA SILVA DA COSTA BOTELHO

Data: 07/07/2025 09:11:41-0300

Verifique em <https://validar.it.gov.br>

Rio Grande - RS
2025

Dedico este trabalho aos meus filhos: Jonas e Gabriela, que são minha maior inspiração e força ao longo desta jornada desafiadora. Cada momento longe de vocês e cada esforço dedicado a este doutorado confirmo que com determinação, resiliência e amor, podemos alcançar nossos sonhos. Vocês são o meu maior motivo para persistir, crescer e acreditar que o futuro sempre reserva algo extraordinário. Esta conquista é tanto de vocês quanto minha. Amo vocês!

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, por me guiar e sustentar ao longo desta jornada, renovando minha fé e força nos momentos mais difíceis e iluminando cada passo do meu caminho.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro por meio da concessão da bolsa de doutorado para a realização deste trabalho.

Agradeço à minha orientadora Silvia Botelho, pelo conhecimento e orientação que foram fundamentais para a realização deste trabalho.

Agradeço também à todos que, em alguma forma, contribuíram para que este trabalho fosse possível. Aos colegas de pesquisa, amigos e familiares que me apoiaram e incentivaram ao longo dessa trajetória.

Por fim, agradeço à todos que torceram por mim e acreditaram na realização deste sonho.

RESUMO

DIAS, Jovania Menezes. **T-Net: um modelo de rede neural de aprendizado profundo para predição de perfis de sedimentos no interior de tubulação a partir de imagens térmicas**. 2025. 120 f. Tese (Doutorado) – Programa de Pós-Graduação em Modelagem Computacional. Universidade Federal do Rio Grande - FURG, Rio Grande.

A sedimentação em tubulações industriais representa um desafio crítico em diversos setores, impactando diretamente a eficiência operacional, a manutenção preditiva e a segurança dos processos industriais. Este estudo apresenta uma metodologia inovadora que integra imagens térmicas e redes neurais convolucionais para a previsão da sedimentação em tubulações industriais, combinando abordagens de aprendizado profundo e sensoriamento térmico. A principal contribuição desta pesquisa está no desenvolvimento de um modelo de CNN treinado a partir de um conjunto de dados híbrido, composto por dados experimentais e sintéticos. Com a utilização de técnicas de aumento de dados (*data augmentation*), o dataset inicial, contendo 22 amostras reais, foi expandido para 4.290 dados de entrada, possibilitando um treinamento mais robusto da rede neural. A metodologia proposta baseia-se na captura de imagens térmicas da superfície da tubulação para mapear a distribuição de temperatura e correlacioná-la com a formação de sedimentos internos. Paralelamente, a modelagem 3D dos perfis sedimentares foi realizada por meio de escaneamento 3D, permitindo a construção de um conjunto de dados detalhado para validação do modelo preditivo. Os experimentos demonstraram que o modelo baseado em imagens térmicas RGB superou aquele treinado com imagens em tons de cinza, evidenciando que a riqueza espectral das imagens térmicas coloridas melhora a precisão da predição da CNN. Essa abordagem possibilita a implementação de um sistema não invasivo e com precisão para monitoramento contínuo da sedimentação, reduzindo custos de manutenção e prevenindo falhas operacionais em tubulações industriais. Embora inicialmente aplicada à indústria de fertilizantes, a metodologia proposta apresenta potencial de aplicação em diversos setores, como indústrias petroquímicas, farmacêuticas e alimentícias, onde a detecção eficiente de sedimentos em tubulações é essencial para o controle de qualidade e otimização dos processos. Os resultados obtidos demonstram que a combinação de visão térmica e aprendizado profundo representa um avanço significativo na monitorização e predição de sedimentação, oferecendo uma solução robusta, escalável e de grande impacto para a indústria. Os resultados ressaltam o potencial desta metodologia para a gestão de tubulações industriais e outros domínios que enfrentam desafios semelhantes, oferecendo uma base para futuros avanços no monitoramento não invasivo de sedimentos e aplicações mais amplas em sistemas industriais.

Palavras-chave: Inspeção não invasiva, imagem térmica, *deep learning*.

ABSTRACT

DIAS, Jovania Menezes. **T-Net: a deep learning neural network model for predicting sediment profiles inside pipelines from thermal images**. 2025. 120 f. Thesis (Doctorate degree) – Programa de Pós-Graduação em Modelagem Computacional. Universidade Federal do Rio Grande - FURG, Rio Grande.

Sedimentation in industrial pipelines represents a critical challenge across various sectors, directly impacting operational efficiency, predictive maintenance, and process safety. This study presents an innovative methodology that integrates thermal imaging and convolutional neural networks (CNNs) to predict sedimentation in industrial pipelines, combining deep learning approaches with thermal sensing. The main contribution of this research lies in the development of a CNN model trained on a hybrid dataset composed of both experimental and synthetic data. By applying data augmentation techniques, the initial dataset, containing 22 real samples, was expanded to 4,290 input data points, enabling more robust neural network training. The proposed methodology is based on capturing thermal images of the pipeline surface to map temperature distribution and correlate it with the formation of internal sediments. In parallel, 3D modeling of sediment profiles was performed through 3D scanning, allowing the construction of a detailed dataset for validating the predictive model. The experiments demonstrated that the model trained on RGB thermal images outperformed the one trained on grayscale images, highlighting that the spectral richness of colored thermal images enhances CNN prediction accuracy. This approach enables the implementation of a non-invasive, high-precision system for continuous sedimentation monitoring, reducing maintenance costs and preventing operational failures in industrial pipelines. Although initially applied to the fertilizer industry, the proposed methodology shows potential applicability in various sectors, such as petrochemical, pharmaceutical, and food industries, where efficient sediment detection in pipelines is essential for quality control and process optimization. The results obtained demonstrate that the combination of thermal vision and deep learning represents a significant advancement in the monitoring and prediction of sedimentation, offering a robust, scalable, and high-impact solution for industry. Furthermore, the findings highlight the potential of this methodology for pipeline management and other domains facing similar challenges, providing a foundation for future advancements in non-invasive sediment monitoring and broader applications in industrial systems.

Keywords: non-invasive inspection, thermal image, deep learning.

LISTA DE FIGURAS

1	Imagem externa real da tubulação industrial. Fonte: Elaborada pela autora (2019)	15
2	Imagem térmica da tubulação industrial com acúmulo de sedimentos. Fonte: Elaborada pela autora (2019)	16
3	Esquema de radiação absorvida e emitida por um corpo negro - Imagem adaptada de (Guerra, 2010).	23
4	Representação da Lei de Fourier em um corpo cilíndrico - Imagem retirada de (Chiwiacowsky, 2002).	23
5	Retângulos concêntricos afetados por distorções do tipo “Barril” (a) e “Almofada” (b) - Imagem retirada de (Lobão, 2004).	26
6	Processo de triangulação para reconstrução 3D. Fonte: Elaborada pela autora (2021)	28
7	Arquitetura básica uma Rede Neural Convolucional (Albelwi and Mahmood, 2017).	30
8	Posicionamento da Câmera térmica para única vista	46
9	Esquema do ambiente controlado em escala. Fonte: Elaborado pela autora	67
10	Visão da configuração real para obtenção da imagem térmica. Fonte: Elaborado pela autora	67
11	Fertilizante NPK granulado. Fonte: Elaborado pela autora	68
12	Fertilizante úmido no interior do tubo.	68
13	Imagem térmica quando a tubulação está vazia. Fonte: Elaborada pela autora.	69
14	Imagem térmica da tubulação com sedimento no interior da tubulação. Fonte: Elaborada pela autora.	70
15	Amostras de sedimentos. Fonte: Elaborado pela Autora	70
16	Visão da configuração real do mapeamento do sedimento. Fonte: Elaborado pela autora	71
17	Sedimento digitalizado. Fonte: Elaborado pela autora	71
18	Reconstrução 3D do sedimento a partir de polígonos. Fonte: Elaborado pela autora	72
19	Exemplificação da relação entre o sedimento e o seu perfil. Fonte: Elaborado pela Autora	73
20	Imagem térmica da tubulação - região de interesse. Fonte: Elaborada pela autora.	73

21	Perfis dos sedimentos elaborados manualmente. Fonte: Elaborado pela Autora	76
22	Aumento de dados, fatiamento da região de interesse. Fonte: Elaborada pela autora.	76
23	Aumento de dados - Uma Fatia da imagem térmica RGB do sedimento. Fonte: Elaborada pela autora.	76
24	Aumento de dados - Uma Fatia da imagem térmica RBG do sedimento (com ruído). Fonte: Elaborada pela autora.	77
25	Aumento de dados - Uma Fatia da imagem térmica RBG do sedimento (com desfoque). Fonte: Elaborada pela autora.	77
26	Aumento de dados - Uma Fatia da imagem térmica RBG do sedimento (com ruído branco). Fonte: Elaborada pela autora.	77
27	Aumento de dados - Uma Fatia da imagem térmica RBG do sedimento (com interferencia). Fonte: Elaborada pela autora.	77
28	Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento. Fonte: Elaborada pela autora.	78
29	Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento (com ruído). Fonte: Elaborada pela autora.	78
30	Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento (com desfoque). Fonte: Elaborada pela autora.	78
31	Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento (com ruído branco). Fonte: Elaborada pela autora.	79
32	Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento (com interferência). Fonte: Elaborada pela autora.	79
33	Visão geral da metodologia proposta.	81
34	Desenho esquemático da arquitetura da rede.	82
35	Comparação entre treino e validação para RGB e tons de cinza. . . .	87
36	Coeficiente de variabilidade dos dados (R^2)	88
37	Interpolação alturas	90
38	Diferenças percentuais das interpolações	90
39	Diferenças percentuais absolutas	91
40	Comparação da interpolação das alturas da saída do processamento das imagens RBG e tons de cinza	92
41	Delta	93
42	Reconstrução 3D de diferentes fatias aleatórias de imagens RBG. . .	94
43	Reconstrução 3D de diferentes fatias aleatórias de imagens em tons de cinza.	96
44	Consolidação do fatiamento para reconstrução do sedimento inteiro .	97

LISTA DE TABELAS

1	Comparação entre diferentes funções de ativação	31
2	Comparação entre Redes de Reconstrução 3D	42
3	Comparação entre redes de reconstrução 3D (b)	45
4	Classificação de similaridade.	55
5	Resumo das técnicas de inspeção de tubulações não invasivas (a). . .	56
6	Resumo das técnicas de inspeção de tubulações não invasivas (b). . .	57
7	Classificação de similaridade utilizando imagens térmicas.	61
8	Comparação de estudos sobre termografia e reconstrução 3D	62
9	Descrição da arquitetura da rede convolucional (Encoder-Decoder) . .	83
10	Descrição da arquitetura do modelo Encoder + MLP	84
11	Descrição da arquitetura para a formação da Saída Final (Concatenação de Saída Auto Encoder e Saída MLP)	84
12	Comparação das partes da Arquitetura da T-Net	84
13	Métricas do Modelo para imagens RGB e Tons de cinza	89
14	Erros de interpolações	92
15	Comparação das métricas de desempenho entre MLP e T-Net	96
16	Limitações, impactos e estratégias de mitigação.	98

LISTA DE ABREVIATURAS E SIGLAS

3D	Tridimensional
CNN	<i>Convolutional Neural Network</i>
END	Ensaio não destrutivo
FLIR	<i>Forward Looking Infra-Red</i>
FOV	<i>Field of View</i>
HFOV	<i>Horizontal field of view</i>
iFOV	<i>Instantaneous Field Of View</i>
IR	Infravermelho
IRTI	<i>Infrared thermal imaging</i>
mFOV	<i>Minimal instantaneous field of view</i>
ML	<i>Machine Learning</i>
NPK	Nitrogênio (N), Fósforo (P) e Potássio (K),
OCR	<i>Optical Character Recognition</i>
RGB	<i>Red, Green e Blue</i>
RM	Ressonância magnética
ROI	<i>Region of interest</i>
TIV	Termografia Infra Vermelha
TTND	Teste de termografia não destrutiva
VFOV	<i>Vertical field of view</i>
VGA	<i>Video Graphics Array</i>

SUMÁRIO

1	Introdução	14
1.1	Objetivo Geral	19
1.1.1	Objetivos específicos	19
1.1.2	Contribuições	20
1.1.3	Publicações	20
2	Referencial Teórico	22
2.1	Lei da Radiação de Stefan-Boltzmann e Lei de Fourier	22
2.2	Termografia	24
2.2.1	Termografia ativa	24
2.2.2	Termografia passiva	25
2.2.3	Termografia Qualitativa	25
2.2.4	Termografia Quantitativa	25
2.2.5	Precisão	25
2.2.6	Distorção Radial	25
2.2.7	Campo de Visão	26
2.2.8	<i>Frame Rate</i>	26
2.3	Laser Scanner 3D	26
2.4	Computação Visual	29
2.4.1	Redes Neurais convolucionais	29
2.4.2	Redes de Reconstrução 3D	34
3	Trabalhos Relacionados	51
3.1	Combinação de imagem infravermelha e informações de geometria 3D	58
4	Uma metodologia para construção de <i>Dataset</i> para ajuste do modelo	65
4.1	Metodologia para a construção do <i>setup</i>	67
4.1.1	Pré-processamento dos Dados	71
4.1.2	Rotulagem dos Dados	72
4.2	Aplicação da metodologia para a obtenção do <i>dataset</i>	73
4.2.1	Aumento de dados	73
4.3	Aplicação de metodologia em um estudo de caso	74
4.3.1	<i>Dataset</i>	75
4.3.2	Rotulagem dos Dados	79
5	T-Net uma rede neural profunda para estimativa do perfil do sedimento	81
5.1	Arquitetura da T-Net	82

6	Resultados e Discussões	86
6.1	Métricas de treino vs validação - imagens RGB e Tons de cinza	87
6.1.1	Comparativo RGB x Tons de cinza	89
6.2	<i>Ground truth</i>	89
6.3	Reconstruções do perfil do sedimento	94
6.4	Limitações do modelo T-Net	97
7	Conclusão e Trabalhos Futuros	99
	Referências	102

1 INTRODUÇÃO

O direito à alimentação e a segurança alimentar são garantidos pelo artigo 25 da Declaração Universal dos Direitos Humanos de 1948 (ONU, 1948) e pela Cúpula Mundial da Alimentação de 1996 (FAO, 1996), respectivamente. Esses documentos definem que toda pessoa tem direito a alimentos nutritivos que atendam às suas necessidades para manter uma vida saudável. Neste século, a demanda por alimentos crescerá, pois está diretamente relacionada ao crescimento da população mundial (Barrett, 2020). Consequentemente, a produção de alimentos deve ser intensificada (Timsina, 2018; D’Odorico et al., 2018; Kopittke et al., 2019), e isto impacta diretamente o plantio que, para sustentar uma maior produção, precisa de solos mais ricos em nutrientes.

Nesse cenário, as indústrias de fertilizantes são demandadas para a preparação de compostos para reposição de nutrientes no solo (Ennab, 2016). Os nutrientes permitem que as plantas completem seus ciclos de vida e promovam a produtividade das culturas (Achaw and Danso-Boateng, 2021). A composição química dos nutrientes depende das necessidades do solo (Möller, 2018; Nair, 2019). Dentre os fertilizantes, um dos mais utilizados é o NPK, que é composto principalmente por Nitrogênio (N), Fósforo (P) e Potássio (K).

O processo de produção de fertilizantes utiliza um grande sistema interligado de equipamentos pesados e dutos, onde os elementos químicos circulam continuamente. Esses elementos são submetidos a vapor úmido e a vapor de amônia em um tambor rotativo chamado granulador (Nuryanto et al., 2020). Este processo produz muita poeira, que se deposita dentro dos dutos (Madanhire et al., 2015). O material acumulado é um problema para o processo produtivo, pois ao interagir com a umidade do granulador, forma-se um sedimento que bloqueia a passagem do componente fertilizante (Huang et al., 2021). O acúmulo de sedimentos deve ser removido periodicamente para preservar a continuidade da produção.

O uso de sensores, comumente utilizados ao longo da tubulação, podem detectar as diferenças de pressão no seu interior indicando a sedimentação no interior da tubulação, sem contudo, informar o perfil sedimentar, nem a localização específica desse sedimento. Além disso, o excesso de poeira e as altas temperaturas no interior da tubulação são

elementos que dificultam o uso de sensores dentro dos dutos (Dua, 2020).

Na prática, a unidade industrial interrompe sua produção para abrir os dutos e realizar a limpeza, onde um funcionário limpa manualmente as tubulações. A Figura 1 exemplifica uma tubulação industrial onde circulam os fertilizantes.



Figura 1: Imagem externa real da tubulação industrial. Fonte: Elaborada pela autora (2019)

O ambiente interno de uma tubulação industrial é confinado e inóspito, trazendo riscos à saúde dos colaboradores que precisam entrar nesses ambientes para realizar a limpeza (Cone, 2015). Os riscos podem ser exemplificados através da exposição a resíduos químicos ou gases tóxicos acumulados, como amônia ou vapores ácidos, comuns na produção de fertilizantes. Além disso, a paralisação da planta industrial impacta diretamente na rentabilidade do setor (Panchal et al., 2018; Syamila, 2020).

A proposição de técnicas capazes de identificar os locais de acúmulo de sedimentos podem minimizar o tempo de inatividade da planta, e subsidiar a tomada de decisão sobre a paralisação da produção para a realização da limpeza. Nesse contexto, uma alternativa é a utilização de técnicas de inspeção não invasiva para mapeamento do perfil de entupimento existente nas estruturas da linha de produção (Madanhire et al., 2015; Alotaibi et al., 2021). Ensaios Não Destrutivos (END) são técnicas utilizadas na inspeção de materiais e equipamentos sem danificá-los, vindo a minimizar as preocupações com a segurança, ao mesmo tempo em que reduzem os custos de manutenção (Alotaibi et al., 2021; Liang et al., 2020). Os END são uma forma de controlar a qualidade de materiais e produtos, que são utilizados em diversos setores, como petroquímico (Sharp and Ersdal, 2021), aeroespacial (Ciampa et al., 2018), naval (Mix, 2005; Shorten, 2018), siderúrgico (Groisman, 2019), entre outros. As principais técnicas de END são: emissão acústica, radiografia, líquido penetrante, partículas magnéticas, ultrassom e termografia (IET, 2020; Meng et al., 2019; Cadelano et al., 2016; Keller et al., 2015).

A termografia infravermelha (TIV) é uma técnica que utiliza radiação infravermelha para adquirir e processar informações térmicas de dispositivos de medição sem a necessidade de contato direto com o elemento analisado (de Luis-Ruiz et al., 2021). Essa aborda-

gem fornece um mapa detalhado das temperaturas superficiais dos objetos (Vollmer and Mollmann, 2018), sendo amplamente utilizada em sistemas de visão computacional baseados em imagens infravermelhas. Essa visão é obtida registrando a intensidade da luz infravermelha emitida ou refletida pela superfície observada.

O Teste de termografia não destrutiva (TTND) é uma aplicação dessa tecnologia, caracterizada por inspecionar rapidamente superfícies e identificar problemas como vazamentos, rachaduras (Dlesk et al., 2022; Ciampa et al., 2018), delaminações (Ratsakou et al., 2020) e descontinuidades internas em materiais ou objetos (Wu et al., 2011). Essa técnica é utilizada devido à sua eficiência e não invasividade.

No entanto, uma imagem térmica 2D, apesar de apresentar o perfil de temperatura planar de um objeto, não revela sua estrutura geométrica. Para superar essa limitação, avanços têm sido realizados no desenvolvimento de modelos térmicos 3D, também conhecidos como termogramas 3D, que combinam imagens RGB visíveis com dados térmicos (Muller and Kroll, 2017; Yang et al., 2018; Chen et al., 2015a; Jia et al., 2017). Uma outra abordagem é a integração de imagens térmicas com *scanners* 3D, como descrito por (Hellstein and Szwedo, 2016).

Essas técnicas permitem reconstruir em 3D, defeitos observados em objetos, oferecendo uma representação detalhada tanto da temperatura quanto da geometria. Essa combinação de dados térmicos e geométricos amplia significativamente as possibilidades de análise, sendo uma ferramenta promissora para aplicações industriais e de inspeção estrutural.

No caso da indústria de fertilizantes, uma imagem térmica externa à tubulação, poderia capturar diferentes assinaturas térmicas associadas a propagação de calor em função das obstruções de sedimentos. Este fato decorre das diferentes características físicas do ar, sedimentos e chapas metálicas. A Figura 2 mostra uma imagem térmica de uma tubulação na indústria de fertilizantes

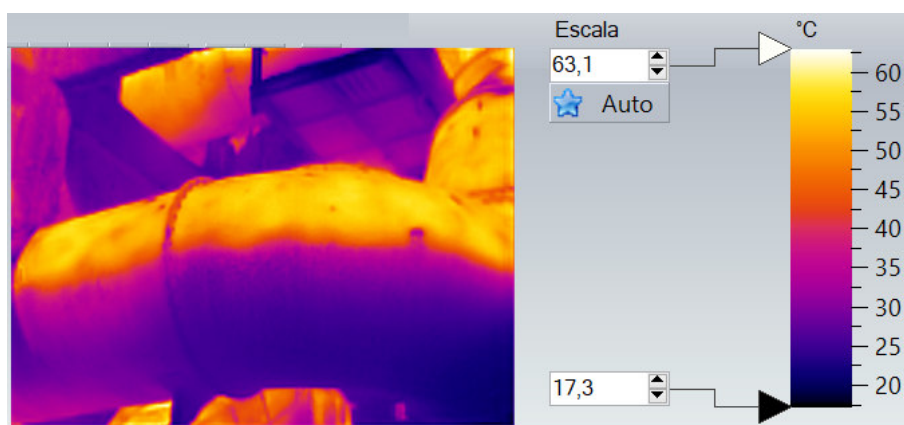


Figura 2: Imagem térmica da tubulação industrial com acúmulo de sedimentos. Fonte: Elaborada pela autora (2019)

Na escala de cores à direita da Figura 2, são representadas as variações de temperatura

ao longo da tubulação, com valores que variam entre 17,3°C (em roxo, indicando áreas mais frias) e 63,1°C (em amarelo e branco, indicando áreas mais quentes). As áreas mais quentes estão concentradas na parte superior da tubulação, onde as cores amarelo e branco predominam, indicando maior dissipação de calor ou maior fluxo térmico naquela região. As áreas mais frias, representadas pelas cores roxo e azul, estão localizadas na parte inferior da tubulação, sugerindo menor fluxo térmico ou possível acúmulo de material sedimentado, que atua como isolante térmico.

A condutividade térmica é uma propriedade dos materiais que determina a capacidade de conduzir calor. Ela quantifica a quantidade de energia térmica que pode ser transmitida através de um material, em função da sua área, da espessura e da diferença de temperatura entre suas faces (Incropera et al., 2002). Materiais com alta condutividade, como os metais, permitem uma rápida dispersão do calor, enquanto aqueles com baixa condutividade, como os isolantes térmicos, retêm a energia e diminuem a transferência de calor (Incropera et al., 2002; Carslaw and Jaeger, 1959; Callister, 2007; Holman, 2010). Essa característica é importante em diversas aplicações, desde o desenvolvimento de sistemas de isolamento e resfriamento em edifícios e dispositivos eletrônicos até a análise de processos industriais, onde variações na condutividade podem indicar diferenças na composição ou na presença de depósitos, como sedimentos em tubulações.

Neste trabalho, parte-se da hipótese de que as diferenças na condutividade térmica entre o fertilizante e os demais materiais presentes no interior da tubulação permitem a extração de um modelo computacional capaz de estimar o perfil de acúmulo de sedimentos a partir de imagens térmicas. Em outras palavras, ao capturar a assinatura térmica da tubulação, é possível identificar as variações de temperatura decorrentes da diferença de condutividade do fertilizante em comparação ao aço e ao ar, possibilitando assim uma melhor estimativa dos perfis de entupimento. Neste cenário a indústria poderia ter melhor estimativa dos perfis de entupimentos nas tubulações.

Assumindo a necessidade de aumentar a produção de alimentos, a intensificação da produção de fertilizantes é natural. Assim, os problemas de tubulações entupidadas devem ser minimizados para maximizar a produção de fertilizantes. Considerando a possibilidade de investigar a viabilidade do uso de imagens térmicas como forma de avaliar as condições internas dessas estruturas, surge a seguinte questão: como o perfil do sedimento pode ser inferido através de uma percepção externa?

Abordagens de mapeamento de imagens térmicas em perfis físico-estruturais de superfície, baseadas na modelagem do fenômeno de propagação de calor nas estruturas da linha de produção (Maldague et al., 1988; Usamentiaga et al., 2014), e sua captura por uma câmera (de Luis-Ruiz et al., 2021; Grechi et al., 2021), são uma alternativa para responder a essas questões. Simultaneamente, o uso de técnicas de aprendizado de máquina também é uma solução viável, pois esses algoritmos estimam padrões usando dados para aprender (Sarker, 2021b).

Dentre os métodos de aprendizado de máquina para processamento de imagens, as técnicas baseadas em aprendizado profundo são atualmente o estado da arte (Sarker, 2021a; Gao et al., 2020; Jain et al., 2020; Zheng et al., 2021; Bai et al., 2024; Lee, 2024; Ilett et al., 2023; Wang et al., 2023a; Liao et al., 2023; Sun et al., 2024; Reddy et al., 2024). Os métodos de aprendizado profundo são frequentemente baseados em redes neurais artificiais dispostas em estruturas convolucionais, cujos parâmetros permitem mapear relações complexas entre dados de entrada (*dataset*) e saída, que são previamente apresentados ao sistema.

A ideia é utilizar um conjunto de dados com perfil de sedimento diversificado como entrada para um algoritmo de aprendizado profundo, capaz de mapear imagem térmica em nível de entupimento. No entanto, não há conjunto de dados públicos de imagens térmicas de linhas de produção de fertilizantes. Dada a dificuldade de acesso ao interior da tubulação e ao quantitativo de amostras necessárias para o treinamento dos modelos de aprendizado profundo, se obter um *dataset* adequado à solução deste problema estabelece-se como um grande desafio ao seu uso. O uso de simuladores para a geração de dados sintéticos (Raghunathan, 2021; Emam, 2020), maquetes (Bao et al., 2020) e algoritmos de aumento de dados (Shorten and Khoshgoftaar, 2019) podem se fazer necessários.

Além disso, surge o desafio de como representar os diferentes perfis de entupimento (saída do modelo). Nesse contexto, nuvens de pontos podem ser utilizadas para representar volumes em um espaço tridimensional (Metzer et al., 2022), com o uso de lasers 3D para anotar pares de dados que correlacionem o entupimento da tubulação com as imagens capturadas. Entretanto, o acesso ao interior das tubulações apresenta desafios significativos, e a interrupção frequente da produção para coletar dados internos não é uma solução viável, limitando a possibilidade de se obter um conjunto de dados *in loco*.

Para contornar essas limitações, propõe-se o uso de um ambiente em escala reduzida, uma maquete, para simular as condições reais de um ambiente industrial. Esses dados sintéticos obtidos em ambiente controlado, podem servir como base para o desenvolvimento de um modelo baseado em aprendizado profundo, apresentando similaridades com o fenômeno real (Nikolenko, 2021), o que os tornariam adequados para treinar modelos capazes de generalizar para situações de ambientes relevantes de fábrica.

Este trabalho propõe um método de inspeção não invasivo que visa estimar o perfil do acúmulo de sedimentos em tubulações industriais, um problema relevante tanto para a manutenção preventiva quanto para a eficiência operacional. A abordagem utiliza uma combinação de imagens térmicas e dados gerados por scanner a laser 3D para capturar informações sobre as variações de temperatura e a geometria interna dos sedimentos. Esses dados heterogêneos são então processados por meio de algoritmos de reconstrução de nuvem de pontos tridimensional, implementados com técnicas de aprendizado profundo, para gerar um modelo computacional do perfil de sedimentos. Assim, o método

transforma dados visuais e espaciais complexos em informações quantitativas que podem indicar a presença e a extensão de obstruções sem a necessidade de intervenções diretas.

Do ponto de vista computacional, o desafio central reside na integração e interpretação de dados provenientes de diferentes fontes, cada uma com suas características específicas, como a resolução e sensibilidade das imagens térmicas e a precisão da nuvem de pontos obtida do *scanners* a laser. Entre os principais trabalhos na área, destacam-se estudos que aplicam termografia infravermelha para monitoramento industrial (Alotaibi et al., 2021; Alfredo Osornio-Rios et al., 2019; Cadelano et al., 2016; Ishikawa et al., 2022), reconstrução 3D de ambientes com sensores ópticos (Wang et al., 2023b; Ma et al., 2023) e técnicas de aprendizado profundo para processamento de imagens complexas (Dosovitskiy et al., 2021). Os desafios incluem a necessidade de pré-processamento e alinhamento rigoroso dos dados, a realização de interpolação precisa para padronizar as nuvens de pontos e o desenvolvimento de modelos que consigam aprender padrões complexos a partir de datasets muitas vezes limitados (Rangel and Soldan, 2014; Vidas et al., 2015; Yang et al., 2018).

O estudo de caso apresentado concentra-se na aplicação deste método na indústria de fertilizantes, onde o acúmulo de sedimentos nas tubulações representa um problema recorrente e crítico para a eficiência do processo produtivo. Contudo, a metodologia proposta é versátil e pode ser adaptada para outras situações de entupimento de tubulações em diferentes setores industriais. Esta proposta insere-se no âmbito do Programa MAI/DAI CNPq, Doutorado Acadêmico em Inovação, o qual busca desenvolver inovação para o setor produtivo brasileiro.

1.1 Objetivo Geral

Desenvolver um modelo computacional capaz de prever, a partir de imagens térmicas 2D, o perfil de sedimentação de fertilizantes no interior de tubulações industriais, possibilitando a representação da distribuição espacial do material acumulado ao longo da tubulação.

1.1.1 Objetivos específicos

- Realizar uma análise crítica do estado da arte em construção de imagem 3D a partir de uma única vista de imagem 2D, com o objetivo de identificar lacunas existentes na literatura e potenciais oportunidades de contribuição científica.
- Propor um método para a geração de um conjunto de dados (*dataset*) para ajuste de modelo baseado em aprendizado de máquina, a partir de uma maquete física que simula, em ambiente controlado, as condições operacionais de uma tubulação industrial com sedimentos.

- Propor um modelo baseado em aprendizado profundo para a reconstrução do perfil do sedimento interno à tubulação a partir da imagem térmica.

1.1.2 Contribuições

As principais contribuições deste trabalho estão listadas a seguir:

- Conjunto de dados público composto por imagens térmicas sintéticas obtidas a partir de uma maquete de tubulações de fertilizantes, capturadas com câmera infravermelha. Cada imagem está associada ao respectivo perfil de sedimentos representado em arquivos de nuvem de pontos. O código da rede neural está disponível em: <https://encurtador.com.br/SBmSj>. O *dataset* está disponível em: <https://encurtador.com.br/Z0PJj>
- Modelo de aprendizado profundo para a predição de perfis de sedimentos a partir de imagens térmicas 2D. O modelo, treinado com dados previamente rotulados, é capaz de identificar variações nos perfis de sedimentação. A metodologia proposta caracteriza-se por ser não invasiva, permitindo a inspeção do perfil de sedimentos no interior de tubulações sem a necessidade de intervenções físicas ou métodos intrusivos, preservando a integridade do sistema industrial.

1.1.3 Publicações

As principais publicações deste trabalho estão listadas a seguir:

Primeira autora

- DIAS, Jovania M; TEIXEIRA, Thiago; EVALD, Paulo J. D. O.; BOTELHO, Silvia S. C.. Prediction of Fertilizer Sedimentation in Pipelines with Thermal Images and Deep Learning. IEEE Transactions on Industrial Informatics - Manuscript No. TII-24-7072. Artigo em revisão.
- DIAS, Jovania M; TEIXEIRA, Thiago; EVALD, Paulo J. D. O.; BOTELHO, Silvia S. C. Um *dataset* de imagens térmicas e nuvens de pontos para análise de sedimentação em tubulações industriais. Revista Junior de Iniciação Científica em Ciências Exatas e Engenharia, v. 1, n. 25, 2025.
Disponível em: <http://sma.c3.furg.br/icceeg/index.php/ICCEeg/article/view/33/5>. Acesso em: 10/02/2025.
- DIAS, Jovania M; TEIXEIRA, Thiago; EVALD, Paulo J. D. O.; BOTELHO, Silvia S. C. Monitoramento inteligente não invasivo de bloqueio em dutos de produção de fertilizantes. Revista Junior de Iniciação Científica em Ciências Exatas e Engenharia. Aceito para publicação

- DIAS, Jovania M.; EVALD, Paulo. J. D.O.; GUTHES, Rafael; DUARTE, Marta A; DREWS, Paulo. L. J.; BOTELHOS, Silvia S.C.. "A non-invasive learning-based method for pipeline overhaul on fertilizer production plants," IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society, 2022, pp. 1-6, doi: 10.1109/IECON49645.2022.9968895.

Co-autora

- DUARTE, Marta ; COCH, Victor ; DIAS, Jovania ; BOTELHO, Silvia ; DUARTE, Nelson ; DREWS, Paulo . Thermographic Non-Invasive Inspection Modelling of Fertilizer Pipelines Using Neural Networks. In: 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2020, Recife/Porto de Galinhas. 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2020. p. 280.
- DUARTE, Marta ; COCH, Victor ; DIAS, Jovania ; BOTELHO, Silvia ; DUARTE, Nelson ; DREWS, Paulo . Non-Invasive Visual Computing Modelling for Obstruction Inspection of Pipes in the Fertilizer Industry. In: 2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE), 2019, Rio Grande. 2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE), 2019. p. 281.

Este trabalho está organizado em sete capítulos. O capítulo 2 apresenta os principais conceitos de termografia e redes neurais convolucionais, conceitos para o melhor entendimento deste trabalho. Em seguida o capítulo 3 discute os trabalhos relacionados com inspeção não invasiva em tubulações industriais e os avanços da termografia 3D para a indústria. Capítulo 4 apresenta a Metodologia para a construção do conjunto de dados. A arquitetura da rede é apresentada no capítulo 5, os resultados são mostrados no capítulo 6. Por fim, no Capítulo 7, são apresentadas as considerações finais deste estudo, bem como uma discussão sobre os trabalhos futuros.

2 REFERENCIAL TEÓRICO

O referencial teórico fundamenta a pesquisa ao abordar conceitos essenciais para a identificação de sedimentos em tubulações industriais por meio de termografia e aprendizado profundo. O modelo computacional proposto tem como finalidade prever o perfil tridimensional de sedimentos acumulados em tubulações industriais a partir de imagens térmicas bidimensionais da tubulação. Para atingir tal finalidade, empregam-se técnicas avançadas de aprendizado profundo, com ênfase no uso de redes neurais convolucionais para extrair e analisar características que representam o gradiente de temperatura das imagens térmicas. O modelo é treinado para aprender a relação entre as propriedades térmicas e o perfil de sedimentos, cuja referência (*ground-truth*) é obtida a partir de dados gerados pelo processo de escaneamento do sedimento.

O presente capítulo apresenta a fundamentação teórica dos principais tópicos pesquisados durante o desenvolvimento deste trabalho. Este capítulo discute os princípios físicos da radiação térmica, os métodos de captura e análise de imagens térmicas, as redes neurais convolucionais aplicadas à reconstrução 3D e as principais abordagens para manutenção preditiva em ambientes industriais.

2.1 Lei da Radiação de Stefan-Boltzmann e Lei de Fourier

A lei de radiação de Stefan-Boltzmann descreve a radiação de corpos (Guerra, 2010). De uma maneira geral, os corpos constantemente emitem e absorvem radiação, não existindo uma forma universal de descrever um padrão de emissão para um corpo qualquer. Todavia, o corpo negro apresenta um padrão conhecido, absorvendo toda a energia que nele é incidida. Na Figura 3, é representado a radiação do corpo negro (Okuno and Yoshimura, 2016).

O conceito do Corpo Negro é um corpo ideal que absorve toda a radiação térmica incidente. É, portanto, um absorvedor perfeito.

A Figura 3 mostra uma representação bidimensional de um corpo absorvendo e emitindo radiação. As setas azuis indicam a radiação absorvida pelo corpo, que é o processo pelo qual a energia radiante é capturada pela superfície. As setas vermelhas representam

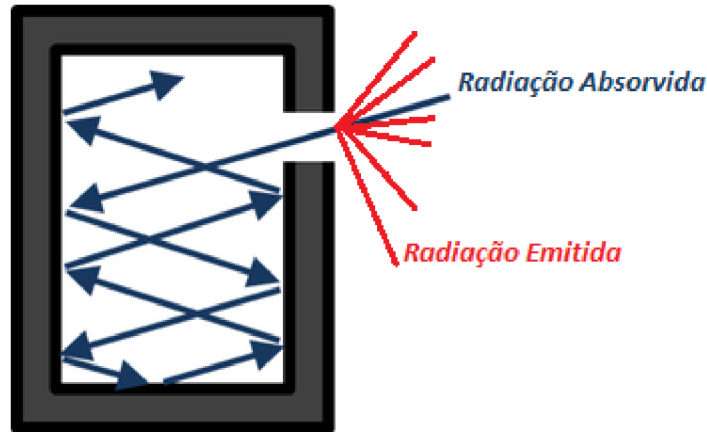


Figura 3: Esquema de radiação absorvida e emitida por um corpo negro - Imagem adaptada de (Guerra, 2010).

a radiação emitida, que ocorre quando o corpo irradia energia térmica para o ambiente externo. Essa imagem é frequentemente usada para descrever o comportamento de um corpo negro, que é um idealizado objeto físico que absorve toda a radiação incidente sem refletir ou transmitir. No contexto de termografia ou sensores infravermelhos, este conceito ajuda a entender como a energia térmica é emitida e medida. Para um corpo negro, a emissividade é igual a 1, ou seja, toda a potência irradiada é igual à absorvida. Os chamados corpos cinzentos têm uma emissividade menor do que 1 e maior do que 0, pois não irradiam toda a energia absorvida e transformam o restante em outras formas de energia: ondas de som, fricção, criação de campos magnéticos e polarização (Kennedy, 2011).

De acordo com Joseph Fourier, a temperatura de um corpo varia linearmente por toda a sua extensão quando este apresenta uma região mais quente que a outra, e o fluxo de calor é proporcional à área da seção transversal do corpo e inversamente proporcional ao seu comprimento (Chiwiacowsky, 2002), como mostrado na Figura 4.

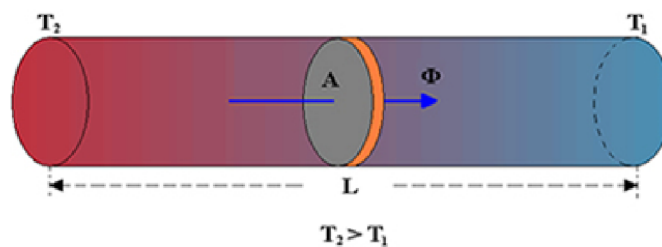


Figura 4: Representação da Lei de Fourier em um corpo cilíndrico - Imagem retirada de (Chiwiacowsky, 2002).

Assim, a lei de Fourier pode ser expressa como

$$\phi = \left(\frac{kA\Delta T}{L} \right), \quad (1)$$

onde ϕ representa o fluxo de calor, k é a condutividade térmica do material, A é a área de seção transversal, ΔT diferença de temperatura e L é a espessura do cilindro.

2.2 Termografia

A imagem térmica 2D, também chamada de termograma, descreve a distribuição de temperatura da superfície de um objeto. O dispositivo que produz a imagem térmica é a câmera térmica, câmera termográfica ou câmera infravermelha. Todas estas nomenclaturas são utilizadas na literatura (Xu et al., 2014; Swamidoss et al., 2021b,a).

A técnica da termografia é utilizada em diversas aplicações, principalmente devido ao seu baixo custo. Sua aplicação na indústria ajuda a melhorar a qualidade de produtos (Feng et al., 2025), na segurança dos colaboradores (Khor et al., 2024b,a) e no aumento da produtividade (Balakrishnan et al., 2022). A técnica de IRTI (*Infrared Thermal Imaging*), que consiste em captação de imagens termográficas, pode ser capaz de identificar falhas em instalações elétricas e mecânicas (Bagavathiappan et al., 2013), sendo utilizadas em aplicações como: monitoramento de estruturas civis (Santos, 2018), componentes elétricos e eletrônicos (Bagavathiappan et al., 2013), deformação, corrosão e soldagem, bem como na inspeção de máquinas, nas áreas nuclear e indústrias aeroespaciais (Santhi, 2017), medicina (Jones, 1998; Kennedy et al., 2009; Dey et al., 2017), agricultura e meio ambiente (Ishimwe et al., 2014; Vadivambal and Jayas, 2011; Lee et al., 2017), cidades inteligentes (Al-Kassir et al., 2005), defesa e de segurança (Bhowmik et al., 2011). A termografia pode ser dividida em termografia ativa e termografia passiva.

2.2.1 Termografia ativa

A termografia ativa utiliza-se de estímulos térmicos externos para gerar contraste térmico no objeto monitorado, destacando regiões de interesse que não são perceptíveis à temperatura ambiente. Essa técnica requer o monitoramento temporal da variação de temperatura na superfície do objeto, permitindo identificar anomalias ou características específicas do material (Bagavathiappan et al., 2013). Por exemplo, ao aquecer uma tubulação com um fluxo de ar quente, a termografia ativa pode identificar áreas com diferentes níveis de sedimentação, pois estas afetam a condução de calor e geram padrões térmicos distintos.

A termografia ativa pode ser subdividida em diferentes técnicas, tais como: Termografia de pulso, *lock-in*, fase de pulso, aquecimento por passo. A termografia de pulso consiste na aplicação de um breve pulso térmico, seguido pela análise da dissipação de calor ao longo do tempo, sendo frequentemente utilizada para detectar delaminações em compósitos. A termografia *lock-in* emprega uma fonte de calor modulada periodicamente, o que a torna particularmente adequada para a inspeção de materiais com espessuras uniformes, como os metais. Por sua vez, a termografia de fase de pulso, similar à técnica de pulso, diferencia-se ao analisar a fase do sinal térmico, proporcionando maior precisão na identificação de defeitos. Por fim, a termografia de aquecimento por passo aplica um estímulo térmico gradual, permitindo a observação de variações térmicas ao longo de um

período mais extenso, sendo comumente empregada na detecção de vazamentos em dutos

2.2.2 Termografia passiva

A termografia passiva não utiliza estímulos térmicos externos (Bagavathiappan et al., 2013) e depende exclusivamente das emissões térmicas naturais do objeto monitorado. Essa abordagem é mais adequada para situações onde o objeto já está em um estado de desequilíbrio térmico devido ao seu próprio funcionamento.

Em relação aos valores de temperatura e aos perfis térmicos, ainda podemos mencionar a termografia qualitativa e quantitativa.

2.2.3 Termografia Qualitativa

A termografia é chamada de qualitativa quando o importante é o perfil térmico e não os valores de temperatura. Isto significa que em um processo de termografia qualitativa o importante é a visualização da zona de interesse e não a escala de temperatura gerada (Santos, 2018). Com base nessa característica, através da leitura qualitativa, pode-se identificar a diferença de temperatura entre os pontos de um equipamentos de forma eficaz, podendo ser útil no diagnóstico de problemas como deficiências de isolamento térmicos, problemas em mancais e transmissões mecânicas, falhas em isolamentos de cabos elétricos, localização de entupimentos em tubulações e outros itens que podem ser reconhecidos através da diferença de temperatura (Nogueira and Reis, 2010).

2.2.4 Termografia Quantitativa

O método quantitativo é, geralmente, o segundo a ser aplicado, pois de maneira geral, o primeiro utilizado é a análise qualitativa. Através do método quantitativo pode-se definir o nível de gravidade da anomalia gerada pelo aquecimento do material (Santos, 2018).

2.2.5 Precisão

Precisão é o grau de concordância de uma medição realizada diversas vezes em condições de repetibilidade, ou seja, mantendo o mesmo sistema de medição, o local onde as medições são realizadas, mesmos procedimentos, operadores e sistemas de medição ou então em reprodutibilidade. A precisão é uma medida de dispersão e geralmente significa um desvio padrão, variância ou coeficiente de variação. A precisão está ligada a um erro aleatório.

2.2.6 Distorção Radial

Atualmente, os dispositivos para obtenção de imagens térmicas são capazes de capturar imagens com grande resolução espacial (Chen et al., 2015b). Porém, fatores como distorção geométrica são significativos no momento da captura de uma imagem. A distorção radial é uma característica encontrada na grande maioria das câmeras comer-

ciais, seu efeito em forma de "barril" e "almofada" (Lobão, 2004), mostrado na Figura 5.

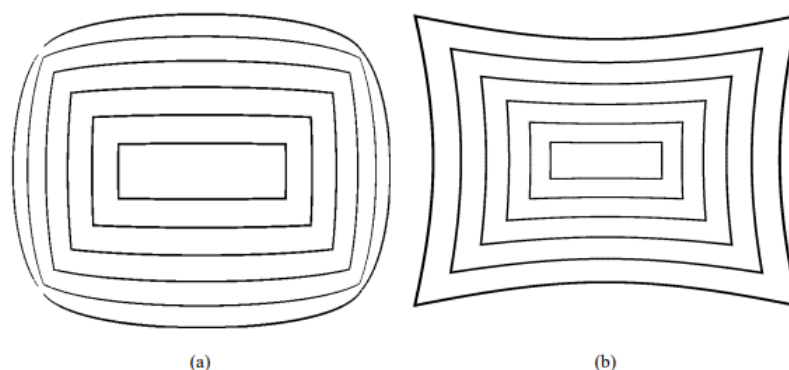


Figura 5: Retângulos concêntricos afetados por distorções do tipo “Barril” (a) e “Almofada” (b) - Imagem retirada de (Lobão, 2004).

2.2.7 Campo de Visão

Outra característica que deve ser levada em conta para especificação de um termo-visor é seu campo de visão, pois através dessa informação, é possível identificar qual o menor objeto que pode ser medido, além de informações sobre qual a distância máxima de medição para o objeto. O campo de visão (FOV) é dividido basicamente em dois: campo de visão instantâneo (iFOV), que apresenta qual a resolução espacial da câmera termográfica, ou seja, quanto menor o iFOV mais distante do objeto, e o campo de visão de medição (mFOV), que determina qual a distância máxima para o objeto a ser medido (Trindade et al., 2012).

2.2.8 Frame Rate

O *frame rate* trata-se de qual a taxa de frequência de quadro da imagem gerada no termo-visor. Câmeras termográficas operam na faixa de 25 Hz e 60 Hz, com maior aplicação em análises de objetos dinâmicos e para gravação de vídeos (Shih et al., 2009).

Além das metodologias aplicadas a câmera termográfica, será necessário realizar a digitalização 3D a laser do perfil de sedimento construído. Na seção a seguir, será apresentado metodologias para tal procedimento de reconstrução 3D.

2.3 Laser Scanner 3D

No setor industrial, a varredura a laser tem sido amplamente utilizada para engenharia e inspeção reversa, tais como: motores, eletrônicos, moldes, engenharia, análise e inspeção de superfície (Lee et al., 2017; Abdel, 2011; Chang et al., 2015a; Pfeifer and Briesse, 2007).

Um *scanner* a laser 3D é um dispositivo capaz de analisar objetos ou ambientes no

mundo real, coletando dados sobre sua forma e até sua cor. Com os dados coletados é possível construir modelos digitais tridimensionais (Abdel, 2011). Tal dispositivo, assim como o *scanner* a laser portátil, pode ser descrito como uma estação total móvel que mede todos os pontos horizontais e verticais de acordo com a direção controlada dos feixes de laser. Para cada ponto medido, sua distância e os ângulos horizontal e vertical em relação ao *scanner* são gravados (Abdel, 2011), sendo possível realizar facilmente o cálculo de coordenadas do espaço em relação à posição do *scanner*.

A tecnologia de varredura a laser 3D é usada principalmente para identificar e modelar a superfície do objeto em estudo e é considerada uma forma rápida de aquisição de dados (Paulus et al., 2014; Pfeifer and Briesse, 2007). A técnica de digitalização 3D não invasiva tem duas abordagens diferentes para a digitalização de um objeto: a técnica ativa e a técnica passiva. Um *scanner* é definido como passivo quando não emite radiação, ou seja, gera uma dependência direta de captação da radiação distribuída no ambiente (Abdel, 2011). Já a técnica ativa emite radiação ou luz e através da detecção, realiza a digitalização do objeto ou ambiente (Abdel, 2011).

Os *scanners* 3D portáteis usam o sistema de triangulação. Na triangulação de sistemas, as coordenadas são obtidas projetando um laser na superfície de um objeto e o feixe refletido é observado pelo campo de visão da câmera acoplada no equipamento. Esta técnica é chamada de triangulação porque o ponto de laser que atinge o objeto, a câmera e o emissor de laser formam um triângulo (Abdel, 2011). Segundo (França et al., 2005), o processo de aquisição 3D por triangulação é feito através da projeção de um feixe de linha vertical produzida por um emissor laser sobre a superfície do objeto a ser capturado. Essa projeção é então capturada pela câmera e a distância até o objeto se dá através de cálculo geométrico.

Os *scanners* baseados no princípio da triangulação possuem uma fonte laser e, no mínimo, um sensor de deslocamento a laser, que captura a luz laser refletida pelo objeto, como mostrado na Figura 6. A geometria entre o laser e o sensor de deslocamento a laser é conhecida e permite o cálculo das coordenadas 3D dos pontos iluminados pela luz laser por intersecção. Este tipo de sistema é adequado para a reconstrução de objetos pequenos e a distância máxima de operação dependerá da base empregada pelo *scanner*, que normalmente é fixa, à uma distância máxima de operação para a qual a precisão é satisfatória.

Ainda, segundo (França et al., 2005), o centro das lentes está localizado na origem com distância focal f até o plano da imagem, sendo a distância entre o projetor e a câmera representada pela linha de base b . A projeção do laser forma o ângulo Θ em relação à linha base e o ponto 3D (x, y, z) no espaço real é projetado no *pixel* da imagem (x', y') . O valor de θ é conhecido pelo sistema e a distância focal f é determinada pela calibração da câmera. Já a linha de base b é calculada posicionando-se o laser de maneira que a sua projeção incida no centro do objeto a ser adquirido, apontando para um objeto a uma

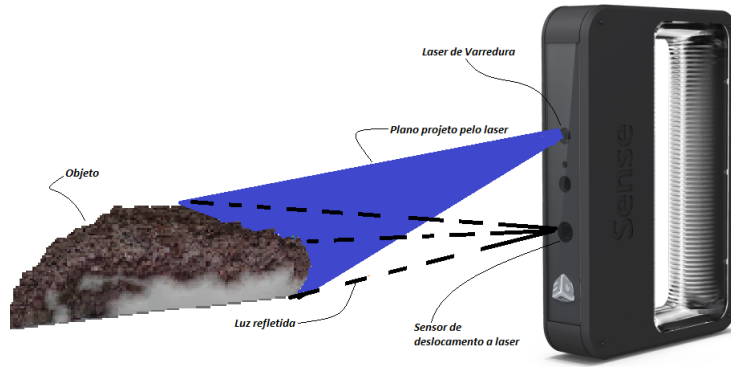


Figura 6: Processo de triangulação para reconstrução 3D. Fonte: Elaborada pela autora (2021)

distância conhecida z_0 .

$$b = \frac{z_0}{\tan \theta}. \quad (2)$$

Um ponto (x, y, z) é então projetado na câmera no espaço (x', y') de forma que

$$x' = \frac{xf}{z}, \quad (3)$$

e

$$y' = \frac{yf}{z}. \quad (4)$$

Inicialmente é feita uma varredura girando-se o laser que estará projetando uma linha vertical, capturando as imagens, armazenando os respectivos valores de θ obtidos pelo sinal de varredura de cada imagem. Determina-se os valores de x' e y' onde o laser foi refletido para cada imagem, e, então, os pontos nas coordenadas reais são calculados da seguinte forma,

$$x = \frac{bx'}{f \cot \theta - x'}, \quad (5)$$

$$y = \frac{by'}{f \cot \theta - x'}, \quad (6)$$

$$z = \frac{bf'}{f \cot \theta - x'}. \quad (7)$$

A principal vantagem de utilizar técnicas de luz estruturadas como essa é seu baixo custo (França et al., 2005).

2.4 Computação Visual

A Computação Visual abrange diferentes abordagens relacionadas à aquisição, processamento e análise de dados visuais, englobando os campos de Visão Computacional, Processamento de Imagens e Computação Gráfica. Esta tecnologia tem como característica interpretar e analisar informações visuais por meio de algoritmos computacionais, permitindo a automação de tarefas que envolvem imagens e vídeos (Posada et al., 2015). De acordo com (Szeliski, 2022), essa área observa imagens digitais e busca reconstruir e compreender suas características, como forma, iluminação e distribuição de cores, utilizando métodos avançados baseados em aprendizado de máquina. Essa tecnologia pode ser utilizada em diferentes aplicações, tais como: reconhecimento de caracteres com sistemas de OCR para extração automática de texto em imagens (Gralinski et al., 2020; Oral et al., 2020; Memon et al., 2020). Reconhecimento de objetos utilizado em sistemas autônomos para identificar diferentes objetos em ambientes dinâmicos (Li et al., 2022; Oliva and Torralba, 2007). Estudo de imagens médicas através de ferramentas de segmentação e classificação para diagnóstico assistido (Ronneberger et al., 2015; Esteva et al., 2021). Veículos e voos Autônomos com sistemas que analisam imagens em tempo real para navegação e segurança (Galvao et al., 2021; Elallid et al., 2022). Reconstrução de objetos 3D utilizada para gerar modelos tridimensionais a partir de imagens bidimensionais capturadas de diferentes ângulos (Fan et al., 2017b; Lian et al., 2019; Duggal and Pathak, 2022). Monitoramento de vigilância para Identificação e rastreamento de movimentos em sistemas de segurança (Thai et al., 2022; Z. et al., 2020). Reconhecimento facial e biométrico usado em autenticação e controle de acesso (Xie et al., 2021; Canedo and Neves, 2019; and, 2021; Saponara et al., 2021).

2.4.1 Redes Neurais convolucionais

As redes neurais convolucionais, conhecidas como *Convolutional Neural Networks* (CNNs ou ConvNets), são um tipo avançado de rede neural utilizada em visão computacional e processamento de imagens (Khan et al., 2020). As CNNs possuem a capacidade de detectar automaticamente *features* ou recursos importantes em imagens, como bordas, texturas, formas, e padrões complexos, sem a necessidade de intervenção humana (Raschka, 2018). Esses recursos são hierarquicamente estruturados, permitindo que as camadas iniciais aprendam características mais simples como linhas e bordas, e as camadas posteriores capturem padrões mais abstratos como objetos completos.

Uma característica das CNNs é o compartilhamento de pesos nos filtros convolucionais, que reduz o número de parâmetros treináveis. Isso não apenas melhora a eficiência computacional, mas também aumenta a capacidade de generalização da rede, reduzindo o risco de *overfitting*. Os estágios de extração de recursos das CNNs são projetados para aprender automaticamente representações relevantes dos dados, explorando informações

espaciais e de canal, bem como a largura e profundidade de suas arquiteturas (Kuo et al., 2020). A saída do modelo depende diretamente das *features* extraídas da imagem de entrada, o que torna a escolha de parâmetros e a arquitetura da rede essenciais para o desempenho.

Os avanços trazidos pelas CNNs estão relacionados a inovações em diversos aspectos, tais como: redes mais profundas e largas permitindo a extração de características mais complexas e um melhor desempenho em tarefas desafiadoras. O desenvolvimento de um modelo de inteligência artificial chamado *Vision Transformer* de 22 bilhões de parâmetros (ViT-22B) (Dehghani et al., 2023), que é uma forma avançada de tecnologia que ajuda computadores a entender e processar imagens de maneira similar ao cérebro humano. Este modelo consegue entender e classificar imagens alcançando resultados excelentes em testes que medem como ele pode identificar diferentes objetos nas imagens. Enquanto as redes de convoluções separáveis em profundidade, também conhecidas como "*depthwise separable convolutions*", reduzem a complexidade computacional das convoluções, permitindo o uso de redes maiores e mais eficientes em dispositivos com recursos limitados como smartphones e dispositivos IoT (Zhou et al., 2020). Redes generativas adversariais (GANs) (Wang et al., 2021) são utilizadas para gerar imagens realistas e podem ser combinadas com CNNs para melhorar o desempenho em diversas tarefas, enquanto a Inteligência Artificial Explicável (XAI) (Speith, 2022) busca desenvolver métodos para tornar as decisões das CNNs mais transparentes e compreensíveis.

2.4.1.1 Arquitetura das Redes Neurais Convolucionais

As CNN se diferem das redes neurais tradicionais pela extração de recursos ou características da imagem de entrada. A imagem é vista como uma matriz bidimensional de pixels. A convolução é a extração das características da imagem através do uso de filtros (Lecun et al., 1998). A arquitetura básica das CNN incluem blocos de convoluções, pooling e camadas totalmente conectadas, conforme representada na figura 7

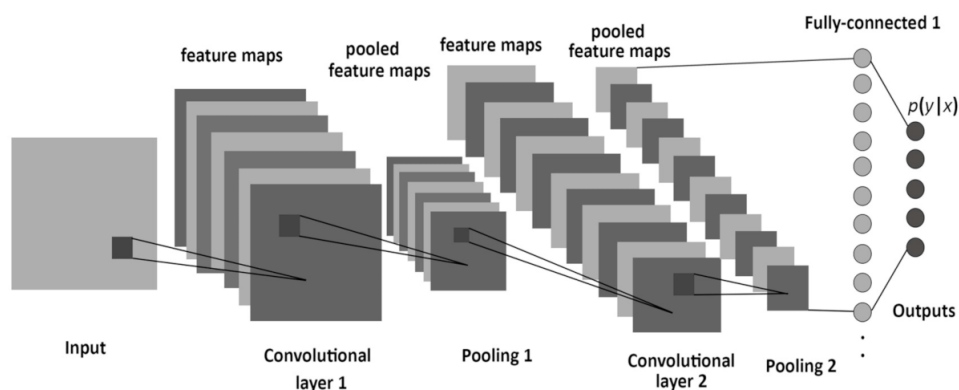


Figura 7: Arquitetura básica uma Rede Neural Convolucional (Albelwi and Mahmood, 2017).

2.4.1.2 Funções de Ativação

As funções de ativação são usadas para gerar relações não lineares entre a entrada e a saída (Dong et al., 2021), isto é, confere uma capacidade matemática da não-linearidade ao processamento. As funções de ativação deliberam o nível de ativação de um neurônio frente a uma determinada entrada. Existem diferentes funções de ativação, tais como: ReLu, Sigmóide, Tangente hiperbólica dentre outras (Ramachandran et al., 2017). A partir da entrada do dado a função sigmoide, gera uma saída no intervalo $[0,1]$. A função sigmoide é utilizada em probabilidades para problemas de classificação. A função Relu (*REctified Linear Unit*) gera saída no intervalo entre $[0, +\infty[$, a saída será $[0]$ somente se a entrada for negativa, para as entradas positivas, então a saída será o próprio valor da entrada (Feng and Lu, 2019). A função *Leaky Rectified Linear Unit*, ou Leaky ReLU, é um tipo de função de ativação baseada em uma ReLU, possui baixos valores se a entrada for negativa (Maas et al., 2013).

A Tabela 1 apresenta as funções de ativação e suas limitações para este estudo

Tabela 1: Comparação entre diferentes funções de ativação

Função de Ativação	Fórmula	Problemas
Sigmoid	$\sigma(x) = \frac{1}{1+e^{-x}}$	Sofre de <i>vanishing gradient</i> , limitando a capacidade de aprendizado em redes profundas. Além disso, gera saídas entre 0 e 1, inadequadas para modelar variações térmicas contínuas.
Tanh	$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	Também sofre de <i>vanishing gradient</i> e, embora seja centrada em zero, limita a amplitude da saída, dificultando a generalização para padrões térmicos variados.
Softmax	$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$	Utilizada para classificações multi-classes, mas inadequada para regressão, que é o foco deste trabalho.
Leaky ReLU	$f(x) = x$ se $x > 0$, $0.01x$ se $x \leq 0$	Evita a desativação total de neurônios, mas pode gerar valores muito pequenos em regiões negativas, o que não é ideal para a segmentação térmica.
ELU (Exponential Linear Unit)	$f(x) = x$ se $x > 0$, $\alpha(e^x - 1)$ se $x \leq 0$	Alto custo computacional

No presente estudo, a função de ativação ReLU (Rectified Linear Unit) foi empregada na rede neural convolucional para extrair características das imagens térmicas e prever a distribuição tridimensional dos sedimentos na tubulação. A escolha da ReLU se deve às suas vantagens computacionais e propriedades matemáticas, que a tornam ideal para

este tipo de aplicação. A ReLU introduz não linearidade sem afetar significativamente o desempenho computacional; evita o problema do vanishing gradient, comum em funções como sigmoid e tanh; permite um treinamento mais rápido, devido à sua computação simples ($\max(0, x)$); facilita o aprendizado de características mais relevantes, uma vez que mantém valores positivos e zera os negativos, melhorando a discriminação dos padrões térmicos.

2.4.1.3 Funções de perda para reconstrução 3D em nuvens de pontos

A função de perda, também denominada função de erro ou função de custo, desempenha um papel fundamental no treinamento de redes neurais, pois quantifica a discrepância entre a saída predita pelo modelo e o valor real de referência, frequentemente chamado de *ground-truth*. A escolha da função de perda depende da natureza do problema e da arquitetura do modelo utilizado. Para problemas de regressão, a Função de Erro Quadrático Médio (*Mean Squared Error* – MSE) é uma escolha comum, enquanto para problemas de classificação a função de entropia cruzada é amplamente empregada (Wang et al., 2018b). Além disso, a camada de saída do modelo deve ser compatível com a função de perda adotada, garantindo que a otimização ocorra de maneira eficiente e alinhada aos objetivos do problema (Yuan et al., 2018; Rezatofighi et al., 2019).

No contexto da reconstrução 3D baseada em nuvens de pontos, diversas métricas de perda têm sido propostas na literatura, cada uma com características distintas para mensurar a similaridade entre a nuvem de pontos gerada pelo modelo e a nuvem de pontos de referência, dentre elas pode-se citar:

- *Chamfer Distance* (CD) é amplamente utilizada para avaliar a dissimilaridade entre duas nuvens de pontos (Lu et al., 2022). No entanto, a métrica tradicional apresenta limitações, como a alta sensibilidade a outliers e a falta de consideração da densidade local dos pontos, o que pode resultar em reconstruções com erros significativos. Para mitigar essas limitações, avanços como a Hyperbolic Chamfer Distance (HyperCD) (Lin et al., 2023) e a Density-Aware Chamfer Distance (DCD) (Wu et al., 2021a) foram desenvolvidos. A HyperCD opera no espaço hiperbólico, atribuindo pesos diferenciados às correspondências de pontos durante a retropropagação, melhorando a suavidade da superfície reconstruída. Já a DCD leva em conta a disparidade na distribuição de densidade entre as nuvens de pontos, otimizando a correspondência entre os conjuntos de dados.
- *Earth Mover's Distance* (EMD) mede a quantidade de trabalho necessária para transformar uma nuvem de pontos em outra, levando em consideração a movimentação e redistribuição dos pontos (Rubner, 2000). Embora a EMD seja mais robusta em relação à densidade dos pontos, ela apresenta um alto custo computacional, tornando-a impraticável para aplicações que envolvem grandes volumes

de dados (Meng et al., 2024).

- *Iterative Closest Point* (ICP) busca minimizar a distância entre os pontos da nuvem de referência e os pontos gerados pelo modelo, ajustando iterativamente a correspondência entre os dois conjuntos (Zhang, 2021). No entanto, o ICP depende fortemente da inicialização dos pontos e pode sofrer com convergência para mínimos locais, especialmente em nuvens de pontos esparsas ou altamente ruidosas.
- *Mean Squared Error* (MSE) é uma métrica amplamente empregada para medir a diferença entre valores previstos e valores reais (Tliba et al., 2022). Essa função é calculada como a média dos quadrados das diferenças entre os valores preditos pelo modelo e os valores reais observados. Uma das vantagens do MSE é ser uma função convexa, garantindo a existência de um mínimo global e facilitando a otimização por métodos baseados em gradiente. Além disso, ao elevar as diferenças ao quadrado, o MSE penaliza erros maiores de maneira mais significativa, tornando-o uma escolha adequada para garantir maior precisão nas previsões.

Neste trabalho, adota-se como função de perda a métrica *Mean Squared Error* (MSE), em razão das características específicas da representação do perfil do sedimento em nuvem de pontos. A nuvem de pontos gerada a partir do escaneamento do sedimento foi discretizada em uma malha quadrangular, sendo as cotas de altura (eixo Z) interpoladas em 800 pontos distribuídos em intervalos fixos. Esse processo confere à nuvem de pontos uma estrutura ordenada, com correspondência direta entre os pontos preditos e os pontos reais, o que torna inadequado o uso de métricas desenvolvidas para conjuntos desordenados, tais com o CD ou EMD. A ordenação implícita decorrente da interpolação permite que a comparação entre as saídas do modelo e os dados de referência seja realizada de forma ponto a ponto. Assim, o MSE mostra-se adequado, pois mede a diferença quadrática média entre os valores preditos e os valores reais, sendo eficiente para tarefas de regressão em espaços contínuos. Além disso, por ser uma função de perda convexa, o MSE favorece a estabilidade durante o treinamento do modelo, contribuindo para uma convergência mais robusta. Dessa forma, sua escolha neste contexto se justifica tanto do ponto de vista teórico quanto computacional.

2.4.1.4 Otimização de parâmetros

Os otimizadores são algoritmos que ajustam os pesos da rede neural durante o treinamento, buscando minimizar a função de perda, para isso são considerados vários aspectos técnicos, dentre eles:

- O tamanho do *Kernel* determina a escala das características que a rede é capaz de extrair. *Kernels* maiores podem capturar informações de maior escala, percebendo características mais abrangentes da entrada, enquanto *kernels* menores focam em

detalhes mais sutis. Ajustar o tamanho do *kernel* influencia diretamente a capacidade da rede de identificar e responder à variações espaciais na entrada. O tamanho do *Kernel* afeta tanto a performance quanto a capacidade de generalização do modelo (Chansong and Supratid, 2021; Romero et al., 2022; Ding et al., 2022).

- Número de Camadas (Profundidade), quanto maior o numero de camadas, maior a capacidade de capturar características mais complexas, mas também aumenta o risco de *overfitting* e eleva o custo computacional (Shafiq and Gu, 2022).
- Taxa de aprendizado (*learning rate*) é um hiperparâmetro importante no treinamento de redes neurais, incluindo CNNs. Ela determina o tamanho do passo que o otimizador (como SGD ou Adam) dá na direção do mínimo da função de perda. A escolha da taxa de aprendizado garante a convergência eficiente e um melhor desempenho do modelo (Li and Yang, 2020). Uma taxa de aprendizado muito alta pode fazer com que o otimizador oscile em torno do mínimo, impedindo a convergência. Uma taxa de aprendizado muito baixa pode levar a uma convergência lenta, exigindo muitas iterações para alcançar um bom resultado. *Learning rate* também influencia a capacidade de generalização do modelo para ter boa performance com os dados de validação, isto é, dados não vistos durante o treinamento (Simoens et al., 2024).
- Regularização são técnicas utilizadas para evitar o sobreajuste (Santos and Papa, 2022)(*overfitting*), isto é, quando o modelo se adapta excessivamente aos dados de treinamento e perde a capacidade de generalização para dados novos. Algumas técnicas comuns incluem: *Dropout* a qual um percentual dos neurônios da rede são desligados durante o treinamento, forçando a rede a aprender representações mais robustas. Regularização L1 e L2 adicionam termos de penalidade à função de perda, penalizando pesos com valores absolutos ou quadrados muito altos, respectivamente. *Batch Normalization* normaliza as ativações de cada camada durante o treinamento, o que pode acelerar a convergência e melhorar o desempenho do modelo. Aumento de Dados (*Data Augmentation*) promove o aumento da quantidade de dados de treinamento através de transformações como rotação, espelhamento, corte e adição de ruído pode ajudar a melhorar a generalização do modelo.

2.4.2 Redes de Reconstrução 3D

Redes Neurais Convolucionais (CNNs) podem ser usadas para realizar a reconstrução 3D de objetos e cenas a partir de uma ou múltiplas imagens. Estas redes usam convoluções para extrair características das imagens. Estas características são usadas para gerar uma representação 3D da cena. Esta abordagem tem a vantagem de ser mais robusta às condições de iluminação, ruídos e objetos invisíveis, pois as redes neurais conseguem

aprender a detetar características essenciais das imagens, mesmo se estas forem alteradas pelas condições de iluminação ou ruídos. Além disso, as redes neurais também são capazes de detectar padrões mais complexos.

O objetivo da reconstrução 3D baseada em imagem é inferir a geometria 3D de objetos ou cenas. Pontos característicos das imagens são observados para que seja feito o alinhamento e reconstrução 3D da imagem do objeto. No entanto, este modelo de reconstrução pode conter falhas de correspondências dos pontos observados motivados por condições de iluminação, textura ou oclusão (Ito et al., 2017).

A reconstrução 3D baseada em uma única imagem requer algum conhecimento prévio, tal como: ângulo do objeto, modelos anteriores, anotações de dados, informações de iluminação fixa (Fan et al., 2017b; Jiang et al., 2018c,a; Fu et al., 2020; Jiang et al., 2018b; Zeng et al., 2018; Mandikal et al., 2019; Li et al., 2020a; Feng et al., 2020; Zou et al., 2024; Liu et al., 2024). Os desafios desta técnica são mencionados por (Fu et al., 2020) e consistem na reconstrução de complexidade de forma de objetos, na incerteza de objetos, na granulação fina, nos requisitos de memória, na seletividade de representações de formas 3D e no conjuntos de dados de treinamento.

A complexidade da forma dos objetos refere-se à dificuldade em reconstruir objetos com geometrias irregulares, superfícies não uniformes e detalhes finos. Objetos do mundo real podem ter formas complexas, com detalhes intrincados, concavidades e convexidades variadas. Quanto mais complexa a estrutura do objeto, mais difícil é prever e gerar um modelo 3D preciso a partir de uma única visão, pois a imagem 2D não captura todas as nuances da geometria 3D. Por exemplo, um cubo é mais fácil de reconstruir do que uma árvore, pois o primeiro possui superfícies planas bem definidas, enquanto a árvore tem galhos e folhas com variações irregulares que dificultam a modelagem precisa.

A incerteza de objetos ocorre na reconstrução 3D a partir de uma única imagem, devido à falta de informações sobre partes ocultas do objeto. Como apenas uma perspectiva é fornecida, a rede neural deve inferir a geometria oculta com base em padrões aprendidos, o que pode gerar ambiguidades. Diferentes formas tridimensionais podem projetar a mesma imagem 2D, tornando a reconstrução incerta. Exemplos incluem objetos com simetria rotacional, onde diferentes orientações podem resultar na mesma projeção; objetos com oclusões, onde partes da estrutura estão escondidas na imagem; e objetos cuja profundidade não pode ser medida com precisão, dificultando a estimativa correta de sua geometria completa.

A granulação fina na reconstrução 3D refere-se à capacidade de capturar e representar detalhes minuciosos da superfície de um objeto, garantindo uma modelagem mais precisa e realista. Esse aspecto trata da preservação de texturas e pequenas irregularidades, como rugas em um rosto humano, padrões em tecidos ou detalhes em peças de objetos complexos. No entanto, a reconstrução desses elementos a partir de uma única imagem apresenta desafios significativos, pois a imagem pode não conter informações suficientes

para inferir com exatidão tais características. A limitação de dados detalhados compromete a fidelidade da reconstrução, exigindo métodos avançados de aprendizado profundo para estimar texturas e superfícies de alta resolução a partir de informações parciais.

A seletividade de representações de formas 3D refere-se ao desafio de escolher a melhor abordagem para modelar objetos 3D, considerando fatores como eficiência computacional, fidelidade visual e compatibilidade com aplicações específicas. Diferentes métodos podem ser utilizados, incluindo nuvens de pontos, mais adequadas para formas irregulares, malhas poligonais, ideais para superfícies suaves, e voxels, que possibilitam a reconstrução de estruturas internas complexas. A escolha da representação impacta diretamente a precisão e aplicabilidade da reconstrução, influenciando desde a renderização detalhada até a otimização para simulações interativas. Esse processo exige técnicas avançadas de aprendizado profundo e conhecimento prévio para garantir uma reconstrução eficiente e realista a partir de uma única imagem, minimizando perdas de informação e otimizando a adaptação para diferentes cenários.

Dentre as redes de reconstrução 3D a partir de uma única imagem, podemos destacar:

- A rede *Point Set Generation* (PSG) (Fan et al., 2017b) propõe um método para reconstrução 3D, a partir de uma única imagem. A PSG em sua arquitetura possui dois ramos de previsão: um ramo com alta flexibilidade na captura de estruturas complexas através das camadas completamente conectadas e outros explorando a continuidade geométrica através do auto-encoder. Considerando uma única imagem em ângulo específico como entrada na PSG, a rede estima a profundidade para as partes visíveis e a geometria do objeto (Fan et al., 2017b). A *Chamfer Distance* (CD) (Wu et al., 2021b) e *Earth Mover's Distance* (EMD) (Hou et al., 2016) são propostas como uma funções de perda.
- (Jiang et al., 2018a) propõe uma abordagem que aprimora a reconstrução 3D a partir de uma única imagem, introduzindo uma função de perda que considera tanto a consistência geométrica quanto a coerência semântica, resultando em modelos tridimensionais mais precisos e realistas.
- (Zeng et al., 2018) propõe um método para reconstrução 3D de nuvens de pontos a partir de uma única imagem monocular. A principal contribuição da pesquisa é a introdução de uma etapa intermediária de estimativa de profundidade, que gera um mapa de profundidade explícito da imagem de entrada. Esse mapa de profundidade é então transformado em uma nuvem de pontos parcial, que posteriormente passa por uma rede de completamento de nuvem de pontos (*point cloud completion*) para produzir uma reconstrução 3D completa. Essa abordagem divide o problema em duas etapas mais simples: a predição da profundidade e o preenchimento da geometria ausente, melhorando a consistência geométrica e a qualidade da nuvem de pontos gerada. O método é validado nos datasets ShapeNet e Pix3D, obtendo

resultados superiores aos métodos anteriores tanto em fidelidade quanto em robustez. Além disso, o uso explícito de profundidade como representação intermediária permite maior interpretabilidade do processo de reconstrução.

- Pix2PC (Choi et al., 2019) propõe uma arquitetura que recebe uma imagem 2D de entrada e uma nuvem de pontos aleatória. A rede gera um vetor de características da imagem, posteriormente deforma a nuvem de pontos de acordo com a representação das características da imagem 2D de entrada. Esta estrutura usa VGG-19 (Simonyan and Zisserman, 2014) para extrair as informações da imagem 2D e, em seguida, calcular o vetor característico para cada ponto da nuvem de pontos. Recursos de cálculos de projeção estatística são usados para guiar a deformação da nuvem do ponto de entrada. A função de perda é *Chamfer Distance* que compara a nuvem de pontos deformada prevista com a nuvem de pontos desejada.
- A rede 3D-PSRNet(Mandikal et al., 2019) recebe uma imagem RGB como entrada e gera em sua saída final uma nuvem de pontos com informações semânticas. A arquitetura da rede utiliza um auto-encoder seguido de camadas totalmente conectadas, em seguida o resultado deste primeiro processamento serve como entrada para uma rede de segmentação, *PointNet* (Lian et al., 2019), que recebe a nuvem de pontos e fornece rótulos semânticos. O processamento conjunto de ambas as tarefas tem um resultado melhor do que o processamento individual.
- A rede 3D-RecGAN++ (Yang et al., 2019) em sua arquitetura utiliza uma estrutura do tipo *Generative Adversarial Networks (GANs)* são um tipo de arquitetura de aprendizado profundo que usa duas redes neurais, uma rede generativa e uma rede discriminadora, para criar novos dados a partir de um determinado conjunto de dados. A rede generativa cria novos dados a partir do conjunto de dados fornecido, enquanto a rede discriminadora avalia os dados gerados e determina se são reais ou falsos. As duas redes são treinadas juntas de forma antagônica, com a rede generativa tentando enganar a rede discriminadora e a rede discriminadora tentando identificar os dados gerados como falsos. As GANs têm sido usadas para criar imagens, textos e músicas realistas, bem como para gerar novos dados para tarefas de aprendizado de máquina. A rede 3D-RecGAN++ sintetiza dados de visualizações 2D, mais a informação de profundidade da image. Estas informações são utilizadas para inferir estruturas 3D do objetos no espaço voxel de alta dimensão. A proposta é combinar os conhecimentos do codificador decodificador 3D e a estrutura de redes GANs.
- 3D-ReConstnet (Li et al., 2020b) recebe como entrada uma imagem 2D de um objeto. O processo de transformação da imagem 2D em uma nuvem de pontos ocorre em três tarefas. A primeira tarefa usa Resnet50 (He et al., 2015) para extrair recur-

sos a partir de imagens de entrada. Estes recursos podem incluir informações sobre a geometria da imagem, suas cores e texturas. Resnet é também chamada de Rede Residual que usa conexões de saída de blocos de camadas que conectam diretamente às entradas dos blocos de camadas seguintes, permitindo que as informações residuais sejam propagadas diretamente para a camada seguinte. Isso permite que a rede aprenda melhor com camadas adicionais, pois as informações residuais são usadas para aprimorar o aprendizado. A segunda tarefa gera um vetor probabilístico a partir dos recursos 2D extraídos. O vetor probabilístico contém informações sobre a posição, orientação e profundidade dos pontos na nuvem de pontos. A terceira tarefa usa o vetor probabilístico para gerar a nuvem de pontos 3D. Os datasets utilizados neste experimento foram ShapeNet e Pix3D.

- DV-Net (Jia et al., 2020) recebe como entrada duas imagens de diferentes perspectivas do mesmo objeto. A arquitetura da rede é composta por três fases: a primeira fase recebe as duas imagens de entrada, utiliza a rede PSG para geração da nuvem de pontos 3D de ambas as imagens de entrada. Na segunda fase é composta por uma estrutura denominada rede de aprendizado de recursos, que é baseada em uma rede PointNet. Na última fase as duas nuvens de pontos são mescladas.
- (Li et al., 2020a) a arquitetura proposta consiste em três módulos principais: (i) extração de características da imagem utilizando ResNet-50, (ii) geração de um vetor de características latentes probabilístico que codifica a geometria do objeto e (iii) geração da nuvem de pontos 3D a partir deste vetor latente. O método é treinado e avaliado nos datasets ShapeNet e Pix3D, mostrando que o modelo é capaz de produzir nuvens de pontos 3D com maior fidelidade e menor distorção em relação às abordagens concorrentes. A principal vantagem da 3D-ReConstNet é a capacidade de preservar melhor as informações geométricas da imagem de entrada, resultando em reconstruções com melhor cobertura e menor perda de detalhes. O artigo também destaca a robustez do método frente a variações de pontos de vista e ruídos nas imagens de entrada.
- (Feng et al., 2020) propõe uma estrutura de reconstrução 3D de mídias porosas a partir de uma única imagem 2D, utilizando aprendizado profundo. O método é baseado em uma arquitetura encoder-decoder convolucional capaz de mapear diretamente uma imagem bidimensional para um volume tridimensional, superando a limitação de técnicas tradicionais que requerem múltiplas imagens ou amostras físicas. A abordagem foi validada com imagens de microtomografia de rochas, demonstrando que o modelo consegue reconstruir adequadamente a estrutura porosa preservando as propriedades geométricas e estatísticas, como porosidade e distribuição espacial. O trabalho destaca ainda a vantagem de ser uma metodo-

logia de reconstrução não destrutiva, automática e eficiente, reduzindo o tempo e o custo normalmente associados à análise de mídias porosas.

- (Müller et al., 2022) apresenta uma nova codificação de entrada para primitivos gráficos neurais que permite o uso de uma rede menor sem sacrificar a qualidade. A codificação é uma tabela de hash multirresolução de vetores de características treináveis, otimizados por descida de gradiente estocástico. A estrutura multirresolução permite que a rede desambigue colisões de hash, resultando em uma arquitetura simples e altamente paralelizável em GPUs modernas. Essa é uma abordagem inovadora para acelerar o treinamento e a inferência de primitivos gráficos neurais, reduzindo significativamente os custos computacionais sem comprometer a qualidade da reconstrução. Os autores demonstram treinamento instantâneo de primitivos gráficos neurais em uma única GPU para diversas tarefas, incluindo imagens gigapixel através do mapeamento de coordenadas 2D para cores RGB em imagens de altíssima resolução.
- (Liu et al., 2024) propõe um método para reconstrução 3D de objetos a partir de uma única imagem, incorporando consciência de partes estruturais do objeto (part-aware). O modelo, chamado Part123, segmenta e organiza a reconstrução em partes semânticas distintas, permitindo a geração de modelos 3D mais detalhados e consistentes com a estrutura real dos objetos. A rede utiliza uma combinação de aprendizado supervisionado com dados sintéticos e reais e introduz uma representação baseada em partes para melhorar a generalização e precisão da reconstrução, mesmo em casos de oclusões ou poses incomuns. O trabalho mostra que a consideração explícita das partes do objeto contribui para uma melhor captura das relações geométricas e semânticas durante a reconstrução 3D, superando métodos anteriores em benchmarks públicos como ShapeNet e Pix3D. Como principais vantagens, o método apresenta maior precisão em detalhes locais e uma melhor preservação da estrutura dos objetos
- (Zou et al., 2024) propõe uma abordagem baseada em um processo de inferência *feed-forward*, que permite a geração de um modelo 3D completo a partir de uma única imagem de entrada, sem necessidade de refinamento iterativo. A metodologia emprega duas redes baseadas em transformadores (*Transformers*), um decodificador de nuvem de pontos e um decodificador triplano, para construir uma representação intermediária híbrida Triplano-Gaussiana. O decodificador de nuvem de pontos é projetado para gerar representações explícitas e estruturadas a partir da imagem de entrada, enquanto o decodificador triplano consulta atributos Gaussianos para cada ponto, permitindo uma modelagem mais detalhada da geometria e aparência do objeto. Essa abordagem resolve desafios inerentes à regressão direta de atributos Gaussianos 3D, que frequentemente sofre de falta de estrutura e

instabilidade na representação. A renderização final é realizada através de *Gaussian Splatting*, utilizando uma MLP (*Multi-Layer Perceptron*) para decodificar os Gaussianos e permitir uma visualização eficiente em tempo real. Os modelos foram treinados em conjuntos de dados 3D de grande escala, garantindo generalização e escalabilidade. Com tempos de inferência significativamente reduzidos e alta fidelidade visual, essa metodologia apresenta um avanço significativo para aplicações em realidade aumentada (AR), realidade virtual (VR), simulações gráficas e geração automatizada de conteúdo 3D, consolidando-se como uma solução eficiente e generalizável para reconstrução 3D a partir de visão única.

- (Nguyen et al., 2024), apresenta um método *Novel Object Pose Estimation (NOPE)* para estimar a pose 3D de objetos a partir de uma única imagem RGB, sem necessidade de modelos CAD ou dados de treinamento específicos para cada objeto. Essa abordagem é especialmente relevante para aplicações como robótica e realidade aumentada, onde novos objetos podem aparecer com frequência e não há tempo ou recursos para coletar dados adicionais. NOPE combina um estimador de pose inicial baseado em redes neurais, que prevê parâmetros de pose (rotação e translação) de forma aproximada usando características visuais extraídas da imagem; um otimizador de pose auto-supervisionado, que refina essa estimativa inicial através de um alinhamento denso entre características da imagem de entrada e projeções de um modelo 3D genérico ou representações aprendidas implicitamente. A principal inovação é o uso de representações 3D genéricas obtidas de grandes conjuntos de dados para modelar a geometria dos objetos de forma agnóstica à classe, permitindo lidar com formas e texturas não vistas anteriormente. O NOPE faz a combinação de predição inicial de rede neural com refinamento auto-supervisionado. Usa de representações 3D genéricas para lidar com formas e aparências diversas.
- (Szymanowicz et al., 2024), apresenta um método de reconstrução 3D em tempo real a partir de uma única imagem RGB, baseado em 3D Gaussian Splatting adaptado para cenários de vista única. Utiliza predição de profundidade e normais para gerar diretamente “splats” gaussianos, evitando a otimização pesada de NeRFs. Reconstrói cenas completas em 0,5 s, mantendo qualidade próxima a métodos state-of-the-art, sendo adequado para AR, robótica e geração de conteúdo 3D interativo.
- (Jiang et al., 2024), apresenta uma abordagem de detecção 3D supervisionada de forma fraca, usando apenas uma imagem monocular e rótulos 2D para treinar. Introduce um pipeline que combina projeções geométricas, restrições físicas e aprendizado de profundidade relativa para estimar posição, dimensão e orientação 3D de objetos.
- (Jang and Agapito, 2024), utiliza método para síntese de novas vistas a partir de

uma única imagem em cenários não controlados. Usa uma rede que combina estimativa de profundidade *depth estimation* confiável, projeção de características para o espaço 3D *feature lifting* e renderização neural adaptada para lidar com oclusões e áreas não vistas. Permite gerar vistas realistas em diferentes ângulos a partir de uma única foto, aplicável em AR, VR e edição de imagens.

- (Cai et al., 2024), foca na reconstrução 3D a partir de raios-X com poucas vistas (sparse-view), incorporando consciência estrutural no processo de aprendizado. Introduce uma rede que integra restrições anatômicas e informações geométricas para melhorar a qualidade de reconstrução, especialmente em áreas com dados limitados. Demonstra avanços significativos em qualidade visual e métricas quantitativas sobre métodos anteriores.

Ao longo dos últimos anos, as abordagens para reconstrução 3D a partir de uma única imagem evoluíram de arquiteturas iniciais, como a Point Set Generation (PSG) (Fan et al., 2017b), voltadas à geração direta de nuvens de pontos usando ramos especializados em captura de geometria e continuidade, para métodos mais sofisticados que incorporam estágios intermediários e representações híbridas. Estratégias como as propostas em (Zeng et al., 2018) e Pix2PC (Choi et al., 2019) passaram a incluir a estimativa explícita de profundidade e o uso de nuvens de pontos parciais com posterior completamente, permitindo maior consistência geométrica. Outras redes, como a 3D-PSRNet (Mandikal et al., 2019) e a 3D-RecGAN++ (Yang et al., 2019), ampliaram a capacidade semântica e o detalhamento volumétrico por meio de segmentação integrada e arquiteturas generativas adversariais. Com o avanço das representações e das funções de perda, modelos como a 3D-ReConstNet (Li et al., 2020b) e o método de (Li et al., 2020a) exploraram codificações probabilísticas para melhor preservar informações geométricas, enquanto abordagens recentes, como Part123 (Liu et al., 2024), introduziram a consciência de partes (part-aware) para reconstruções mais detalhadas e consistentes. Ao mesmo tempo, trabalhos como o de (Zou et al., 2024) inovaram com representações triplano-gaussianas e decodificadores híbridos, unindo a estrutura explícita de nuvens de pontos à eficiência de renderizações por Gaussian Splatting. No campo de aplicações correlatas, métodos como NOPE (Nguyen et al., 2024) expandiram a inferência para estimativa de pose 3D de objetos inéditos, enquanto (Szymanowicz et al., 2024) e (Jang and Agapito, 2024) exploraram reconstruções rápidas e síntese de novas vistas, respectivamente, a partir de visão única. Finalmente, soluções como (Cai et al., 2024) evidenciam a aplicação dessas técnicas em domínios especializados, como a reconstrução estrutural a partir de raios-X com vistas esparsas. Essa trajetória demonstra uma transição de arquiteturas monolíticas para pipelines modulares, multimodais e orientados a representações intermediárias, resultando em maior precisão, velocidade e adaptabilidade a diferentes cenários e domínios.

A seguir apresenta-se um comparativo entre as redes de reconstrução 3D menciona-

Tabela 2 – Continuação da página anterior

Autor	Entrada	Saída	Dataset	Aplicação Principal	Pontos Fortes	Limitações
(Szymanowicz et al., 2024)	Única imagem RGB de cena	Reconstrução 3D densa usando Gaussian Splatting	CO3D, DTU, Tanks and Temples	Reconstrução 3D ultrarrápida para AR/VR e robótica	Velocidade muito alta ($\sim 0,5s$), qualidade próxima ao <i>state-of-the-art</i> , sem otimização pesada	Menor precisão geométrica que métodos mais lentos; dependente de estimativa de profundidade acurada
(Jiang et al., 2024)	Única imagem monocular de cena	Caixas delimitadoras 3D (pose, dimensão, posição)	KITTI, nuScenes	Detecção 3D com supervisão fraca	Não requer rótulos 3D, usa apenas anotações 2D; desempenho competitivo	Menor precisão em distâncias grandes; limitado a categorias de treino
(Jang and Agapito, 2024)	Única imagem RGB	Novas vistas sintéticas da cena	LLFF, Spaces, CO3D	Síntese de novas vistas (<i>novel view synthesis</i>)	Funciona em cenários não controlados, lida bem com oclusões	Qualidade reduzida em regiões não vistas; dependente de estimativa de profundidade precisa
(Cai et al., 2024)	Poucas imagens de raio-X (<i>sparse-view</i>)	Volume 3D reconstruído	Conjuntos médicos de raio-X	Reconstrução 3D médica com dados limitados	Usa restrições anatômicas para melhor qualidade; bom em dados escassos	Aplicação restrita a anatomias conhecidas; não testado em cenários industriais

Na Tabela 3 apresentamos a disponibilidade do código no repositório github.

Tabela 3: Comparação entre redes de reconstrução 3D (b)

Autor/Rede	Github
(Fan et al., 2017b) PSG (Point Set Generation)	https://github.com/fanhqme/PointSetGeneration
(Jiang et al., 2018a) GAL	-
(Zeng et al., 2018)	-
(Choi et al., 2019) Pix2PC	https://github.com/bharat-b7/IPNet
(Mandikal et al., 2019)	https://github.com/val-iisc/3d-psrnet
(Li et al., 2020b)	-
(Jia et al., 2020) DV-Net	-

A reconstrução tridimensional a partir de uma única imagem representa um dos desafios mais complexos dentro da visão computacional. No entanto, ao se tratar da identificação de sedimentos no interior de tubulações industriais, a problemática torna-se ainda mais desafiadora, pois a reconstrução não pode ser diretamente comparada às abordagens tradicionais utilizadas para objetos visíveis. A proposta deste trabalho envolve a inferência de um volume oculto, utilizando exclusivamente dados térmicos capturados externamente à tubulação. A principal limitação imposta pela vista única na reconstrução térmica de perfis de sedimentos reside na impossibilidade de acesso direto à geometria interna do duto. A identificação de sedimentos baseia-se na inferência do perfil do material depositado a partir da distribuição do calor e das diferenças de condutividade térmica entre os componentes internos da tubulação. Isso significa que a reconstrução não se fundamenta na captura direta da forma do sedimento, mas sim em uma estimativa inferida a partir da resposta térmica da superfície externa do duto.

No contexto deste estudo, a imagem térmica obtida a partir da base da tubulação apresenta uma vantagem estratégica: é a única configuração de vista em que se captam simultaneamente as respostas térmicas de três elementos fundamentais: o metal da tubulação, o sedimento depositado e o ar. Essa disposição permite que o modelo receba informações térmicas mais diretas e menos ambíguas sobre as diferentes camadas envolvidas, favorecendo uma interpretação mais robusta, conforme apresentado na Figura 8a. Em contraste, se a aquisição térmica fosse realizada a partir da lateral da tubulação, as camadas entre metal, sedimento e ar poderiam variar em número e disposição ao longo da linha de visão da câmera. Isso resultaria em sobreposições térmicas mais complexas e em gradientes menos definidos, aumentando significativamente o risco de ambiguidade na interpretação do perfil interno, conforme apresentado na Figura 8b. Nessas condições, o modelo teria de lidar com múltiplas interfaces térmicas, zonas de transição irregulares e possíveis interferências causadas por geometrias parciais ou sobrepostas, elevando o nível de dificuldade da reconstrução.

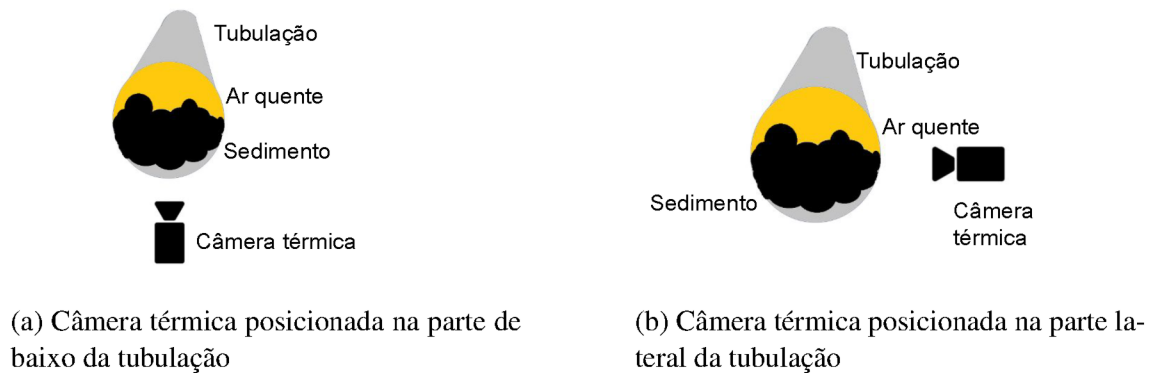


Figura 8: Posicionamento da Câmera térmica para única vista

Essa abordagem impõe restrições à acurácia do modelo, pois a inferência da distribuição do perfil do sedimento depende de fatores como: propagação térmica dentro da tubulação, considerando que diferentes materiais possuem diferentes coeficientes de condutividade térmica, resultando em padrões de aquecimento distintos; influência das condições operacionais, como temperatura do sedimento, umidade, fluxo de ar e tempo de sedimentação, que podem alterar a resposta térmica registrada; e capacidade da câmera térmica em detectar variações sutis de temperatura para identificar nuances na distribuição do calor que indicam a presença e a espessura do sedimento. Pequenas diferenças de leitura podem ser decisivas para a correta inferência do perfil do depósito.

Dessa forma, a reconstrução térmica a partir de vista única não deve ser tratada como um problema clássico de reconstrução 3D, mas sim como um problema de inferência de distribuição volumétrica baseada em perfis térmicos. A principal distinção está no fato de que a reconstrução 3D convencional busca recuperar a geometria visível de um objeto, enquanto o problema aqui abordado requer a predição da forma interna de um sedimento sem visibilidade direta. Essa diferença metodológica posiciona a proposta em um paradigma distinto, no qual a interpretação precisa das assinaturas térmicas constitui o elemento central para o sucesso da reconstrução do perfil do sedimento. Essa mudança de perspectiva abre espaço para soluções que priorizam a utilização de informações térmicas externas de forma otimizada, reduzindo a dependência de métodos intrusivos e de instrumentação interna à tubulação. Ao compreender a reconstrução como um problema de inferência volumétrica baseada em perfis térmicos, torna-se possível desenvolver modelos que aliem precisão na interpretação das respostas térmicas a uma maior viabilidade operacional, estabelecendo as bases para estratégias de monitoramento mais acessíveis e menos invasivas. Nesse contexto, a proposta desta tese busca minimizar a necessidade de hardware invasivo, adotando uma abordagem não destrutiva e de baixo custo operacional para o monitoramento industrial, o que representa um avanço significativo para a indústria ao possibilitar o desenvolvimento de sistemas preditivos eficientes para detecção de sedimentos.

2.4.2.1 Dataset

Um *dataset* consiste em um conjunto estruturado de dados, organizado de forma sistemática para viabilizar o treinamento, teste e a validação de modelos de aprendizado de máquina, incluindo redes neurais convolucionais (CNNs). A qualidade e a representatividade são importantes, pois influenciam a capacidade do modelo de generalizar padrões e realizar previsões precisas. Em aplicações que envolvem processamento de imagens, como reconhecimento de padrões e reconstrução tridimensional, a curadoria do *dataset* deve considerar a diversidade das amostras, a presença de ruídos e a adequação dos rótulos, assegurando que o modelo aprenda características relevantes para a tarefa proposta. Cada *dataset* é projetado para uma tarefa específica, contendo exemplos relevantes para o problema em questão. Os dados podem variar em formato, incluindo imagens (Wu et al., 2014), vídeos (Soomro et al., 2012), textos (Wang et al., 2018a), áudios (Panayotov et al., 2015), modelos CAD (Chang et al., 2015b), entre outros. A curadoria, o balanceamento e a qualidade dos dados são elementos importantes para garantir que os modelos sejam treinados de maneira eficiente e ética (Stando et al., 2023). A qualidade dos dados pode ser exemplificada como a presença de dados ruidosos ou corrompidos, isto pode afetar o aprendizado da rede e prejudicar a capacidade de generalização do modelo.

Além disso, o avanço de técnicas como *data augmentation* e *transfer learning* tem possibilitado maximizar o desempenho dos modelos mesmo em cenários de escassez de dados.

A técnica de *Data augmentation* (aumento de Dados) é utilizada para expandir artificialmente a diversidade de um conjunto de dados. Em muitos cenários, a coleta de dados pode ser um processo dispendioso e demorado. É nesse contexto que a técnica de *data augmentation* se torna importante. Seu principal objetivo é melhorar a generalização do modelo, reduzindo a dependência de um grande volume de dados rotulados e minimizando problemas como *overfitting*. Essa técnica gera variações dos dados existentes, ampliando sua representatividade sem a necessidade de novas coletas.

O *Data Augmentation* consiste em transformações matemáticas e geométricas aplicadas a amostras do conjunto de dados, tais como: rotação, espelhamento, translação, variação de brilho, contraste, ruído, dentre outras. A introdução de variações mantém a semântica original que possibilitam um treinamento mais robusto. As diferentes transformações podem ser aplicadas a imagens, textos, áudio, vídeos e séries temporais, dependendo do domínio da aplicação. O conhecimento do domínio da aplicação é fundamental para escolher as técnicas de *data augmentation* mais adequadas. As transformações aplicadas devem ser relevantes para a tarefa em questão e que não alterem a classe ou o significado dos dados. É importante encontrar um equilíbrio entre a quantidade de dados aumentados e a qualidade dos dados originais, evitando a criação de exemplos irreais ou que possam prejudicar o aprendizado.

Para assegurar que as amostras sintéticas preservem características fundamentais do

domínio de aplicação e sejam eficazes no treinamento de redes neurais convolucionais, a seleção da técnica de aumento de dados deve considerar as especificidades da tarefa proposta. Abordagens baseadas em *embeddings* semânticos, como a apresentada por (Wang et al., 2024), visam alinhar a distribuição dos dados sintéticos com a dos dados reais por meio da captura de representações latentes. Essa estratégia permite a adaptação do espaço de características do modelo generativo ao domínio-alvo, reduzindo discrepâncias entre os dados sintéticos e os dados de teste. No entanto, para este trabalho, o foco central foi a melhoria da qualidade das imagens térmicas utilizadas na predição de sedimentação em tubulações industriais.

Nesse contexto, a escolha da *Super-Resolution Generative Adversarial Network* (SRGAN) (Wang et al., 2018c; Ledig et al., 2017) foi motivada por sua capacidade de gerar imagens sintéticas com maior nível de detalhamento, garantindo a preservação dos padrões térmicos essenciais para a identificação dos depósitos. A SRGAN emprega uma arquitetura baseada em redes adversariais generativas para aprimorar a resolução espacial das imagens, permitindo que a rede neural convolucional capture com maior precisão as variações térmicas associadas à presença de sedimentos. Esse processo não apenas amplia o conjunto de dados disponível, mas também reduz a perda de informações relevantes causada por limitações na resolução das imagens originais. Dessa forma, a estratégia adotada contribui diretamente para o aperfeiçoamento do modelo preditivo, viabilizando uma solução mais robusta e eficiente para o monitoramento não invasivo de tubulações industriais.

A utilização da SRGAN para aprimoramento das imagens térmicas sintéticas representa um avanço na qualidade dos dados utilizados para a predição de sedimentação em tubulações industriais. No entanto, para validar a eficácia do modelo proposto, é fundamental contar com um conjunto de dados robusto que permita avaliar e comparar seu desempenho com abordagens já estabelecidas na literatura. Nesse sentido, diversos *datasets* voltados à reconstrução 3D têm sido amplamente empregados em pesquisas para aprendizado de representações tridimensionais a partir de imagens bidimensionais. Esses *datasets* fornecem um referencial para a modelagem e validação de redes neurais aplicadas à reconstrução de estruturas geométricas complexas, sendo essenciais para a generalização de modelos baseados em aprendizado profundo.

Dentre os *datasets* 3D encontrados na literatura pode-se citar o trabalho de ShapeNet (Chang et al., 2015b) (<https://shapenet.org/>) que propõe um repositório de objetos do cotidiano como cadeiras, mesas, aviões entre outros. A proposta possui 55 categorias de objetos com aproximadamente 51.300 itens em modelo CAD 3D. Outro trabalho é o ModelNet (Wu et al., 2014) (<https://modelnet.cs.princeton.edu/ModelNet>) o qual apresenta diferentes classes de objetos, totalizando 127.915 modelos CAD, 662 categorias de elementos, 10 categorias com orientação anotada. Também tem-se o PASCAL3D+ (Xiang et al., 2014) (<https://cvgl.stanford>.

edu/projects/pascal3d.html) o qual possui imagens em diferentes poses de objetos no mundo real, contendo 12 categorias de objetos anotados com informações de pose (azimute, elevação e distância da câmera). O ScanObjectNN (Uy et al., 2019) (<https://hkust-vgd.github.io/scanobjectnn/>), foi criado para avaliar a capacidade de aprender e classificar objetos 3D a partir de nuvens de pontos. A proposta é baseada na base de dados ScanNet, que consiste em cenas de objetos 3D organizados em ambientes 3D reais. Cada objeto é registrado em sua cena de origem e com um fundo neutro. Por fim, o Pix3D (Sun et al., 2018) (<http://pix3d.csail.mit.edu/>) contém uma coleção de objetos do mundo real incluindo móveis, utensílios de cozinha e brinquedos, de várias fontes. Cada um representado por uma única imagem RGB e um modelo de forma 3D.

A principal lacuna entre os datasets apresentados (como ShapeNet, ModelNet, PASCAL3D+, ScanObjectNN e Pix3D) e um dataset de imagens térmicas reside na natureza dos dados e nas informações que eles representam. Enquanto os primeiros são baseados em modelos CAD, nuvens de pontos e imagens RGB capturadas no espectro visível, fornecendo informações sobre a geometria, textura e cor dos objetos, os datasets de imagens térmicas contêm dados relativos à distribuição de temperatura. Essa diferença implica que os padrões extraídos dos dados visuais tradicionais podem não se aplicar diretamente ao domínio térmico, que depende de características específicas do infravermelho, como variações de calor e dissipação térmica, fundamentais para aplicações em inspeção industrial, por exemplo, na detecção de sedimentação. Portanto, existe uma lacuna significativa quanto à disponibilidade de conjuntos de dados que combinem a riqueza da informação geométrica 3D com os sinais térmicos, necessários para treinar modelos de aprendizado profundo capazes de operar eficientemente nesse domínio específico.

2.4.2.2 *Representação de saída de nuvem de pontos*

A escolha da representação de saída neste trabalho foi fundamentada na utilização de nuvens de pontos para descrever o perfil do sedimento presente no interior de tubulações industriais, a partir das cotas preditas pela rede neural, cotas que compõem o perfil do sedimento. Essa decisão está alinhada com as necessidades específicas do problema, uma vez que o objetivo principal não é reconstruir toda a superfície interna da tubulação, mas sim representar com precisão o contorno do acúmulo de sedimentos, ou seja, o seu perfil longitudinal.

As nuvens de pontos, amplamente utilizadas em tarefas de reconstrução 3D (Fan et al., 2017b; Lin et al., 2018), consistem em um conjunto de coordenadas tridimensionais (X, Y, Z) que descrevem pontos discretos no espaço. Neste contexto, as cotas (Z) estimadas pela rede neural correspondem à altura do sedimento ao longo de posições conhecidas da tubulação (X), sendo o eixo (Y) representado pelo diâmetro da tubulação. Essa abordagem apresenta vantagens significativas para o problema abordado, pois per-

mite gerar uma representação leve e eficiente da geometria do perfil sedimentar.

Por outro lado, abordagens baseadas em malhas triangulares (*meshes*) ou em *voxels* são mais adequadas para reconstruções completas de superfícies e volumes. As malhas triangulares, por exemplo, criam superfícies contínuas conectando os pontos em polígonos, resultando em modelos geométricos mais detalhados (Hu et al., 2021). No entanto, para o caso específico deste trabalho, que busca apenas representar o perfil do sedimento e não a superfície completa da tubulação, o uso de malhas tornaria o processo mais custoso computacionalmente sem ganhos significativos na qualidade da informação necessária.

Da mesma forma, a modelagem baseada em *voxels* gera representações volumétricas discretas e regulares da cena tridimensional (Liu and Liu, 2021). Apesar de sua utilidade em tarefas que requerem operações volumétricas ou segmentação espacial detalhada, a elevada demanda de memória e processamento torna a abordagem inviável para representar perfis unidimensionais ou bidimensionais como o caso da seção transversal do sedimento.

Assim, a representação por nuvem de pontos se mostrou mais adequada ao propósito da tese, pois proporciona uma descrição direta e eficiente do perfil de sedimentação ao longo da tubulação. Essa escolha permite ainda que as cotas inferidas possam ser facilmente interpretadas. Além disso, a simplicidade da nuvem de pontos facilita a comparação quantitativa com dados obtidos via escaneamento 3D, como foi adotado neste trabalho.

Para encerrar o capítulo, é fundamental destacar a importância dos métodos investigados e sua relevância para a solução proposta nesta pesquisa. A análise comparativa dos diferentes modelos e técnicas de processamento de imagens térmicas para a predição de sedimentação em tubulações industriais demonstra que a abordagem baseada em redes neurais convolucionais apresenta um potencial significativo para superar as limitações dos métodos tradicionais. A reconstrução do perfil do sedimento a partir de imagens térmicas bidimensionais, aliada a estratégias de aumento de dados e otimização de hiperparâmetros, mostra-se como uma alternativa viável e eficiente para estimar o perfil dos sedimentos sem a necessidade de sensores intrusivos. Além disso, os desafios inerentes à obtenção e curadoria do conjunto de dados podem ser mitigados por meio de técnicas de aumento de dados e aprimoramento da resolução das imagens, garantindo um melhor desempenho do modelo preditivo. Os resultados obtidos evidenciam que a combinação de termografia e aprendizado profundo pode fornecer uma ferramenta inovadora para o monitoramento preditivo de tubulações industriais, abrindo novas perspectivas para aplicações em manutenção preditiva e otimização de processos industriais.

3 TRABALHOS RELACIONADOS

O objetivo desta seção é apresentar um levantamento bibliográfico envolvendo técnicas não invasivas utilizadas para inspeção e detecção de obstruções em dutos. Abordar o estado da arte em imagens térmicas para aplicações industriais e reconstrução 3D a partir de imagens térmicas 2D.

As políticas de manutenção podem ser categorizadas em corretivas, preventivas, preditivas e proativas (Iqbal et al., 2016). As políticas de manutenção corretivas aguardam a ocorrência da falha para reparo posterior. Com melhorias nas técnicas de inspeção, foi desenvolvida a manutenção preditiva e proativa ou baseada em riscos. As rotinas de manutenção são essenciais em qualquer setor para garantir operações confiáveis de ativos. A manutenção se propõe a prosseguir adequadamente com a operação, reduzindo os tempos de falha e reduzindo os custos de reparo (Alotaibi et al., 2021).

A Inspeção Não Invasiva (INI) ou Ensaio Não Destrutivo (END), tornou-se uma ferramenta fundamental nas estratégias modernas de manutenção industrial (Alotaibi et al., 2021). O END é definido como um exame, teste ou avaliação realizado em qualquer tipo de objeto sem alterar sua integridade física para determinar a ausência ou presença de condições que possam afetar a utilidade de um objeto (Hellier, 2020). Na literatura, existem algumas técnicas para realizar o monitoramento e análise não invasiva do material depositado ao longo das tubulações. De acordo com (Wegelin, 2014), as formas de avaliação podem ser divididas em técnicas de hardware e software. As técnicas de software utilizam sensores presentes na tubulação para monitoramento operacional, que fornecem informações como pressão, vazão e temperatura. As técnicas de hardware usam sensores específicos para identificar falhas que são exemplificadas pelo uso de sensores acústicos, ópticos e ultrassônicos.

A análise não invasiva pode ser dividida em técnicas baseadas em contato e sem contato (Alotaibi et al., 2021). As técnicas baseadas em contato usam ferramentas como vibração, corrente alternada, tensão e análise de fluxo (Alotaibi et al., 2021; Wegelin, 2014). Técnicas sem contato usam laser, imagem, acústica, termografia e radar (Alotaibi et al., 2021). São elementos que podem ser usados externamente ao tubo, sem entrar em contato com ele.

Existem também técnicas híbridas, combinando os dois métodos. Um método baseado em ultrassom para inspeção de dutos foi proposto em (Piao et al., 2022), onde o acúmulo de lodo dentro de dutos industriais pode ser estimado em tempo real, sem interromper o processo químico em andamento. Em geral, transdutores ultrassônicos foram instalados nas paredes externas das tubulações. Com isso, foi identificado o volume de material indesejado no interior dos tubos, analisando a amplitude de onda longitudinal entre transdutores posicionados em lados opostos de uma tubulação. Este caso é um exemplo de uma combinação de classificação de tipo de hardware com contato. Hardware porque foi usado um sensor específico e com contato porque ficou preso no tubo.

Um projeto piloto em ambiente industrial controlado, onde foi orientado a verificar a abordagem utilizando quimiometria acústica no processo de granulação industrial, foi apresentado em (Halstensen et al., 2006). Nesta abordagem, sensores não intrusivos foram utilizados para capturar as características de vibração do sistema, que são basicamente o ruído produzido pelos equipamentos nos processos de fabricação e transporte de ureia. Essa técnica pode identificar a formação de grumos no fundo do reator, permitindo um alerta antecipado ao operador, que pode tomar medidas para retornar o granulador à operação nominal.

Outra abordagem não invasiva foi apresentada em (Zhao et al., 2021), onde foi proposto um sistema de medição de vazão de fertilizante baseado em capacitância. Os capacitores foram posicionados em lados opostos da tubulação de fertilizantes, o que permitiu detectar informações de fluxo de massa e velocidade do material, além de identificar quando há bloqueios.

A obstrução do duto também foi analisada no (Lile et al., 2012), onde foi observada a relação entre os níveis de bloqueio dentro de uma tubulação e o padrão de vibração do corpo vibratório da tubulação. Mudanças nos parâmetros de vibração foram notadas à medida que o tamanho do bloco variava. Quanto maior o bloqueio, mais turbulento era o fluxo. A velocidade da água dentro do tubo aumentou à medida que a área de fluxo diminuiu. Mudanças na velocidade e pressão da água dentro da tubulação aumentaram a vibração da tubulação circular.

Para a indústria, a temperatura é um indicador viável do estado operacional dos equipamentos. A termografia infravermelha (TRI) detecta a energia infravermelha emitida por um objeto, que é traduzida em uma temperatura aparente do objeto monitorado (Fang and Maldague, 2020). Assim, a TRI é uma técnica comumente utilizada como identificador não invasivo para equipamentos e sistemas industriais (Fang and Maldague, 2022).

Devido ao avanço na tecnologia de câmeras infravermelhas e sua progressiva redução de custos, a imagem térmica tem sido amplamente utilizada para inspeção de aplicações elétricas (Manohar and di Scalea, 2014), mecânicas (Alfredo Osornio-Rios et al., 2019; Cadelano et al., 2016) e outras. Um método baseado em termografia para estimar águas residuais em dutos foi apresentado em (Ishikawa et al., 2022). Este estudo mostrou que a

temperatura na superfície da água está diretamente relacionada à altura do efluente dentro de uma tubulação. Este método foi comparado experimentalmente com um método de estimativa numérica e mostrou uma estimativa de água adequada.

Recentemente, a fusão de informações geométricas e térmicas recebeu atenção para estender a imagem térmica de bidimensional (2D) para tridimensional (3D). A vantagem de mesclar imagens é a capacidade de combinar dados de temperatura com dados geométricos (Rangel and Soldan, 2014; Vidas et al., 2015; Yang et al., 2018). Um framework para mesclar modelos CAD 3D com imagens infravermelhas foi apresentado em (Meng et al., 2019). Neste modelo de avaliação não destrutiva, foi feita uma combinação de termografia ativa e imagem CAD 3D através de interpolação bilinear. Um experimento foi realizado para identificar compósitos revestidos curvos com defeitos subsuperficiais. Pontos artificiais foram marcados como referências para obter uma matriz de mapeamento de dados de imagem térmica (2D) para o modelo de coordenadas espaciais 3D digital. Um modelo digital 3D foi transformado em um novo sistema de coordenadas e então reprojado exatamente na imagem do recurso de onda térmica. A mesma relação de modelos CAD 3D com imagens térmicas também foi usada no (Sels et al., 2019a). Este modelo permitiu identificar automaticamente falhas no modelo CAD 3D, que foram mapeadas em imagem térmica 2D. Essa metodologia apresentou avanços relevantes na literatura. No entanto, sua desvantagem é a necessidade de pré-alinhar manualmente as imagens.

Outro método foi proposto em (Saifi et al., 2015), onde a termografia infravermelha foi aplicada para determinar espessuras de escala em tubos de aço, enquanto um método de elementos finitos 3D foi implementado para identificar incrustações de carbonato de cálcio nas paredes internas de tubos de água. Os resultados mostraram que a resposta térmica da imagem do tubo aumenta com o aumento da incrustação.

Por fim, no que diz respeito à manutenção industrial, um sistema de Realidade Aumentada para alinhar automaticamente informações virtuais e temperatura em qualquer objeto 3D foi apresentado em (Ortega et al., 2021). Este sistema utiliza um sensor RGB-D e uma câmera infravermelha para realizar esta tarefa em tempo real.

As tecnologias previamente discutidas são classificadas como não intrusivas, o que significa que sua aplicação em ambiente industrial não requer a interrupção do processo produtivo nem impõe tempo de inatividade na linha de produção. Para uma análise mais detalhada, as Tabelas 5 e 6 apresentam uma comparação entre as metodologias utilizadas para detectar obstruções em tubulações.

Além das informações já comparadas, será adicionada uma coluna de classificação de similaridade com a abordagem desta tese, referente ao uso de imagens térmicas para a identificação de perfis de sedimentos internos na tubulação. Essa classificação será realizada em uma escala de 1 a 5, onde 1 indica baixa similaridade e 5 indica alta similaridade. O critério para essa pontuação está apresentado na Tabela 4 e será baseado nos seguintes aspectos:

- O método faz uso de câmeras infravermelhas para capturar variações de temperatura na superfície da tubulação.
- A técnica é capaz de detectar bloqueios ou acúmulos de material dentro da tubulação.
- O método permite estimar a espessura ou altura do material sedimentado.
- A abordagem gera uma representação tridimensional do acúmulo interno na tubulação.
- O processamento de dados inclui a aplicação de CNNs para análise e predição da sedimentação.

Para cada critério atendido, a técnica receberá 1 ponto, e o somatório dessas pontuações resultará na nota final de similaridade. Dessa forma, será possível comparar quantitativamente as diferentes abordagens em relação à solução proposta, evidenciando seus pontos fortes e limitações. As Tabelas 5 e 6 incluirão vantagens e desvantagens de cada método, permitindo uma avaliação crítica das características de cada abordagem para a inspeção não invasiva de tubulações industriais.

Tabela 4: Classificação de similaridade.

Autores/Aspectos	Utiliza imagens térmicas	Identifica obstrução no interior da tubulação	Identifica a altura da obstrução no interior da tubulação	Reconstrução em 3D da obstrução no interior da tubulação	Utiliza redes neurais convolucionais	Similaridade
(Piao et al., 2022)	0	1	1	0	0	2
(Ishikawa et al., 2022)	1	1	1	0	0	3
(Zhao et al., 2021)	0	1	0	0	0	1
(Meng et al., 2019)	1	0	0	0	0	1
(Lile et al., 2012)	0	1	0	0	0	1
(Halstensen et al., 2006)	0	1	0	0	0	1
(Saifi et al., 2015)	1	1	0	0	0	2

Tabela 5: Resumo das técnicas de inspeção de tubulações não invasivas (a).

Referência	Classe (1-5)	Tecnologia	Identificação	Metodologia	Observações
(Piao et al., 2022)	2	Ultrassom	Lodo	Instalação de transdutores ultrassônicos em paredes externas de dutos, posicionados em lados opostos, analisando amplitude de onda longitudinal entre transdutores.	Exige transdutores fixados nas paredes da tubulação, a termografia infravermelha pode ser realizada externamente, sem a necessidade de contato físico. Enquanto o ultrassom mede pontos específicos, a termografia pode mapear áreas extensas rapidamente.
(Ishikawa et al., 2022)	3	Termografia	Desperdício de água	Uso da termografia para estimar resíduos de água em tubulações, mostrando que a temperatura na superfície da água está diretamente relacionada à altura da água residual dentro da tubulação.	Este estudo utiliza a termografia para estimar águas residuais em dutos, mostrando uma relação direta entre a temperatura da superfície da água e a altura do efluente. Embora não se concentre diretamente em sedimentos sólidos, a técnica é aplicável à detecção de variações no interior das tubulações.
(Zhao et al., 2021)	1	Capacitores	Sedimento	Sistema de medição de fluxo de fertilizante baseado em capacitância, posicionando capacitores em lados opostos de uma tubulação para detectar fluxo de massa e velocidade do material.	Estes estudos utilizam técnicas como quimiometria acústica e capacitância para monitorar processos industriais e detectar obstruções. Embora relevantes para o monitoramento de tubulações, eles não utilizam termografia infravermelha e têm uma aplicação mais geral
(Meng et al., 2019)	1	Termografia	Compósitos	Uso da termografia para identificar compósitos revestidos curvos com defeitos subsuperficiais.	Estes estudos focam na fusão de imagens térmicas com modelos CAD 3D para inspeção de falhas em materiais. Embora a termografia seja utilizada, a aplicação é mais voltada para a detecção de danos em materiais do que para a identificação de sedimentos em tubulações.

Tabela 6: Resumo das técnicas de inspeção de tubulações não invasivas (b).

Referência	Classe	Tecnologia	Identificação	Metodologia	Observações
(Lile et al., 2012)	1	Vibração	Obstrução	Análise da mudança de vibração da tubulação, com base nas mudanças na velocidade e pressão da água dentro da tubulação.	A análise de vibração demonstra que existe uma relação direta entre o tamanho do bloqueio e os padrões de vibração. No entanto, não identifica formato do elemento que está bloqueando a tubulação
(Halstensen et al., 2006)	1	Ultrassom	Formação de grumos no fundo do reator	Quimiometria acústica, sensores não intrusivos para capturar as características de vibração do sistema.	Estes estudos utilizam técnicas como quimiometria acústica e capacitância para monitorar processos industriais e detectar obstruções. Embora relevantes para o monitoramento de tubulações, eles não utilizam termografia infravermelha e têm uma aplicação mais geral
(Saifi et al., 2015)	2	Termografia	Incrustação	Uso de termografia para identificar incrustações dentro de tubulações.	O estudo aplica diretamente a termografia infravermelha para determinar espessuras de escala em tubos de aço e utiliza um método de elementos finitos 3D para identificar incrustações de carbonato de cálcio. Portanto, está relacionado à identificação de sedimentos.

Com base na análise comparativa das diferentes técnicas de inspeção não invasiva, observa-se que a termografia infravermelha, em particular sua aplicação na identificação de obstruções em tubulações, apresenta vantagens significativas em relação a outras abordagens. Enquanto métodos baseados em ultrassom, vibração e capacitância demonstram eficácia na detecção de anomalias estruturais e variações de fluxo, eles carecem da capacidade de fornecer informações detalhadas sobre a distribuição espacial dos sedimentos no interior da tubulação. Além disso, técnicas que utilizam modelos CAD 3D e fusão de imagens térmicas são amplamente aplicáveis em inspeções estruturais, mas não necessariamente voltadas para o monitoramento interno de tubulações.

3.1 Combinação de imagem infravermelha e informações de geometria 3D

O avanço das técnicas de inspeção e monitoramento industrial tem impulsionado o desenvolvimento de metodologias que integram múltiplas fontes de dados para aprimorar a análise de componentes e estruturas. Entre essas abordagens, a fusão de imagens infravermelhas com informações geométricas 3D tem sido explorada como uma estratégia para melhorar a precisão da caracterização de objetos e materiais. Essa integração possibilita não apenas a captura das propriedades térmicas dos elementos inspecionados, mas também a representação espacial detalhada de suas geometrias, permitindo uma avaliação mais completa e precisa do objeto sob análise (Kovács et al., 2020; Luo et al., 2019).

Diferentes estudos têm demonstrado que a utilização conjunta de sensores térmicos e técnicas de reconstrução 3D, contribuem significativamente para a identificação e diagnóstico de falhas estruturais, desgaste de superfícies e variações térmicas associadas a diferentes materiais (Schramm et al., 2014). A partir dessa fusão de informações, é possível gerar modelos 3D enriquecidos por dados térmicos, frequentemente denominados termogramas 3D. Esses modelos possibilitam a visualização integrada da distribuição de temperatura e da morfologia do objeto, favorecendo a detecção de anomalias, a estimativa da profundidade de defeitos subsuperficiais e a análise do comportamento térmico de superfícies complexas.

Além disso, metodologias que combinam sensores de imagem térmica com sistemas de escaneamento 3D, como LiDAR e câmeras estéreo, vêm sendo aplicadas em diferentes domínios, incluindo inspeção de estruturas industriais, monitoramento de processos térmicos e controle de qualidade em manufatura avançada. A combinação dessas tecnologias amplia a aplicabilidade da termografia infravermelha ao permitir um alinhamento preciso entre a distribuição térmica capturada e a geometria do objeto, tornando possível a reconstrução digital de superfícies e a extração de métricas quantitativas para análise automatizada (Schramm et al., 2014).

Um dispositivo portátil, chamado HeatWave, composto por uma câmera termográfica

infravermelha, uma câmera colorida e um sensor de alcance pode ser usado para medir e representar a temperatura da superfície em modelos 3D (Vidas et al., 2015). Este dispositivo captura dados para posterior processamento e alinhamento das imagens, gerando um termograma 3D, utilizado por um operador para inspecionar um objeto.

A fusão de imagens térmicas com nuvens de pontos de um scanner foi apresentada em (Zhu et al., 2021; Hellstein and Szwedo, 2016), que permite colorir a nuvem de pontos adquirida. Esta combinação foi utilizada para avaliar o consumo de energia e as fugas de energia de um edifício em (Zhu et al., 2021), onde foram identificados pontos-chave nas imagens para que posteriormente pudessem ser alinhados por um algoritmo de determinação de correspondência. O resultado é uma representação suave de nuvens de pontos térmicos. Em (Dua, 2020; Duarte et al., 2019), uma câmera térmica foi usada para mapear os pixels de sedimentos em tubos. A imagem térmica foi processada em uma rede Multi-Layer Perceptron (MLP) para identificar o sedimento interno na tubulação.

Termografia 3D foi usada para identificar compósitos através de diagnósticos não destrutivos de estruturas como barcos, aviões e turbinas eólicas em (Hellstein and Szwedo, 2016). Isso foi feito com um algoritmo de fusão de nuvens chamado KinfuLargeScale (Bondarev et al., 2013; Izadi et al., 2011; Rusu and Cousins, 2011), que usa uma GPU que mapeia temperaturas usando algoritmos de renderização e ray tracing. O mapeamento foi realizado em tempo real durante a visualização da imagem.

A combinação de modelos CAD com imagens térmicas também foi apresentado por (Sels et al., 2019b). Foi utilizado o algoritmo PwP3D (Prisacariu and Reid, 2012) para mapear o modelo CAD e sua imagem térmica, a fim de obter-se uma posição 3D de um defeito.

Uma rede de segmentação semântica (U-Net), foi utilizada para detectar defeitos através de uma sequência de imagens termográficas obtidas por termografia ativa em (Müller et al., 2022). Em seguida a rede LSTM foi usada para identificar aspectos da forma espacial e permitir a reconstrução de defeitos. A profundidade e a geometria dos defeitos foram identificadas usando unidades recorrentes fechadas (GRUs), que é uma variação do LSTM (Houdt et al., 2020), em amostras de materiais compósitos via termografia ativa em (Fang and Maldague, 2020).

A imagem térmica foi usada em estudos do sono para identificar a postura esquelética do paciente, mesmo com oclusões visuais, em (Transue et al., 2019). Nesse trabalho, redes neurais convolucionais foram utilizadas para caracterizar e identificar a volumetria térmica e indicar a estimativa da postura esquelética correspondente.

As tecnologias apresentadas estão relacionadas ao uso da termografia na manutenção industrial e o uso de termografia para identificar postura esquelética em pacientes. Para uma análise comparativa, a Tabela 8 apresenta uma comparação entre as metodologias utilizadas.

Além das informações já comparadas, será adicionada uma coluna de classificação

de similaridade com a abordagem baseada em imagens térmicas para a identificação de sedimentos na tubulação. Essa classificação será realizada em uma escala de 1 a 5, onde 1 indica baixa similaridade e 5 indica alta similaridade. O critério para essa pontuação está apresentado na Tabela 7 e será baseado nos seguintes aspectos:

- O método faz uso de câmeras infravermelhas para capturar variações de temperatura no objeto de estudo.
- A técnica é capaz de detectar objetos oclusos
- O método permite estimar a espessura ou altura do material ocluso.
- A abordagem gera uma representação 3D do objeto ocluso
- O processamento de dados inclui a aplicação de CNNs para análise e predição do objeto ocluso.

Para cada critério atendido, a técnica receberá 1 ponto, e o somatório dessas pontuações resultará na nota final de similaridade. Dessa forma, será possível comparar quantitativamente as diferentes abordagens em relação à solução proposta, evidenciando seus pontos fortes e limitações. A Tabela 7 inclui vantagens e desvantagens de cada método, permitindo uma avaliação crítica das características de cada abordagem para a inspeção não invasiva de tubulações industriais.

Tabela 7: Classificação de similaridade utilizando imagens térmicas.

Autores/Aspectos	Utiliza imagens térmicas	Identifica material ocluso	Identifica a altura do material ocluso	Reconstrução em 3D da oclusão	Utiliza redes neurais convolucionais	Similaridade
(Müller et al., 2022)	1	0	0	0	1	2
(Zhu et al., 2021)	1	0	0	0	0	1
(Schramm et al., 2014)	1	0	0	0	0	1
(Dua, 2020)	1	1	1	0	0	3
(Sels et al., 2019b)	1	0	0	0	0	1
(Transue et al., 2019)	1	1	0	0	1	3

Tabela 8: Comparação de estudos sobre termografia e reconstrução 3D

Referencia	Classe	Dispositivo	Metodologia	Identificação	Observação
(Müller et al., 2022)	2	Câmera térmica	Segmentação térmica de defeitos em estruturas usando aprendizado profundo.	Defeitos em materiais compósitos.	Uso de aprendizado profundo melhora a precisão, mas requer grande volume de dados anotados. Trata objetos visíveis.
(Zhu et al., 2021)	1	Câmera térmica + scanner 3D	Uso de scanner 3D para mapear superfícies e fusão com imagem térmica.	Fugas térmicas em edifícios.	Método que requer alinhamento cuidadoso entre imagem térmica e geometria 3D. Reconstrução 3D de fachadas de edifícios
(Schramm et al., 2014)	1	Câmera térmica + sensores adicionais	Fusão de imagem térmica com outros sensores para criar modelos 3D térmicos.	Objetos térmicos visíveis.	Apenas objetos visíveis foram considerados na análise térmica.
(Dua, 2020)	3	Câmera térmica	Uso de MLP para identificar padrões térmicos em tubulações.	Sedimentos em tubulações.	Oclusão interna nas tubulações detectada pela variação térmica.
(Sels et al., 2019b)	1	Câmera térmica + modelo CAD	Uso de modelos CAD para combinar geometrias conhecidas com dados térmicos.	Defeitos em modelos CAD.	Depende de CADs conhecidos para alinhamento térmico, dificultando o uso em ambientes desconhecidos.
(Transue et al., 2019)	3	Câmera térmica + redes neurais convolucionais	Mapeamento volumétrico de pacientes por imagens térmicas e CNNs.	Postura esquelética em pacientes.	Aplicação voltada para humanos, mas ilustra potencial para reconhecimento de formas térmicas.

A abordagem adotada por (Dua, 2020) baseia-se na utilização de janelas ou *kernels* da imagem térmica como entrada para a rede Multi-Layer Perceptron (MLP). Cada kernel é uma região específica da imagem térmica, cujos valores de pixels são processados como um vetor de entrada para a MLP. A saída da rede é composta por um único neurônio, cuja ativação representa a altura do sedimento naquela região específica da tubulação. Essa estratégia permite que o modelo aprenda a relação entre as variações de temperatura e a altura dos sedimentos, mas apresenta limitações em termos de generalização e captura de padrões espaciais globais. Como cada *kernel* é tratado de forma independente, a abordagem não considera a correlação entre regiões adjacentes da imagem térmica, o que pode resultar em inconsistências na reconstrução do perfil de sedimentação. Além disso, a MLP não explora a estrutura espacial das imagens, ao contrário das redes neurais convolucionais (CNNs), que extraem características hierárquicas e preservam relações espaciais entre os pixels.

A partir da análise comparativa das metodologias de inspeção não invasiva, observa-se que a combinação de imagens térmicas com informações geométricas 3D tem se destacado como uma abordagem promissora para aplicações industriais. A literatura revisada demonstra que técnicas baseadas em termografia infravermelha são eficazes na identificação de variações térmicas associadas a obstruções, sedimentos e defeitos estruturais, permitindo uma análise detalhada sem a necessidade de contato direto com os elementos inspecionados. Entretanto, a limitação inerente das imagens térmicas em representar a geometria tridimensional dos objetos monitorados evidencia a necessidade de fusão com técnicas de reconstrução 3D, como o uso de modelos CAD, nuvens de pontos ou escaneamento volumétrico.

A avaliação das metodologias também evidencia que a aplicação de redes neurais convolucionais (CNNs) na análise térmica pode aprimorar a interpretação dos padrões térmicos, permitindo inferências mais robustas sobre a presença e a distribuição de sedimentos dentro de tubulações industriais. No entanto, desafios permanecem, especialmente em relação à oclusão de elementos dentro da tubulação, à calibração precisa dos sensores térmicos e à necessidade de volumes de dados anotados para o treinamento de modelos baseados em aprendizado profundo.

Com base nesse levantamento, conclui-se que há um grande potencial na exploração de metodologias híbridas que combinem termografia infravermelha e reconstrução geométrica 3D para aprimorar a precisão na identificação e monitoramento de obstruções em tubulações industriais. Os avanços na fusão de dados térmicos e geométricos podem proporcionar soluções mais robustas e adaptáveis às complexidades do ambiente industrial, contribuindo para o desenvolvimento de sistemas de manutenção preditiva mais eficientes e autônomos.

Assim, este trabalho se fundamenta na integração dessas tecnologias, propondo uma solução inovadora baseada na análise térmica e reconstrução 3D para detecção e

caracterização de sedimentos, com potencial aplicação em diversas indústrias que dependem do monitoramento contínuo de suas infraestruturas.

A estimativa do nível de entupimento interno de tubulações industriais representa um desafio significativo, especialmente devido à complexidade dos fenômenos envolvidos na deposição de sedimentos e às dificuldades inerentes ao monitoramento interno. Diante desse cenário, esta tese propõe uma abordagem inovadora baseada na combinação de termografia infravermelha externa e redes neurais convolucionais (CNNs) para inferir a distribuição dos sedimentos no interior das tubulações.

Para viabilizar essa abordagem, a CNN deve ser treinada para estabelecer uma relação entre imagens térmicas capturadas externamente e diferentes perfis de sedimentação interna. O processo de treinamento envolverá a otimização dos hiperparâmetros da rede neural, de modo que o modelo aprenda a mapear variações térmicas na superfície da tubulação para as correspondentes alturas dos depósitos internos. Essa estratégia possibilita inferências automatizadas sem a necessidade de modelagem física detalhada dos fenômenos de deposição. No entanto, a precisão do modelo depende diretamente da disponibilidade de um conjunto de dados anotados que representem adequadamente a diversidade de padrões térmicos associados às diferentes configurações de sedimentos.

A obtenção de um conjunto de dados representativo constitui um dos principais desafios desta pesquisa. Entre as dificuldades enfrentadas, destacam-se a aquisição de dados anotados, para que a CNN aprenda padrões confiáveis, é necessário um conjunto de imagens térmicas associadas a medições precisas da altura do sedimento no interior da tubulação. Como essa informação não é diretamente observável, requer técnicas auxiliares para anotação, como sensores de referência ou modelos geométricos. Dados sintéticos são utilizados para complementar conjuntos de dados reais e aumentar a variabilidade das amostras. No entanto, assegurar que as amostras sintéticas reproduzam com fidelidade as características térmicas do ambiente industrial real é um desafio técnico relevante. O modelo treinado deve ser capaz de generalizar para diferentes condições operacionais sem superajustar-se a um conjunto específico de dados. Para isso, o *dataset* deve ser suficientemente amplo e diversificado nos perfis de sedimentação.

Para superar esses desafios, esta tese propõe uma metodologia estruturada para a construção do conjunto de dados, incluindo um *setup* experimental que simula as condições da planta industrial em ambiente controlado. Esse ambiente permitirá a coleta sistemática de imagens térmicas enquanto se monitora tanto o comportamento térmico da tubulação quanto a geometria dos depósitos internos. A partir desse *setup*, será possível gerar um *dataset* robusto para o treinamento da CNN, viabilizando uma solução eficiente para o monitoramento não invasivo de sedimentação em tubulações industriais.

4 UMA METODOLOGIA PARA CONSTRUÇÃO DE *DATA-SET* PARA AJUSTE DO MODELO

Tendo em vista a dificuldade de inferir o nível de entupimento interno da tubulação, seja através de sensores *in loco*, seja através de modelos preditivos de escoamento e depósito, esta tese propõe o uso de termografia externa associada a Redes Neurais Artificiais (RNA) para estimar o perfil de depósito de sedimento interno. Para tal as RNAs deverão ser treinadas de forma a mapear imagens térmicas em diferentes níveis de cotas de perfis de sedimento. O treinamento resultará no ajuste dos hiperparâmetros da rede que conduzirão a estimativas assertivas.

Se por um lado o uso de RNA reduz a necessidade de entendimento físico do fenômeno de depósito, por outro exige a necessidade, para treinamento da rede, de um conjunto de dados anotados suficientes para representar a diversidade de possibilidades de perfis de depósitos (*ground thuth*) e sua representação térmica externa a ser capturada pela câmera.

Obter este conjunto de dados é um desafio, tendo em vista que o acesso ao interior da tubulação em ambientes industriais é extremamente restrito devido a questões de segurança, limitações operacionais e custos elevados. As tubulações geralmente transportam materiais em condições adversas, como altas temperaturas, pressão elevada ou ambientes corrosivos, tornando inviável a inserção de sensores ou a realização de inspeções manuais diretas.

Além disso, a necessidade de interromper o processo produtivo para realizar medições internas gera impactos econômicos consideráveis. Outra dificuldade adicional reside na geometria confinada e na presença de obstruções, o que limita o alcance de equipamentos convencionais de medição, como réguas ou dispositivos ópticos.

Por fim, a variabilidade dos perfis de sedimentação ao longo do tempo e das condições operacionais torna difícil capturar uma amostragem representativa e abrangente apenas com medições pontuais ou em situações controladas. Estas restrições justificam a busca por soluções alternativas, como a abordagem proposta nesta tese, baseada na análise de imagens térmicas externas e aprendizado de máquina.

Nesta tese propõe-se uma metodologia para a obtenção de datasets que permitam o

ajuste (aprendizado) dos hiperparâmetros do modelo neural.

As tubulações do ambiente industrial estão localizadas em ambiente interno e também em ambiente externo. Estas tubulações estão sujeitas aos intempéries do tempo, da umidade, do sol, da poeira, das mudanças de temperatura, além contaminação, corrosão, incrustações externas e vibrações. Estas são algumas das variáveis que compõem o ambiente real. Devido as dificuldades de modelar os fenômenos físicos envolvidos neste ambiente, uma alternativa é utilizar um ambiente em escala e controlado, a fim de estimar a relação das imagens térmicas capturadas e o perfil de sedimento interno na tubulação. A ideia é fazer uso de uma maquete em escala que reproduza os fenômenos que ocorrem na planta e permita aferir o comportamento térmico externo à tubulação e a geometria interna do depósito.

Assim, o *setup* para a criação do dataset proposto será baseado em uma maquete que reproduzirá em ambiente controlado as condições da planta, podendo ser monitorada tanto o comportamento térmico da tubulação quanto o perfil geométrico do depósito. A maquete se constituirá de uma tubulação em aço galvanizado revestido por tinta térmica da cor preta. Este revestimento bloqueia o reflexo gerado pelo infravermelho emitido da câmera termográfica. Uma vez que, segundo a lei de Stefan Boltzmann, discutida no Capítulo 2.1, um corpo negro é definido como um meio ou substância que absorve toda energia incidente sobre ele, nenhuma parte da radiação incidente é refletida ou transmitida. Desta forma, a assinatura térmica da tubulação obterá um grau de confiabilidade e maior precisão.

No interior desta tubulação será depositado o fertilizante com diferentes perfis geométricos de depósito. Estes perfis são aferidos por laser de alta precisão, já que a maquete pode ser facilmente monitorada. O interior da tubulação com a presença do sedimento é aquecido, e uma câmera térmica captura o perfil térmico externo dos fenômenos envolvidos.

O sedimento deve ser elaborado com fertilizante real proveniente da planta industrial, garantindo a representatividade das condições reais de operação. Devem ser criados os perfis de sedimentos, que reflitam as condições observadas no ambiente industrial real, baseando-se em relatos fornecidos por trabalhadores da indústria de fertilizantes. Para esta tese, esses relatos destacaram padrões típicos de sedimentação, como acúmulos em áreas de baixa circulação, formação de camadas compactas e distribuições irregulares ao longo do comprimento da tubulação. Devido à influência da gravidade, os sedimentos tendem a se depositar na parte inferior da tubulação. Essa abordagem permite uma diversidade dos perfis encontrados em operações industriais, proporcionando uma base para o desenvolvimento de métodos de análise e predição baseados em visão computacional. Para a validação e anotação dos perfis sedimentares gerados experimentalmente, é utilizado um *scanner a laser 3D*, capaz de capturar a geometria do sedimento com alta precisão. Esse procedimento permite a correlação entre as variações térmicas observadas

externamente e a distribuição volumétrica dos sedimentos no interior do duto.

A seguir é detalhada a metodologia para o *setup* envolvendo tanto o projeto da maquete, como as tecnologias para captura de imagem térmica.

4.1 Metodologia para a construção do *setup*

Propõe-se um modelo em escala que reproduza as condições de depósito de fertilizantes em tubulações industriais que possa ser percebido por câmera térmica. A Figura 9 apresenta um esquema da solução proposta, enquanto a Figura 10 apresenta a imagem real da tubulação com o soprador térmico posicionado em uma extremidade da tubulação e a câmera térmica posicionada na parte inferior à tubulação.

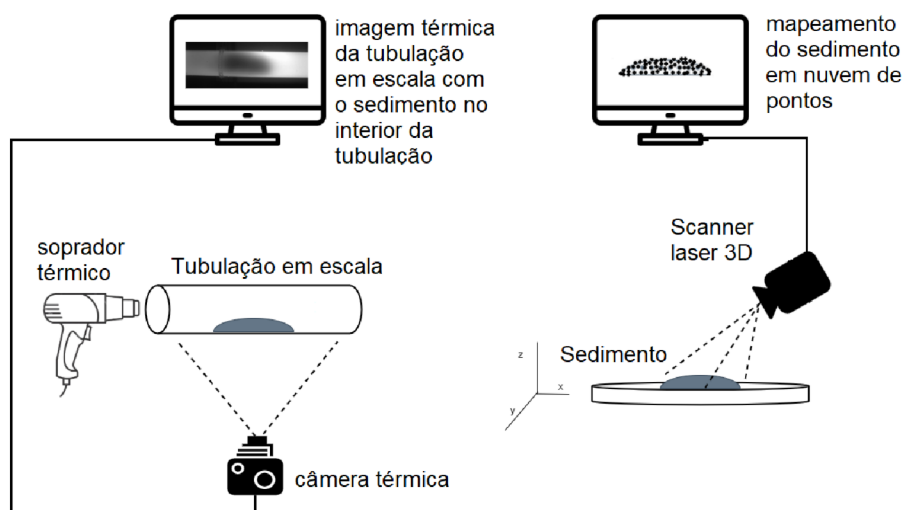


Figura 9: Esquema do ambiente controlado em escala. Fonte: Elaborado pela autora



Figura 10: Visão da configuração real para obtenção da imagem térmica. Fonte: Elaborado pela autora

O *setup* experimental baseia-se na geração de um fluxo de calor que percorre uma tubulação em escala, simulando as condições térmicas presentes no ambiente industrial. Utiliza-se uma mistura de fertilizante NPK da planta industrial, conforme a Figura 11.



Figura 11: Fertilizante NPK granulado. Fonte: Elaborado pela autora

No interior da tubulação, aloja-se um perfil de sedimento composto por uma mistura de 300g de fertilizante NPK e 200ml de água, replicando os padrões de umidade observados na planta industrial. O fertilizante e a água são misturados manualmente gerando um bloco de sedimento. Este bloco tem 30 cm de comprimento, largura igual ao diâmetro do tubo e alturas variáveis, conforme a Figura 12



Figura 12: Fertilizante úmido no interior do tubo.

Ao inserir uma amostra na tubulação, a extremidade da tubulação oposta ao soprador

é fechada. O soprador térmico é posicionado em uma extremidade do tubo para simular o fluxo de ar quente em seu interior, operando a uma temperatura de 60°C. A extremidade oposta do tubo é selada para permitir que o sistema atinja a estabilidade térmica, o que ocorre aproximadamente 40 minutos após o início do aquecimento pelo soprador térmico. A estabilização térmica é monitorada e confirmada usando o software da câmera térmica.

Obtem-se a imagem termográfica apontando a câmera para o duto totalmente preto fosco. A imagem térmica da base do sedimento, o qual consideramos o ângulo 0° entre base do sedimento e o posicionamento da câmera. A câmera térmica está estrategicamente posicionada para capturar a imagem externa da parte inferior do tubo, onde o sedimento tende a se acumular, de acordo com a configuração real apresentada na Figura 10. Essa decisão do posicionamento da câmera é baseada em relatos de trabalhadores da indústria de fertilizantes que observaram que, devido à gravidade, o material tende a se depositar predominantemente no fundo da tubulação. A câmera está posicionada a uma distância de aproximadamente 80 cm da tubulação, e essa distância é verificada usando o software da câmera, que indica automaticamente quando ocorrem distorções geométricas devido à proximidade excessiva. Esse procedimento garante que as imagens térmicas registradas preservem a fidelidade geométrica necessária para análises posteriores.

Dessa forma, o experimento permite observar como o acúmulo de material interfere no fluxo de calor e na distribuição térmica ao longo da tubulação. Essa abordagem possibilita correlacionar padrões térmicos externos com a presença e a altura do sedimento interno à tubulação.

Após a estabilização térmica, as assinaturas térmicas da superfície externa da tubulação são capturadas usando a câmera térmica.

A Figura 13 mostra um exemplo de uma imagem termográfica infravermelha do modelo em escala do duto vazio. A Figura 14 mostra o mesmo duto com sedimento de fertilizante alocado.

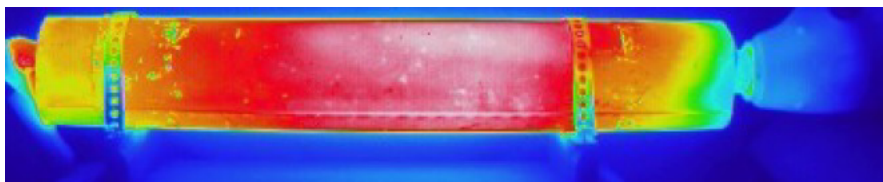


Figura 13: Imagem térmica quando a tubulação está vazia. Fonte: Elaborada pela autora.

É possível observar na Figura 14 que a temperatura da tubulação diminui com a presença do sedimento em seu interior, o que leva à percepção de que o fertilizante atua como barreira térmica. Isso indica que quando a temperatura no duto diminui, ocorre um acúmulo indesejável de sedimento, que pode levar a uma situação adversa, ao ponto de afetar toda a operação de produção de fertilizantes.

Após a conclusão da aquisição da imagem, desliga-se o soprador térmico. Aguarda-se

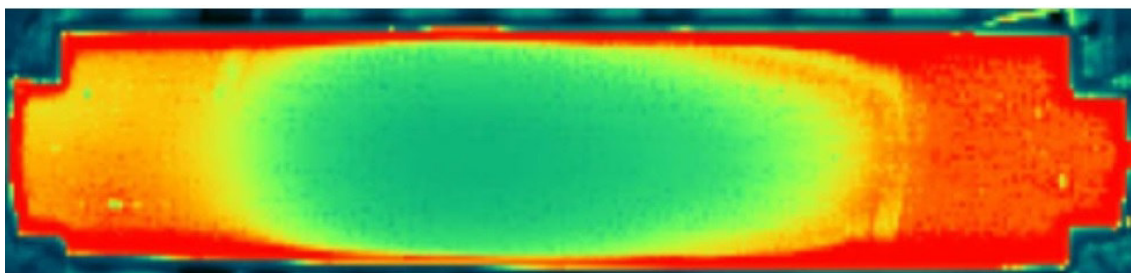


Figura 14: Imagem térmica da tubulação com sedimento no interior da tubulação. Fonte: Elaborada pela autora.

o resfriamento da tubulação e do sedimento. Destaca-se que um ponto do *setup* é permitir o acesso ao depósito formado na tubulação. Nesta proposta, as extremidades da tubulação são abertas, desta forma, após se aglutinar o sedimento de forma sólida, este pode ser facilmente retirado do interior da tubulação.

A Figura 15 apresenta uma exemplificação de um conjunto de amostras em temperatura ambiente, já removidas do processo de umidificação, aglutinação e imageamento térmico. Diferentes amostras são criadas de forma a representar diferentes perfis sedimentares.



Figura 15: Amostras de sedimentos. Fonte: Elaborado pela Autora

Após o resfriamento, remove-se o sedimento do interior da tubulação. O mapeamento do sedimento é feito com um scanner 3D, conforme Figura 16. O scanner está posicionado com um ângulo de 15° a 30cm de distancia da mesa giratória onde está o sedimento a ser mapeado.

O mapeamento é acompanhado pelo software do scanner. O scaneamento permite a reconstrução detalhada da estrutura do sedimento, conforme a Figura 17.

Após a realização do escaneamento de todas as faces do sedimento de fertilizante é gerado um arquivo em formato *.obj*, com faces e vértices do sedimento. Este arquivo de saída contém uma lista de vértices do polígono, que são em torno de 88.000 vértices, Dependendo do volume do sedimento, este arquivo poderá ter maior ou menor quantidade de vértices. A representação é dada pela seguinte correspondência: coordenada x indica a posição do sedimento no sentido longitudinal da tubulação, a coordenada y indica o diâmetro da tubulação, e a coordenada z indica a altura do sedimento. A quantidade

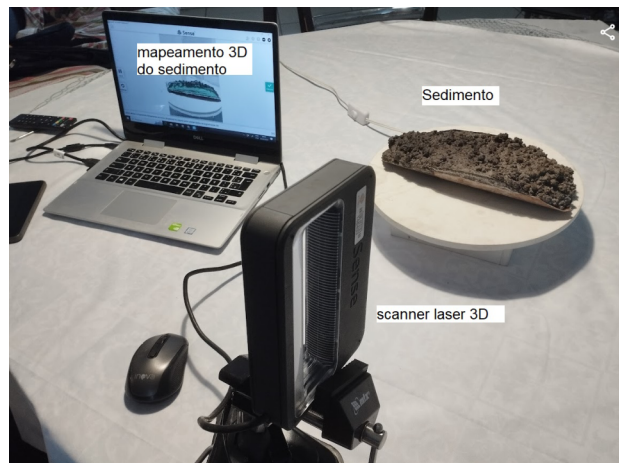


Figura 16: Visão da configuração real do mapeamento do sedimento. Fonte: Elaborado pela autora



Figura 17: Sedimento digitalizado. Fonte: Elaborado pela autora

de vértices mapeados pelo scanner impacta diretamente na quantidade de parâmetros necessários na rede neural. Para este exemplo de 88.000 vértices, seriam necessários 88.000 saídas na rede. Cada vértice é um ponto tridimensional e cada face é um triângulo formado por três pontos, ou seja, a superfície é formada por polígonos. Os vértices são as coordenadas $(x; y; z)$ e as faces são listas de três vértices cada.

A Figura 18 apresenta uma perspectiva do conceito sobre polígonos gerados para reconstrução tridimensional do objeto, onde é possível observar que o sedimento é formado por uma malha tridimensional constituída de diversos polígonos.

4.1.1 Pré-processamento dos Dados

Os dados coletados são pré-processados para garantir consistência e usabilidade. As imagens térmicas são normalizadas, e ruídos reduzidos para melhorar a visibilidade das características relevantes.

No processamento do arquivo no formato .obj, que contém a saída do escaneamento 3D do sedimento, é necessário realizar uma filtragem dos dados para manter apenas as informações referentes aos vértices. Esses vértices representam as coordenadas espaciais

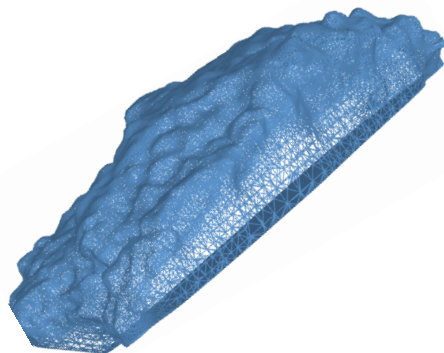


Figura 18: Reconstrução 3D do sedimento a partir de polígonos. Fonte: Elaborado pela autora

que definem a geometria do sedimento. No entanto, para garantir a consistência dos dados e a compatibilidade com o modelo preditivo, não se utiliza a totalidade dos vértices presentes no arquivo original. Em vez disso, aplica-se uma técnica de amostragem que extrai um número fixo de vértices distribuídos ao longo do perfil, representando de forma padronizada a quantidade de cotas (altura) do sedimento em cada amostra. Essa padronização é fundamental para alinhar as informações geométricas com os dados térmicos, permitindo o treinamento eficiente e coerente da rede neural.

4.1.2 Rotulagem dos Dados

A rotulagem dos dados é uma etapa importante no desenvolvimento do modelo preditivo proposto, uma vez que permite associar cada imagem térmica capturada à sua respectiva representação geométrica em forma de perfil de sedimento. Para isso, cada amostra física digitalizada por meio do scanner 3D, possui sua correspondente imagem térmica. Com essa abordagem, cada imagem térmica da superfície externa da tubulação, obtida com a presença de uma amostra de sedimento específica, é associada a um vetor numérico que representa a altura do perfil correspondente em seu interior. Esse emparelhamento de imagem e vetor de cotas forma a base do processo supervisionado de aprendizado, no qual a rede é treinada a prever o perfil do sedimento localizado no interior da tubulação a partir da imagem térmica externa à tubulação. A precisão dessa rotulagem é importante para a qualidade do modelo, já que erros nessa correspondência impactam diretamente na capacidade da rede de generalizar para novos exemplos. O processo garante que cada amostra térmica esteja vinculada a uma única resposta geométrica padronizada, assegurando consistência nos dados utilizados para treinamento, validação e teste do modelo. Esse processo assegura o alinhamento adequado entre os dados térmicos e os perfis geométricos.

4.2 Aplicação da metodologia para a obtenção do *dataset*

A partir do setup, elabora-se diferentes perfis de sedimento de forma manual, a fim de representar diferentes estágios de obstrução na tubulação. A Figura 19 apresenta uma exemplificação da relação entre o sedimento e o seu respectivo perfil.

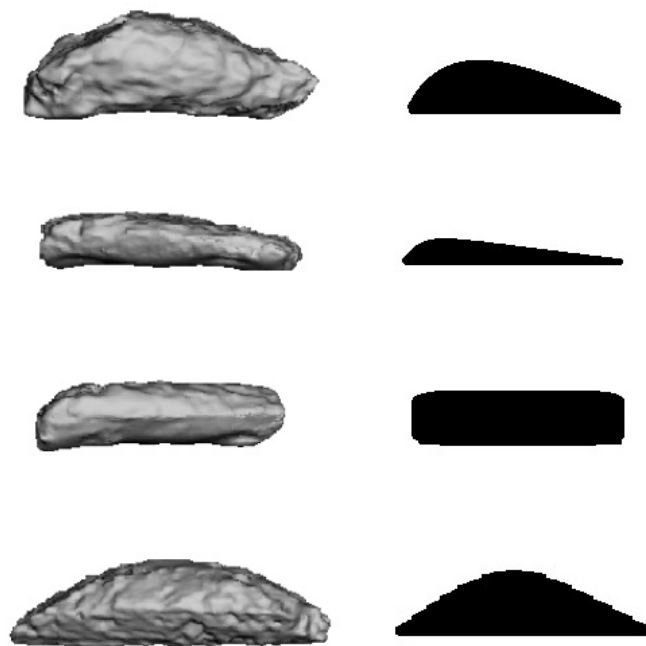


Figura 19: Exemplificação da relação entre o sedimento e o seu perfil. Fonte: Elaborado pela Autora

Na imagem térmica, estabelece-se a região de interesse do estudo, conforme está representada na Figura 20 onde está localizado o sedimento de fertilizantes.

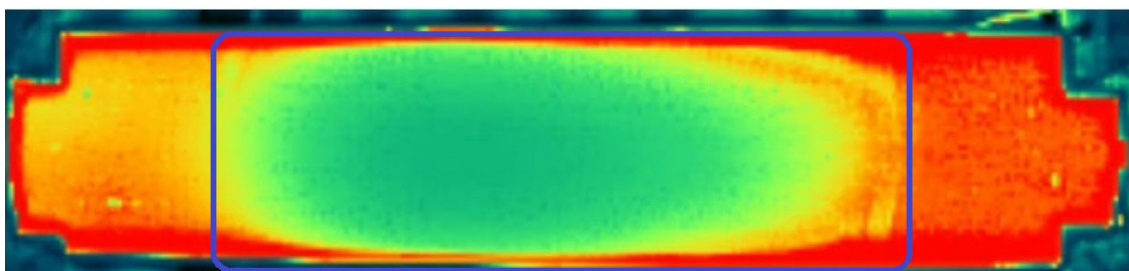


Figura 20: Imagem térmica da tubulação - região de interesse. Fonte: Elaborada pela autora.

4.2.1 Aumento de dados

Para garantir maior robustez ao modelo preditivo e compensar a limitação de um número reduzido de amostras reais, aplica-se técnicas de aumento de dados (data aug-

mentation) nas imagens térmicas. Inicialmente, as imagens térmicas passam por recortes de regiões de interesse (ROI), focando na área inferior da tubulação, onde ocorre o acúmulo de sedimentos, conforme Figura 20. Em seguida, é aplicada a técnica de super-resolução por redes neurais (ESRGAN – Enhanced Super Resolution Generative Adversarial Networks), utilizando o modelo RRDB-ESRGAN-x4. Esse processo resulta em imagens com maior resolução espacial, permitindo ao modelo capturar mais detalhes das variações térmicas.

Além da super-resolução, são aplicados filtros (Gonzalez and Woods, 2002) de ruído, desfoque e interferência, simulando condições adversas encontradas em ambientes industriais reais, como vibração mecânica, reflexos térmicos e interferência eletromagnética.

A adição de ruído e desfoque às imagens térmicas simula condições reais de captura, onde a qualidade da imagem pode ser afetada por vibrações, movimento da câmera ou limitações do equipamento. Pequenos deslocamentos da câmera podem resultar em imagens borradas. O ruído branco é uma perturbação aleatória na intensidade dos pixels, introduzida para simular ruídos térmicos no sensor da câmera infravermelha. Essas condições podem afetar a qualidade das imagens capturadas por sensores infravermelhos em campo. A aplicação desses ruídos de forma controlada permite simular tais variações sem comprometer a integridade dos dados, tornando o modelo mais resiliente a ruídos operacionais. Essas técnicas de aumento de dados replicam, de forma sintética, a diversidade de situações que o sistema pode encontrar durante a operação em ambientes industriais, melhorando sua capacidade de generalização e reduzindo a dependência de coleta massiva de dados reais.

4.3 Aplicação de metodologia em um estudo de caso

A metodologia proposta foi aplicada na construção do dataset que foi utilizado para o ajuste do modelo proposto nesta tese. A metodologia envolveu um conjunto de dispositivos existentes no laboratório.

A seguir, cada um dos componentes do sistema é especificado, apresentando as tecnologias envolvidas na respectiva etapa.

- Equipamento de captura termográfica InfiRay P2 Pro - para aquisição das imagens termográficas externas à tubulação, optou-se por utilizar câmera térmica InfiRay P2 Pro, com resolução de 256×192 pixels e tamanho de pixel de 12 μm (micrômetro), a câmera oferece alta sensibilidade térmica. Opera a uma taxa de quadros de 25 Hz, cobre uma faixa de temperatura entre -20°C e 550°C , com precisão de $\pm 2^{\circ}\text{C}$ e suporta diversas paletas de cores para análise térmica
- Scanner portátil a Laser 3D - com finalidade de obter a anotação (*ground truth*) dos dados para treinamento, busca-se extrair o perfil de sedimento em formato di-

gital 3D. Para tal propõe-se o uso da tecnologia de escaneamento a laser 3D. Um equipamento que em suas características estruturais, contém: laser de varredura, sensor de deslocamento e câmera.

- Soprador Térmico - utilizado para o estabelecimento do fluxo de ar quente dentro da tubulação. O soprador térmico usado é o modelo *Skil 8003*, que pode operar em 3 níveis de temperatura e com uma variedade de acessórios para acoplamento na saída de ar. Este equipamento realiza um papel de extrema importância, conseguindo atingir temperaturas equivalentes ao ambiente estrutural real.
- Tubulação em escala - para reproduzir em escala o ambiente de fábrica, projetou-se uma maquete de tubulação. A tubulação apresenta as seguintes dimensões: 100 cm de comprimento e 10 cm de diâmetro. Esta maquete da tubulação é baseada em regiões retas. Neste dataset não foram avaliadas as tubulações curvas.

4.3.1 Dataset

Foram elaboradas 22 amostras de sedimentos, os perfis de sedimentos representam estágios variados de obstrução na tubulação, identificando acumulação em diferentes dimensões e níveis, conforme relatado pelos trabalhadores da indústria. A representação dos padrões de perfis estão mostrados na Figura 21, para cada padrão foi identificado uma amostra correspondente.

No entanto 22 amostras não são suficientes para generalização no processamento de uma rede neural. Com isso utilizou-se técnicas de aumento de dados. Inicialmente a região de interesse das imagens térmicas tinham resolução 280 x 780 pixels e foi aplicada a ESRGAN nas imagens para aumentar a resolução dos pixels usando o modelo RRDB-ESRGAN-x4. O tamanho final das imagens após a aplicação do ESRGAN é de 1120 x 3120 pixels.

Cada uma das 22 regiões de interesse foram divididas em 15 partes menores, conforme apresentado na Figura 22, resultando em 330 amostras (fatias). Cada fatia, exemplificada na Figura 23 ficou com resolução 1120 x 208 *pixels*.

O processo de segmentação das imagens não foi utilizado apenas como uma abordagem para ampliar o volume de dados disponíveis para o treinamento da rede neural, mas também para permitir que o modelo aprenda padrões térmicos locais com maior precisão. No entanto, mesmo com esse refinamento, as imagens térmicas ainda estão sujeitas a interferências ambientais, que podem comprometer a qualidade dos dados e impactar a confiabilidade das predições. Portanto, é importante considerar os desafios impostos pelas condições reais de operação, como vibrações mecânicas, interferências eletromagnéticas e variações térmicas inesperadas, que influenciam diretamente a captura das imagens e a eficácia do modelo na identificação dos sedimentos.

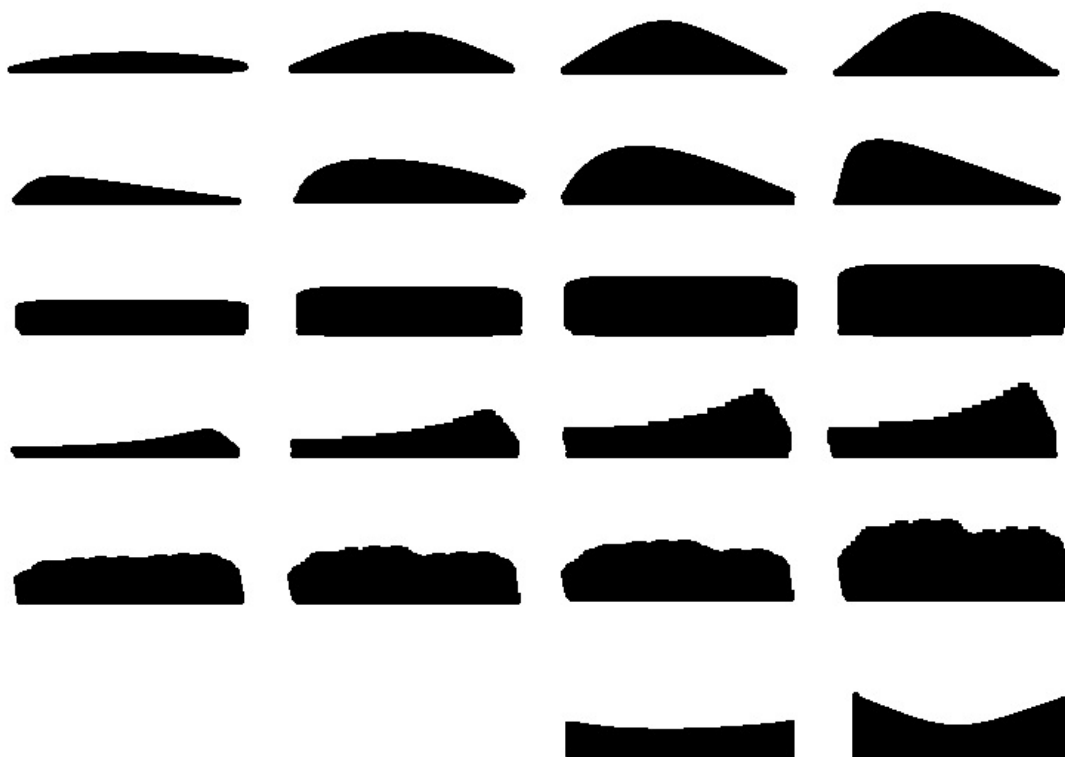


Figura 21: Perfis dos sedimentos elaborados manualmente. Fonte: Elaborado pela Autora

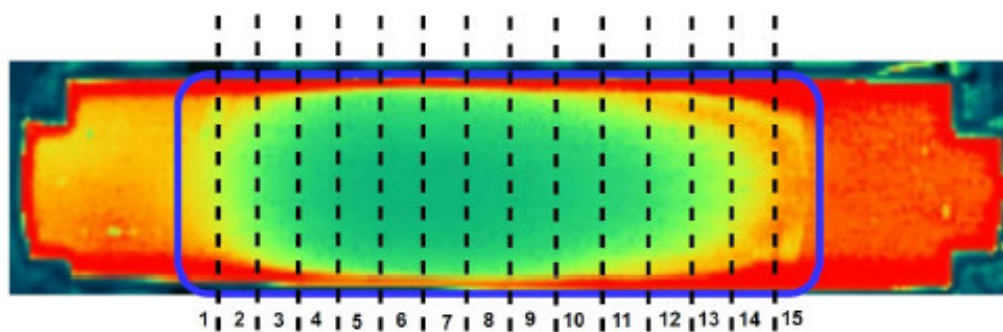


Figura 22: Aumento de dados, fatiamento da região de interesse. Fonte: Elaborada pela autora.



Figura 23: Aumento de dados - Uma Fatia da imagem térmica RGB do sedimento. Fonte: Elaborada pela autora.

Diante desse cenário, a aplicação de filtros de ruído, desfoque, ruído branco e interferências artificiais desempenham um papel importante na simulação das condições reais da indústria, contribuindo para um aumento de dados que melhora a robustez do modelo na identificação dos padrões térmicos característicos da sedimentação.

Para ampliar o conjunto de dados e atender algumas das diversidades possíveis na captura das imagens térmicas, foram aplicados quatro filtros nas imagens (ruído, desfoque, ruído branco, interferência), conforme as figuras 24, 25, 26 e 27 respectivamente, cada um com três intensidades.

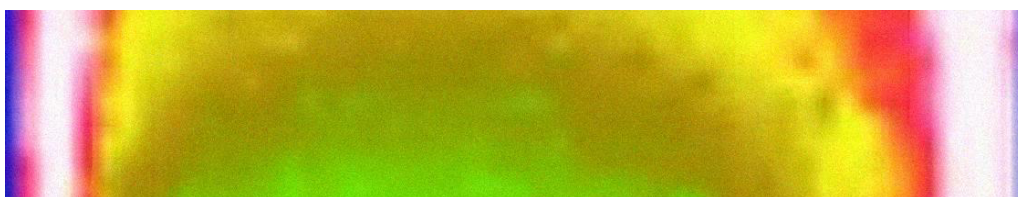


Figura 24: Aumento de dados - Uma Fatia da imagem térmica RGB do sedimento (com ruído). Fonte: Elaborada pela autora.

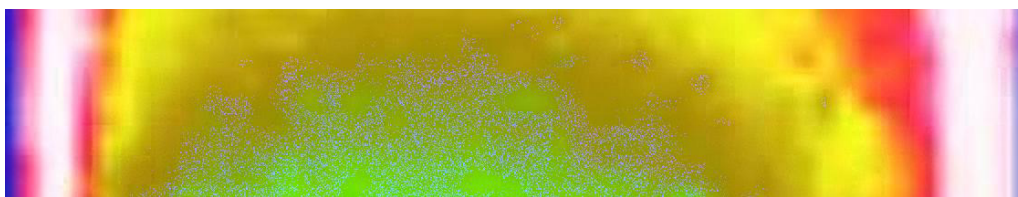


Figura 25: Aumento de dados - Uma Fatia da imagem térmica RGB do sedimento (com desfoque). Fonte: Elaborada pela autora.

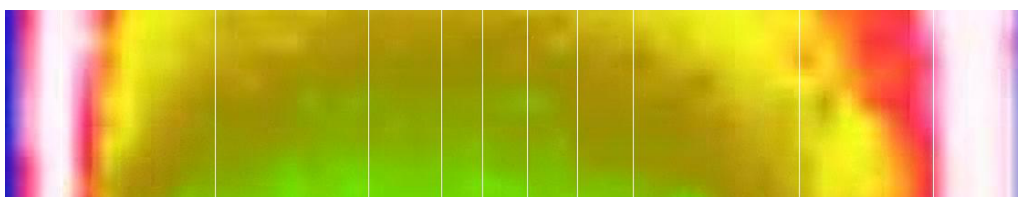


Figura 26: Aumento de dados - Uma Fatia da imagem térmica RGB do sedimento (com ruído branco). Fonte: Elaborada pela autora.



Figura 27: Aumento de dados - Uma Fatia da imagem térmica RGB do sedimento (com interferência). Fonte: Elaborada pela autora.

Contudo, a aplicação desses filtros deve ser realizada de forma controlada, evitando a degradação excessiva das imagens e garantindo que as informações essenciais para a

predição sejam preservadas. Em ambientes sujeitos a vibrações, a qualidade da captura pode ser prejudicada pelo deslocamento da câmera ou pela movimentação da própria tubulação. Assim, a introdução de ruídos sintéticos auxilia no treinamento do modelo, tornando-o mais resiliente às variações encontradas em condições reais de operação.

Com o objetivo de otimizar a entrada de dados para a rede neural convolucional, as imagens RGB pré-processadas foram convertidas para escala de cinza, permitindo uma abordagem mais simplificada. Após essa conversão, as mesmas técnicas de aumento de dados foram aplicadas às imagens em tons de cinza, garantindo a coerência no treinamento do modelo e ampliando a diversidade do conjunto de dados utilizado no aprendizado da CNN.

As imagens em tons de cinza estão apresentadas nas Figuras 28, 29, 30, 31, 32, referentes a imagem original, imagem com ruído, imagem com desfoque, imagem com ruído branco, imagem com interferência, respectivamente.



Figura 28: Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento. Fonte: Elaborada pela autora.



Figura 29: Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento (com ruído). Fonte: Elaborada pela autora.



Figura 30: Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento (com desfoque). Fonte: Elaborada pela autora.

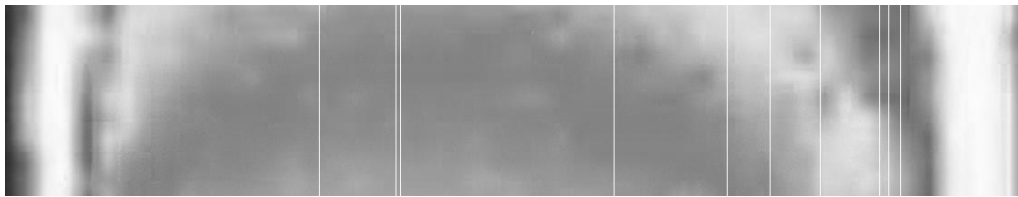


Figura 31: Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento (com ruído branco). Fonte: Elaborada pela autora.



Figura 32: Aumento de dados - Uma Fatia da imagem térmica tom de cinza do sedimento (com interferência). Fonte: Elaborada pela autora.

Esse processo expandiu o conjunto de dados da seguinte forma $15 \text{ cortes} \times (1 \text{ fatia original} + 4 \text{ filtros} \times 3 \text{ parâmetros diferentes nos filtros}) = 195 \text{ fatias distintas}$. $22 \text{ amostras} \times 195 \text{ fatias de cada amostra} = 4290 \text{ fatias}$. Resultando em 4290 amostras distintas RGB e 4.290 amostras distintas em tons de cinza a fim de melhorar a robustez e reduzir o risco de *overfitting*.

4.3.2 Rotulagem dos Dados

A nuvem de pontos gerada pelo scanner (arquivo de vértices) é segmentada em 15 partes para manter a proporção entre as 15 fatias térmicas. Cada fatia da nuvem de pontos é interpolada para um tamanho fixo de 800 pontos, padronizando os dados e garantindo alinhamento entre os conjuntos. Esses rótulos servem como *ground truth*, permitindo o treinamento da rede neural na tarefa de mapear características das imagens térmicas às alturas dos sedimentos.

A correspondência entre a fatia da imagem térmica e a respectiva fatia do arquivo de coordenadas são utilizadas no treinamento do modelo. A entrada é o fragmento da imagem térmica e o *ground-truth* é a coordenada z do respectivo fragmento do arquivo de coordenadas.

O sistema experimental segue a metodologia proposta para replicar, de forma controlada, as condições reais de sedimentação de fertilizantes em tubulações industriais, permitindo a análise térmica e geométrica dos depósitos. A combinação da câmera térmica, do scanner a laser 3D e do soprador térmico possibilita a aquisição de dados sobre a distribuição dos sedimentos e suas respectivas assinaturas térmicas. Assim, a estrutura do experimento garante a obtenção de informações relevantes para o desenvolvimento de soluções inovadoras voltadas ao monitoramento não invasivo de tubulações industriais.

Neste capítulo foram apresentados os materiais e a metodologia para elaboração dos

dados a serem aprendidos pelo modelo. A seguir apresentaremos a arquitetura da rede neural convolucional proposta como modelo para obtenção do perfil do entupimento.

5 T-NET UMA REDE NEURAL PROFUNDA PARA ESTIMATIVA DO PERFIL DO SEDIMENTO

Este trabalho propõe um método não invasivo baseado no processamento de imagens térmicas de sedimentos dentro de uma tubulação. Este processamento é realizado através de um modelo computacional baseado no processamento da imagem. Através da visão computacional será inferido o perfil de obstrução do tubo. A Figura 33 apresenta a visão geral do modelo.

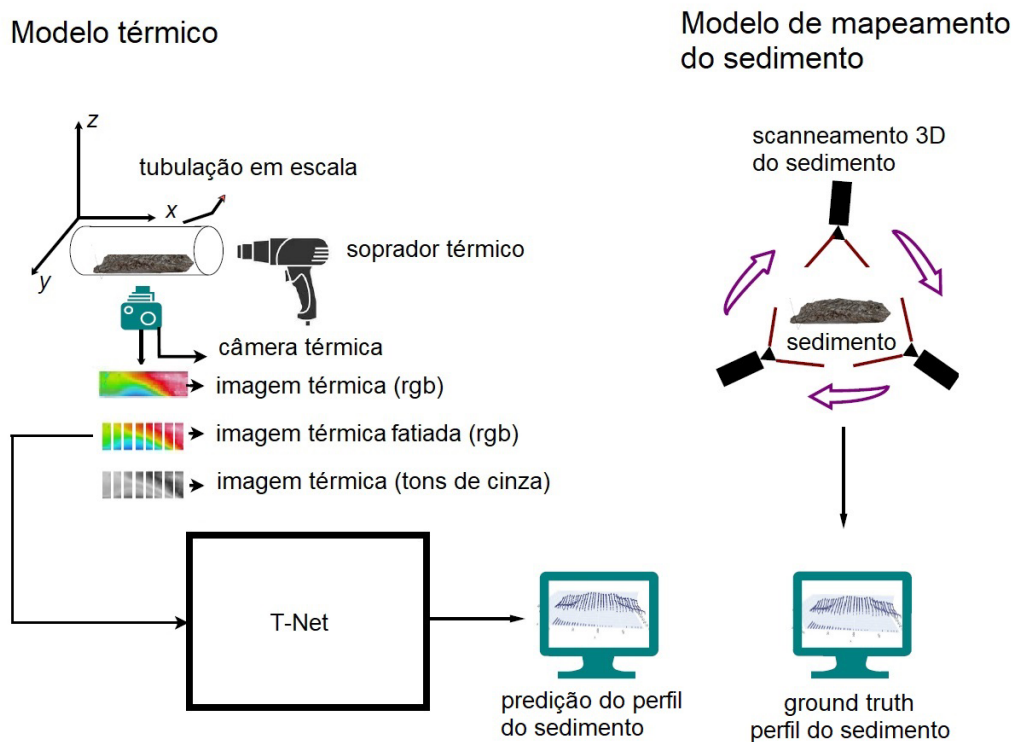


Figura 33: Visão geral da metodologia proposta.

5.1 Arquitetura da T-Net

A arquitetura da T-Net é inspirada na rede *Point Set Generation* (PSG) (Fan et al., 2017a). A PSG usa dados sintéticos para reconstrução 3D de objetos visíveis e a reconstrução 3D é exibida em uma nuvem de pontos. Na T-Net, as imagens térmicas são a entrada da rede, a reconstrução do sedimento ocluso à tubulação é apresentado em cotas, conforme esquematizado na Figura 34

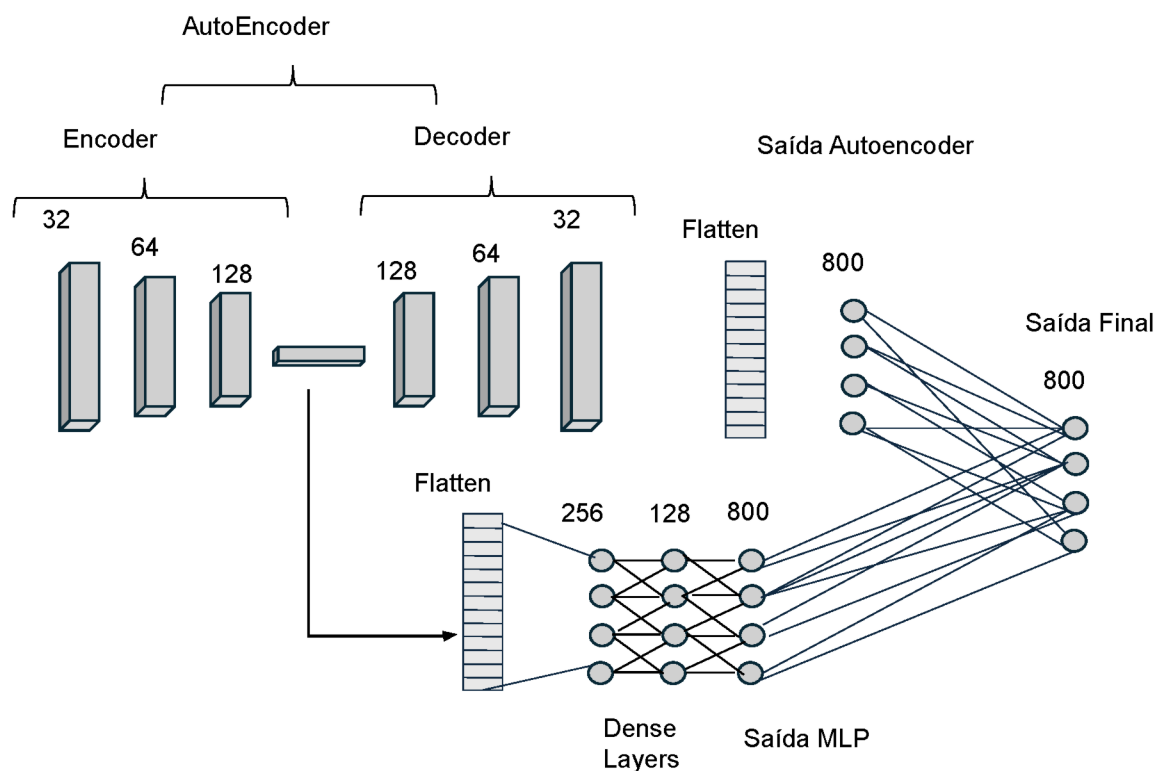


Figura 34: Desenho esquemático da arquitetura da rede.

Este modelo é baseado, inicialmente, em uma arquitetura *autoencoder* (*encoder-decoder*), onde a entrada é uma imagem com dimensões (Altura da imagem, Largura da imagem, Quantidade de canais de cores). O *encoder* (Codificador) é uma rede convolucional que extrai representações hierárquicas da imagem. O *Encoder* é responsável por extrair características relevantes das imagens de entrada, reduzindo progressivamente a dimensionalidade e capturando representações latentes significativas. O *decoder* (decodificador) reconstrói a informação a partir das características extraídas.

Na arquitetura proposta, os dados de saída do *encoder* convolucional seguem dois fluxos de processamento distintos e complementares. O primeiro fluxo direciona os mapas de ativação para o *decoder*, cuja função é reconstruir as características espaciais relevantes da imagem térmica original, preservando a estrutura e a morfologia do perfil térmico associado ao sedimento. O segundo fluxo encaminha as mesmas representações latentes para uma rede MLP (Multi-Layer Perceptron), responsável por realizar um mapeamento direto dessas características para valores numéricos que representam o perfil geométrico do se-

dimento. A estratégia de bifurcação da informação permite explorar simultaneamente a capacidade de reconstrução espacial do *autoencoder* e a eficiência de modelagem vetorial da MLP. Ao final do processamento, as saídas do *decoder* e da MLP são concatenadas, resultando em uma representação que combina informações estruturais e numéricas para a estimativa do perfil de sedimentação.

Com o objetivo de apresentar de maneira clara e estruturada a arquitetura da rede neural desenvolvida e a estratégia de bifurcação do processamento, apresenta-se três tabelas específicas Tabela 9, Tabela 10 e Tabela 11. Essas tabelas descrevem detalhadamente a composição da rede, incluindo o modelo baseado em codificador-decodificador convolucional, o codificador associado a uma multilayer perceptron (MLP) e o processo de concatenação responsável pela formação da saída final em 800 pontos. Cada estrutura está apresentada especificando-se as camadas, os tamanhos das entradas e saídas, os parâmetros dos kernels, as funções de ativação aplicadas e as operações de pooling realizadas. Tal organização visa apresentar a modelagem proposta, favorecendo a reprodutibilidade do estudo. A Tabela 9 descreve as informações do processamento do *autoencoder*, a Tabela 10 apresenta as informações do *encoder* seguido da MLP, a Tabela 11 apresenta as informações de concatenação da saída do *autoencoder* com a saída da MLP. As informações de "(Altura, Largura, Canais)" apresentadas nas colunas de "Entrada" e "Saída" nas Tabelas 9, 10 e 11 referem-se as dimensões da imagem de entrada.

Tabela 9: Descrição da arquitetura da rede convolucional (Encoder-Decoder)

Camada	Entrada	Kernel	Função de Ativação	Max Pooling	Saída
<i>Input Layer</i>	(Altura, Largura, Canais)	–	–	–	(Altura, Largura, Canais)
1ª Convolução	(Altura, Largura, Canais)	3×3×32	ReLU	–	(Altura, Largura, 32)
1º Pooling	(Altura, Largura, 32)	–	–	2×2	(Altura/2, Largura/2, 32)
2ª Convolução	(Altura/2, Largura/2, 32)	3×3×64	ReLU	–	(Altura/2, Largura/2, 64)
2º Pooling	(Altura/2, Largura/2, 64)	–	–	2×2	(Altura/4, Largura/4, 64)
3ª Convolução	(Altura/4, Largura/4, 64)	3×3×128	ReLU	–	(Altura/4, Largura/4, 128)
3º Pooling	(Altura/4, Largura/4, 128)	–	–	2×2	(Altura/8, Largura/8, 128)
1ª Deconvolução	(Altura/8, Largura/8, 128)	3×3×128	ReLU	–	(Altura/8, Largura/8, 128)
<i>BatchNormalization.1</i>	(Altura/8, Largura/8, 128)	–	–	–	(Altura/8, Largura/8, 128)
2ª Deconvolução	(Altura/8, Largura/8, 128)	3×3×64	ReLU	–	(Altura/8, Largura/8, 64)
<i>BatchNormalization.2</i>	(Altura/8, Largura/8, 64)	–	–	–	(Altura/8, Largura/8, 64)
3ª Deconvolução	(Altura/8, Largura/8, 64)	3×3×32	ReLU	–	(Altura/8, Largura/8, 32)
<i>BatchNormalization.3</i>	(Altura/8, Largura/8, 32)	–	–	–	(Altura/8, Largura/8, 32)
<i>Flatten</i>	(Altura/8, Largura/8, 32)	–	–	–	(<i>Flattened Vector</i>)
<i>Dense (Output Layer)</i>	(<i>Flattened Vector</i>)	–	–	–	800

A apresentação detalhada das tabelas que descrevem as camadas da rede neural convolucional, seu fluxo de processamento e as respectivas configurações técnicas, torna-se importante contextualizar essas informações em termos de desempenho e implicações práticas para o modelo proposto. Para isso, a Tabela 12 compara criticamente as vantagens e desvantagens de cada componente da arquitetura: o encoder, o decoder e a MLP. Essa comparação permite avaliar, de maneira estruturada, como cada parte contribui para

Tabela 10: Descrição da arquitetura do modelo Encoder + MLP

Camada	Entrada	Kernel	Função de Ativação	Max Pooling	Saída
<i>Input Layer</i>	(Altura, Largura, Canais)	–	–	–	(Altura, Largura, Canais)
1ª Convolução	(Altura, Largura, Canais)	3×3×32	ReLU	–	(Altura, Largura, 32)
1º Pooling	(Altura, Largura, 32)	–	–	2×2	(Altura/2, Largura/2, 32)
2ª Convolução	(Altura/2, Largura/2, 32)	3×3×64	ReLU	–	(Altura/2, Largura/2, 64)
2º Pooling	(Altura/2, Largura/2, 64)	–	–	2×2	(Altura/4, Largura/4, 64)
3ª Convolução	(Altura/4, Largura/4, 64)	3×3×128	ReLU	–	(Altura/4, Largura/4, 128)
3º Pooling	(Altura/4, Largura/4, 128)	–	–	2×2	(Altura/8, Largura/8, 128)
<i>Flatten</i>	(Altura/8, Largura/8, 128)	–	–	–	(Flattened Vector)
<i>Dense 1 (MLP)</i>	(<i>Flattened Vector</i>)	–	ReLU	–	256
Dropout (50%)	256	–	–	–	256
<i>Dense 2 (MLP)</i>	256	–	ReLU	–	128
<i>Dense 3 (Output Layer)</i>	128	–	-	–	800

Tabela 11: Descrição da arquitetura para a formação da Saída Final (Concatenação de Saída Auto Encoder e Saída MLP)

Camada	Entrada	Kernel	Função de Ativação	Pooling	Saída
<i>Concatenate</i>	(800, 800)	–	–	–	(2 × 800)
<i>Dense (Final Output)</i>	(2 × 800)	–	-	–	800

o objetivo final de reconstrução e predição do perfil de sedimentos, evidenciando os pontos fortes e as limitações associadas à escolha desta configuração híbrida.

Tabela 12: Comparação das partes da Arquitetura da T-Net

Arquitetura	Vantagens	Desvantagens
Autoencoder (Encoder-Decoder)	- Preserva informações espaciais - Boa reconstrução da estrutura térmica - Captura padrões complexos	- Maior complexidade e custo computacional - Exige mais dados para treinamento
Encoder + MLP	- Simples e eficiente - Menor número de parâmetros - Requer menos dados para treinamento	- Perde estrutura espacial - Dificuldade na modelagem de relações geométricas
Autoencoder + MLP (Fusão de Saídas)	- Combina a reconstrução espacial com modelagem vetorial - Melhor capacidade preditiva -	- Modelo mais pesado e complexo - Maior tempo de treinamento e consumo de memória

Para concluir este capítulo, a arquitetura proposta, denominada T-Net, representa uma abordagem para a estimativa do nível de sedimentos em tubulações industriais a partir de imagens térmicas. A combinação entre *autoencoders* e redes densas permite explorar tanto a estrutura espacial das imagens quanto a modelagem vetorial das características

térmicas capturadas. A análise comparativa das diferentes arquiteturas evidencia o compromisso entre a preservação da informação espacial e a eficiência computacional. O próximo capítulo apresentará a implementação do modelo, detalhando os processos de treinamento, validação e avaliação dos resultados, consolidando a metodologia proposta.

6 RESULTADOS E DISCUSSÕES

A rede neural foi treinada no Google Colaboratory (Google Colab) que é um ambiente de computação em nuvem para execução de modelos de aprendizado de máquina. Os dados foram processados em GPU modelo T4 com 51 GB de RAM e 15GB de VRAM.

O modelo foi treinado utilizando tanto imagens em RGB quanto em tons de cinza, permitindo uma análise comparativa do impacto das diferentes representações de cor no desempenho da rede. Essa abordagem possibilitou avaliar se a redução da complexidade das imagens, ao remover informações de cor, afeta a capacidade do modelo de identificar padrões térmicos e estimar o nível de sedimentação na tubulação. Os resultados dos treinamentos são apresentados por meio de métricas quantitativas, possibilitando a comparação entre as duas abordagens e destacando qual delas proporciona maior precisão na reconstrução e inferência dos perfis de sedimentos.

O primeiro treinamento da T-Net foi realizado com 4290 imagens RGB e o segundo treinamento foram 4290 imagens tons de cinza. Para ambos os treinos foram utilizadas 60% das imagens para treinamento, 20% das imagens para validação e 20% das imagens para teste, selecionadas aleatoriamente. Foram processadas em 100 épocas, *batch size* 8, otimizador Adam, learning rate no valor de 0.0001.

As métricas utilizadas para verificar as diferenças entre os valores previstos pela T-Net e o *ground-truth* foram o Mean Square Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Coeficiente de Determinação (R^2) que são calculadas respectivamente da seguinte forma:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - y_i)^2, \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - y_i)^2}, \quad (9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - y_i|. \quad (10)$$

e

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

onde n , y e y_i são o número total de amostras, valor de referência (*ground truth*) e valor da amostra, respectivamente. Assim, o erro é dado pela diferença entre y e y_i . A métrica R^2 , também conhecida como coeficiente de determinação, é uma importante medida de desempenho para avaliar a qualidade do modelo de aprendizado de máquina. No contexto da rede T-Net, ela pode ser descrita como uma métrica adicional utilizada para verificar o quão bem os valores preditos pela rede se aproximam dos valores reais (*ground-truth*). O R^2 quantifica a proporção da variabilidade dos dados que é explicada pelo modelo, isto é importante no processo de validação do aprendizado da rede. Neste contexto, um valor de R^2 próximo de 1 indica que o modelo é capaz de explicar a maior parte da variabilidade dos dados e, portanto, tem uma alta capacidade preditiva. Por outro lado, valores de R^2 próximos de 0 sugerem que o modelo não consegue explicar bem a variabilidade dos dados, indicando possíveis deficiências no treinamento ou no modelo utilizado. Em conjunto com as métricas como RMSE, MSE e MAE, o R^2 fornece uma visão mais completa sobre a qualidade do modelo no processo de predição da sedimentação com base em imagens térmicas.

6.1 Métricas de treino vs validação - imagens RBG e Tons de cinza

Na Figura 35 são apresentadas as informações de função de perda entre o treino e a validação das imagens RBG e tons de cinza.

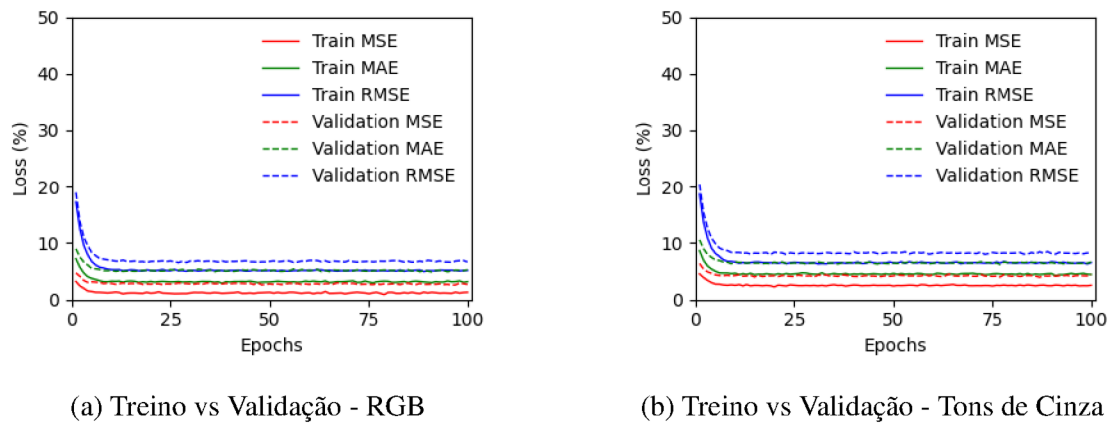


Figura 35: Comparação entre treino e validação para RGB e tons de cinza.

Na Figura 35 (a) e (b), o eixo X (horizontal) representa o número de épocas, variando de 0 a 100, enquanto o eixo Y (vertical) indica o valor da perda (*Loss*) em percentual (%). As linhas sólidas correspondem ao desempenho no conjunto de treinamento, sendo a linha vermelha associada ao erro quadrático médio (MSE – Mean Squared Error), a linha verde ao erro absoluto médio (MAE – Mean Absolute Error) e a linha azul à raiz do

erro quadrático médio (RMSE – Root Mean Squared Error). As linhas tracejadas, por sua vez, representam os mesmos erros, mas no conjunto de validação.

Observa-se que, no início do treinamento (época 0), os valores de perda são elevados, caindo rapidamente durante as primeiras 20 épocas. Após esse período inicial, as curvas de perda tendem a se estabilizar. Nota-se ainda que os erros no conjunto de validação permanecem ligeiramente superiores aos do conjunto de treinamento, o que sugere a ocorrência de um leve overfitting. O RMSE apresenta valores mais altos do que as demais métricas, uma vez que é derivado do MSE. De modo geral, todas as métricas (MSE, MAE e RMSE) mostram uma tendência de queda acentuada nas primeiras épocas, estabilizando-se posteriormente em valores baixos, o que indica que o modelo aprendeu os padrões presentes nos dados. Além disso, o processamento das imagens RGB proporcionou um treinamento mais eficiente e resultou em um menor erro geral, possivelmente devido à maior riqueza de informações contidas nesse formato de dado.

Na Figura 36 são apresentadas as informações do coeficiente de variabilidade das imagens RGB e tons de cinza

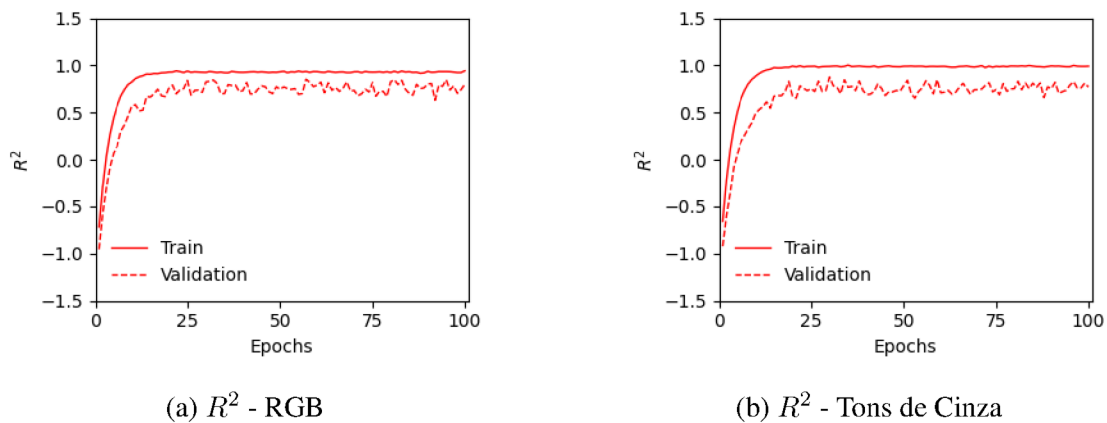


Figura 36: Coeficiente de variabilidade dos dados (R^2)

Na Figura 36 (a) e (b), o eixo X (horizontal) representa o número de épocas, variando de 0 a 100, enquanto o eixo Y (vertical) apresenta o coeficiente de determinação R^2 , utilizado para medir a qualidade do ajuste do modelo. A linha sólida vermelha corresponde ao valor de R^2 no conjunto de treinamento, e a linha tracejada vermelha representa o R^2 no conjunto de validação. No início do treinamento (época 0), os valores de R^2 são negativos, indicando que o modelo ainda não foi treinado adequadamente. Nas primeiras 20 épocas, observa-se uma rápida melhoria nos valores de R^2 , seguida por uma tendência de estabilização. O R^2 do treinamento se aproxima de 1, sugerindo um excelente ajuste aos dados de treinamento. Já o R^2 da validação também apresenta melhoria significativa, embora com maior oscilação e valores ligeiramente inferiores aos do treinamento.

6.1.1 Comparativo RGB x Tons de cinza

A Tabela 13 apresenta de forma consolidada os resultados obtidos para cada configuração de entrada RGB e Tons de cinza.

Tabela 13: Métricas do Modelo para imagens RGB e Tons de cinza

Métrica	RGB (%)	Tons de cinza (%)
MAE	4.91765	6.21163
MSE	2.59747	3.96941
RMSE	6.55544	7.95289
R ²	0.89318	0.87645

Os resultados mostram que o modelo obteve um desempenho superior utilizando imagens em RGB. O erro absoluto médio (MAE) e o erro quadrático médio (MSE) foram menores para imagens coloridas, indicando que as predições foram mais precisas e com desvios médios menores. Isso sugere que a informação adicional dos três canais (R, G, B) ajudou o modelo a capturar padrões térmicos e geométricos de forma mais eficiente. O RMSE também foi inferior para o caso RGB, mostrando que grandes erros foram menos frequentes em comparação ao treinamento com imagens em tons de cinza. Além disso, o coeficiente de determinação (R^2) atingiu 89,3% para imagens RGB, contra 87,6% para tons de cinza, reforçando que a presença da informação de cor melhora a capacidade preditiva do modelo. Apesar da vantagem em termos de precisão, é importante considerar que o processamento de imagens RGB demanda maior custo computacional, devido ao tratamento de três canais ao invés de apenas um.

6.2 Ground truth

Durante o processamento dos arquivos de nuvem de pontos obtidos via escaneamento 3D dos sedimentos, cada perfil foi segmentado em 15 fatias longitudinais. A fim de garantir uniformidade estrutural para as análises subsequentes, cada fatia foi submetida a um processo de interpolação, padronizando a quantidade de amostras em 800 pontos por fatia. Essa abordagem visou minimizar variações na densidade de pontos ao longo dos perfis e assegurar a compatibilidade dimensional entre os dados geométricos e as imagens térmicas utilizadas como entrada no modelo de aprendizado profundo.

O gráfico ilustrado na Figura 37 apresenta os perfis de altura de 330 fatias (22 amostras x 15 fatias), comparando as alturas originais (linha vermelha) com alturas interpoladas usando três métodos: Linear (linha tracejada azul), Cúbico (linha tracejada roxa) e Vizinho Mais próximo "Nearest" (linha tracejada verde). Essa metodologia de interpolação foi importante para viabilizar o desenvolvimento do modelo, normalizando os dados de saída do modelo.

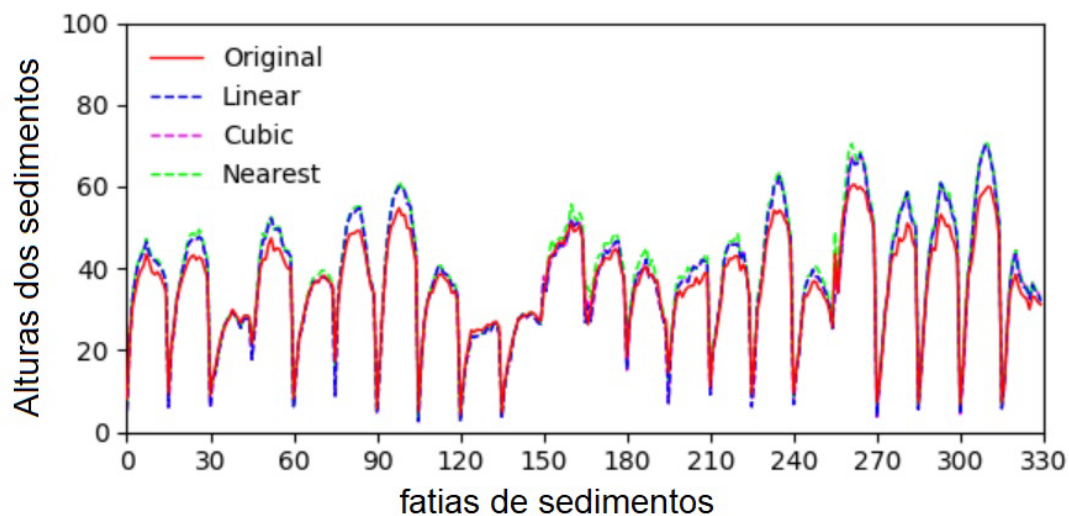


Figura 37: Interpolação alturas

A interpolação linear (linha azul) mostra um alinhamento próximo com os dados originais, mantendo uma tendência consistente com pequenos desvios em picos e vales. A interpolação cúbica (linha roxa) segue de perto os dados originais, oferecendo uma curva mais suave, particularmente nos pontos de inflexão. Parece reduzir o ruído e fornecer uma aproximação mais refinada. Enquanto a interpolação *Nearest* (linha verde) introduz mudanças em forma de degrau, especialmente em torno de picos e vales, devido à sua natureza discreta de atribuição de valores com base no ponto mais próximo.

O gráfico apresentado na Figura 38 exibe as diferenças percentuais ($\Delta\%$) entre os valores de alturas originais e aqueles obtidos de três métodos de interpolação: Linear (linha azul), Cúbico (linha tracejada roxa) e *Nearest* (linha tracejada verde) em 330 fatias.

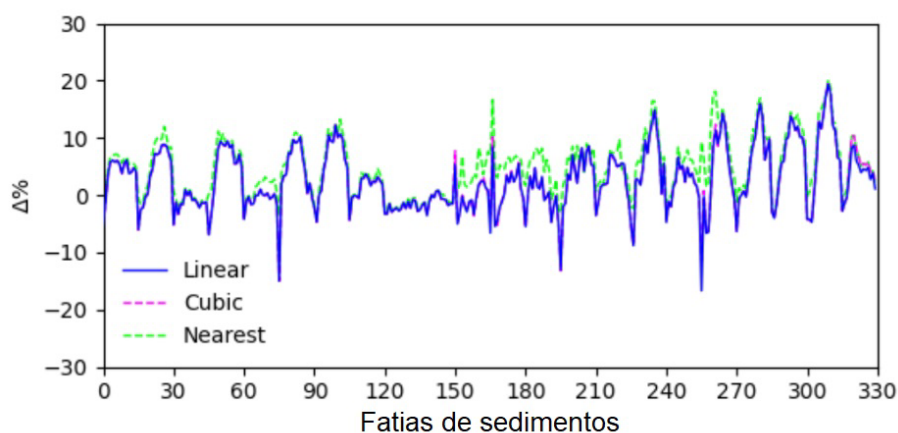


Figura 38: Diferenças percentuais das interpolações

O eixo y representa o desvio percentual dos valores interpolados dos valores originais. Os valores positivos indicam que a interpolação superestimou as alturas originais,

enquanto valores negativos indicam subestimação. A interpolação linear (linha azul) mostra o comportamento mais estável, com desvios agrupados em torno de 0%. Isso sugere que a interpolação linear aproxima os dados originais com erros mínimos. Enquanto a interpolação cúbica (linha tracejada roxa) também mostra pequenos desvios, mas com transições ligeiramente mais suaves em comparação à interpolação linear. Os desvios permanecem mais próximos de 0%, indicando precisão, especialmente em torno de transições complexas. A interpolação *Nearest* (linha tracejada verde) exibe desvios maiores e mais abruptos em comparação com os outros métodos. Ela tem picos e vales mais altos, perceptíveis em torno de mudanças abruptas nos dados originais, devido à sua natureza discreta. As interpolações linear e cúbica são melhores para manter a precisão, com a cúbica superando ligeiramente a linear em suavidade. A interpolação *Nearest* introduz mais variabilidade e é menos confiável para previsão de altura. O gráfico destaca a importância de selecionar o método de interpolação apropriado dependendo da precisão desejada e dos requisitos da aplicação.

A Figura 39 apresenta as diferenças percentuais absolutas ($|\Delta|%$), entre os valores de altura originais e os valores interpolados usando três métodos: linear (linha azul), cúbico (linha tracejada roxa) e *Nearest* (linha tracejada verde) em 330 fatias.

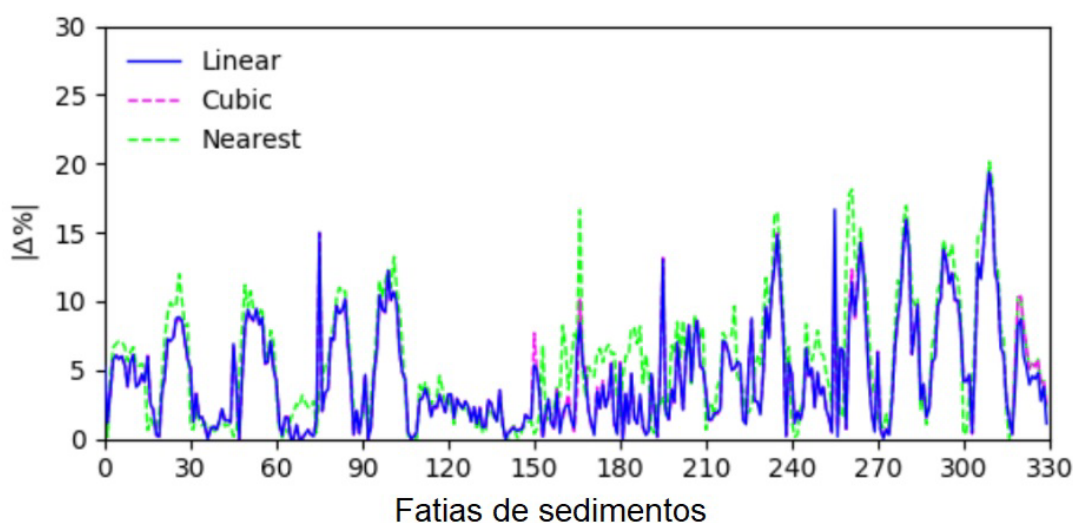


Figura 39: Diferenças percentuais absolutas

O eixo y representa a magnitude das diferenças percentuais, independentemente da direção (positiva ou negativa), enfatizando a escala de desvio dos valores originais. A interpolação linear (linha azul) demonstra desvios consistentes e relativamente pequenos, com a maioria dos valores permanecendo abaixo de 10%. Ele mantém uma tendência suave, mostrando variação mínima entre as fatias. Enquanto a interpolação cúbica (linha tracejada roxa) exibe desvios pequenos, sobrepondo-se ao método linear em regiões com dados estáveis. O método cúbico também fornece transições mais suaves, destacando sua

precisão na captura de detalhes mais finos. A interpolação *Nearest* (linha tracejada verde) mostra os maiores desvios, com picos pronunciados excedendo 20% em algumas fatias. Isso indica que a interpolação *Nearest* tem dificuldade para aproximar os dados originais com precisão, particularmente em regiões com transições abruptas ou alta variabilidade. O gráfico confirma que as interpolações lineares e cúbicas são adequadas para tarefas que exigem precisão, com a cúbica fornecendo resultados mais suaves.

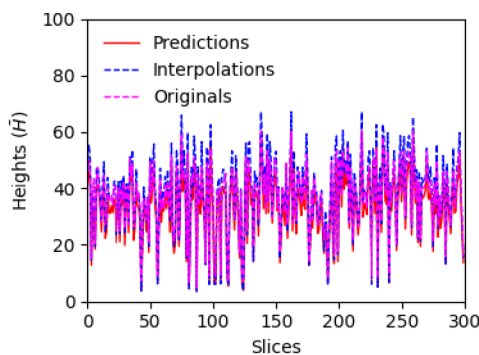
A tabela 14 apresenta as médias das diferenças percentuais e também dos erros de interpolação (MSE, RMSE e MAE) de todas as cotas para as 330 fatias

Tabela 14: Erros de interpolações

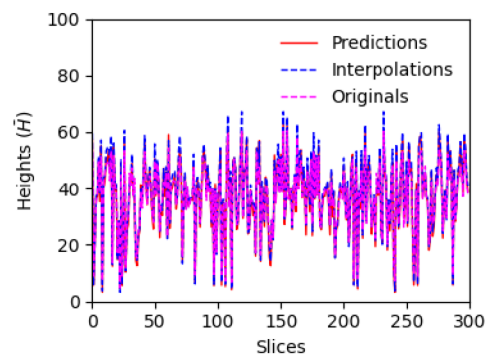
Métricas	Interpolações		
	Linear	Cubica	Nearest
Erros de interpolação (em elevação/altura)			
MAE	2.58271	2.57999	3.11618
MSE	11.46077	11.36087	15.65578
RMSE	3.38538	3.37059	3.95674
Erros de interpolação (em percentual)			
MAE	3.83262%	3.82610%	4.62127%
MSE	17.00727%	16.84807%	23.21740%
RMSE	5.02374%	4.99856%	5.86781%

Analisando as métricas de erro, pode-se afirmar que a interpolação cúbica mostrou-se mais precisa para esta aplicação. Além disso, a interpolação cúbica e a interpolação linear apresentaram valores de erro muito semelhantes.

A Figura 40 apresenta os resultados comparativos das médias das alturas de cada fatia, comparando a média da altura da fatia original, da fatia interpolada e da fatia predita pela rede.



(a) Alturas - RGB



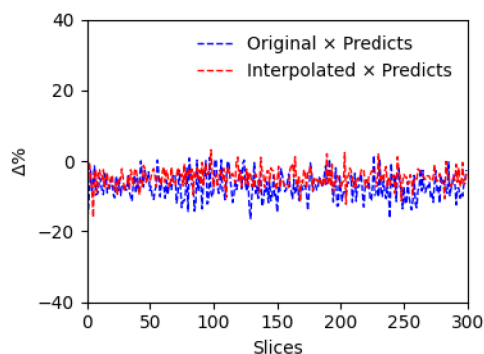
(b) Alturas - Tons de Cinza

Figura 40: Comparação da interpolação das alturas da saída do processamento das imagens RGB e tons de cinza

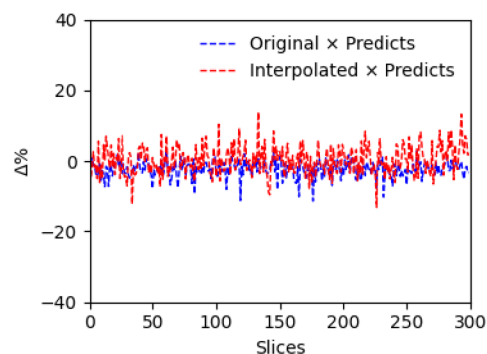
Na Figura 40 (a) e (b), o eixo X (horizontal) representa os índices das fatias (*Slices*), variando de 0 a 330. Eixo Y (vertical) representa os valores de altura, variando de 0 a 100, onde 100 representa o diâmetro da tubulação em escala. A linha Vermelha (sólida) representa os valores preditos pelo modelo. Enquanto a linha azul (tracejada) representa os valores interpolados pela interpolação cúbica. A linha rosa (tracejada) representa os valores originais do conjunto de dados.

As curvas original e interpolada parecem bastante sobrepostas, sugerindo que a interpolação reconstrói bem os dados reais. A curva de predição segue a tendência das outras duas curvas, mas apresenta menor variabilidade, podendo indicar uma suavização excessiva ou uma perda de detalhes finos. As predições apresentam menor erro em relação aos dados originais e maior erro em relação aos dados interpolados, indicando que o modelo pode estar mais alinhado com a distribuição dos dados reais do que com os valores gerados pela interpolação. Apesar das oscilações, o erro percentual permanece controlado, demonstrando um desempenho consistente na predição da sedimentação. O padrão oscilatório representam as variações naturais do fenômeno estudado.

A Figura 41 apresenta as diferenças no erro percentual ($\Delta\%$) entre as médias das alturas de cada fatia do sedimento (*slices*). A comparação dos percentuais é feita com as médias das fatias originais, fatias interpoladas e fatias preditas.



(a) Delta - RGB



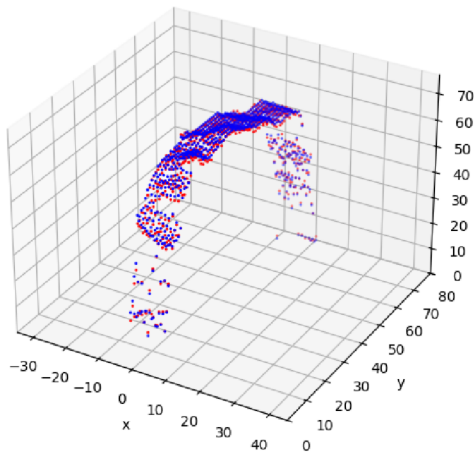
(b) Delta - Tons de Cinza

Figura 41: Delta

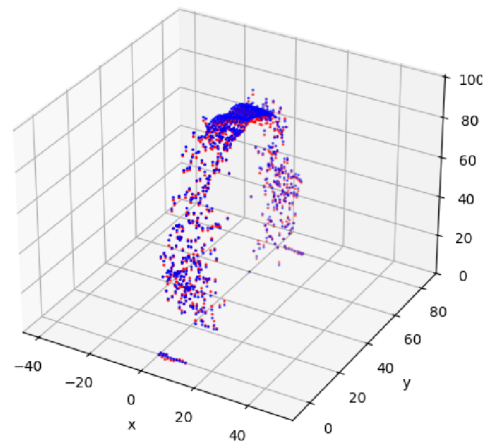
A diferença percentual entre os valores originais e previstos (linha azul tracejada) e entre os interpolados e previstos (linha vermelha tracejada) mostra um padrão semelhante nos dois casos. A Figura 41b a curva vermelha (Interpolated x Predicts) apresenta variações ligeiramente maiores do que a curva azul (Original x Predicts), sugerindo que o modelo se ajusta melhor aos dados originais do que aos interpolados. Isso sugere que o modelo tem uma maior variabilidade no erro ao processar imagens em tons de cinza.

6.3 Reconstruções do perfil do sedimento

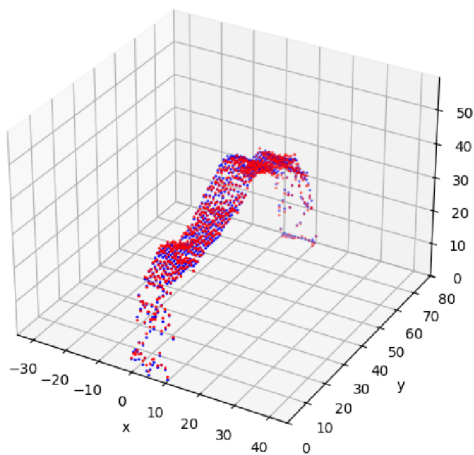
As Figuras 42a, 42b, 42c, 42d são as representações 3D do conjunto de dados de imagens RGB selecionados aleatoriamente para o *dataset* de teste da T-Net. As coordenadas x e y foram fixadas, a coordenada x representa a fração longitudinal da tubulação e a coordenada y representa o diâmetro da tubulação. A representação 3D apresenta a comparação entre a coordenada z da saída da rede neural, na cor vermelha, comparada com a coordenada z da fatia escaneada do sedimento, na cor azul.



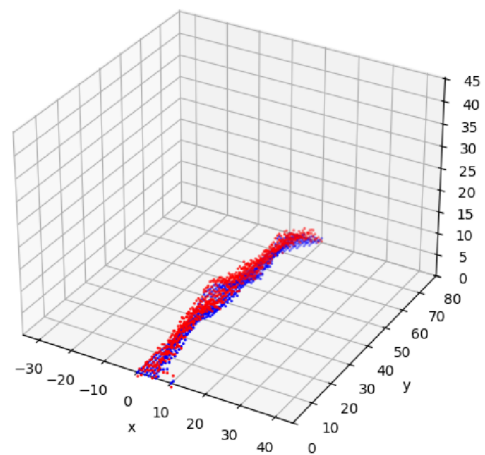
(a) Reconstrução 3D de imagem RGB - fatia aleatória 1



(b) Reconstrução 3D de imagem RGB - fatia aleatória 2



(c) Reconstrução 3D de imagem RGB - fatia aleatória 3



(d) Reconstrução 3D de imagem RGB - fatia aleatória 4

Figura 42: Reconstrução 3D de diferentes fatias aleatórias de imagens RGB.

Os gráficos exibem nuvens de pontos tridimensionais nos quais os pontos vermelhos representam os pontos preditos pelo modelo, enquanto os pontos azuis representam o *ground truth*, ou seja, os valores reais das alturas do sedimento. Em todos os gráficos, há

uma correspondência geral entre os pontos vermelhos (preditos) e os pontos azuis (*ground truth*). Isso sugere que o modelo está conseguindo reconstruir a estrutura do sedimento de forma razoável.

Em algumas regiões, os pontos vermelhos parecem mais dispersos do que os azuis, indicando maiores desvios nas previsões. Em geral estas dispersões estão nas laterais dos sedimentos, as quais sofrem maior influência do aquecimento do metal da tubulação. Possíveis melhorias incluem refinamento da arquitetura do modelo para melhorar a acurácia em áreas de maior dispersão. Aplicação de pós-processamento nas previsões para reduzir o espalhamento de pontos. Análise dos dados de entrada para verificar se há padrões específicos que afetam negativamente as previsões.

As Figuras 43a, 43b, 43c, 43d são as representações 3D do conjunto de dados de imagens em tons de cinza selecionados aleatoriamente para o *dataset* de teste da CNN. As coordenadas x e y foram fixadas, a coordenada x representa a fração longitudinal da tubulação e a coordenada y representa o diâmetro da tubulação. A representação 3D apresenta a comparação entre a coordenada z da saída da rede neural, na cor vermelha, comparada com a coordenada z da fatia escaneada do sedimento, na cor azul.

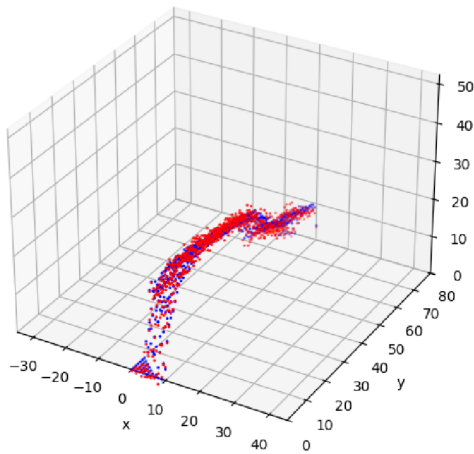
A estrutura geral para reconstrução foi realizada satisfatoriamente, tanto para imagens RGB quanto para imagens em tons de cinza, apresentados pelo alinhamento entre os pontos preditos e o *ground truth*. Isso sugere que ambos os modelos capturam a morfologia do sedimento de forma coerente. As imagens foram selecionadas aleatoriamente. Na exemplificação, as reconstruções de imagens RGB apresentadas nas Figuras 42a, 42b, 42c, 42d apresentam uma maior dispersão em algumas áreas, especialmente nas partes superiores da curva de sedimentação. As Figuras 43a, 43b, 43c, 43d representação 3D dos perfis de sedimentos de imagens em tons de cinza, parecem apresentar uma melhor compactação dos pontos preditos, indicando um menor erro em algumas regiões.

A Figura 44 apresenta uma reconstrução completa do perfil do sedimento, através da consolidação das partes fatiadas, validando o método baseado em imagens térmicas e redes neurais convolucionais (CNNs) para prever o perfil de sedimentação em tubulações industriais sem a necessidade de abertura ou intervenção física.

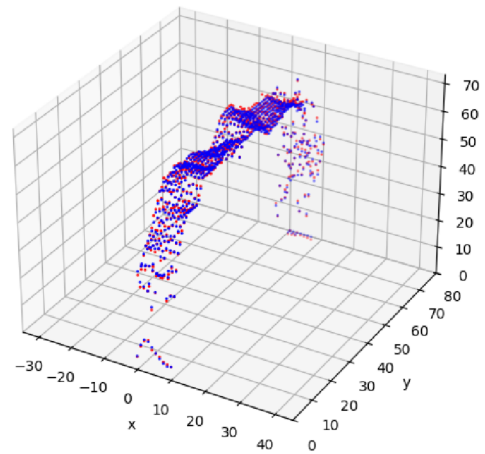
Possíveis melhorias incluem filtragem de ruído nos dados RGB para reduzir dispersões. Ajustes no modelo de interpolação para diminuir discrepâncias nas regiões superiores. Combinação de RGB e Grayscale para testar um modelo híbrido que aproveite os benefícios de ambos.

A tabela 15 apresenta um comparativo das métricas de erro entre o modelo T-Net e a rede MLP apresentada por (Dua, 2020). As métricas são referentes ao processamento das imagens RGB.

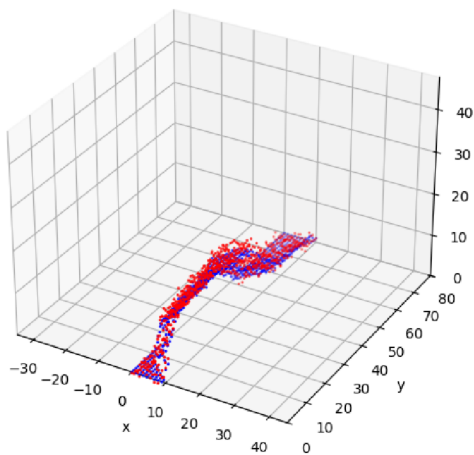
As métricas MAE, MSE, RMSE do processamento MLP apresentam erros menores, mostrando previsões mais próximas em termos absolutos, e penalizando menos os erros grandes. Enquanto o Coeficiente de determinação (R^2) do modelo O T-Net apresenta um



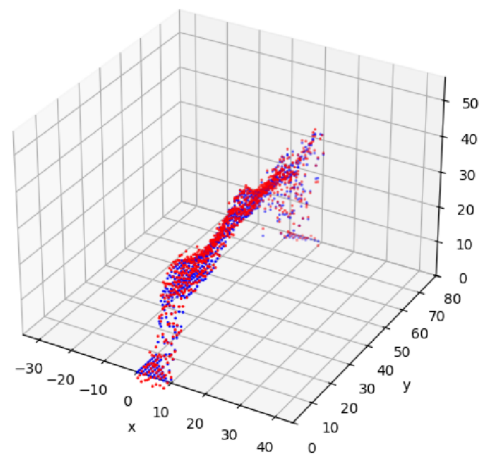
(a) Reconstrução 3D de imagem tons de cinza - fatia aleatória 1



(b) Reconstrução 3D de imagem tons de cinza - fatia aleatória 2



(c) Reconstrução 3D de imagem tons de cinza - fatia aleatória 3



(d) Reconstrução 3D de imagem tons de cinza - fatia aleatória 4

Figura 43: Reconstrução 3D de diferentes fatias aleatórias de imagens em tons de cinza.

Tabela 15: Comparação das métricas de desempenho entre MLP e T-Net

Métrica	MLP	T-Net (tons de cinza)
MAE (%)	2.71767	6.21163
MSE (%)	0.27353	3.96941
RMSE (%)	5.17997	7.95289
R^2	0.77027	0.87645

valor maior (0.89318), o que significa que ele explica melhor a variabilidade dos dados, mesmo tendo erros absolutos maiores. Logo o MLP de (Dua, 2020) é mais estável nos erros médios (boa para previsões globais com pouco desvio). Enquanto o modelo T-Net captura melhor a estrutura dos dados (maior R^2), mas com maior dispersão e erros médios absolutos maiores.

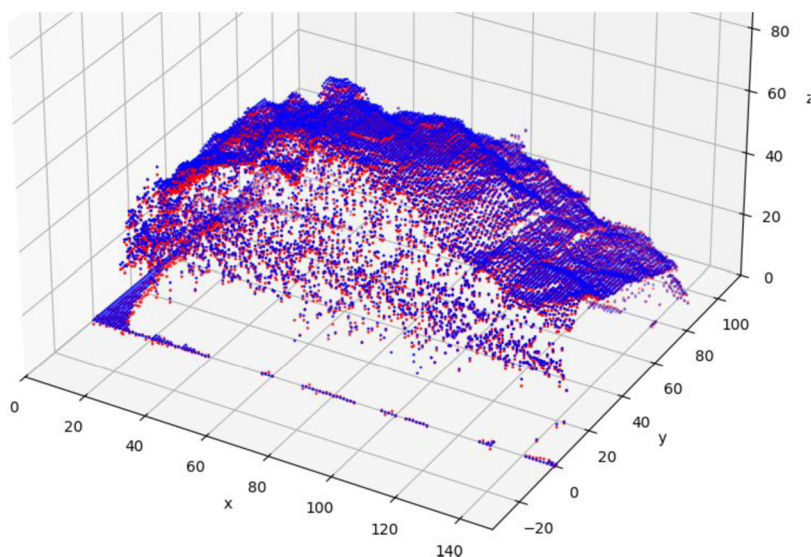


Figura 44: Consolidação do fatiamento para reconstrução do sedimento inteiro

6.4 Limitações do modelo T-Net

Apesar dos avanços alcançados na predição do perfil de sedimentos a partir de imagens térmicas em ambiente controlado, é necessário reconhecer as limitações do modelo proposto ao extrapolá-lo para situações industriais reais. A transição da maquete para aplicações em campo envolve desafios associados à complexidade dos fenômenos físicos, variabilidade operacional e limitações de calibração do sistema. Para tornar a análise mais estruturada, a Tabela 16 apresenta uma síntese das principais limitações identificadas, categorizadas em termos de sua origem, descrição técnica e implicações práticas para a aplicação em ambientes industriais. Esta análise é importante para orientar a adaptação do modelo a diferentes cenários de inspeção de tubulações.

A análise das limitações destacadas evidencia que, embora a metodologia proposta tenha demonstrado resultados promissores em ambiente controlado, sua aplicação em cenários industriais exige etapas adicionais de calibração, adaptação e validação. A variabilidade das condições operacionais, a complexidade dos perfis de sedimentação reais e as características específicas de cada planta industrial impõem novos desafios que precisam ser abordados para garantir a generalização do modelo.

Neste capítulo, foram apresentados os resultados obtidos com a aplicação do modelo proposto para a predição do perfil de sedimentos em tubulações industriais, utilizando imagens térmicas associadas a redes neurais convolucionais. Inicialmente, analisaram-se as curvas de perda durante o treinamento e a validação, medidas pelas métricas MSE, MAE e RMSE, bem como o coeficiente de determinação (R^2), os quais evidenciaram a capacidade do modelo em generalizar os dados e capturar os padrões de sedimentação. Na sequência, foi realizada uma comparação entre os modelos treinados com imagens RGB e

Tabela 16: Limitações, impactos e estratégias de mitigação.

Limitação	Descrição	Impacto	Estratégia de mitigação
diferenças de condições operacionais	variações de temperatura, fluxo, umidade e vibrações nos ambientes industriais.	redução da precisão e falhas de detecção em condições adversas.	calibração em campo e coleta de dados reais para ajuste do modelo (fine-tuning).
escala e dissipação térmica	diferença no comportamento térmico entre maquete reduzida e escala real.	alteração na resposta térmica, diminuindo correspondência com o modelo.	ajuste dos hiperparâmetros e uso de simulações físicas para adaptação de escala.
heterogeneidade dos sedimentos	variabilidade na composição dos sedimentos (umidade, granulometria, contaminantes).	diminuição da capacidade de generalização da rede neural.	expansão do dataset com diferentes tipos de sedimentos e uso de <i>transfer learning</i> .
dependência de configurações óticas	variações de distância, ângulo e configuração da câmera no ambiente industrial.	distorções nas imagens capturadas, prejudicando a inferência.	desenvolvimento de protocolo de posicionamento e calibração adaptativa da câmera.
ruído e interferências ambientais	presença de ruídos térmicos, vibrações e ruído eletromagnético em campo.	introdução de artefatos nas imagens, impactando a detecção de padrões.	aplicação de aumento de dados simulando ruído e desenvolvimento de filtros de pré-processamento.
limitação da representação térmica	diferenças sutis de condutividade podem ser insuficientes para detecção confiável.	redução da sensibilidade sob baixo contraste térmico.	combinação de imagens térmicas com dados de vibração ou espectroscópicos (análise multimodal).
necessidade de recalibração	adaptação limitada do modelo a novos cenários e configurações.	necessidade de treinamento adicional, aumentando tempo e custos.	implantação de aprendizado federado e recalibração periódica automática.

imagens em escala de cinza, observando-se uma ligeira superioridade do modelo baseado em imagens RGB, o que sugere que as informações dos canais de cores contribuem para um melhor desempenho preditivo. Posteriormente, avaliou-se o desempenho dos métodos de interpolação utilizados para padronizar os dados do perfil sedimentar, destacando-se que a interpolação cúbica resultou nos menores erros absolutos e percentuais. Com a aplicação das técnicas de interpolação, procedeu-se à reconstrução do perfil do sedimento, a partir dos pontos gerados de forma padronizada, permitindo uma representação mais consistente e comparável entre amostras distintas. Por fim, foram discutidas as limitações identificadas no modelo proposto, indicando caminhos para possíveis melhorias futuras.

7 CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho apresentou uma abordagem inovadora para a inspeção não invasiva de sedimentação em tubulações industriais, explorando a combinação de imagens térmicas e redes neurais convolucionais. O trabalho atendeu aos seus objetivos, propondo e validando um modelo capaz de inferir o perfil interno de sedimentos a partir de dados térmicos externos, sem necessidade de intervenção física ou interrupção do processo industrial. A análise crítica do estado da arte evidenciou que grande parte dos estudos existentes concentra-se na reconstrução 3D de objetos visíveis, não considerando cenários onde o alvo está oculto, como no caso da sedimentação em tubulações, o que fundamentou a relevância da pesquisa desenvolvida.

Para permitir o treinamento do modelo, foi elaborada uma metodologia para a construção de um conjunto de dados sintéticos, utilizando uma maquete física representativa de uma tubulação industrial. O experimento controlado possibilitou a aquisição de imagens térmicas rotuladas com perfis geométricos obtidos por escaneamento tridimensional dos sedimentos. Este conjunto de dados, posteriormente disponibilizado publicamente, representa uma contribuição importante para a comunidade científica, pois oferece uma base inédita para pesquisas futuras no tema de inspeção térmica e modelagem de sedimentação.

O modelo de rede neural desenvolvido baseou-se em uma arquitetura que combina autoencoder e MLP (Multi-Layer Perceptron), permitindo tanto a preservação de características espaciais extraídas da imagem térmica quanto a abstração de alto nível das informações latentes. As saídas dessas duas estruturas foram concatenadas para gerar a predição final do perfil de sedimentação. As métricas de desempenho MAE, MSE, RMSE e R^2 demonstraram que a rede neural proposta foi capaz de capturar adequadamente os padrões de sedimentação, generalizando bem para dados de validação.

A interpolação das nuvens de pontos para 800 vértices fixos mostrou-se relevante para padronizar a representação dos perfis de sedimento, reduzindo variabilidades desnecessárias e melhorando a qualidade do treinamento. A utilização de interpolação cúbica se destacou pela minimização dos erros absolutos e percentuais. A reconstrução dos perfis após a interpolação confirmou que a metodologia preservou as principais características

geométricas dos sedimentos, validando a escolha dos métodos de pré-processamento.

Na comparação entre o processamento de imagens RGB e tons de cinza, observou-se que as imagens RGB proporcionaram ligeiras melhorias nos erros de predição, indicando que a informação adicional dos canais de cores podem capturar variações térmicas mais sutis. No entanto, o uso de imagens em tons de cinza também mostrou ser uma alternativa viável, com vantagem de menor custo computacional, aspecto relevante para futuras aplicações em dispositivos embarcados.

Apesar dos avanços, o estudo reconhece limitações importantes. A geração dos dados em ambiente controlado, embora necessária para esta fase exploratória, restringe a diversidade de padrões térmicos e materiais encontrados em condições reais de operação. A escala reduzida da maquete e a composição simplificada do sedimento são fatores que precisarão ser calibrados e ajustados para aplicações industriais em larga escala. A extrapolação direta para sistemas reais exige novas etapas de validação, considerando variações de temperatura, umidade, composição do fertilizante, entre outros fatores ambientais.

Além disso, a generalização do modelo para tubulações de diferentes geometrias, materiais ou configurações operacionais será um desafio técnico relevante. A adaptação do modelo poderá envolver técnicas de transferência de aprendizado e novos conjuntos de dados de calibração.

Com base nos resultados obtidos, diversas direções para trabalhos futuros são sugeridas. Primeiramente, investigar o impacto de fatores ambientais, como flutuações de temperatura ambiente, variações de umidade e mudanças na composição dos fertilizantes, sobre a qualidade das imagens térmicas capturadas e a acurácia da predição. Estratégias como a compensação térmica ou a fusão de sensores poderiam ser incorporadas para fortalecer a robustez da metodologia.

Outra linha de evolução relevante é a modelagem temporal da sedimentação, uma vez que o acúmulo de material ocorre progressivamente. A utilização de arquiteturas voltadas a séries temporais, como LSTMs (Long Short-Term Memory) ou TCNs (Temporal Convolutional Networks), permitiria não apenas detectar a sedimentação atual, mas também prever sua evolução, antecipando necessidades de manutenção antes da ocorrência de bloqueios críticos.

Adicionalmente, a fusão multimodal de dados emerge como um caminho promissor. A combinação de imagens térmicas com dados espectrais ou de vibração pode permitir uma análise mais rica, capaz de distinguir entre diferentes tipos de depósitos e detectar sedimentação em estágios iniciais. O uso de aprendizado federado também é uma perspectiva para o futuro, permitindo o treinamento colaborativo de modelos em diversas plantas industriais, respeitando a privacidade dos dados e aumentando a escalabilidade da solução.

Por fim, a incorporação de técnicas de Inteligência Artificial Explicável (XAI) será

importante para a adoção em ambientes industriais críticos. A capacidade de interpretar as decisões do modelo e justificar suas predições é um requisito para atender normas de segurança, auditorias e confiabilidade operacional.

Em suma, este trabalho representa um avanço significativo na aplicação de técnicas de aprendizado profundo para o monitoramento preditivo de dutos industriais, demonstrando a viabilidade de se utilizar termografia associada a redes neurais para inferir o perfil de sedimentação de maneira não invasiva, eficiente e com potencial de integração em sistemas industriais futuros.

REFERÊNCIAS

- (2020). Guest editorial: Non-destructive testing. *Electronics Letters*, 56(19):972–973.
- (2020). *Thermographic Non-Invasive Inspection Modelling of Fertilizer Pipelines Using Neural Networks*. IEEE.
- Abdel, M. (2011). 3d laser scanners: history, applications, and future. *Assiut University. October*.
- Achaw, O.-W. and Danso-Boateng, E. (2021). Fertilizer technology. In *Chemical and Process Industries*, pages 107–133. Springer.
- Al-Kassir, A. R., Fernandez, J., Tinaut, F., and Castro, F. (2005). Thermographic study of energetic installations. *Applied Thermal Engineering*, 25(2-3):183–190.
- Albelwi, S. and Mahmood, A. (2017). A framework for designing the architectures of deep convolutional neural networks. *Entropy*, 19(6):242.
- Alfredo Osornio-Rios, R., Antonino-Daviu, J. A., and de Jesus Romero-Troncoso, R. (2019). Recent industrial applications of infrared thermography: A review. *IEEE Transactions on Industrial Informatics*, 15(2):615–625.
- Alotaibi, M., Honarvar Shakibaei Asli, B., and Khan, M. (2021). Non-invasive inspections: A review on methods and tools. *Sensors*, 21(24).
- and, H. C. (2021). Deep learning algorithms based fingerprint authentication: Systematic literature review. *Journal of Artificial Intelligence and Systems*, 3(1):157–197.
- Bagavathiappan, S., Lahiri, B., Saravanan, T., Philip, J., and Jayakumar, T. (2013). Infrared thermography for condition monitoring—a review. *Infrared Physics & Technology*, 60:35–55.
- Bai, Y., Wong, L., and Twan, T. (2024). Survey on fundamental deep learning 3d reconstruction techniques.

- Balakrishnan, G. K., Yaw, C. T., Koh, S. P., Abedin, T., Raj, A. A., Tiong, S. K., and Chen, C. P. (2022). A review of infrared thermography for condition-based monitoring in electrical energy: Applications and recommendations. *Energies*, 15(16):6000.
- Bao, H., Dinh, N., Lin, L., Youngblood, R., Lane, J., and Zhang, H. (2020). Using deep learning to explore local physical similarity for global-scale bridging in thermal-hydraulic simulation. *Annals of Nuclear Energy*, 147:107684.
- Barrett, C. B. (2020). Overcoming global food security challenges through science and solidarity. *American Journal of Agricultural Economics*, 103(2):422–447.
- Bhowmik, M. K., Saha, K., Majumder, S., Majumder, G., Saha, A., Sarma, A. N., Bhattacharjee, D., Basu, D. K., and Nasipuri, M. (2011). Thermal infrared face recognition—a biometric identification technique for robust security system. *Reviews, refinements and new ideas in face recognition*, 7:113–138.
- Bondarev, E., Heredia, F., Favier, R., Ma, L., and de With, P. H. (2013). On photo-realistic 3D reconstruction of large-scale and arbitrary-shaped environments. In *2013 IEEE 10th Consumer Communications and Networking Conference (CCNC)*, pages 621–624.
- Cadelano, G., Bortolin, A., Ferrarini, G., Molinas, B., Giantin, D., Zonta, P., and Bison, P. (2016). Corrosion detection in pipelines using infrared thermography: Experiments and data processing methods. *Journal of Nondestructive Evaluation*, 35(3).
- Cai, Y., Wang, J., Yuille, A., Zhou, Z., and Wang, A. (2024). Structure-aware sparse-view x-ray 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11174–11183.
- Callister, W. D. (2007). *Materials Science and Engineering: An Introduction*. Wiley, New York, 7 edition.
- Canedo, D. and Neves, A. J. R. (2019). Facial expression recognition using computer vision: A systematic review. *Applied Sciences*, 9(21):4678.
- Carslaw, H. S. and Jaeger, J. C. (1959). *Conduction of Heat in Solids*. Oxford University Press, Oxford, 2 edition.
- Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., et al. (2015a). Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*.
- Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., Xiao, J., Yi, L., and Yu, F. (2015b). Shapenet: An information-rich 3D model repository.

Chansong, D. and Supratid, S. (2021). Impacts of kernel size on different resized images in object recognition based on convolutional neural network. In *2021 9th International Electrical Engineering Congress (iEECON)*, pages 448–451.

Chen, C.-Y., Yeh, C.-H., Chang, B. R., and Pan, J.-M. (2015a). 3D reconstruction from IR thermal images and reprojective evaluations. *Mathematical Problems in Engineering*, 2015:1–8.

Chen, C.-Y., Yeh, C.-H., Chang, B. R., and Pan, J.-M. (2015b). 3d reconstruction from ir thermal images and reprojective evaluations. *Mathematical Problems in Engineering*, 2015.

Chiwiacowsky, L. D. (2002). Uso da função de transferência em problemas de condução do calor com a lei de fourier modificada.

Choi, S., Nguyen, A.-D., Kim, J., Ahn, S., and Lee, S. (2019). Point cloud deformation for single image 3d reconstruction. In *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE.

Ciampa, F., Mahmoodi, P., Pinto, F., and Meo, M. (2018). Recent advances in active infrared thermography for non-destructive testing of aerospace components. *Sensors*, 18(2).

Cone, D. C. (2015). Confined space and limited access situations. In *Emergency Medical Services*, pages 373–376. John Wiley & Sons, Ltd.

de Luis-Ruiz, J. M., Sedano-Cibrián, J., Pérez-Álvarez, R., Pereda-García, R., and Malagón-Picón, B. (2021). Metric contrast of thermal 3D models of large industrial facilities obtained by means of low-cost infrared sensors in UAV platforms. *International Journal of Remote Sensing*, 43(2):457–483.

Dehghani, M., Djolonga, J., Mustafa, B., Padlewski, P., Heek, J., Gilmer, J., Steiner, A., Caron, M., Geirhos, R., Alabdulmohsin, I., Jenatton, R., Beyer, L., Tschannen, M., Arnab, A., Wang, X., Riquelme, C., Minderer, M., Puigcerver, J., Evci, U., Kumar, M., van Steenkiste, S., Elsayed, G. F., Mahendran, A., Yu, F., Oliver, A., Huot, F., Basting, J., Collier, M. P., Gritsenko, A., Birodkar, V., Vasconcelos, C., Tay, Y., Mensink, T., Kolesnikov, A., Pavetić, F., Tran, D., Kipf, T., Lučić, M., Zhai, X., Keysers, D., Harmsen, J., and Houlsby, N. (2023). Scaling vision transformers to 22 billion parameters.

Dey, N., Ashour, A. S., and Althoupety, A. S. (2017). Thermal imaging in medical science. In *Recent Advances in Applied Thermal Imaging for Industrial Applications*, pages 87–117. IGI Global.

Ding, X., Zhang, X., Han, J., and Ding, G. (2022). Scaling up your kernels to 31x31: Revisiting large kernel design in cnns. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11963–11975.

Dlesk, A., Vach, K., and Pavelka, K. (2022). Photogrammetric co-processing of thermal infrared images and rgb images. *Sensors*, 22(4).

D’Odorico, P., Davis, K. F., Rosa, L., Carr, J. A., Chiarelli, D., Dell’Angelo, J., Gephart, J., MacDonald, G. K., Seekell, D. A., Suweis, S., et al. (2018). The global food-energy-water nexus. *Reviews of geophysics*, 56(3):456–531.

Dong, S., Wang, P., and Abbas, K. (2021). A survey on deep learning and its applications. *Computer Science Review*, 40:100379.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*.

Duarte, M., Coch, V., Dias, J., Botelho, S., Duarte, N., and Drews, P. (2019). Non-invasive visual computing modelling for obstruction inspection of pipes in the fertilizer industry. In *2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE)*, pages 281–286.

Duggal, S. and Pathak, D. (2022). Topologically-aware deformation fields for single-view 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1536–1546.

Elallid, B. B., Benamar, N., Hafid, A. S., Rachidi, T., and Mrani, N. (2022). A comprehensive survey on the application of deep and reinforcement learning approaches in autonomous driving. *Journal of King Saud University - Computer and Information Sciences*, 34(9):7366–7390.

Emam, K. E. (2020). Seven ways to evaluate the utility of synthetic data. *IEEE Security & Privacy*, 18(4):56–59.

Ennab, H. (2016). Effect of organic manures, biofertilizers and npk on vegetative growth, yield, fruit quality and soil fertility of eureka lemon trees (citrus limon (l.) burm). *Journal of Soil Sciences and Agricultural Engineering*, 7(10):767–774.

Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A., Liu, Y., Topol, E., Dean, J., and Socher, R. (2021). Deep learning-enabled medical computer vision. *npj Digital Medicine*, 4(1).

- Fan, H., Su, H., and Guibas, L. (2017a). A point set generation network for 3D object reconstruction from a single image. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2463–2471.
- Fan, H., Su, H., and Guibas, L. J. (2017b). A point set generation network for 3d object reconstruction from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 605–613.
- Fang, Q. and Maldague, X. (2020). A method of defect depth estimation for simulated infrared thermography data with deep learning. *Applied Sciences*, 10(19).
- Fang, Q. and Maldague, X. (2022). Coded thermal wave imaging based defect detection in composites using neural networks. *International Journal of Engineering*, 35(1).
- FAO (1996). Declaração de roma sobre a segurança alimentar mundial e plano de acção da cimeira mundial da alimentação. url<https://www.fao.org/3/w3613p/w3613p00.htm>.
- Feng, C., Ji, J., Xu, S., Hou, L., Cui, G., Lan, H., Wei, F., Yang, J., and Yang, W. (2025). Scalable thermal interface materials with close-packed structure and high through-plane thermal conductivity. *Polymer Composites*.
- Feng, J. and Lu, S. (2019). Performance analysis of various activation functions in artificial neural networks. *Journal of Physics: Conference Series*, 1237(2):022030.
- Feng, J., Teng, Q., Li, B., He, X., Chen, H., and Li, Y. (2020). An end-to-end three-dimensional reconstruction framework of porous media from a single two-dimensional image based on deep learning. *Computer Methods in Applied Mechanics and Engineering*, 368:113043.
- França, J. G. D., Gazziro, M. A., Ide, A. N., and Saito, J. H. (2005). A 3d scanning system based on laser triangulation and variable field of view. In *IEEE International Conference on Image Processing 2005*, volume 1, pages I–425. IEEE.
- Fu, K., Peng, J., He, Q., and Zhang, H. (2020). Single image 3d object reconstruction based on deep learning: A review. *Multimedia Tools and Applications*, 80(1):463–498.
- Galvao, L. G., Abbod, M., Kalganova, T., Palade, V., and Huda, M. N. (2021). Pedestrian and vehicle detection in autonomous vehicle perception systems—a review. *Sensors*, 21(21):7267.
- Gao, Z., Zhang, Y., and Li, Y. (2020). Extracting features from infrared images using convolutional neural networks and transfer learning. *Infrared Physics & Technology*, 105:103237.

Gonzalez, R. C. and Woods, R. E. (2002). *Digital Image Processing*. Prentice-Hall, 2nd edition.

Gralinski, F., Stanislawek, T., Wróblewska, A., Lipinski, D., Kaliska, A., Rosalska, P., Topolski, B., and Biecek, P. (2020). Kleister: A novel task for information extraction involving long documents with complex layout. *CoRR*, abs/2003.02356.

Grechi, G., Fiorucci, M., Marmoni, G. M., and Martino, S. (2021). 3D thermal monitoring of jointed rock masses through infrared thermography and photogrammetry. *Remote Sensing*, 13(5):957.

Groysman, A. (2019). *Nondestructive Testing and Corrosion Monitoring*, chapter 9, pages 261–409. John Wiley & Sons, Ltd, New York, USA.

Guerra, R. (2010). Verificação experimental da lei de stefan-boltzmann.

Halstensen, M., de Bakker, P., and Esbensen, K. H. (2006). Acoustic chemometric monitoring of an industrial granulation production process—a pat feasibility study. *Chemometrics and Intelligent Laboratory Systems*, 84(1-2):88–97.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition.

Hellier, C. (2020). *Handbook of Nondestructive Evaluation*, 3E. McGraw-Hill Education.

Hellstein, P. and Szwedo, M. (2016). 3D thermography in non-destructive testing of composite structures. *Measurement Science and Technology*, 27(12):124006.

Holman, J. P. (2010). *Heat Transfer*. McGraw-Hill, New York, 10 edition.

Hou, L., Yu, C., and Samaras, D. (2016). Squared earth mover’s distance-based loss for training deep neural networks. *CoRR*, abs/1611.05916.

Houdt, G. V., Mosquera, C., and Nápoles, G. (2020). A review on the long short-term memory model. *Artificial Intelligence Review*, 53(8):5929–5955.

Hu, T., Wang, L., Xu, X., Liu, S., and Jia, J. (2021). Self-supervised 3d mesh reconstruction from single images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6002–6011.

Huang, J., Yang, L., Zhang, J., Hu, H., and Liu, G. (2021). Granular fertilizer mass flow measurement and vehicle experiments based on microwave doppler method. *Applied Engineering in Agriculture*, 37(1):135–147.

- Ilett, T. P., Yuval, O., Ranner, T., Cohen, N., and Hogg, D. C. (2023). 3d shape reconstruction of semi-transparent worms. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12565–12575.
- Incropera, F. P., DeWitt, D. P., Bergman, T. L., and Lavine, A. S. (2002). *Fundamentals of Heat and Mass Transfer*. Wiley, New York, 5 edition.
- Iqbal, H., Tesfamariam, S., Haider, H., and Sadiq, R. (2016). Inspection and maintenance of oil & gas pipelines: a review of policies. *Structure and Infrastructure Engineering*, 13(6):794–815.
- Ishikawa, M., Kawai, Y., Ishigaki, H., Ogawa, K., and Nishino, H. (2022). Active thermography inspection of residual water in steel pipes: Detection and water height estimation. *Nuclear Engineering and Design*, 386:111566.
- Ishimwe, R., Abutaleb, K., Ahmed, F., et al. (2014). Applications of thermal imaging in agriculture—a review. *Advances in remote Sensing*, 3(03):128.
- Ito, S., Kaneko, N., Yoshida, T., and Sumi, K. (2017). Detection of defective regions in 3D reconstruction to support image acquisition. *SICE Journal of Control, Measurement, and System Integration*, 10(5):402–409.
- Izadi, S., Davison, A., Fitzgibbon, A., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., and Freeman, D. (2011). KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*, New York, New York, USA. ACM Press.
- Jain, R., Gupta, M., Taneja, S., and Hemanth, D. J. (2020). Deep learning based detection and analysis of COVID-19 on chest x-ray images. *Applied Intelligence*, 51(3):1690–1700.
- Jang, W. and Agapito, L. (2024). Nvist: In the wild new view synthesis from a single image with transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10181–10193.
- Jia, T., Wang, H., Chen, D., Hou, M., Wu, C., Tu, M., Jiang, Y., and Zhang, S. (2017). 3D temperature distribution model based on thermal infrared image. *Journal of Sensors*, 2017:1–10.
- Jia, X., Yang, S., Peng, Y., Zhang, J., and Chen, S. (2020). Dv-net: Dual-view network for 3d reconstruction by fusing multiple sets of gated control point clouds. *Pattern Recognition Letters*, 131:376 – 382.
- Jiang, L., Shi, S., Qi, X., and Jia, J. (2018a). Gal: Geometric adversarial loss for single-view 3d-object reconstruction. In *European Conference on Computer Vision*.

- Jiang, L., Shi, S., Qi, X., and Jia, J. (2018b). Gal: Geometric adversarial loss for single-view 3d-object reconstruction. In Ferrari, V., Hebert, M., Sminchisescu, C., and Weiss, Y., editors, *Computer Vision – ECCV 2018*, pages 820–834, Cham. Springer International Publishing.
- Jiang, L., Zhang, J., Deng, B., Li, H., and Liu, L. (2018c). 3d face reconstruction with geometry details from a single image. *IEEE Transactions on Image Processing*, 27(10):4756–4770.
- Jiang, X., Jin, S., Lu, L., Zhang, X., and Lu, S. (2024). Weakly supervised monocular 3d detection with a single-view image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10508–10518.
- Jones, B. F. (1998). A reappraisal of the use of infrared thermal image analysis in medicine. *IEEE transactions on medical imaging*, 17(6):1019–1027.
- Keller, M., Popelin, A.-L., Bousquet, N., and Remy, E. (2015). Nonparametric estimation of the probability of detection of flaws in an industrial component, from destructive and nondestructive testing data, using approximate bayesian computation. *Risk Analysis*, 35(9):1595–1610.
- Kennedy, C. F. G. (2011). risk analysis paralysis. *Aerospace America*, page 29.
- Kennedy, D. A., Lee, T., and Seely, D. (2009). A comparative review of thermography as a breast cancer screening technique. *Integrative cancer therapies*, 8(1):9–16.
- Khan, A., Sohail, A., Zahoor, U., and Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, 53(8):5455–5516.
- Khor, W., Chen, Y. K., Roberts, M., and Ciampa, F. (2024a). Automated detection and classification of concealed objects using infrared thermography and convolutional neural networks. *Scientific Reports*, 14(1).
- Khor, W., Chen, Y. K., Roberts, M., and Ciampa, F. (2024b). Non-contact, portable, and stand-off infrared thermal imager for security scanning applications. *AIP Advances*, 14(4).
- Kopittke, P. M., Menzies, N. W., Wang, P., McKenna, B. A., and Lombi, E. (2019). Soil and the intensification of agriculture for global food security. *Environment international*, 132:105078.
- Kovács, P., Lehner, B., Thummerer, G., Mayr, G., Burgholzer, P., and Huemer, M. (2020). Deep learning approaches for thermographic imaging. *Journal of Applied Physics*, 128(15):155103.

- Kuo, W., Angelova, A., Lin, T.-Y., and Dai, A. (2020). Mask2cad: 3d shape prediction by learning to segment and retrieve.
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., and Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4681–4690.
- Lee, K. B., Cheon, S., and Kim, C. O. (2017). A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes. *IEEE Transactions on Semiconductor Manufacturing*, 30(2):135–142.
- Lee, Y. (2024). Three-dimensional dense reconstruction: A review of algorithms and datasets. *Sensors*, 24(18):5861.
- Li, B., Zhang, Y., Zhao, B., and Shao, H. (2020a). 3d-ReConstnet: A single-view 3d-object point cloud reconstruction network. *IEEE Access*, 8:83782–83790.
- Li, B., Zhang, Y., Zhao, B., and Shao, H. (2020b). 3d-ReConstnet: A single-view 3d-object point cloud reconstruction network. *IEEE Access*, 8:83782–83790.
- Li, J. and Yang, X. (2020). A cyclical learning rate method in deep learning training. In *2020 International Conference on Computer, Information and Telecommunication Systems (CITS)*, pages 1–5.
- Li, M., Hu, L., Xiong, Z., Zhang, B., Pan, P., and Liu, D. (2022). Recurrent dynamic embedding for video object segmentation. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE.
- Lian, Y., Feng, T., and Zhou, J. (2019). A dense pointnet++ architecture for 3d point cloud semantic segmentation. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 5061–5064. IEEE.
- Liang, Z., Liu, B., Xie, M., and Parlikad, A. K. (2020). Condition-based maintenance for long-life assets with exposure to operational and environmental risks. *International Journal of Production Economics*, 221:107482.
- Liao, T., Zhang, X., Xiu, Y., Yi, H., Liu, X., Qi, G.-J., Zhang, Y., Wang, X., Zhu, X., and Lei, Z. (2023). High-fidelity clothed avatar reconstruction from a single image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8662–8672.

- Lile, N. L. T., Jaafar, M. H. M., Roslan, M. R., and Azmi, M. S. M. (2012). Blockage detection in circular pipe using vibration analysis. *International Journal on Advanced Science, Engineering and Information Technology*, 2(3):252.
- Lin, C.-H., Kong, C., and Lucey, S. (2018). Learning efficient point cloud generation for dense 3d object reconstruction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1).
- Lin, F., Yue, Y., Hou, S., Yu, X., Xu, Y., Yamada, K. D., and Zhang, Z. (2023). Hyperbolic chamfer distance for point cloud completion. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 14549–14560.
- Liu, A., Lin, C., Liu, Y., Long, X., Dou, Z., Guo, H.-X., Luo, P., and Wang, W. (2024). Part123: Part-aware 3d reconstruction from a single-view image.
- Liu, F. and Liu, X. (2021). Voxel-based 3d detection and reconstruction of multiple objects from a single image. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W., editors, *Advances in Neural Information Processing Systems*, volume 34, pages 2413–2426. Curran Associates, Inc.
- Lobão, F. S. (2004). Calibração com mapeamento das distorções geométricas aplicada a um sistema de visão estéreo. *Distrito Federal: Universidade De Brasilia. Faculdade De Tecnologia*.
- Lu, J., Li, Z., Bai, J., and Yu, Q. (2022). Oriented and directional chamfer distance losses for 3d object reconstruction from a single image. *IEEE Access*, 10:61631–61638.
- Luo, Q., Gao, B., Woo, W., and Yang, Y. (2019). Temporal and spatial deep learning network for infrared thermal defect detection. *NDT & E International*, 108:102164.
- Ma, D., Fang, H., Wang, N., Pang, G., Li, B., Dong, J., and Jiang, X. (2023). A low-cost 3d reconstruction and measurement system based on structure-from-motion (sfm) and multi-view stereo (mvs) for sewer pipelines. *Tunnelling and Underground Space Technology*, 141:105345.
- Maas, A. L., Hannun, A. Y., Ng, A. Y., et al. (2013). Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, volume 30, page 3. Atlanta, Georgia, USA.
- Madanhire, I., Mugwindiri, K., and Mbohwa, C. (2015). Enhancing cleaner production application in fertilizer manufacturing: case study. *Clean Technologies and Environmental Policy*, 17(3):667–679.

- Maldague, X., Krapez, J. C., Cielo, P., and Poussart, D. (1988). Processing of thermal images for the detection and enhancement of subsurface flaws in composite materials. In Chen, C. H., editor, *Signal Processing and Pattern Recognition in Nondestructive Evaluation of Materials*, pages 257–285, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Mandikal, P., Navaneet, K. L., and Babu, R. V. (2019). 3d-psrnet: Part segmented 3d point cloud reconstruction from a single image. In *Lecture Notes in Computer Science*, pages 662–674. Springer International Publishing.
- Manohar, A. and di Scalea, F. L. (2014). Modeling 3d heat flow interaction with defects in composite materials for infrared thermography. *NDT & E International*, 66:1–7.
- Memon, J., Sami, M., Khan, R. A., and Uddin, M. (2020). Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR). *IEEE Access*, 8:142642–142668.
- Meng, G., Zhou, R., Liu, L., Liang, P., Liu, F., Chen, D., Niemier, M., and Hu, X. S. (2024). Efficient approximation of earth mover’s distance based on nearest neighbor search.
- Meng, X., Wang, Y., Liu, J., and He, W. (2019). Nondestructive inspection of curved clad composites with subsurface defects by combination active thermography and three-dimensional (3d) structural optical imaging. *Infrared Physics & Technology*, 97:424–431.
- Metzer, G., Hanocka, R., Giryas, R., Mitra, N. J., and Cohen-Or, D. (2022). Z2p: Instant visualization of point clouds. *Computer Graphics Forum*, 41(2):461–471.
- Mix, P. E. (2005). *Introduction to Nondestructive Testing*.
- Möller, K. (2018). Soil fertility status and nutrient input–output flows of specialised organic cropping systems: a review. *Nutrient Cycling in Agroecosystems*, 112(2):147–164.
- Muller, A. O. and Kroll, A. (2017). Generating high fidelity 3-d thermograms with a handheld real-time thermal imaging system. *IEEE Sensors Journal*, 17(3):774–783.
- Müller, T., Evans, A., Schied, C., and Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41(4):1–15.
- Müller, D., Netzelmann, U., and Valeske, B. (2022). Defect shape detection and defect reconstruction in active thermography by means of two-dimensional convolutional neural network as well as spatiotemporal convolutional lstm network. *Quantitative InfraRed Thermography Journal*, 19(2):126–144.

Nair, K. P. (2019). Soil fertility and nutrient management. In *Intelligent soil management for sustainable agriculture*, pages 165–189. Springer.

Nguyen, V. N., Groueix, T., Ponimatkin, G., Hu, Y., Marlet, R., Salzmann, M., and Lepetit, V. (2024). Nope: Novel object pose estimation from a single image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 17923–17932.

Nikolenko, S. I. (2021). *Synthetic Data for Deep Learning*.

Nogueira, J. C. B. and Reis, V. P. (2010). Emprego da termografia na inspeção preditiva. *Bolsista de Valor: Revista de divulgação do Projeto Universidade Petrobras e IF Fluminense*, 1:29–35.

Nuryanto, R., Nurcahyo, R., and Farizal (2020). Using failure and repair data for performance evaluation of npk fertilizer production line. *2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA)*, pages 404–408.

Okuno, E. and Yoshimura, E. M. (2016). *Física das radiações*. Oficina de Textos.

Oliva, A. and Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, 11(12):520–527.

ONU (1948). Declaração universal dos direitos humanos. [urlhttps://www.unicef.org/brazil/declaracao-universal-dos-direitos-humanos](https://www.unicef.org/brazil/declaracao-universal-dos-direitos-humanos).

Oral, B., Emekligil, E., Arslan, S., and Eryiğit, G. (2020). Information extraction from text intensive and visually rich banking documents. *Information Processing & Management*, 57(6):102361.

Ortega, M., Ivorra, E., Juan, A., Venegas, P., Martínez, J., and Alcañiz, M. (2021). Mantra: An effective system based on augmented reality and infrared thermography for industrial maintenance. *Applied Sciences*, 11(1):385–411.

Panayotov, V., Chen, G., Povey, D., and Khudanpur, S. (2015). Librispeech: An ASR corpus based on public domain audio books. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE.

Panchal, D., Mangla, S. K., Tyagi, M., and Ram, M. (2018). Risk analysis for clean and sustainable production in a urea fertilizer industry. *International Journal of Quality & Reliability Management*, 35(7):1459–1476.

Paulus, S., Schumann, H., Kuhlmann, H., and Léon, J. (2014). High-precision laser scanning system for capturing 3d plant architecture and analysing growth of cereal plants. *Biosystems Engineering*, 121:1–11.

- Pfeifer, N. and Briese, C. (2007). Laser scanning—principles and applications. In *GeoSiberia 2007-International Exhibition and Scientific Congress*, pages cp–59. European Association of Geoscientists & Engineers.
- Piao, C., Kim, S. H., Lee, J. K., Choi, W. G., and Kim, Y. Y. (2022). Non-invasive ultrasonic inspection of sludge accumulation in a pipe. *Ultrasonics*, 119:106602.
- Posada, J., Toro, C., Barandiaran, I., Oyarzun, D., Stricker, D., De Amicis, R., Pinto, E. B., Eisert, P., Döllner, J., and Vallarino, I. (2015). Visual computing as a key enabling technology for industrie 4.0 and industrial internet. *IEEE computer graphics and applications*, 35(2):26–40.
- Prisacariu, V. A. and Reid, I. D. (2012). PWP3D: Real-time segmentation and tracking of 3D objects. *International Journal of Computer Vision*, 98(3):335–354.
- Raghunathan, T. E. (2021). Synthetic data. *Annual Review of Statistics and Its Application*, 8(1):129–140.
- Ramachandran, P., Zoph, B., and Le, Q. V. (2017). Searching for activation functions.
- Rangel, J. and Soldan, S. (2014). 3d thermal imaging: Fusion of thermography and depth cameras. In *Proceedings of the 2014 International Conference on Quantitative InfraRed Thermography*. QIRT Council.
- Raschka, S. (2018). MLxtend: Providing machine learning and data science utilities and extensions to python’s scientific computing stack. *Journal of Open Source Software*, 3(24):638.
- Ratsakou, A., Skarlatos, A., Reboud, C., and Lesselier, D. (2020). Shape reconstruction of delamination defects using thermographic infrared signals based on an enhanced canny approach. *Infrared Physics & Technology*, 111:103527.
- Reddy, P., Elezi, I., and Deng, J. (2024). G3dr: Generative 3d reconstruction in imagenet. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9655–9665.
- Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., and Savarese, S. (2019). Generalized intersection over union.
- Romero, D. W., Brountjes, R.-J., Tomczak, J. M., Bekkers, E. J., Hoogendoorn, M., and van Gemert, J. C. (2022). Flexconv: Continuous kernel convolutions with differentiable kernel sizes.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation.

- Rubner, Y. (2000). The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121.
- Rusu, R. B. and Cousins, S. (2011). 3D is here: Point cloud library (pcl). In *2011 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–4.
- Saifi, A., Amiri, A. E., Elhassnaoui, A., Obbadi, A., Errami, Y., and Sahnoun, S. (2015). Thermography model for detecting the scale thickness in water pipes. In *3rd International Renewable and Sustainable Energy Conference (IRSEC)*, pages 1–4.
- Santhi, V. (2017). Recent advances in applied thermal imaging for industrial applications.
- Santos, C. F. G. D. and Papa, J. a. P. (2022). Avoiding overfitting: A survey on regularization methods for convolutional neural networks. *ACM Comput. Surv.*, 54(10s).
- Santos, Y. F. (2018). A utilização da termografia como técnicas de manutenção preditiva em indústria de fios—estudo de caso.
- Saponara, S., Elhanashi, A., and Zheng, Q. (2021). Recreating fingerprint images by convolutional neural network autoencoder architecture. *IEEE Access*, 9:147888–147899.
- Sarker, I. H. (2021a). Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, 2(6).
- Sarker, I. H. (2021b). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3).
- Schramm, S., Osterhold, P., Schmoll, R., and Kroll, A. (2014). Generation of large-scale 3d thermograms in real-time using depth and infrared cameras. In *International Conference on Quantum Infrared Thermography*, pages 1–8.
- Sels, S., Verspeek, S., Ribbens, B., Bogaerts, B., Vanlanduit, S., Penne, R., and Steenackers, G. (2019a). A cad matching method for 3d thermography of complex objects. *Infrared Physics & Technology*, 99:152–157.
- Sels, S., Verspeek, S., Ribbens, B., Bogaerts, B., Vanlanduit, S., Penne, R., and Steenackers, G. (2019b). A CAD matching method for 3D thermography of complex objects. *Infrared Physics & Technology*, 99:152–157.
- Shafiq, M. and Gu, Z. (2022). Deep residual learning for image recognition: A survey. *Applied Sciences*, 12(18):8972.
- Sharp, J. and Ersdal (2021). *Inspection Methods for Offshore Structures Underwater*, chapter 4, pages 79–140. John Wiley Sons, Ltd.

- Shih, K.-J., Lin, Y.-H., and Hsiung, P.-A. (2009). Quality-of-service optimization for multiple multimedia tasks in real-time embedded systems. In *2009 10th International Symposium on Pervasive Systems, Algorithms, and Networks*, pages 474–479. IEEE.
- Shorten, C. and Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*.
- Shorten, D. C. (2018). *Introduction to Nondestructive Testing and Condition Monitoring*, pages 1–7. John Wiley Sons, Ltd.
- Simoens, J., Verguts, T., and Braem, S. (2024). Learning environment-specific learning rates. *PLoS computational biology*, 20(3):e1011978.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition.
- Soomro, K., Zamir, A. R., and Shah, M. (2012). Ucf101: A dataset of 101 human actions classes from videos in the wild.
- Speith, T. (2022). A review of taxonomies of explainable artificial intelligence (xai) methods. In *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency*, pages 2239–2250.
- Stando, A., Cavus, M., and Biecek, P. (2023). The effect of balancing methods on model behavior in imbalanced classification problems.
- Sun, S., Zhao, C., Guo, Y., Wang, R., Huang, X., Chen, Y. V., and Ren, L. (2024). Behind the veil: Enhanced indoor 3d scene reconstruction with occluded surfaces completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12744–12753.
- Sun, X., Wu, J., Zhang, X., Zhang, Z., Zhang, C., Xue, T., Tenenbaum, J. B., and Freeman, W. T. (2018). Pix3d: Dataset and methods for single-image 3d shape modeling. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE.
- Swamidoss, I. N., Almelhi, O. K., Mansoori, A. A. A., Amro, A. B., Marzooqi, A. M. A., and Sayadi, S. (2021a). Robust stereo infrared camera calibration for 3d depth estimation and measurement. In Valenta, C. R., Shaw, J. A., and Kimata, M., editors, *SPIE Future Sensing Technologies 2021*. SPIE.
- Swamidoss, I. N., Amro, A. B., and Sayadi, S. (2021b). Systematic approach for thermal imaging camera calibration for machine vision applications. *Optik*, 247:168039.
- Syamila, A. I. (2020). Health and safety risk analysis in the fertilizer industry. *Indian Journal of Public Health Research & Development*, 11(6).

Szeliski, R. (2022). *Computer Vision*. 2^o edition.

Szymanowicz, S., Rupprecht, C., and Vedaldi, A. (2024). Splatter image: Ultra-fast single-view 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10208–10217.

Thai, P., Alam, S., Lilith, N., and Nguyen, B. T. (2022). A computer vision framework using convolutional neural networks for airport-airside surveillance. *Transportation Research Part C: Emerging Technologies*, 137:103590.

Timsina, J. (2018). Can organic sources of nutrients increase crop yields to meet global food demand? *Agronomy*, 8(10):214.

Tliba, M., Chetouani, A., Valenzise, G., and Dufaux, F. (2022). Point cloud quality assessment using cross-correlation of deep features. In *Proceedings of the 2nd Workshop on Quality of Experience in Visual Multimedia Applications, MM '22*, page 63–68. ACM.

Transue, S., Nguyen, P., Vu, T., and Choi, M.-H. (2019). Volumetric reconstruction of thermal-depth fused 3D models for occluded body posture estimation. *Smart Health*, 11:29–44.

Trindade, P. M., Facco, D. S., and Pereira Filho, W. (2012). Aula 1-sensoriamento remoto: evolução histórica e princípios físicos.

Usamentiaga, R., Venegas, P., Guerediaga, J., Vega, L., Molleda, J., and Bulnes, F. (2014). Infrared thermography for temperature measurement and non-destructive testing. *Sensors*, 14(7):12305–12348.

Uy, M. A., Pham, Q.-H., Hua, B.-S., Nguyen, D. T., and Yeung, S.-K. (2019). Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 1588–1597.

Vadivambal, R. and Jayas, D. S. (2011). Applications of thermal imaging in agriculture and food industry—a review. *Food and bioprocess technology*, 4(2):186–199.

Vidas, S., Moghadam, P., and Sridharan, S. (2015). Real-time mobile 3d temperature mapping. *IEEE Sensors Journal*, 15(2):1145–1152.

Vollmer, M. and Mollmann, K.-P. (2018). *Infrared Thermal Imaging*. Wiley-VCH.

Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., and Bowman, S. R. (2018a). Glue: A multi-task benchmark and analysis platform for natural language understanding.

Wang, H., Guo, J., Yan, D., Quan, W., and Zhang, X. (2018b). Learning 3d keypoint descriptors for non-rigid shape matching. In *ECCV*.

Wang, J., Yoon, J. S., Wang, T. Y., Singh, K. K., and Neumann, U. (2023a). Complete 3d human reconstruction from a single incomplete image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8748–8758.

Wang, N., Dong, J., Fang, H., Li, B., Zhai, K., Ma, D., Shen, Y., and Hu, H. (2023b). 3d reconstruction and segmentation system for pavement potholes based on improved structure-from-motion (sfm) and deep learning. *Construction and Building Materials*, 398:132499.

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., and Loy, C. C. (2018c). ESRGAN: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision Workshops (ECCVW)*.

Wang, Y. O., Chung, Y., Wu, C. H., and De la Torre, F. (2024). Domain gap embeddings for generative dataset augmentation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, page 28684–28694. IEEE.

Wang, Z., She, Q., and Ward, T. E. (2021). Generative adversarial networks in computer vision: A survey and taxonomy. *ACM Computing Surveys*, 54(2):1–38.

Wegelin, F. A. (2014). *SISTEMA DE DETECC AO DE VAZAMENTOS EM DUTOS UTILIZANDO FILTRAGEM ADAPTATIVA*. PhD thesis, Universidade Federal do Rio de Janeiro.

Wu, C., Wang, W., Yuan, Q., Li, Y., Zhang, W., and Zhang, X. (2011). Infrared thermography non-destructive testing of composite materials. *Advanced Materials Research*, 291-294:1307–1310.

Wu, T., Pan, L., Zhang, J., Wang, T., Liu, Z., and Lin, D. (2021a). Density-aware chamfer distance as a comprehensive metric for point cloud completion.

Wu, T., Pan, L., Zhang, J., Wang, T., Liu, Z., and Lin, D. (2021b). Density-aware chamfer distance as a comprehensive metric for point cloud completion. *CoRR*, abs/2111.12702.

Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., and Xiao, J. (2014). 3d shapenets: A deep representation for volumetric shapes.

Xiang, Y., Mottaghi, R., and Savarese, S. (2014). Beyond pascal: A benchmark for 3d object detection in the wild. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*.

- Xie, J., Zheng, Y., Du, R., Xiong, W., Cao, Y., Ma, Z., Cao, D., and Guo, J. (2021). Deep learning-based computer vision for surveillance in ITS: Evaluation of state-of-the-art methods. *IEEE Transactions on Vehicular Technology*, 70(4):3027–3042.
- Xu, C., Xie, J., Chen, G., and Huang, W. (2014). An infrared thermal image processing framework based on superpixel algorithm to detect cracks on metal surface. *Infrared Physics & Technology*, 67:266–272.
- Yang, B., Rosa, S., Markham, A., Trigoni, N., and Wen, H. (2019). Dense 3d object reconstruction from a single depth view. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(12):2820–2834.
- Yang, M.-D., Su, T.-C., and Lin, H.-Y. (2018). Fusion of infrared thermal image and visible image for 3d thermal model reconstruction using smartphone sensors. *Sensors*, 18(7):2003.
- Yuan, W., Khot, T., Held, D., Mertz, C., and Hebert, M. (2018). PCN: Point completion network. In *2018 International Conference on 3D Vision (3DV)*. IEEE.
- Z., A. Y., R., D. K., Aadarsh, S., and G., H. K. (2020). Border surveillance system using computer vision. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE.
- Zeng, W., Karaoglu, S., and Gevers, T. (2018). Inferring point clouds from single monocular images by depth intermediation.
- Zhang, Z. (2021). *Iterative Closest Point (ICP)*, pages 718–720. Springer International Publishing, Cham.
- Zhao, J., Wang, X., Tian, H., Lu, Y., Guo, C., and Liu, H. (2021). A fertilizer discharge detection system based on point cloud data and an efficient volume conversion algorithm. *Computers and Electronics in Agriculture*, 185:106131.
- Zheng, X., Zheng, S., Kong, Y., and Chen, J. (2021). Recent advances in surface defect inspection of industrial products using deep learning techniques. *The International Journal of Advanced Manufacturing Technology*, 113(1-2):35–58.
- Zhou, D., Hou, Q., Chen, Y., Feng, J., and Yan, S. (2020). Rethinking bottleneck structure for efficient mobile network design. In Vedaldi, A., Bischof, H., Brox, T., and Frahm, J.-M., editors, *Computer Vision – ECCV 2020*, pages 680–697, Cham. Springer International Publishing.
- Zhu, J., Xu, Y., Ye, Z., Hoegner, L., and Stilla, U. (2021). Fusion of urban 3D point clouds with thermal attributes using MLS data and TIR image sequences. *Infrared Physics & Technology*, 113:103622.

Zou, Z.-X., Yu, Z., Guo, Y.-C., Li, Y., Liang, D., Cao, Y.-P., and Zhang, S.-H. (2024). Triplane meets gaussian splatting: Fast and generalizable single-view 3d reconstruction with transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10324–10335.