

UNIVERSIDADE FEDERAL DO RIO GRANDE - FURG
CENTRO DE CIÊNCIAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO
CURSO DE MESTRADO EM ENGENHARIA DE COMPUTAÇÃO

Dissertação de Mestrado

Classificação de Tipos de Água em Imagens Monoculares Subaquáticas

Edwilson Silva Vaz Junior

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal do Rio Grande - FURG, como requisito parcial para a obtenção do grau de Mestre em Engenharia de Computação.

Orientador: Prof. Dr. Paulo Lilles Jorge Drews Junior

Rio Grande, 2024

Ficha Catalográfica

V393c Vaz Junior, Edwilson Silva.

Classificação de tipos de água em imagens monoculares subaquáticas / Edwilson Silva Vaz Junior. – 2024.

51 f.

Dissertação (mestrado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-Graduação em Computação, Rio Grande/RS, 2024.

Orientador: Dr. Paulo Lilles Jorge Drews Junior.

1. Classificação de tipos de água 2. Imagens monoculares subaquáticas 3. Modelo de Jerlov I. Drews Junior, Paulo Lilles Jorge II. Título.

CDU 004

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

Banca examinadora:



Documento assinado digitalmente
PAULO LILLES JORGE DREWS JUNIOR
Data: 30/07/2024 22:08:41-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Paulo Lilles Jorge Drews Junior (C3-FURG)



Documento assinado digitalmente
THIAGO LOPES TRUGILLO DA SILVEIRA
Data: 31/07/2024 11:53:44-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Thiago Lopes Trugillo da Silveira (DINF-UFRGS)



Documento assinado digitalmente
LEONARDO RAMOS EMMENDORFER
Data: 05/08/2024 14:02:33-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Leonardo Ramos Emmendorfer (CT-UFSM)



Documento assinado digitalmente
EDUARDO NUNES BORGES
Data: 10/08/2024 09:58:09-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Eduardo Nunes Borges (C3-FURG)

*Dedico este trabalho a meus pais, que apesar
de todas dificuldades, sempre me incentivaram
e estiveram do meu lado durante essa caminhada.*

AGRADECIMENTOS

A Deus por me dar força, saúde mental e resiliência.

Aos meus pais, Edwilson Silva Vaz e Andréia da Roza Vaz, por nunca desistirem, me incentivarem e serem meus melhores amigos em todos os momentos.

Aos meus avôs Orion Lima Vaz (*in memorian*) e Osvaldir Silveira da Roza (*in memorian*) e ao meu bisavô Ítalo Silva (*in memorian*), meus maiores incentivadores e exemplos.

Às minhas avós Amália Rodrigues da Roza e Edi Maria Silva Vaz por terem ajudado a me educar e me criar.

Aos meus irmãos Elik Vaz e Kevyn da Roza por serem ombros amigos em bons e maus momentos.

Aos meus afilhados Pablo da Roza e Évylin da Roza e ao meu sobrinho Pedro Henrique Vaz por botarem sorriso no meu rosto nos momentos de dificuldade.

Ao meu grande orientador Paulo Drews Junior pela insistência, persistência, incentivo e vontade de fazer acontecer, além de todos os conselhos, amizade, parceria e conhecimento compartilhado.

Aos meus tios, primos, amigos e demais familiares.

E a todos que de alguma forma colaboraram direta ou indiretamente nesta etapa, meu sincero muito obrigado.

RESUMO

VAZ JUNIOR, Edwilson Silva. **Classificação de Tipos de Água em Imagens Monoculares Subaquáticas**. 2024. 51 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Computação. Universidade Federal do Rio Grande - FURG, Rio Grande.

O crescimento da capacidade computacional possibilitou um avanço no desenvolvimento de novos métodos na área de processamento de imagens. Consequentemente, aumentam-se os estudos em imagens afetadas por meios participativos, em especial por imagens subaquáticas. Embora com grande aplicação tanto em ecologia quanto industrial, o alto grau de absorção da luz faz o processamento de imagens subaquáticas um grande desafio. Para métodos que utilizam imagens monoculares, entender o contexto de onde aquela cena está submersa pode ser crucial. Em especial, para imagens subaquáticas, das classificações pré-existentes na literatura, poder estimar qual tipo de água a imagem está inserida é uma tarefa árdua. A falta de grandes conjuntos de imagens subaquáticas reais, principalmente capturadas em diferentes tipos de água, como águas oceânicas e costeiras, dificulta uma pesquisa mais aprofundada na área. Para tal, este trabalho apresenta um *dataset* com imagens de diferentes resoluções, tamanhos e tipos de água para poder testar diferentes metodologias. O *dataset* proposto conta ainda com referência do tipo de água que a imagem está alocada. Como foco principal do trabalho, dois métodos de classificação de tipos de água são propostos. O Classificador Baseado em Modelo Físico (CLAFIS) baseia-se na otimização de classificação de uma abordagem clássica, utilizando uma modelagem física baseada na literatura. O Classificador Baseado em Modelo de Cor (CLACOR) é um método de classificação a qual se baseia em uma série de análises empíricas focando na coloração da imagem, em especial na região onde se tem apenas água. São apresentados resultados quantitativos comparando os dois métodos propostos com a literatura e apresentando resultados qualitativos apresentando e detalhando imagens e estimações dos algoritmos, bem como casos falhos dos métodos. CLAFIS e CLACOR performam melhor que a literatura segundo as métricas utilizadas.

Palavras-chave: Classificação de tipos de água, Imagens monoculares subaquáticas, Modelo de Jerlov.

ABSTRACT

VAZ JUNIOR, Edwilson Silva. **Water Type Classification in Monocular Underwater Images**. 2024. 51 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Computação. Universidade Federal do Rio Grande - FURG, Rio Grande.

The growth of computational capacity has enabled advancement in the development of new methods in the field of image processing. Consequently, there is an increase in studies on images affected by participatory media, especially underwater images. Although widely applicable in both ecology and industry, the high degree of light absorption makes underwater image processing a significant challenge. For methods using monocular images, understanding the context of where the scene is submerged can be crucial. In particular, for underwater images, given pre-existing classifications in the literature, estimating the type of water the image is immersed in is a daunting task. The lack of large sets of real underwater images, mainly captured in different types of water, such as oceanic and coastal waters, hinders further research in the field. To address this, this work presents a dataset with images of different resolutions, sizes, and types of water to test different methodologies. The proposed dataset also includes a reference to the type of water the image is allocated to. As the main focus of the work, two methods for classifying types of water are proposed. The Physics-Based Classifier (CLAFIS) is based on optimizing the classification of a classical approach, using a physically based modeling based on the literature. The Color Model-Based Classifier (CLACOR) is a classification method based on a series of empirical analyses focusing on image coloring, especially in the region where only water is present. Quantitative results comparing the two proposed methods with the literature are presented, as well as qualitative results presenting and detailing images and algorithm estimates, as well as cases of method failures. CLAFIS and CLACOR outperform the literature according to the metrics used.

Keywords: water type classification, monocular underwater images, Jerlov model.

LISTA DE FIGURAS

1	Diferentes tipos de meios participativos. Sendo: (a) chuva, (b) água, (c) fuligem de solda, (d) névoa, (e) neve e (f) poeira.	13
2	Mapa mundi segundo classificações de tipo de água propostas por Jerlov [1]. Fonte: Johnson <i>et al.</i> [2].	19
3	Os tipos de água propostos por Jerlov [1] agrupados em 6 classificações propostas por Uplavikar <i>et al.</i> [3].	20
6	Diagrama elucidativo do primeiro passo para escolha do <i>cluster</i> com apenas água.	24
7	<i>Cluster</i> de área de apenas água. Na ordem da esquerda para direita: imagem fonte, região com apenas água de acordo com Berman <i>et al.</i> [4] e região selecionada pelo método <i>KMeans</i> como apenas água do CLACOR. A região água aparece na mesma cor da imagem fonte.	26
8	Da esquerda para direita: imagem fonte, saída do algoritmo <i>KMeans</i> e imagem final pós-otimização.	27
9	Diagrama demonstrativo de fluxo para seleção do <i>cluster</i> de área com apenas água em imagens com canal azul dominante.	28
10	Imagens de exemplo que possuem média global de cores descartada como critério de escolha. A imagem da linha superior tem sua média global de cores descartada em função do espaço HSV, enquanto a imagem inferior em função do canal vermelho.	31
11	Diagrama demonstrativo de fluxo para seleção do <i>cluster</i> de área com apenas água em imagens com canal verde dominante.	32
12	Diagrama demonstrativo de fluxo para seleção do <i>cluster</i> de área com apenas água em imagens com canal vermelho dominante.	34
13	Resultados obtidos para CLACOR e CLAFIS utilizando imagens dos tipos de água oceânicas do DRTA.	39
14	Resultados obtidos para CLACOR e CLAFIS utilizando imagens dos tipos de água oceânicas do DRTA.	41
15	Resultados obtidos para CLACOR e CLAFIS utilizando imagens dos tipos de água costeiras do DRTA.	42
16	Resultados obtidos para CLACOR e CLAFIS utilizando imagens dos tipos de água costeiras do DRTA.	43
17	Casos falhos onde imagens são capturadas para o fundo do mar e superfície respectivamente. A esquerda, a imagem original. A direita, o <i>cluster</i> selecionado pelo algoritmo CLACOR.	44

18	Casos falhos de escolha errada de <i>cluster</i> pelo algoritmo CLACOR. A esquerda, a imagem original. A direita, o <i>cluster</i> selecionado. . . .	44
----	--	----

LISTA DE TABELAS

1	Resultados quantitativos utilizando o <i>dataset</i> DRTA.	36
2	Matriz de confusão de CLAFIS segundo a classificação de Jerlov [1] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método CLAFIS. 36	
3	Matriz de confusão de CLACOR segundo a classificação de Jerlov [1] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método CLACOR.	37
4	Matriz de confusão de Berman <i>et al.</i> [4] segundo a classificação de Jerlov [1] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método Berman <i>et al.</i> [4].	37
5	Matriz de confusão de CLAFIS segundo a classificação de Uplavikar <i>et al.</i> [3] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método CLAFIS.	37
6	Matriz de confusão de CLACOR segundo a classificação de Uplavikar <i>et al.</i> [3] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método CLACOR.	38
7	Matriz de confusão de Berman <i>et al.</i> [4] segundo a classificação de Uplavikar <i>et al.</i> [3] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método Berman <i>et al.</i> [4].	38
8	Resultados detalhados da Figura 13	40
9	Resultados detalhados da Figura 14	40
10	Resultados detalhados da Figura 15	40
11	Resultados detalhados da Figura 16	40

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Motivação	14
1.2	Objetivos e Contribuições	14
1.3	Estrutura do Trabalho	15
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	Modelo de formação de imagem	16
2.2	Definição do problema	17
3	TRABALHOS RELACIONADOS	18
3.1	Conclusões	21
4	METODOLOGIA	22
4.1	Classificação Baseada em Modelo Físico - CLAFIS	22
4.2	Classificação Baseada em Modelo de Cor - CLACOR	23
4.2.1	Critérios de escolha para a região com água	27
5	RESULTADOS	35
5.1	<i>Dataset</i> de Referência do Tipo de Água	35
5.2	Resultados quantitativos	35
5.3	Resultados qualitativos	36
5.3.1	Tipos de água oceânicas	38
5.3.2	Tipos de água costeiras	40
5.3.3	Casos falhos	42
6	CONSIDERAÇÕES FINAIS	45
6.1	Conclusões	45
6.2	Trabalhos futuros	45
	REFERÊNCIAS	47

1 INTRODUÇÃO

Nos últimos tempos, inúmeros métodos e técnicas foram desenvolvidos para extrair informações da imagem, seja através de métodos de processamento de imagem ou visão computacional. Apesar das grandes capacidades criadas, quase todos métodos são desenvolvidos para meios não-participativos, ou seja, se assume que câmera, bem como a cena, está em um ambiente de ar puro. Assume-se, portanto, que os raios de luz não sofrem alteração no seu caminho da cena até o sensor da câmera. Apesar desta inferência, muitos meios, os chamados meios participativos, interferem direta e indiretamente na direção dos raios de luz, bem como em suas respectivas intensidades de raio luminoso.

O meio participativo ocasiona em distorção e absorção da luz antes que esta chegue à câmera. O problema causado pelos meios participativos se torna relevante uma vez que, devido à distorção e absorção da luz, o contraste da imagem é reduzido como um todo. Além disso, há o efeito de modificação de cor, ou seja, há alteração na coloração da imagem em relação à coloração da cena. Ambos fatores impactam diretamente na redução de visibilidade da cena observada. Esta degradação torna-se mais relevante de acordo com a distância da cena com a câmera. Uma vez que esta distância aumenta, menor a visibilidade dos objetos na cena.

Apesar da maior parte das técnicas de processamento de imagem adotadas na literatura não levarem em conta os meios participativos, problemas do mundo real envolvem meios os quais interferem nas imagens capturadas. Em geral, os métodos que levam em consideração os meios participativos restauram a imagem de entrada (sob os efeitos do meio) com o intuito de remover os efeitos causados pelo meio, possibilitando que, posteriormente, se aplique algum outro algoritmo para aplicação específica. Outros métodos criam novos algoritmos de aplicação os quais levam em consideração os efeitos gerados pelo meio. Como exemplo de aplicação, pode-se tomar métodos os quais executam mapeamento [5], voltados para utilização em câmeras de vigilância [6] [7], aplicados a robôs [8] [9] e para auxiliar a tomada de decisão de veículos autônomos [10]. Nota-se que alguns destes métodos podem ou não estar submersos em um meio participativo.

São chamados de meios participativos quaisquer meios que, de alguma forma, interferem na direção ou intensidade dos raios de luz, afetando a imagem capturada. Na lite-



Figura 1: Diferentes tipos de meios participativos. Sendo: (a) chuva, (b) água, (c) fuligem de solda, (d) névoa, (e) neve e (f) poeira.

ratura, alguns meios tratados são água (imagens subaquáticas) [11] [12] [13] [14], névoa [15] [16] [17], chuva (gotas de chuva) [18] [19] [20], poeira [21] [22] [23] [24], neve [25] e fuligem de solda [26]. A Figura 1 ilustra alguns dos meios participativos existentes.

Apesar de uma grande diversidade de possíveis problemas do mundo real, são mais comuns trabalhos os quais focam em imagens subaquáticas e imagens afetadas por névoa. Isso se deve ao fato da relevância dos assuntos, os quais possuem uma alta quantidade de possíveis aplicações. Em um contexto geral, tratar imagens subaquáticas é um desafio maior, dado o alto nível de absorção e grande variância no tamanho das partículas.

Visto que a qualidade da imagem é de suma importância para as aplicações, como nos casos de visão computacional, quando a cena está submersa a meios participativos, o desafio de qualidade de imagem adquirida se torna grande. Em se tratando de imagens subaquáticas, o desafio é ainda maior, dadas as características de absorção de luz na água.

Apesar da absorção da luz e perdas de informação, as imagens subaquáticas são realmente importantes se pensarmos que aproximadamente 70% do planeta Terra é composto por oceanos. Além disso, nos mares e oceanos estão contidos inimagináveis diferentes tipos de ecossistemas, os quais, muitas vezes pela incapacidade humana ou mesmo robótica, não podem ser explorados.

Um dos maiores desafios ao tratar imagens subaquáticas é entender a distância que

os objetos observados na cena estão em relação à câmera. Uma vez que a qualidade da cena embaixo d'água é baixa, medir a distância da câmera até um objeto é extremamente difícil. Para que se calcule esta distância, em geral utilizam-se métodos os quais geram os chamados mapas de profundidade. Estes mapas indicam o quão próximo ou distante cada região da cena está em relação à câmera. Alguns trabalhos propuseram a criação de *datasets* os quais possuem mapa de profundidade de cada imagem. Em geral, estes *datasets* são criados em ambientes controlados (chamados *indoor*) e são utilizados sensores capazes de mensurar a distância de cada objeto presente na cena [27] [28]. Via de regra, os *datasets* chamados RGBD (com inclusão do mapa de profundidade) não são capturados em cenas afetadas por algum meio participativo, dada a incapacidade de precisão para estimar distância quando o ambiente não é controlado.

Em boa parte dos trabalhos, para que se estime o mapa de profundidade de uma cena e possa ser feita a reconstrução da mesma, retirando os efeitos causados pela água, é necessário entender o ambiente em que a cena se encontra. Visto que existe uma infinidade de variáveis ao que tange a ambientes subaquáticos, diferentes tipos de águas podem ser observados ao redor do mundo. Dentre os diferentes tipos de água, características como coloração e turbidez mudam arduamente de um ambiente para outro, o que dificulta algoritmos genéricos de reconstrução de imagem a obterem bons resultados. Compreender o cenário, em especial o tipo de água, sua coloração, profundidade e características se torna crucial para muitos casos, apesar de sua complexidade.

Com base no previamente colocado, é possível afirmar que, para cenas subaquáticas, entender o ambiente em que se encontram, e naturalmente seu tipo de água, é uma tarefa árdua. A partir disso, a proposta deste trabalho passa por desenvolver uma metodologia para a classificação de tipos de água em imagens subaquáticas. O foco deste trabalho se dará em imagens monoculares, as quais demandam menor capacidade de *hardware* e proporcionam maior aplicabilidade.

1.1 Motivação

1.2 Objetivos e Contribuições

O objetivo geral do trabalho proposto é a criação de um algoritmo capaz de classificar imagens monoculares subaquáticas. Objetiva-se entender as características da água e, a partir disso, estimar o tipo de água dada uma classificação.

A seguir apresentam-se os objetivos específicos do trabalho.

1. Avaliar métodos estado da arte. Como base deste trabalho, é fundamental ter conhecimento do que há proposto na literatura. Deseja-se, portanto, conhecer e avaliar os diferentes e mais recentes métodos propostos para classificação de imagens monoculares subaquáticas.

2. Avaliar e criar *dataset*. Dado o contexto subaquático e o foco deste trabalho em trabalhar com imagens reais (não utilizar imagens sintéticas), busca-se soluções na literatura para obter *datasets* com os mais variados contextos de cenas subaquáticas, seja por coloração da água, turbidez ou objetos presentes. Além disso, busca-se *datasets* subaquáticos reais com referência, os quais são largamente difíceis de serem encontrados.
3. Construir dois métodos capazes de classificar o tipo de água contido na imagem de entrada. Como objetivo principal do trabalho, objetiva-se, dada uma imagem subaquática colorida, estimar qual, a partir de uma classificação pré-determinada, é o tipo de água contida naquela cena.

Este trabalho tem como principais contribuições o desenvolvimento de metodologias para classificação de imagens monoculares, destacando o desenvolvimento de dois algoritmos, o CLAFIS e o CLACOR, que serão detalhados no Capítulo 3. Os métodos e resultados obtidos no desenvolvimento deste trabalho foram parcialmente publicados em:

1. Vaz-Jr *et al.* [29] - Nesse trabalho publicado no IEEE LARS/SBR (Qualis-CC B1), com foco em estimação de profundidade, foi apresentado o método CLAFIS como parte da resolução do macro problema, bem como alguns resultados obtidos com ele.
2. Toledo *et al.* [30] - Nesse trabalho publicado no IEEE LARS/SBR (Qualis-CC B1), realizado em conjunto com o aluno de graduação Everson Fagundes de Toledo, foi apresentado uma versão inicial método CLACOR, apresentando resultados preliminares de tal versão, bem como comparativos com outros métodos.

1.3 Estrutura do Trabalho

Este trabalho está separado por capítulo e seções. O Capítulo 2 traz os trabalhos relacionados a este trabalho e apresentando parte da literatura voltada para classificação de tipos de água. O Capítulo 3 consiste na Metodologia do trabalho, onde são apresentados os métodos de classificação propostos. Já o Capítulo 4 mostra os resultados obtidos, trazendo informações de resultados quantitativos, qualitativos e sobre o *dataset* criado e utilizado por esse trabalho. Por fim, o Capítulo 5 é destinado à apresentação da conclusão deste trabalho e os trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Modelo de formação de imagem

Quando a câmera é imersa em um meio participativo, a imagem capturada sofre distorções. Estas distorções são causadas por uma série de fatores que envolvem a cena observada, o comportamento da luz e as partículas do meio. Para entender este comportamento como um todo, os trabalhos de Jaffe [31] e McGlamery [32] são a base para os estudos desde então, em especial para imagens subaquáticas. Seus modelos são baseados na propagação da luz e são compostos por três componentes principais, sendo eles: componente direta (I_d), *forward-scattering* (I_{fs}) e *backscattering* (I_{bs}), como mostra a Eq. (1), onde I é a imagem capturada:

$$I = I_d + I_{fs} + I_{bs}. \quad (1)$$

A componente direta é a luz direta advinda do objeto observado na cena, luz a qual não sofre por nenhum tipo de espalhamento ou anomalia. O efeito de espalhamento é chamado de *forward-scattering*, o qual gera efeitos de borramento na imagem capturada devido às deflexões da luz em diferentes ângulos. O *forward-scattering* costuma ser negligenciado, uma vez que o *backscattering* é o principal motivo da degradação de contraste na imagem [33]. A Eq. (2) apresenta a transmissão do meio (t), ou seja, quanto o meio interfere na imagem. A Eq. (3) apresenta a formularização da componente direta. A variável β é o coeficiente de atenuação, d é a distância entre a câmera e a cena (profundidade) e J é a radiância da cena, conforme seguem as equações:

$$t = e^{-\beta d}, \quad (2)$$

$$I_d = Jt. \quad (3)$$

O *backscattering* deve-se às partículas presentes na água. Estas partículas possuem a capacidade de modificar a rota usual da luz e, por conseguinte, degradar a imagem capturada. Quanto mais próxima a câmera de um objeto (ou cena), menor a incidência

de *backscattering*. Porém, distâncias maiores entre cena e câmera, considerada a profundidade da cena acarretam em grande incidência deste componente e portanto pouca visibilidade da cena real. A Eq. (4) apresenta o comportamento do *backscattering*:

$$I_{bs} = A(1 - e^{-\beta d}) = A(1 - t), \quad (4)$$

onde A é a luz do ambiente no infinito. Este termo é dependente do comprimento de onda (da cor da luz) e da profundidade da água, sendo considerada a distância entre a superfície da água e a câmera. A partir da Eq. (4) pode-se notar com clareza a afirmação supracitada: o *backscattering* cresce de acordo com a distância entre a cena e a câmera, ou seja, quanto maior esta distância, menor a visibilidade da cena na imagem e maior o *backscattering*.

Apesar de sua existência no ambiente subaquático, o *forward-scattering* costuma ser negligenciado dada sua irrelevância se comparado ao *backscattering*. Uma simplificação para a Eq. (1) é apresentada na Eq. (5), a qual é considerada o modelo de formação de imagem subaquática:

$$I(x) = I_d + I_{bs} = J(x)t(x) + A(1 - t(x)), \quad (5)$$

onde x representa as coordenadas de cada pixel e $I(x)$ é a imagem obtida. Nota-se que A independe de x , uma vez que é considerado que a iluminação é homogênea nesta premissa adotada. Além disso, é possível associar que quando a transmissão tende a 1, a imagem não foi capturada em um meio participativo ou o meio insignificamente afeta a imagem capturada.

2.2 Definição do problema

Dado o contexto de imagens subaquáticas e a necessidade para certas aplicações da remoção dos efeitos causados por esse meio, o problema baseia-se na **classificação de tipos de água em imagens monoculares**. Entender o tipo de água em que se encontra a cena é base para muitos métodos e trabalhos e, se bem estimado, pode auxiliar na tomada de decisão de um método ou algoritmo, bem como colaborar na restauração desta imagem.

3 TRABALHOS RELACIONADOS

Este capítulo apresenta uma revisão de literatura sobre classificação de tipos de água em imagens submersas. Visto que essa área é limitada na literatura, estudou-se diferentes aplicações, mesmo que em trabalhos que não possuem como objetivo final a classificação, mas que em algum passo do seu processamento classificam o tipo de água. O trabalho proposto foca em entender o contexto subaquático onde a imagem está inserida, o que é uma tarefa árdua, em especial para imagens monoculares. Apesar de ser uma busca complexa, alguns trabalhos se propõem a entender o ambiente da imagem e nestes trabalhos que essa proposta se baseia.

Como base no estudo de tipos de água, Jerlov [1] é uma grande referência para o campo, uma vez que é tido com um dos precursores no ramo de classificação de tipos de água e especialmente pelo seu trabalho publicado em 1976 e que ainda é utilizado no estado da arte. À época, Jerlov viajou o mundo medindo coeficientes de atenuação dos mares através de um turbidímetro, definindo assim os tipos de água com base em suas características ópticas. A Figura 2, retirada de Johnson *et al.* [2], apresenta os tipos de água em mar aberto (tipos I a III) dos oceanos de acordo com a classificação de Jerlov.

Berman *et al.* [4] apresentam uma abordagem para a restauração de imagens subaquáticas, considerando o tipo de água conforme definido por Jerlov [1]. O método estima dois parâmetros globais essenciais: a razão de atenuação azul-vermelho e azul-verde. O estudo inclui uma tabela comparativa que representa um objeto completamente branco observado em diferentes tipos de água propostos por Jerlov [1], em profundidades variando de 1 a 20 metros. Além disso, são fornecidos coeficientes de atenuação aproximados para os diversos tipos de água.

Na implementação de Berman *et al.* [4], o processo de restauração é realizado iterativamente, com a aplicação do método para cada um dos tipos de água propostos por Jerlov [1]. Berman *et al.* restauram a mesma imagem subaquática dez vezes, variando os coeficientes conforme a classificação de Jerlov [1]. O resultado final é selecionado com base na premissa de Grey-World, conforme definido por Buchsbaum [34], escolhendo a saída mais otimizada dentre as dez possíveis.

O trabalho proposto por Akkaynak *et al.* [35] estabelece os comprimentos de onda

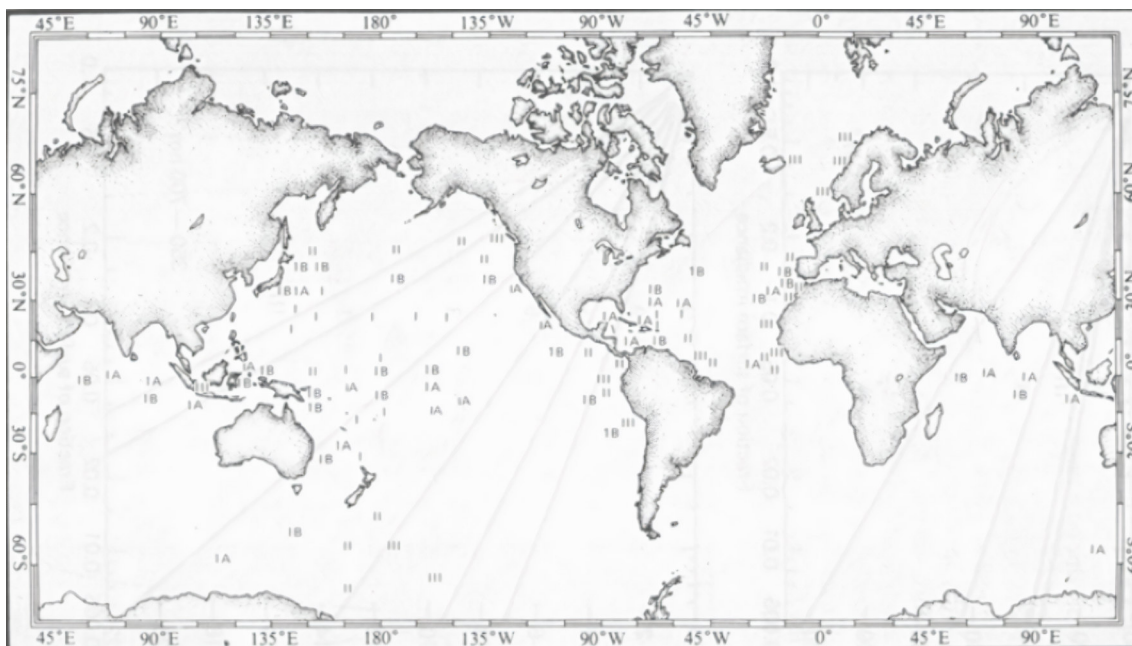


Figura 2: Mapa mundi segundo classificações de tipo de água propostas por Jerlov [1]. Fonte: Johnson *et al.* [2].

cruciais a serem considerados na avaliação das propriedades ópticas da água, fornecendo os coeficientes de atenuação associados a cada tipo de água identificado por Jerlov. A validação desse enfoque é realizada por meio de experimentos conduzidos em duas massas de água distintas, o Mar Vermelho e o Mar Mediterrâneo.

Posteriormente, Akkaynak *et al.* [36] apresentam uma abordagem para a remoção consistente da água em imagens subaquáticas. O principal objetivo desta pesquisa é possibilitar a remoção da água dessas imagens, permitindo que sejam processadas por algoritmos desenvolvidos para o tratamento de imagens em ambientes terrestres.

Uplavikar *et al.* [3] propõem a criação de um classificador robusto para os diversos tipos de água, visando lidar com a variabilidade dos ambientes marinhos. A estratégia adotada envolve a aplicação de *Domain-Adversarial Learning* (aprendizado de domínio adversário) e a utilização de um *dataset* sintético para compensar a escassez de conjuntos de dados do mundo real com *ground truth* (imagens de referência). O modelo incorpora duas redes neurais convolucionais, uma dedicada à geração de imagens subaquáticas claras e outra à classificação dessas imagens de acordo com os tipos de água propostos por Jerlov [1].

Uma contribuição significativa do trabalho de Uplavikar *et al.* [3] é a reorganização dos 10 tipos de água sugeridos por Jerlov [1] em 6 grupos distintos. Essa reclassificação agrupa os tipos I, IA e IB em um primeiro conjunto, os tipos II e III em um segundo, e os tipos 1 e 3 em um terceiro. Os tipos 5, 7 e 9 são mantidos em seu formato original. Esse processo é implementado para amplificar a discrepância entre os tipos de água, aprimorando assim o desempenho do classificador. O esquema de agrupamento proposto por

esse estudo é visualmente representado na Figura 3.

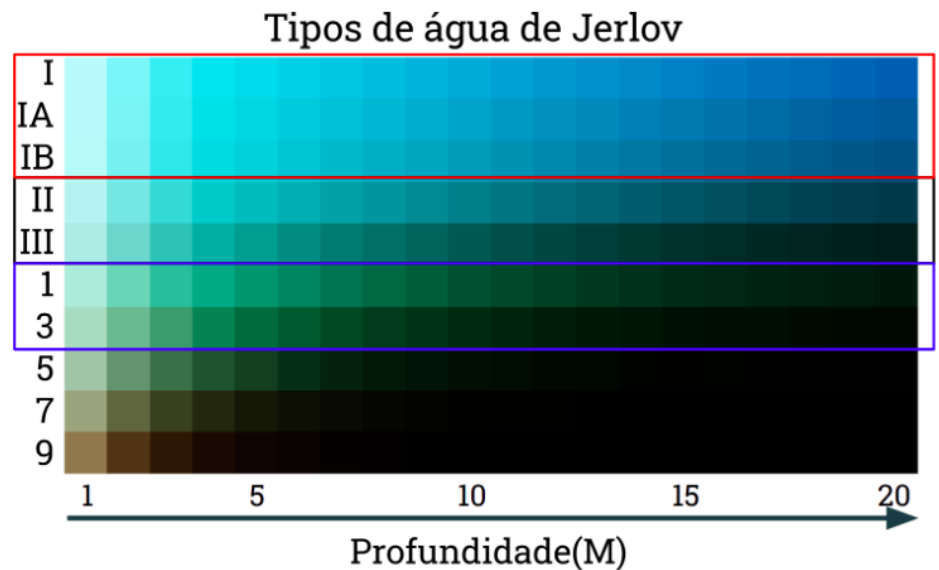


Figura 3: Os tipos de água propostos por Jerlov [1] agrupados em 6 classificações propostas por Uplavikar *et al.* [3].

Islam *et al.* [37] apresentam o primeiro conjunto de dados em grande escala voltado para a segmentação semântica de imagens subaquáticas, denominado SUIM (Segmentation of Underwater Imagery). Esse conjunto de dados compreende mais de 1500 imagens, cada uma delas com anotações de *pixels* para oito categorias distintas de objetos: peixes, recifes, plantas aquáticas, destroços, ruínas, mergulhadores humanos, robôs e o fundo do mar.

Carlevaris-Bianco *et al.* [13] propõem uma abordagem para mitigar os efeitos do espalhamento da luz em imagens subaquáticas. Carlevaris-Bianco *et al.* aproveitam a marcante discrepância na atenuação entre os três canais de cor da imagem na água para realizar uma estimativa da profundidade na cena. Kalaiarasi *et al.* [38] propõem um método sistemático o qual analisa imagens subaquáticas utilizando aprendizado profundo. Com foco em redes neurais, o trabalho visa detectar objetos em imagens submersas.

Aas *et al.* [39] desenvolveram um mapa dos tipos de água de Jerlov nas águas dos mares nórdicos e regiões adjacentes. O estudo abrangeu diversas estações, situadas entre as latitudes 54° e 82° Norte e longitudes 31° e 49° Oeste, revelando que o Atlântico e as águas polares geralmente exibem características oceânicas dos tipos II e III. No entanto, as condições ópticas podem se modificar durante a proliferação de algas, transitando para os tipos costeiros 1, 3 e 5.

Bekerman *et al.* [40] têm como objetivo efetuar uma estimativa das propriedades ópticas da água, abrangendo os coeficientes de dispersão e atenuação da luz, fundamentais para modelos de restauração de imagens subaquáticas com base em princípios físicos. Em contrapartida às abordagens convencionais, se estima os coeficientes de atenuação por

meio da distribuição de cores na imagem e procura calcular a cor da *veiling light* (luz ambiente) presente na cena. Uma característica distintiva do método de Bekerman *et al.* é a não adoção dos coeficientes propostos por Jerlov [1], bem como a independência da identificação de *pixels* de fundo na imagem para o seu funcionamento.

Desai *et al.* [41] propõem a geração de imagens subaquáticas sintéticas considerando o modelo de formação de imagem. Como objetivo principal, através de redes neurais (CGAN), a ideia baseia-se na restauração de imagens subaquáticas degradadas e a geração de conjuntos de imagens sintéticas baseando-se nos coeficientes de Jerlov [1]. Utilizando um método baseado em modelo e com base na classificação de Jerlov, Raihan *et al.* [42] criou um *dataset* com referência e simulando diferentes tipos de água em diferentes profundidades.

Liu *et al.* [43] apresentam um modelo simplificado para resolver problemas de degradação em imagens subaquáticas. O método utiliza os coeficientes de atenuação dos tipos de água de Jerlov para simular imagens degradadas de 10 tipos diferentes de água. O algoritmo ainda propõe uma rede neural codificador-decodificador para aprimorar suas imagens subaquáticas a partir das imagens simuladas.

Williamson *et al.* [44] visam estudar os perfis de profundidade para os tipos de água de Jerlov. Utilizando uma fonte geograficamente diversificada de dados, os autores analisam quantitativamente e extraem o perfil de profundidade de cada imagem. Começando de cima para baixo na tabela de Jerlov, os autores descrevem o impacto da claridade em cada tipo de água e suas profundidades.

3.1 Conclusões

Embora um trabalho recente apresente o extensivo desenvolvimento de algoritmos de classificação de imagens subaquáticas [45], essa classificação é focada em objetos e estruturas submersas. A classificação do tipo de água, é bastante relevante em aplicações práticas de algoritmos de restauração e mesmo classificação de objetos, permitindo um conhecimento das limitações e até facilitando a estimativa de incerteza.

Assim, a literatura ainda tem bastante espaço para novos algoritmos focados em entender a água e seus diferentes tipos, como os métodos propostos neste trabalho.

4 METODOLOGIA

Este capítulo apresenta os métodos propostos neste trabalho, bem como os conceitos implementados. O capítulo está dividido em seções, as quais apresentam e detalham os dois classificadores de tipos de água propostos neste trabalho.

4.1 Classificação Baseada em Modelo Físico - CLAFIS

Em 1976, Jerlov [1] criou, através de experimentos em águas do mundo todo, uma classificação de 10 tipos de água. Os tipos de água foram classificados através da captura de imagens subaquáticas em distintos mares, focando em um objeto completamente branco e variando a distância do objeto com a câmera de 1 a 20 metros, como apresentado na Figura 4. Além disso, cada tipo de água teve seu coeficiente de atenuação médio estimado, sendo apresentado em função do comprimento de onda (λ) na Figura 5. A classificação dos tipos de água de Jerlov é subdividida em águas oceânicas (I, IA, IB, II, III) e águas costeiras (1, 3, 5, 7, 9). Para entender a classificação de Jerlov, pode-se utilizar a Eq. (6), adaptada de Roser *et al.* [46]:

$$A_{\infty}(wd) = TI_0e^{-\beta wd}, \quad (6)$$

onde T é o coeficiente de transmissão na superfície da água e I_0 é a intensidade atmosférica na superfície. Considerando, portanto, que $T \approx 1$, a luz é branca ($I_0 = 1$) e nos experimentos de Jerlov foi utilizado uma superfície igualmente branca na aquisição das imagens, a Eq. (6) simplifica-se à Eq. (7):

$$A_{\infty}(wd) = e^{-\beta wd}, \quad (7)$$

onde β é estimada através de um turbidímetro, medindo os valores diretamente na água, e wd é a distância entre a câmera e a superfície da água. A_{∞} é a mesma luz ambiente representada no Capítulo 1 como A .

× ×

× ×

Apesar de utilizada em métodos de restauração de imagem, a classificação dos tipos

de água de Jerlov, pode ser um tanto quanto complexa. Isso deve-se ao fato de que alguns tipos de água assemelham-se, como é o caso dos tipos I, IA e IB. Além disso, boa parte dos *datasets* presentes na literatura, além de eventualmente serem erroneamente classificados pela possível semelhança entre as águas, podem conter imagens capturadas em distintos locais, fazendo com que este *dataset* torne-se híbrido, em termos de classificação de Jerlov, ou seja, seja composto de imagens capturadas em diferentes tipos de água. Os casos dos *datasets* [37], [47], [48] são exemplos de conjuntos de imagens as quais variam sua classificação de tipo de água, segundo Jerlov.

A classificação do tipo de água, através do CLAFIS, é feita baseando-se na metodologia aplicada em [49], onde inicialmente estima-se o chamado *veiling-light* ou luz ambiente da imagem (apresentado na Eq. (7)). Dada uma imagem de entrada, aplica-se uma equalização de histograma local e posteriormente a geração de um mapa de bordas, onde busca-se regiões suavizadas e sem textura. A maior região da imagem que se adequa no quesito procurado é considerada a luz ambiente.

Como explicado anteriormente, Berman *et al.* [49] baseiam-se nos experimentos propostos por Jerlov [1] para classificar o tipo de água de uma dada imagem de entrada. Porém, o método de Berman *et al.*, além de tomar como base a luz ambiente da imagem, necessita estimar a transmissão, reconstruir a cena para os 10 tipos de água e a partir do melhor resultado, segundo a premissa de *Grey World*, definir o tipo de água. Este método propõe uma otimização para a classificação do tipo de água.

Após estimar a luz ambiente, A_∞ , minimiza-se a Eq. (7). Otimiza-se a profundidade da água wd para cada tipo de água, segundo a classificação de Jerlov [1]. A otimização é baseada em limites, portanto, utilizou-se o método proposto por Coleman [50]. Estendeu-se a profundidade da água para um máximo de $100m$, dada a capacidade limitada de iluminação natural em maiores profundidades para alguns tipos de água, ou seja, onde o modelo é válido. A Eq. (8) apresenta a otimização proposta. A otimização proposta permite a classificação do tipo de água de forma muito mais rápida que em [49], além de também estimar a profundidade que a imagem de entrada está empregada, conforme segue:

$$\min_{\beta, wd} \sum_{c \in R, G, B} A_c - e^{-\beta_c wd}. \quad (8)$$

4.2 Classificação Baseada em Modelo de Cor - CLACOR

Também com base na proposta de Jerlov [1] para classificação dos tipos de água, o presente método utilizou uma tabela contendo um objeto completamente branco, visualizado em 10 tipos distintos de água de Jerlov, em profundidades variando de 1 a 50 metros e os valores dos canais de cor dos tipos de água presentes na tabela foram utilizados como base para a classificação.

Para um estudo mais aprofundado, as imagens utilizadas nesse trabalho foram majoritariamente coletadas de três conjuntos de dados: Li *et al.* [48], Islam *et al.* [37] e Drews *et al.* [51]. A ideia desse *dataset* misto é suprir a escassez de imagens dos tipos de água localizados nas camadas mais profundas da tabela.

Para o estudo, diferentes imagens subaquáticas foram selecionadas aleatoriamente dos conjuntos de dados mencionados e analisadas em relação aos seus espaços de cores RGB e HSV. Com base nas análises realizadas, alguns padrões foram identificados, os quais serviram de base para o desenvolvimento de um algoritmo com o objetivo de classificar uma imagem de entrada em um dos tipos de água e selecionar alguns pontos para determinar a profundidade em que a cena foi capturada, além do tipo de água.

Inicialmente, a entrada do algoritmo CLACOR consiste na tabela de classificação contendo os tipos de água e a profundidade, bem como a imagem a ser classificada (imagem de entrada). Em seguida, são calculadas as médias dos canais de cor da imagem de entrada separadamente nos espaços HSV e RGB. A informação referente ao espaço HSV é utilizada para auxiliar na seleção da região da imagem que contém apenas água.

Para a seleção do canal dominante, é escolhida a componente do espaço RGB com o maior valor de média. Essa componente é então denominada de "canal dominante". O processo pode ser visualizado na Figura 6.

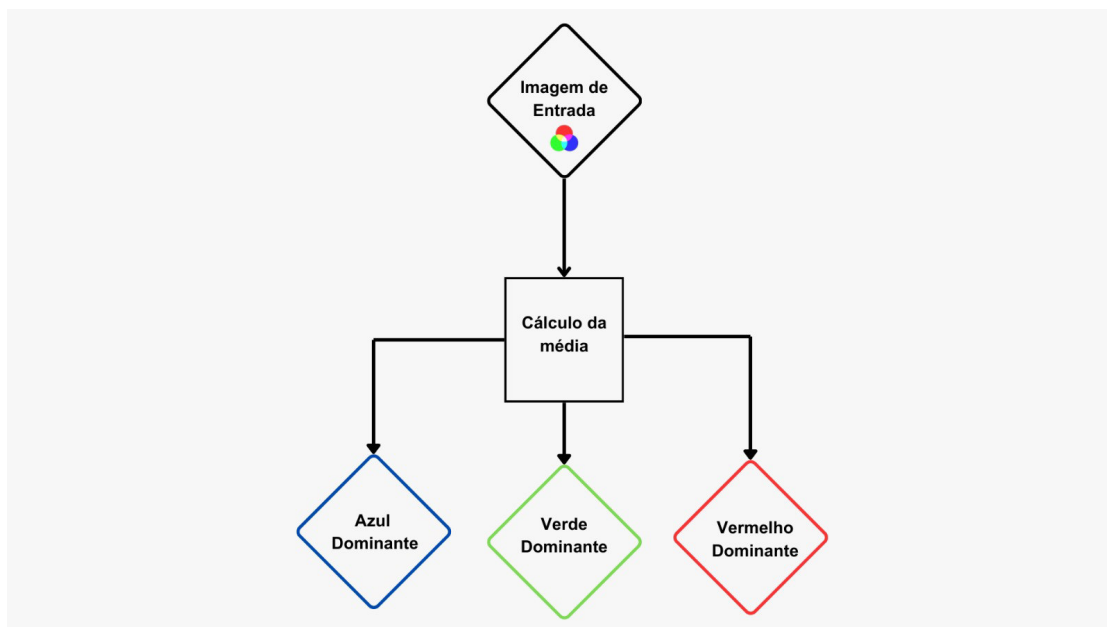


Figura 6: Diagrama elucidativo do primeiro passo para escolha do *cluster* com apenas água.

Esse é o primeiro passo para a classificação da imagem, uma vez que foram estudadas as propriedades dos tipos de água na tabela em relação aos seus canais RGB. A partir desses estudos, observou-se que os tipos de água localizados entre I e II possuem o canal azul como dominante. O tipo III apresenta os canais azul e verde em quantidades exatamente

iguais. Os tipos 1, 3 e 5 possuem o canal verde como dominante, enquanto os tipos 7 e 9 possuem o canal vermelho como dominante.

Com base nessas observações, a imagem é tratada de acordo com seu canal de cor dominante e é classificada em um dos tipos de água que seguem a lógica da média de seus *pixels*, com algumas exceções.

O algoritmo *KMeans* [52] é uma técnica de clusterização amplamente utilizada em diversas áreas da ciência e da engenharia. Ele é usado para dividir um conjunto de dados em K grupos, onde K é um parâmetro definido pelo usuário. O algoritmo funciona iterativamente para atribuir cada ponto de dados a um dos K grupos com base nas médias dos pontos em cada grupo. O objetivo do algoritmo é minimizar a soma dos quadrados das distâncias entre cada ponto de dados e seu grupo atribuído.

No contexto deste trabalho, o algoritmo *KMeans* foi aplicado para agrupar os *pixels* da imagem de entrada em um número pré-definido de grupos. Isso permite a identificação de regiões na imagem com características semelhantes, como a presença de água ou objetos submersos. Além disso, as coordenadas no plano X, Y da imagem são salvas para uso posterior, enquanto os agrupamentos são baseados apenas nos canais de cor RGB, permitindo uma melhor segmentação da imagem.

O método do cotovelo é usado para determinar o número ideal de *clusters*, ou grupos, para uma imagem específica. Esse método envolve a execução do algoritmo *KMeans* com diferentes quantidades de *clusters*, buscando um equilíbrio entre a homogeneidade de cada *cluster* e uma boa diferenciação entre eles. O algoritmo *KMeans* calcula a distância de cada observação até o centro do *cluster* ao qual ela pertence e o objetivo é minimizar essa distância. Matematicamente, busca-se minimizar a soma dos quadrados intra-*clusters*, que é um parâmetro chamado de *inertia* e é calculado pelo algoritmo *KMeans*, disponibilizado pela biblioteca *SKLearn.Cluster* na linguagem *Python*. A quantidade ideal de *clusters* é determinada pelo ponto de inflexão na curva formada pela relação entre o número de *clusters* e a *inertia*, que é o ponto mais distante da reta que une o menor e o maior número de *clusters* usados no método do cotovelo.

Para a aplicação do algoritmo *KMeans*, é necessário definir um parâmetro que determine os valores iniciais para os centros dos *clusters*. É importante salientar que o *KMeans* é sensível à inicialização, ou seja, definir centros iniciais inadequados pode resultar em resultados insatisfatórios. Devido a esse problema, Peña *et al.* [53] realizaram um estudo empírico comparando quatro métodos de inicialização para o algoritmo. A implementação do algoritmo mencionada anteriormente oferece um parâmetro para otimizar a escolha dos centros iniciais.

Após a inicialização com um vetor contendo todos os *pixels* da imagem de entrada com seus respectivos valores nos canais de cor no espaço RGB, e a definição do número de *clusters* ideal para aquela imagem específica pelo método do cotovelo, bem como a otimização da inicialização dos centros dos *clusters* através do parâmetro *init++* dis-

ponível na versão utilizada, o algoritmo *KMeans* agrupa cada ponto do vetor de entrada no *cluster* cujo centro mais se assemelha ao referido ponto.

A saída consiste em um vetor que contém os centros utilizados para o agrupamento, sendo que cada centro apresenta a média dos canais R, G e B de seus respectivos pontos. Considera-se que um dos *clusters* representados pelos referidos centros contém apenas água e o objetivo primário do algoritmo proposto é identificar tal *cluster*. A abordagem descrita tem como finalidade selecionar a região que contém água independentemente de sua posição ou orientação na imagem. Como exemplo, o método de Berman *et al.* [4], selecionou a parte superior da imagem como região de interesse e, posteriormente, buscou por similaridade de *pixels* nesta região para descobrir o tipo de água. No entanto, esse método não se mostrou eficiente em imagens que contêm objetos ou uma fonte de alta luminosidade na superfície. A Figura 7 ilustra dois exemplos de escolha da região com água, um de acordo com o método de Berman *et al.* [4] e outro de acordo com o método proposto.

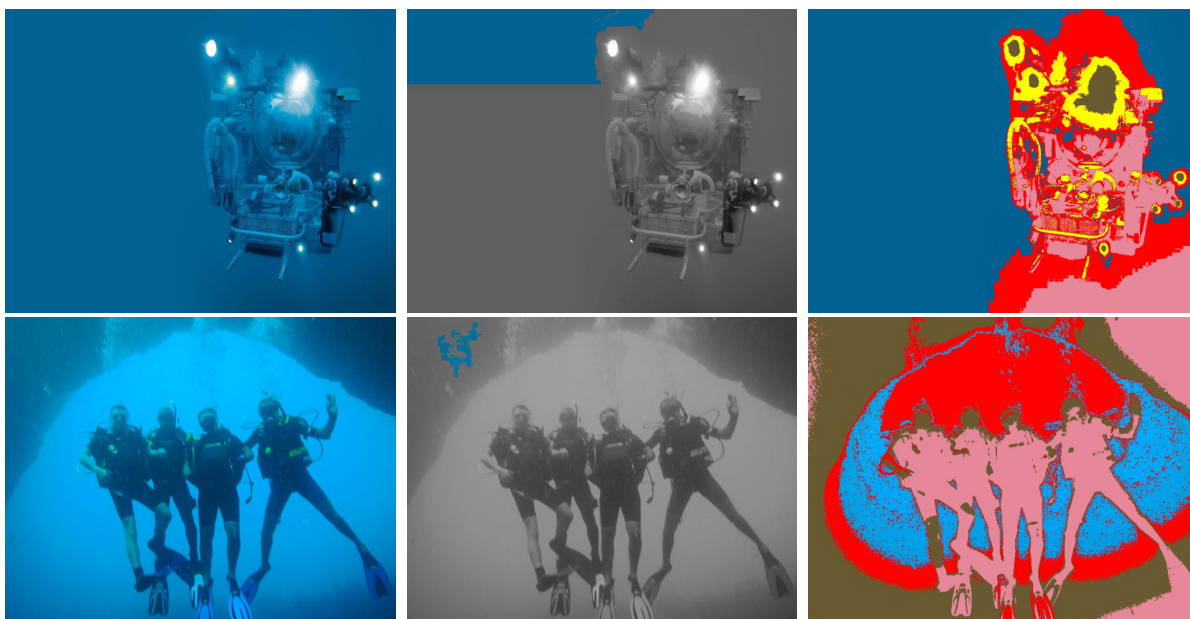


Figura 7: *Cluster* de área de apenas água. Na ordem da esquerda para direita: imagem fonte, região com apenas água de acordo com Berman *et al.* [4] e região selecionada pelo método *KMeans* como apenas água do CLACOR. A região água aparece na mesma cor da imagem fonte.

O algoritmo *KMeans* também retorna o vetor de entrada preenchido com números inteiros de 1 a N , onde N é o número de *clusters* gerados pelo algoritmo. Cada *pixel* recebe um número que indica a qual grupo ele pertence. Essa informação é usada para criar a imagem de saída com os *clusters*. O grupo que contém apenas a camada de água permanece na cor original, enquanto os demais grupos recebem cores predefinidas de acordo com o número atribuído a eles.

Foi implementado um processo de otimização da clusterização gerada pelo *KMeans*.

Esse processo envolve o cálculo do desvio padrão de cada canal de cor, no espaço RGB, do *cluster* identificado pelo *KMeans* como sendo a camada de água. Em seguida, os *pixels* desse *cluster* que estão a mais de um desvio padrão de distância do centro do *cluster* em um dos canais de cor são excluídos.

A Figura 8 apresenta algumas imagens de entrada utilizadas nos experimentos, bem como suas saídas clusterizadas pelo algoritmo *KMeans* e a saída final utilizada para a classificação após o processo de otimização da clusterização descrito.

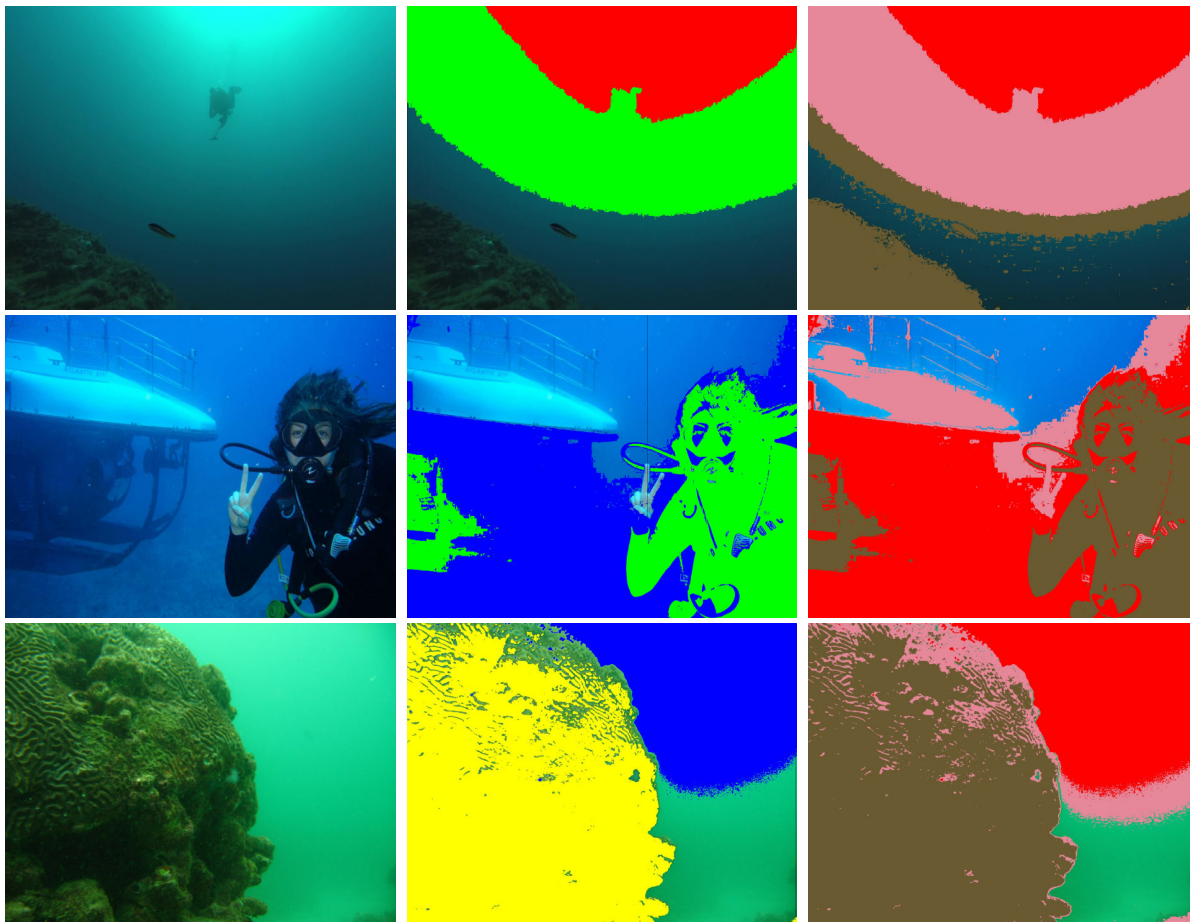


Figura 8: Da esquerda para direita: imagem fonte, saída do algoritmo *KMeans* e imagem final pós-otimização.

4.2.1 Critérios de escolha para a região com água

Conforme mencionado previamente, a abordagem de processamento de imagem varia de acordo com o canal dominante da mesma. Nas próximas seções, serão detalhados os procedimentos adotados para cada tipo de imagem.

4.2.1.1 Canal Azul Dominante

A grande maioria dos conjuntos de dados disponíveis é composta por imagens em que o canal azul é dominante, exigindo, assim, uma abordagem mais complexa no método

proposto. A Figura 9 exemplifica, através de um diagrama, o fluxo para seleção do *cluster* quando o canal azul é o canal dominante na imagem.

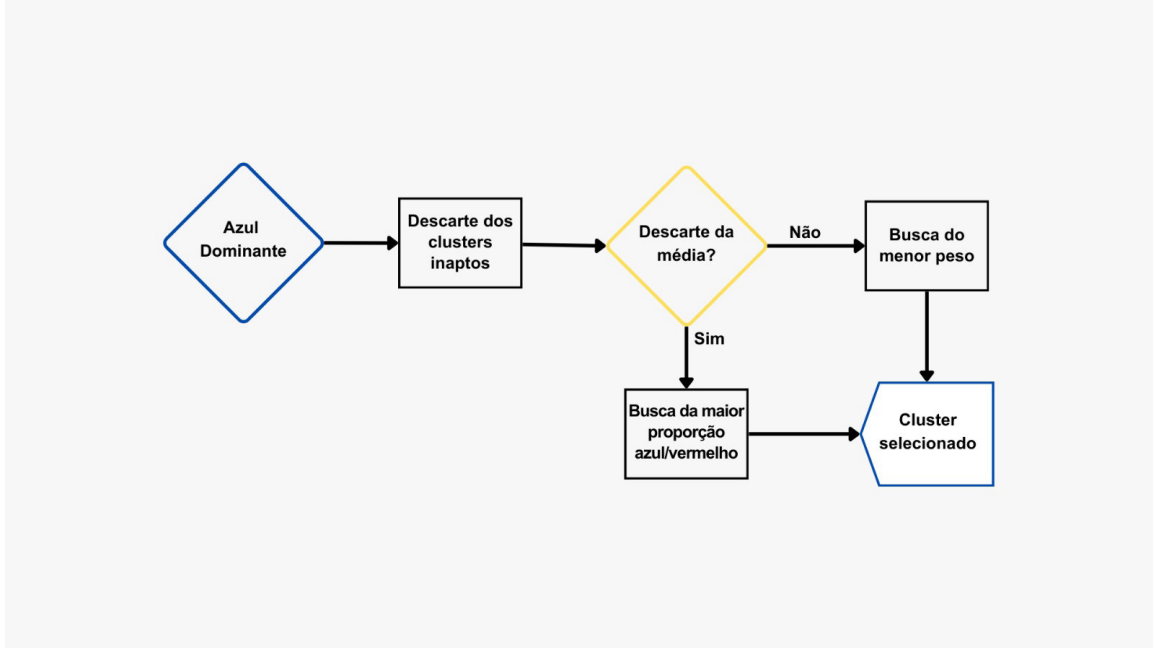


Figura 9: Diagrama demonstrativo de fluxo para seleção do *cluster* de área com apenas água em imagens com canal azul dominante.

Para as imagens em que o canal azul é dominante, uma variável é utilizada para determinar o peso atribuído a cada centro retornado pelo algoritmo de *KMeans*, a fim de escolher o centro que representa o *cluster* de apenas água. Busca-se o centro com o menor peso e inicialmente, este peso é composto por:

$$Peso_{KM} = abs(mA_c - mA_I) - 5 \times (mA_c/mV_c), \quad (9)$$

considerando que mA_c representa a média dos *pixels* do canal azul do centro retornado, mA_I representa a média dos *pixels* do canal azul da imagem, e mV_c representa a média dos *pixels* do canal vermelho do centro retornado, a fórmula proposta busca selecionar o centro com a média do canal azul mais próxima da média do canal azul da imagem e, ao mesmo tempo, com maior proporção do canal azul sobre o canal vermelho do próprio centro. Essa abordagem foi definida com base nos resultados do estudo de uma tabela com classificações e dos padrões observados em um grande número de imagens subaquáticas com canal azul dominante.

Para que um centro seja considerado apto para seleção, ele deve possuir as seguintes características:

1. $mA_c > mA_I/2$, onde a média do canal azul do centro deve ser maior que a metade da média do canal azul da imagem

2. $mA_c < mA_I \times 2$, onde a média do canal azul do centro deve ser menor que o dobro da média do canal azul da imagem
3. $mA_c > mVd_c$, onde a média do canal azul do centro deve ser maior que a média do canal verde do centro

Considera-se que mVd_c seja a média dos valores de *pixel* no canal verde do centro retornado e mantém-se as demais variáveis mencionadas anteriormente. A finalidade do critério apresentado é identificar e descartar pontos que se afastam do padrão observado, conhecidos como *outliers*. São eliminados nesta etapa os pontos com alta intensidade no canal azul devido a possíveis efeitos de iluminação na cena, assim como os pontos com intensidade muito baixa decorrentes de objetos. O espaço de cor HSV é utilizado para auxiliar na identificação da região com água, permitindo descartar o uso das médias da imagem em situações específicas:

1. $Sm_I > Vm_I \times 2.5$, onde a média no canal S da imagem deve ser maior que a média no canal V da imagem multiplicada por 2.5,
2. $Sm_I > Hm_I$, onde a média no canal S da imagem deve ser maior que a média no canal H da imagem.

Considerando Sm_I como a média do canal S da imagem, Vm_I como a média do canal V e Hm_I como a média do canal H.

Neste espaço de cor HSV, o canal H (*hue*) representa a matiz, ou seja, a tonalidade da cor, abrangendo todo o espectro de cores, do vermelho ao violeta com adição do magenta. O canal S (*saturation*) representa a saturação, ou seja, a quantidade da cor assumida pelo canal H. Quanto menor esse valor, mais próxima da cor cinza será a imagem, e quanto maior, mais próxima da cor do canal H. Por fim, o canal V (*value*) define o brilho da imagem, sendo que um valor mais alto indica uma imagem mais clara.

Na conversão matemática do espaço RGB para o espaço HSV, o canal V do HSV é obtido como o canal de maior intensidade da imagem no espaço RGB. Em outras palavras, o canal V do espaço HSV contém o canal dominante que estamos buscando. Os canais H e S podem ser calculados usando as seguintes equações, considerando valores do RGB de 0 a 1:

$$H = 60 \times ((G - B)/(max - min)), \text{ se } max = R \wedge G \geq B, \quad (10)$$

$$H = 60 \times ((R - G)/(max - min)) + 360, \text{ se } max = R \wedge G < B, \quad (11)$$

$$H = 60 \times ((B - R)/(max - min)) + 120, \text{ se } max = G, \quad (12)$$

$$H = 60 \times ((R - G)/(max - min)) + 240, \text{ se } max = B, \quad (13)$$

$$S = (max - min)/max. \quad (14)$$

Sendo H , S e V os canais de cor no espaço HSV, e R , G e B os canais de cor no espaço RGB, as variáveis max e min representam os canais de maior e menor intensidade no espaço RGB, respectivamente.

Conclui-se, portanto, que quando o canal S do espaço HSV em imagens de canal dominante azul é amplamente superior ao canal V (o próprio canal azul no espaço RGB), isso indica uma grande concentração de cor que não é do canal dominante na cena. Isso pode ser causado pela presença de um objeto de cor diferente do azul, ocupando uma região considerável da imagem e afetando significativamente a média dos *pixels*. Portanto, a média dos canais de cor no espaço RGB da imagem é desconsiderada como critério de escolha nesse caso.

Outra situação em que ocorre o descarte de uso de médias é quando um valor alto para o canal vermelho da imagem é observado, sendo esse valor superior a 175 na escala de 0 a 255 do canal. Isso pode indicar que a imagem foi pré-processada, com alguma restauração no canal vermelho, uma vez que esse canal é atenuado rapidamente na água e não é comum ter um valor elevado. Além disso, pode ser uma imagem com muita luminosidade, com uma grande região branca devido à luz da superfície que elevou o valor do canal vermelho. Em último caso, pode ser uma cena capturada próxima à superfície, onde o canal vermelho foi pouco atenuado. Nesses casos, a média dos canais de cor no espaço RGB da imagem também é desconsiderada como critério de escolha.

A Figura 10 representa duas imagens em que a média global de *pixels* foi descartada na hora da escolha da região com água. A primeira imagem foi descartada com base no critério adotado para o espaço de cor HSV, enquanto a segunda imagem foi descartada com base no critério do canal vermelho.

A escolha dos centros aptos para identificação de regiões com água, quando a média dos canais de cor no espaço RGB é descartada, é baseada exclusivamente na proporção entre os canais azul e vermelho. Busca-se o maior valor nessa razão, calculando a média dos *pixels* do canal azul e do canal vermelho dos centros retornados e comparando-os. Essa proporção, representada como mA_c/mV_c , é utilizada como critério de seleção dos centros aptos para identificar regiões com água na imagem. Considera-se mA_c a média dos *pixels* do canal azul do centro e mV_c a média dos *pixels* do canal vermelho do centro.

4.2.1.2 Canal verde dominante

Quando se trata de imagens com canal dominante verde, o critério de escolha é baseado no ponto com maior intensidade de verde diretamente, com a exclusão de *outliers*. Inicialmente, o valor do canal verde do primeiro centro retornado é armazenado em uma variável. Em seguida, os demais centros são percorridos, comparando o valor do seu canal verde com o valor do maior verde atualmente armazenado. Ao final do processo, o

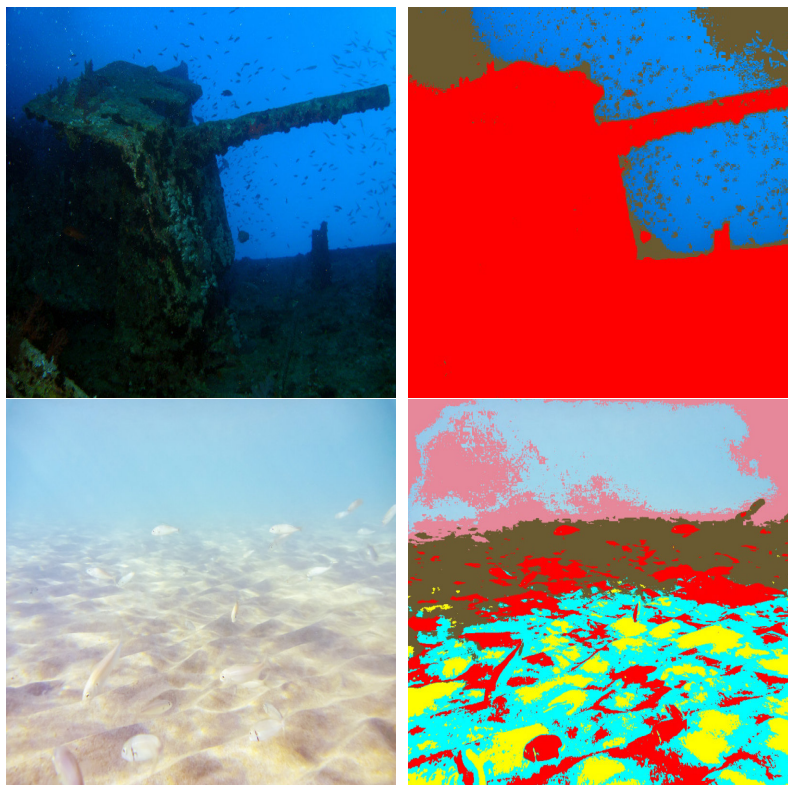


Figura 10: Imagens de exemplo que possuem média global de cores descartada como critério de escolha. A imagem da linha superior tem sua média global de cores descartada em função do espaço HSV, enquanto a imagem inferior em função do canal vermelho.

valor armazenado corresponderá ao do *cluster* com a maior intensidade de verde no canal verde, sendo escolhido como o centro apto para identificar regiões com água na imagem.

Imagens em que o canal verde é dominante geralmente apresentam uma intensidade menor nesse canal em comparação com o canal azul em imagens com dominância de azul. Essa informação é obtida a partir de uma tabela de referência e do estudo de imagens subaquáticas com essa tonalidade. No processo de escolha do centro apto para identificar regiões com água, caso o valor do canal verde do centro selecionado seja igual ou superior a 150 na escala de 0 a 255, esse centro é descartado como um *outlier*. Esse processo é repetido considerando apenas os centros remanescentes, e ao final, o centro não descartado com o maior valor no canal verde do espaço RGB é escolhido como o que contém água. A Figura 11 ilustra esse processo de forma simplificada.

4.2.1.3 Canal vermelho dominante

As imagens com dominância de cor vermelha são menos frequentes e geralmente são registradas em profundidades menores em comparação com imagens que apresentam dominância de cor verde ou azul. Isso ocorre devido à rápida perda de visibilidade em águas com tonalidade vermelha, como é o caso dos tipos 7 e 9 representados na Figura 4. Ao aplicar o algoritmo *KMeans* a essas imagens, é comum obter três grupos distintos,

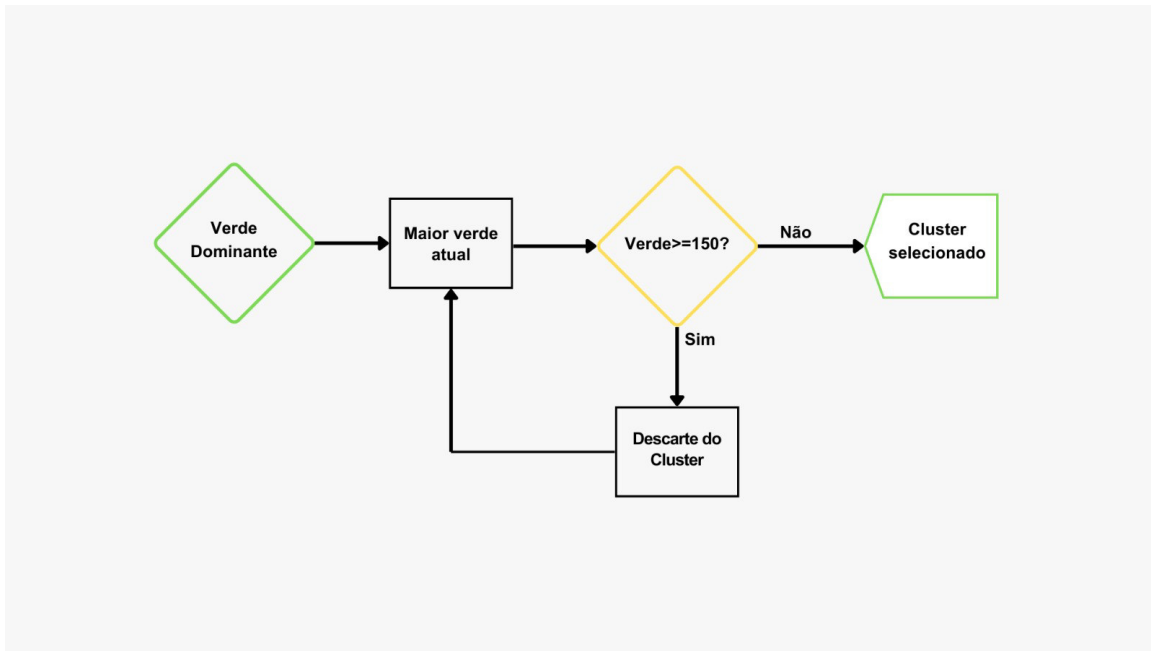


Figura 11: Diagrama demonstrativo de fluxo para seleção do *cluster* de área com apenas água em imagens com canal verde dominante.

correspondendo a iluminação, objetos e água.

Com base nesse padrão observado, a abordagem proposta consiste em excluir os centros com as intensidades mais altas e mais baixas no canal vermelho dessas imagens. O centro com intensidade intermediária no canal vermelho é selecionado como a região de água. No caso de divisão em quatro ou mais grupos, o centro com menor intensidade no canal vermelho é escolhido como a região de água após a exclusão dos centros mais e menos intensos. O processo é ilustrado de forma simplificada na Figura 12.

O algoritmo desenvolvido para o processamento do conjunto de dados proposto por Islam et al. [37] requer como entrada a imagem original juntamente com a máscara de segmentação correspondente, onde a região de interesse (fundo) é representada pela cor preta.

A primeira etapa consiste na comparação entre a imagem e sua máscara, com o objetivo de substituir os *pixels* pretos na máscara pelos *pixels* correspondentes na imagem original. Dessa forma, a máscara é ajustada para refletir a cor original da água na imagem.

Em seguida, é adotada uma estratégia similar à otimização de *clusters* do algoritmo KMeans, conforme mencionado anteriormente. A média dos grupos de *pixels* representativos da água é calculada, servindo como centro de cada grupo. Posteriormente, é calculado o desvio padrão dessa distribuição. *pixels* que se encontram a uma distância superior a um desvio padrão da média, em pelo menos um dos canais do espaço de cor RGB, são excluídos do grupo final.

Após a seleção e otimização da região da imagem contendo apenas água, o algoritmo procede com a classificação propriamente dita. Inicialmente, é conduzido um teste para

determinar a inclusão ou não do canal vermelho na classificação da imagem para um determinado tipo de água.

Para realizar essa comparação, é empregada a Distância Euclidiana entre os valores dos canais de cor no espaço RGB do centro do *cluster* otimizado selecionado como representativo da água e os valores correspondentes aos tipos de água e profundidades da tabela de referência. A distância euclidiana é calculada pela seguinte equação:

$$Dist = \sqrt{(mA_t - mA_{ce})^2 + (mVd_t - mVd_{ce})^2 + (mV_t - mV_{ce})^2}, \quad (15)$$

sendo $Dist$ a distância entre a cor do tipo e profundidade atual da tabela e a cor do centro do *cluster* otimizado escolhido como representativo da água, mA a média dos *pixels* no canal de cor azul, mVd a média dos *pixels* no canal verde e mV a média dos *pixels* no canal vermelho. A marcação t representa os valores na tabela e a marcação ce representa os valores do centro escolhido.

A tabela completa é percorrida, variando os valores dos parâmetros marcados com t na equação apresentada. A imagem é então classificada no tipo e profundidade em que esses valores se aproximam mais, ou seja, o tipo e profundidade na tabela que minimizam o valor de $Dist$.

O processo é conduzido levando em consideração os três canais de cor do espaço RGB, ou apenas os canais azul e verde, se o valor do canal vermelho for inferior a 50 na escala de 0 a 255. Nesse caso, o seguinte termo da equação apresentada é desconsiderado: $(mV_t - mV_{ce})^2$.

A exclusão do canal vermelho quando o valor é baixo ocorre devido à similaridade matemática no cálculo da Distância Euclidiana entre as águas, onde o canal azul é pronunciado e o vermelho é baixo, e as águas onde os valores nos três canais são semelhantes.

Depois de classificada a imagem, a etapa final do algoritmo envolve a seleção da maior coluna d'água dentro do *cluster* que contém a água. Um *pixel* é retirado do centro. Os valores dos canais RGB do *pixel* selecionado são comparados com todas as 50 profundidades do tipo de água em que a imagem foi classificada, utilizando a Eq. (15). No entanto, essa comparação é restrita ao *pixel* escolhido e às profundidades do tipo em que a imagem foi classificada na tabela. Isso permite a definição da profundidade da cena em um ponto específico. Nesse processo, não há descarte de nenhum dos três canais de cor.

Ao final da execução, o algoritmo gera dois arquivos de texto. O primeiro arquivo contém o nome da imagem, os valores médios dos canais de cor do *cluster* otimizado escolhido como representativo da água e o tipo de água em que a imagem foi classificada, que é o objetivo principal do trabalho. O segundo arquivo também inclui o nome da imagem, o tipo em que ela foi classificada e o *pixel* selecionado para a definição da profundidade. As informações do *pixel* incluem suas coordenadas, seus valores nos canais RGB, a profundidade definida e o coeficiente de distância, que é o menor valor para

a variável $Dist$ na Eq. (15), utilizando o *pixel* escolhido e as profundidades do tipo de água em que a imagem foi classificada.

Visualmente, o algoritmo retorna a imagem original dividida com o *cluster* da água utilizando a cor original da imagem de entrada e destacando a coluna da qual o *pixel* escolhido foi retirado, representada na cor branca.

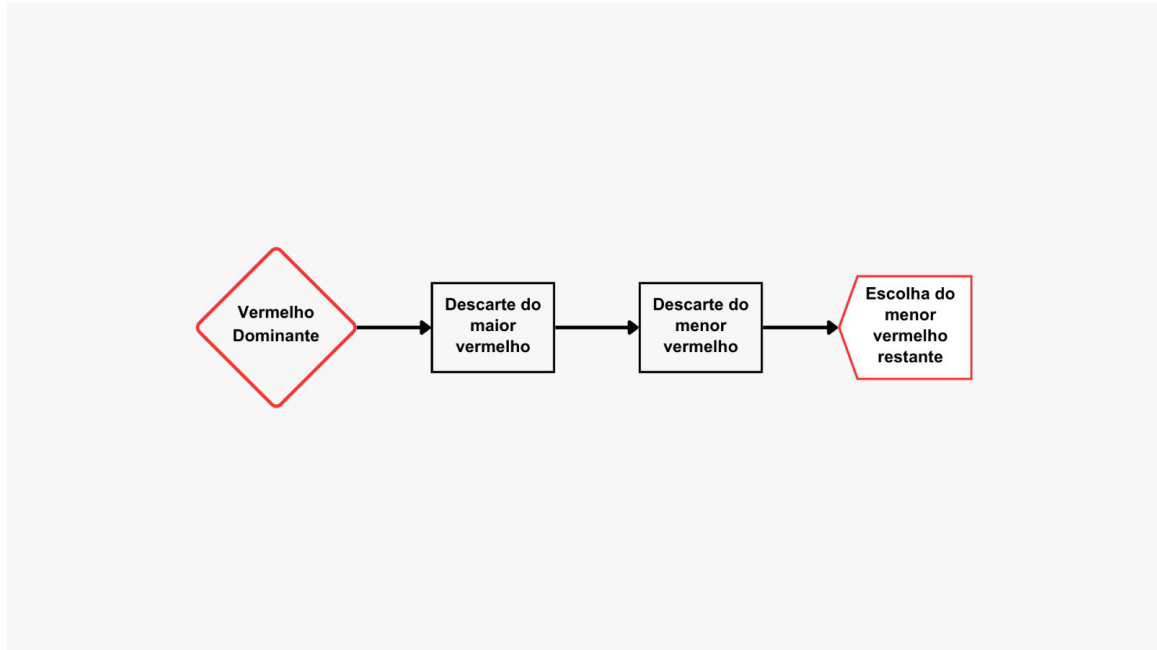


Figura 12: Diagrama demonstrativo de fluxo para seleção do *cluster* de área com apenas água em imagens com canal vermelho dominante.

5 RESULTADOS

5.1 *Dataset de Referência do Tipo de Água*

Um fator preponderante para o trabalho proposto é se ter um *dataset* confiável e com imagens que condizam com o tema abordado, uma vez que este conjunto de imagens validará os resultados obtidos pelos métodos propostos. Para tal, foi gerado um *dataset* de referência denominado DRTA, *Dataset de Referência do Tipo de Água*. Este *dataset* é composto por imagens do *dataset* proposto por Drews *et al.* [51] e imagens do *dataset* proposto por Islam *et al.* [37]. Este conjunto, que inicialmente era composto de um número maior de imagens, foi classificado por três pessoas, as quais, baseadas na classificação dos tipos de água de Jerlov [1], inferiam qual tipo de água e qual profundidade está aquela cena.

Em casos de divergência entre os três avaliadores, a imagem era retirada do *dataset*. Em caso de concordância dos três avaliadores, a imagem permanecia no *dataset*. Desta forma, o DRTA ficou composto contendo 89 imagens subaquáticas reais entre oceânicas e costeiras, permitindo uma vasta diversidade de dados e possibilidade de teste dos algoritmos. Foi necessária a geração deste *dataset* bem como a validação dos avaliadores para que um *ground truth* fosse gerado e comparado. Na literatura não encontrou-se *datasets* deste gabarito, uma vez que os poucos conjuntos de imagens com referência são em ambiente controlado (e não em casos reais) ou estão inseridos apenas em águas oceânicas.

Além da necessidade de se ter um *dataset* que abranja diferentes tipos de água para que possa-se testar os algoritmos, é fundamental que estas imagens possuam diferentes tamanhos, resoluções e características. Quanto maior diversidade do *dataset*, maior a capacidade de testar a versatilidade de cada método.

5.2 Resultados quantitativos

Utilizando o *Dataset de Referência do Tipo de Água*, esta seção apresenta os resultados quantitativos obtidos pelos dois classificadores propostos. A Tabela 1 apresenta o quantitativo de acerto entre os métodos: CLAFIS, CLACOR e Berman *et al.* [4].

Analisando os dados apresentados na Tabela 1, observa-se que, para a classificação

Tabela 1: Resultados quantitativos utilizando o *dataset* DRTA.

Classificação	Berman <i>et al.</i> [4]	CLAFIS	CLACOR	Total
Jerlov [1]	6	30	62	89
Uplavikar <i>et al.</i> [3]	17	63	78	89

de tipos de água de Jerlov, o CLACOR obteve 69,66% de acurácia, enquanto o CLAFIS obteve 33,7% e Berman *et al.* obteve apenas 6,74% de acertos. Por outro lado, considerando a classificação de tipos de água de Uplavikar *et al.*, a taxa de acerto do CLACOR sobe para 87,64%, o CLAFIS atinge 70,79% de acurácia e Berman *et al.* 19,1%.

Em mais detalhes, as Tabelas 2, 3, 4, 5, 6 e 7 mostram as matrizes de confusão de cada método, sendo as Tabelas 2, 3 e 4 aplicadas à classificação de Jerlov e as Tabelas 5, 6 e 7 aplicadas à classificação de Uplavikar *et al.* As matrizes demonstram a relação entre o tipo real de água, baseado na referência anotada, nas linhas e o tipo de água estimado por cada método. Os valores são inteiros, indicando a quantia de imagens do *Dataset* de Referência do Tipo de Água que cada método estimou entre erros e acertos.

Tabela 2: Matriz de confusão de CLAFIS segundo a classificação de Jerlov [1] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método CLAFIS.

I	18	4	3	0	0	0	0	0	0	0
IA	6	3	2	0	0	0	0	0	0	0
IB	4	2	1	1	1	0	1	1	0	0
II	3	0	3	3	5	1	2	0	0	0
III	0	0	0	4	3	1	1	0	0	0
1	5	0	1	1	2	2	0	0	0	0
3	0	0	0	1	0	3	0	0	0	0
5	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0
	I	IA	IB	II	III	1	3	5	7	9

As tabelas apresentam as matrizes de confusão que demonstram os resultados obtidos utilizando o *dataset* DRTA. Na primeira coluna de cada matriz estão os tipos de água referência, ou seja, os tipos de água anotados pelos avaliadores. Na última linha estão os tipos de água estimados pelo classificador. Em outras palavras, as linhas indicam o *ground truth* e as colunas indicam quantas vezes o método estimou aquele tipo de água. O melhor método é aquele que possui maior somatório na diagonal principal da matriz.

5.3 Resultados qualitativos

Para analisar os resultados obtidos pelos métodos qualitativamente, dividiu-se os resultados pelos tipos de água: oceânicas e costeiras. Esta maneira de enxergar e separar os

Tabela 3: Matriz de confusão de CLACOR segundo a classificação de Jerlov [1] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método CLACOR.

I	25	0	0	0	0	0	0	0	0	0
IA	5	2	4	0	0	0	0	0	0	0
IB	1	1	7	2	0	0	0	0	0	0
II	0	0	0	16	1	0	0	0	0	0
III	0	0	0	2	6	1	0	0	0	0
1	0	0	0	0	2	3	2	4	0	0
3	0	0	0	0	0	0	3	1	0	0
5	0	0	0	0	0	0	1	0	0	0
7	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0
	I	IA	IB	II	III	1	3	5	7	9

Tabela 4: Matriz de confusão de Berman *et al.* [4] segundo a classificação de Jerlov [1] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método Berman *et al.* [4].

I	3	1	1	3	1	2	2	12	0	0
IA	1	0	1	1	1	0	2	5	0	0
IB	0	1	1	0	0	2	1	6	0	0
II	1	0	1	0	2	0	3	10	0	0
III	0	0	0	4	1	1	3	0	0	0
1	1	0	0	1	1	0	0	8	0	0
3	0	1	0	0	0	0	0	3	0	0
5	0	0	0	0	0	0	0	1	0	0
7	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0
	I	IA	IB	II	III	1	3	5	7	9

Tabela 5: Matriz de confusão de CLAFIS segundo a classificação de Uplavikar *et al.* [3] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método CLAFIS.

I	43	2	1	1	0	0
III	6	15	5	0	0	0
1	6	4	5	0	0	0
5	0	0	1	0	0	0
7	0	0	0	0	0	0
9	0	0	0	0	0	0
	I	III	1	5	7	9

tipos de água é uma forma mais objetiva de expressar a diferença, os problemas enfrentados e os resultados em cada grande grupo de tipos, uma vez que estes diferem-se em características.

Tabela 6: Matriz de confusão de CLACOR segundo a classificação de Uplavikar *et al.* [3] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método CLACOR.

I	45	2	0	0	0	0
III	0	25	1	0	0	0
1	0	2	8	5	0	0
5	0	0	1	0	0	0
7	0	0	0	0	0	0
9	0	0	0	0	0	0
	I	III	1	5	7	9

Tabela 7: Matriz de confusão de Berman *et al.* [4] segundo a classificação de Uplavikar *et al.* [3] utilizando imagens do DRTA. As linhas indicam as anotações feitas pelos avaliadores e as colunas indicam as predições do método Berman *et al.* [4].

I	9	6	9	23	0	0
III	2	7	7	10	0	0
1	2	2	0	11	0	0
5	0	0	0	1	0	0
7	0	0	0	0	0	0
9	0	0	0	0	0	0
	I	III	1	5	7	9

5.3.1 Tipos de água oceânicas

Para os grupos de tipos de água oceânicos, que se encontram entre os tipos I e III, os resultados dos algoritmos propostos condizem com o esperado visualmente. Isso, em CLACOR, pode-se notar, por exemplo, pelo baixo valor no coeficiente de distância entre o *cluster* selecionado na imagem e a tabela de referência de Jerlov [1] ou Uplavikar *et al.* [3]. Para demonstrar os resultados qualitativamente, as Figuras 13 e 14 mostram imagens de entrada de imagens em águas oceânicas, bem como o *cluster* selecionado pelo método CLACOR e o tipo de água, dada uma profundidade, que o algoritmo estimou. As imagens nas Figuras 13 e 14 foram retiradas do DRTA.

As Tabelas 8 e 9 detalham os resultados das Figuras 13 e 14 respectivamente. As tabelas mostram cinco imagens, o valor de referência de cada uma delas, segundo anotação feita no DRTA, o valor estimado por CLACOR, a profundidade empregada, o resultado estimado por CLAFIS e sua profundidade.

As imagens selecionadas para compor as Figuras 13 e 14 têm como objetivo ilustrar a similaridade entre os tipos de água oceânicos, fundamentando, assim, a adoção do agrupamento por Uplavikar *et al.* [3] na classificação proposta por Jerlov [1]. Essa similaridade é também a razão das discrepâncias nos resultados quantitativos entre os dois métodos de classificação, conforme apresentado na Tabela 1.

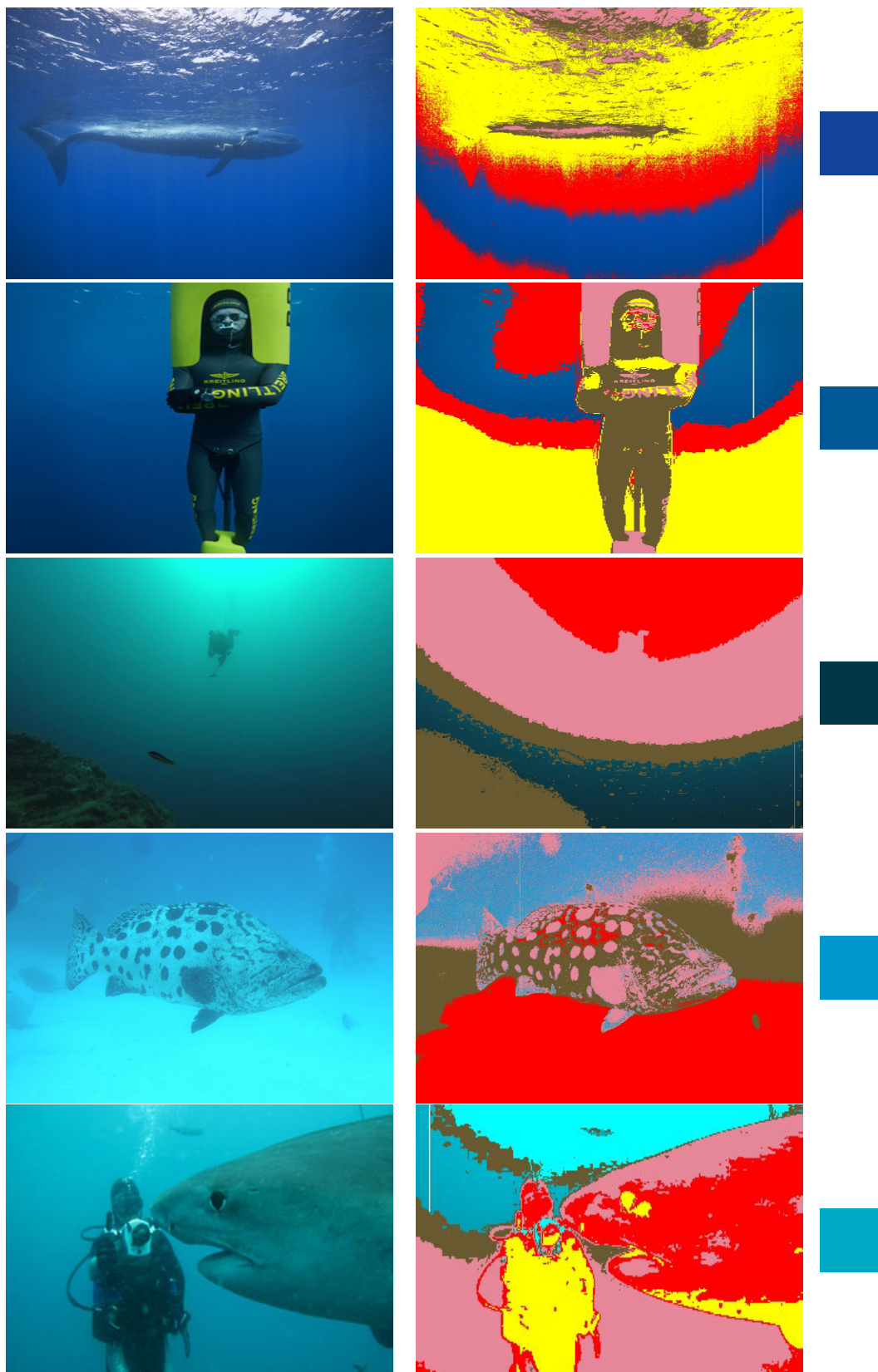


Figura 13: Resultados obtidos para CLACOR e CLAFIS utilizando imagens dos tipos de água oceânicas do DRTA.

Tabela 8: Resultados detalhados da Figura 13

Imagem	Referência	CLACOR	Profundidade (m)	CLAFIS	Profundidade (m)
1	I	I	28	I	21,5
2	IA	IA	21	I	22,17
3	II	II	21	I	6,03
4	I	I	12	I	9,36
5	IB	IB	8	I	10,65

Tabela 9: Resultados detalhados da Figura 14

Imagem	Referência	CLACOR	Profundidade (m)	CLAFIS	Profundidade (m)
1	IB	IB	21	II	8,72
2	I	I	28	I	26,23
3	I	I	25	IA	21,02
4	II	II	10	I	33,12
5	I	I	12	IA	15,3

5.3.2 Tipos de água costeiras

Considerando as imagens em águas costeiras, lida-se com uma limitação de *datasets*, principalmente referentes aos tipos de água 7 e 9 na tabela de Jerlov. Apesar de o *dataset* contemplar imagens submersas a águas do tipo costeiras, o conjunto é limitado a poucas imagens. Apesar disso, conseguiu-se obter resultados visualmente condizentes com a realidade, apesar do coeficiente de distância de CLACOR um pouco mais elevado. Os resultados são apresentados nas Figuras 15 e 16 e valores detalhados mostrados nas Tabelas 10 e 11.

Tabela 10: Resultados detalhados da Figura 15

Imagem	Referência	CLACOR	Profundidade (m)	CLAFIS	Profundidade (m)
1	1	1	3	I	7,85
3	3	3	3	1	3,59
4	3	3	5	1	6,92

Tabela 11: Resultados detalhados da Figura 16

Imagem	Referência	CLACOR	Profundidade (m)	CLAFIS	Profundidade (m)
1	1	1	4	III	6,22
2	3	3	2	1	4,08
3	3	3	3	1	4,83

A Figura 15 apresenta três casos onde o canal de verde é bem mais acentuado que os outros canais de cor. Para esses casos, nota-se visualmente que a profundidade estimada não é exatamente a profundidade da imagem, porém é a profundidade (e tipo) que o algoritmo encontrou como mais próxima da realidade. Já na Figura 16 são apresentados

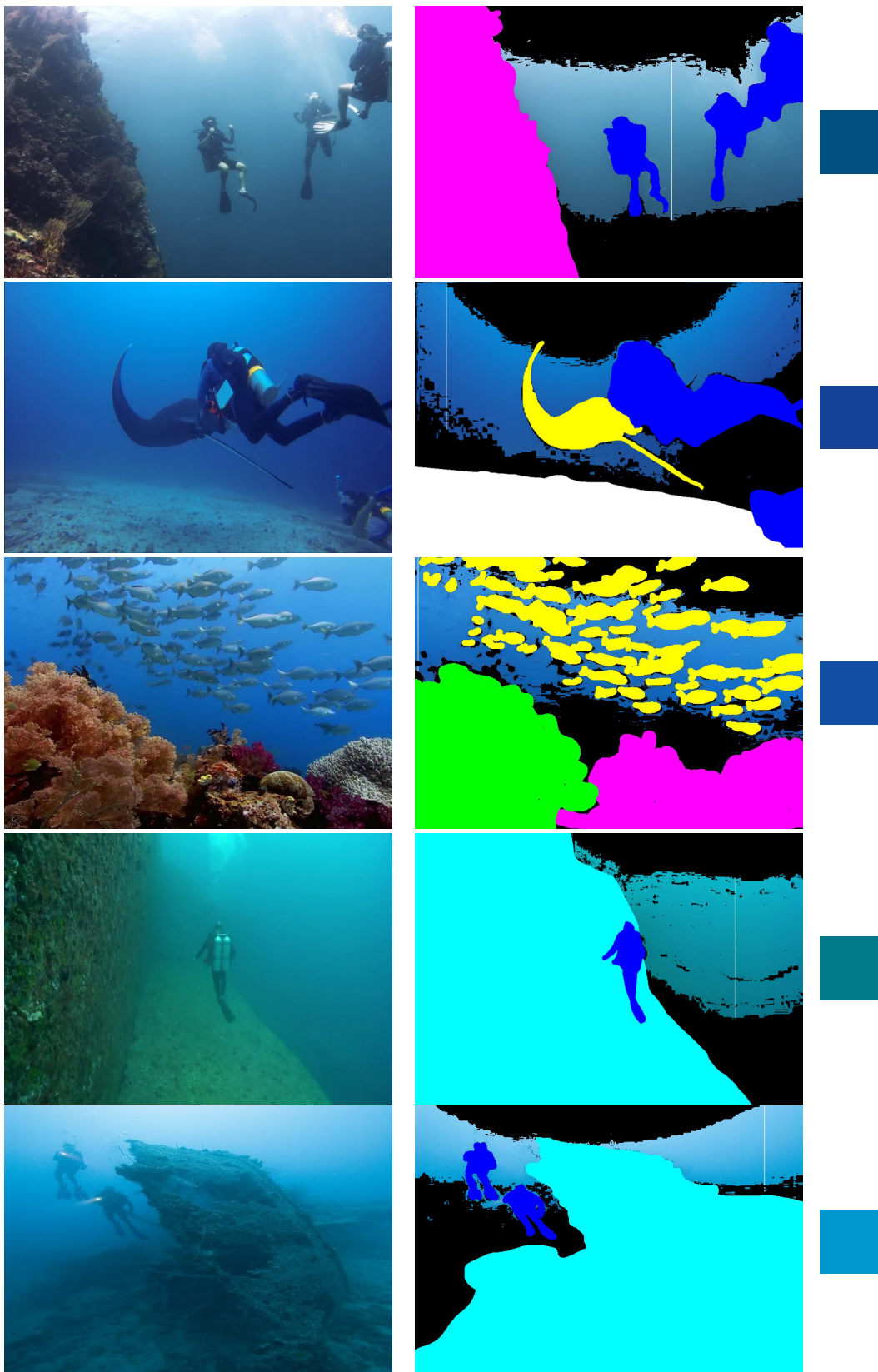


Figura 14: Resultados obtidos para CLACOR e CLAFIS utilizando imagens dos tipos de água oceânicas do DRTA.

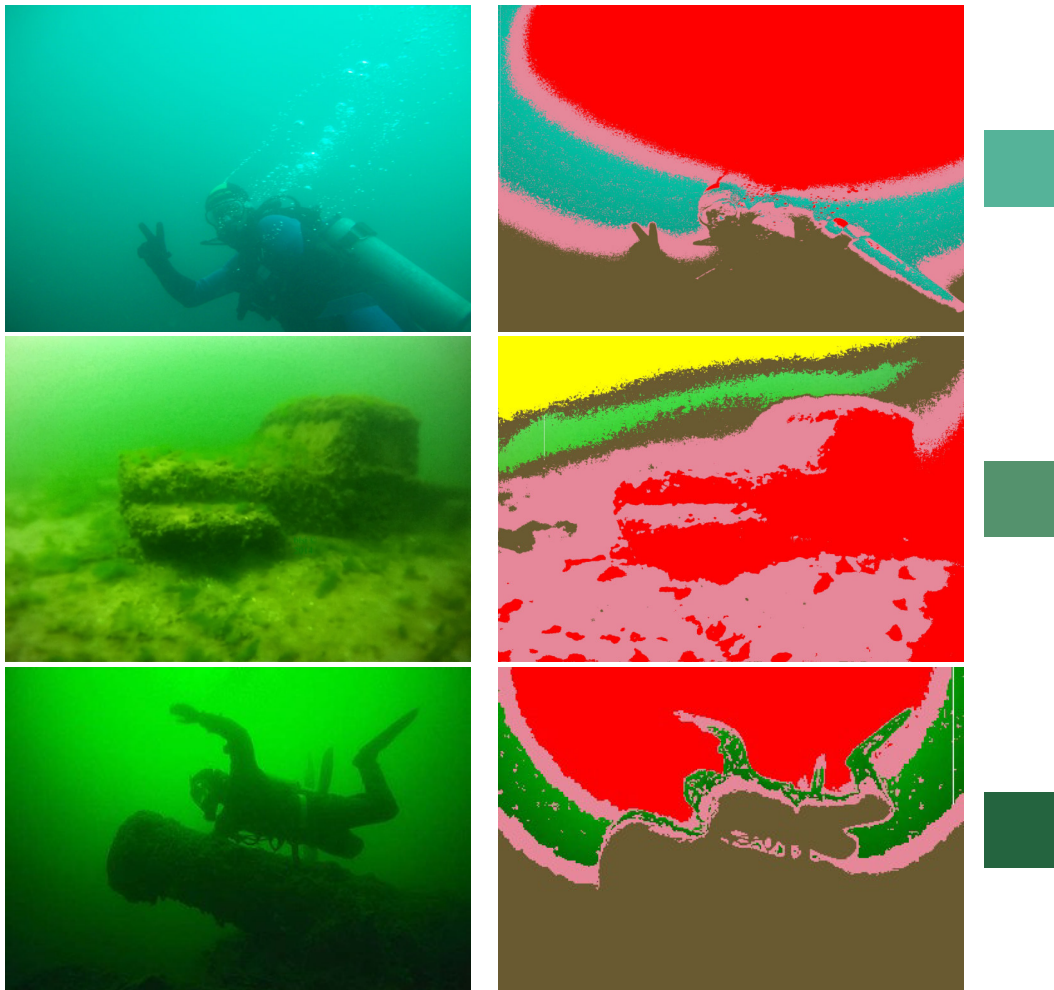


Figura 15: Resultados obtidos para CLACOR e CLAFIS utilizando imagens dos tipos de água costeiras do DRTA.

exemplos que se diferem bastante em coloração, porém com profundidade muito parecida. Isso mostra que em águas costeiras a atenuação dos canais de cor é mais significativa e portanto há maior perda de visibilidade.

5.3.3 Casos falhos

Os casos falhos podem ser ocasionados por dois fatores principais: imagens erroneamente capturadas e falha de reconhecimento da região de apenas água na imagem por parte do algoritmo. Como imagens inadequadas, pode-se considerar imagens capturadas na vertical, ou seja, apontando para a superfície da água ou para o fundo do mar, apontando apenas para um objeto ou em profundezas altas o suficiente para ter baixa ou nenhuma visibilidade da cena. A Figura 17 apresenta exemplos.

Em relação a erros de detecção de algoritmo, a Figura 18 exemplifica dois casos onde, para o método CLACOR foram selecionados como apenas água os *clusters* erroneamente. Apesar de os *clusters* terem sido gerado de maneira correta em ambos os casos, a seleção foi falha. Em ambas imagens da Figura 18, nota-se que há uma variação considerável

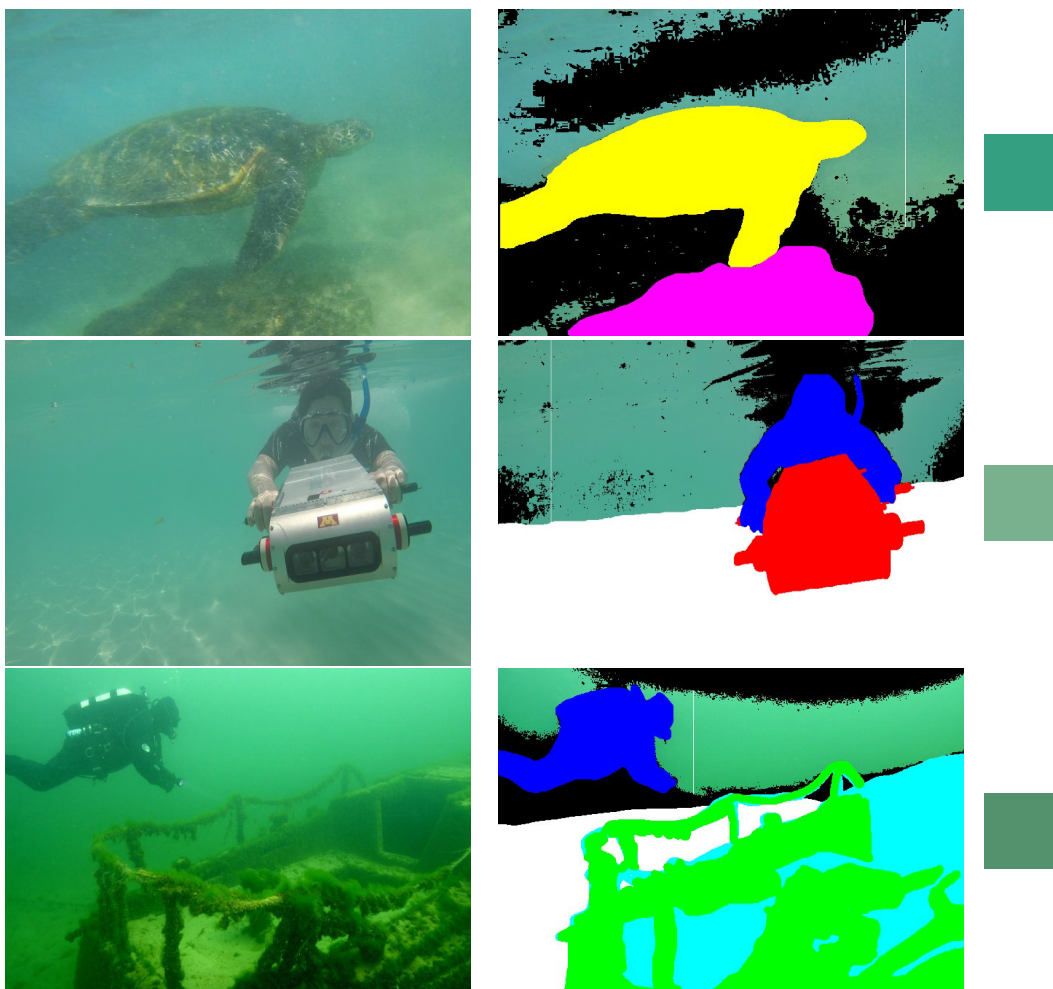


Figura 16: Resultados obtidos para CLACOR e CLAFIS utilizando imagens dos tipos de água costeiras do DRTA.

entre colorações nas imagens, principalmente na primeira imagem nesta Figura. Portanto, na seleção entre os *clusters*, o algoritmo acabou selecionando casos onde objetos estão inseridos e não apenas água.



Figura 17: Casos falhos onde imagens são capturadas para o fundo do mar e superfície respectivamente. A esquerda, a imagem original. A direita, o *cluster* selecionado pelo algoritmo CLACOR.

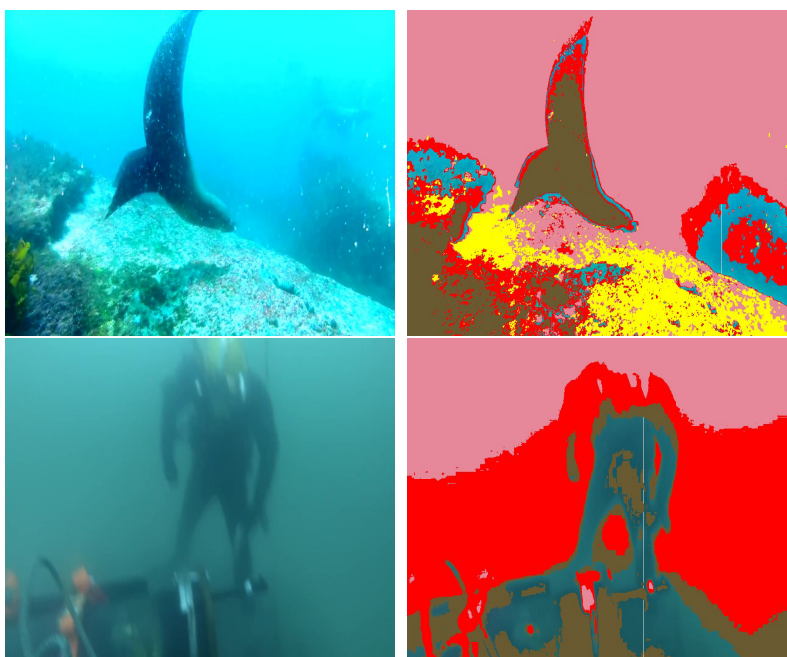


Figura 18: Casos falhos de escolha errada de *cluster* pelo algoritmo CLACOR. A esquerda, a imagem original. A direita, o *cluster* selecionado.

6 CONSIDERAÇÕES FINAIS

6.1 Conclusões

O objetivo deste trabalho é apresentar propostas para solucionar o problema de classificação de tipos de água em imagens subaquáticas monoculares. Neste trabalho, foram construídos dois métodos, um baseado em modelo físico (CLAFIS) e outro baseado em modelo de cor. Apesar de terem sido adotados métodos mais clássicos e empíricos, em comparação ao método de Berman *et al.* [4] os resultados são promissores. Além disso, foi criado um *dataset* denominado *Dataset de Referência de Tipos de Água (DRTA)*, o qual possui 89 imagens subaquáticas monoculares de diferentes resoluções, tamanhos e submersas a diferentes tipos água, de forma a poder testar a amplitude dos métodos.

Os resultados obtidos mostram uma larga eficácia do método CLACOR utilização ambas classificações de Jerlov [1] e Uplavikar *et al.* [3]. Pouco abaixo em eficácia em relação ao método CLACOR, o método CLAFIS obteve ótimos resultados em comparação ao estado da arte até então. Os métodos foram construídos com estudo aprofundado da literatura, bem como observações empíricas que ajudaram a construir o algoritmo CLACOR e a gerar um algoritmo capaz de interpretar as observações anotadas.

Em suma, o trabalho rendeu duas publicações como contribuição, bem como dois métodos que podem ser úteis para métodos de restauração de imagens, detecção de objetos, mapas de profundidade e tantas outras aplicações subaquáticas, uma vez que entender o tipo de água pode ser requisito básico ou uma informação importante para o algoritmo a ser aplicado. Além disso, há a contribuição de um *dataset* contendo a referência de tipo de água anotada, o qual pode ser utilizado para testar e validar outros métodos.

6.2 Trabalhos futuros

Como trabalhos futuros, fica a sugestão de aplicação de redes neurais para aprender a classificar tipos de água. Apesar de aplicações como esta já serem utilizadas, costumemente estes métodos são difíceis de conseguir classificar tipos de águas costais, principalmente pela falta de um *dataset* contendo esse tipo de água. Portanto, outro trabalho futuro é a ampliação do DRTA contendo mais imagens e mais tipos de água com

suas respectivas referências, possibilitando estudos mais aprofundados e até aplicação de redes neurais.

REFERÊNCIAS

- [1] JERLOV, N. G. *Marine optics*. [S.l.]: Elsevier, 1976. v. 14.
- [2] JOHNSON, L. J. The underwater optical channel. *Dept. Eng., Univ. Warwick, Coventry, UK, Tech. Rep*, 2012.
- [3] UPLAVIKAR, P. M.; WU, Z.; WANG, Z. All-in-one underwater image enhancement using domain-adversarial learning. In: *CVPR Workshops*. [S.l.: s.n.], 2019. p. 1–8.
- [4] BERMAN, D. et al. Underwater single image color restoration using haze-lines and a new quantitative dataset. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, 2020.
- [5] DEBATS, S. R. et al. A generalized computer vision approach to mapping crop fields in heterogeneous agricultural landscapes. *Remote Sensing of Environment*, Elsevier, v. 179, p. 210–221, 2016.
- [6] ZHANG, L. et al. A super-resolution reconstruction algorithm for surveillance images. *Signal Processing*, Elsevier, v. 90, n. 3, p. 848–859, 2010.
- [7] YAO, Y.; ABIDI, B. R.; ABIDI, M. A. Extreme zoom surveillance: System design and image restoration. *Journal of Multimedia*, v. 2, n. 1, p. 20–31, 2007.
- [8] TORRES-MÉNDEZ, L. A.; DUDEK, G. Color correction of underwater images for aquatic robot inspection. In: SPRINGER. *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*. [S.l.], 2005. p. 60–73.
- [9] QIAO, X. et al. Visibility enhancement for underwater robots based on an improved underwater light model. *Journal of Robotics and Mechatronics*, Fuji Technology Press Ltd., v. 30, n. 5, p. 781–790, 2018.
- [10] MANZANILLA, A. et al. Autonomous navigation for unmanned underwater vehicles: Real-time experiments using computer vision. *IEEE Robotics and Automation Letters*, IEEE, v. 4, n. 2, p. 1351–1356, 2019.

- [11] ANCUTI, C. et al. Enhancing underwater images and videos by fusion. *Proceedings of IEEE Conference of Computer Vision and Pattern Recognition*, p. 81–88, 2012.
- [12] LI, C.-Y. et al. Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior. *IEEE Transactions on Image Processing*, IEEE, v. 25, n. 12, p. 5664–5677, 2016.
- [13] CARLEVARIS-BIANCO, N.; MOHAN, A.; EUSTICE, R. M. Initial results in underwater single image dehazing. In: IEEE. *Oceans 2010 Mts/IEEE Seattle*. [S.l.], 2010. p. 1–8.
- [14] CHIANG, J. Y.; CHEN, Y.-C. Underwater image enhancement by wavelength compensation and dehazing. *IEEE Transactions on Image Processing*, IEEE, v. 21, n. 4, p. 1756–1769, 2011.
- [15] ANCUTI, C.; ANCUTI, C. O. Single image dehazing by multi-scale fusion. *IEEE Transactions on Image Processing*, v. 22, n. 8, p. 3271–3282, 2013.
- [16] HE, K.; SUN, J.; TANG, X. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 33, n. 12, p. 2341–2353, 2010.
- [17] FATTAL, R. Single image dehazing. *ACM transactions on graphics (TOG)*, ACM, v. 27, n. 3, p. 72, 2008.
- [18] YOU, S. et al. Adherent raindrop modeling, detection and removal in video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 38, n. 9, p. 1721–1733, 2016.
- [19] YOU, S. et al. Raindrop detection and removal from long range trajectories. *Asian Conference on Computer Vision*, 2014.
- [20] EIGEN, D.; KRISHNAN, D.; FERGUS, R. Restoring an image taken through a window covered with dirt or rain. *IEEE International Conference on Computer Vision*, 2013.
- [21] ZHOU, C.; LIN, S. Removal of image artifacts due to sensor dust. *Proceedings of IEEE Conference of Computer Vision and Pattern Recognition*, 2007.
- [22] GU, J. et al. Removing image artifacts due to dirty camera lenses and thin occluders. *ACM SIGGRAPH Asia*, 2009.
- [23] YANG, P. et al. Dust particle detection in surveillance video using salient visual descriptors. *Computers and Electrical Engineering*, 2016.

- [24] LI, C.; ZHOU, K.; LIN, S. Removal of dust artifacts in focal stack image sequences. *International Conference on Pattern Recognition*, 2012.
- [25] PEI, S.-C.; TSAI, Y.-T.; LEE, C.-Y. Removing rain and snow in a single image using saturation and visibility features. In: IEEE. *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. [S.l.], 2014. p. 1–6.
- [26] VAZ-JR, E. S.; DREWS-JR, P.; STEFFENS, C. Restoration of images affected by welding fume. *Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2018.
- [27] SILBERMAN DEREK HOIEM, P. K. N.; FERGUS, R. Indoor segmentation and support inference from rgb-d images. In: *ECCV*. [S.l.: s.n.], 2012.
- [28] CORDTS, M. et al. The cityscapes dataset for semantic urban scene understanding. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 3213–3223.
- [29] VAZ, E. S.; TOLEDO, E. F. de; DREWS, P. L. Underwater depth estimation based on water classification using monocular image. In: IEEE. *2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE)*. [S.l.], 2020. p. 1–6.
- [30] TOLEDO, E. F. de; VAZ, E. S.; DREWS, P. L. J. Water classification based on underwater monocular image. In: *2021 Latin American Robotics Symposium (LARS), 2021 Brazilian Symposium on Robotics (SBR), and 2021 Workshop on Robotics in Education (WRE)*. [S.l.: s.n.], 2021. p. 282–287.
- [31] JAFFE, J. S. Computer modeling and the design of optimal underwater imaging systems. *IEEE Journal of Oceanic Engineering*, IEEE, v. 15, n. 2, p. 101–111, 1990.
- [32] MCGLAMERY, B. A computer model for underwater camera systems. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Ocean Optics VI*. [S.l.], 1980. v. 208, p. 221–231.
- [33] SCHECHNER, Y. Y.; KARPEL, N. Recovery of underwater visibility and structure by polarization analysis. *IEEE Journal of oceanic engineering*, IEEE, v. 30, n. 3, p. 570–587, 2005.
- [34] BUCHSBAUM, G. A spatial processor model for object colour perception. *Journal of the Franklin institute*, Elsevier, v. 310, n. 1, p. 1–26, 1980.
- [35] AKKAYNAK, D. et al. What is the space of attenuation coefficients in underwater computer vision? In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2017. p. 4931–4940.

- [36] AKKAYNAK, D.; TREIBITZ, T. Sea-thru: A method for removing water from underwater images. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2019. p. 1682–1691.
- [37] ISLAM, M. J. et al. Semantic segmentation of underwater imagery: Dataset and benchmark. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. [S.l.: s.n.], 2020. p. 1769–1776.
- [38] KALAIARASI, G. et al. A deep learning approach to detecting objects in underwater images. *CYBERNETICS AND SYSTEMS*, TAYLOR & FRANCIS INC 530 WALNUT STREET, STE 850, PHILADELPHIA, PA 19106 USA, 2023.
- [39] AAS, E. et al. Optical water types of the nordic seas and adjacent areas. *Oceanologia*, Elsevier, v. 55, n. 2, p. 471–482, 2013.
- [40] BEKERMAN, Y.; AVIDAN, S.; TREIBITZ, T. Unveiling optical properties in underwater images. In: IEEE. *2020 IEEE International Conference on Computational Photography (ICCP)*. [S.l.], 2020. p. 1–12.
- [41] DESAI, C. et al. Ruig: Realistic underwater image generation towards restoration. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2021. p. 2181–2189.
- [42] RAIHAN, A. J.; ABAS, P. E.; SILVA, L. C. D. A simulated water type dataset (swtd) based on jerlov water types for underwater image quality analysis. *Journal of Advances in Information Technology Vol*, v. 12, n. 4, 2021.
- [43] LIU, Y. et al. Model-based underwater image simulation and learning-based underwater image enhancement method. *Information*, MDPI, v. 13, n. 4, p. 187, 2022.
- [44] WILLIAMSON, C. A.; HOLLINS, R. C. Depth profiles of jerlov water types. *Limnology and Oceanography Letters*, Wiley Online Library, v. 8, n. 5, p. 781–788, 2023.
- [45] MITTAL, S.; SRIVASTAVA, S.; JAYANTH, J. P. A survey of deep learning techniques for underwater image classification. *IEEE Transactions on Neural Networks and Learning Systems*, v. 34, n. 10, p. 6968–6982, 2023.
- [46] ROSER, M.; DUNBABIN, M.; GEIGER, A. Simultaneous underwater visibility assessment, enhancement and improved stereo. In: IEEE. *2014 IEEE International Conference on Robotics and Automation (ICRA)*. [S.l.], 2014. p. 3840–3847.
- [47] CHAVEZ, A. G. et al. Caddy underwater stereo-vision dataset for human–robot interaction (hri) in the context of diver activities. *Journal of Marine Science and Engineering*, v. 7, n. 1, 2019. ISSN 2077-1312.

- [48] LI, C. et al. An underwater image enhancement benchmark dataset and beyond. *IEEE Transactions on Image Processing*, v. 29, p. 4376–4389, 2020.
- [49] BERMAN, D. et al. Underwater single image color restoration using haze-lines and a new quantitative dataset. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.
- [50] COLEMAN, T.; LI, Y. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, v. 6, n. 2, p. 418–445, 1996.
- [51] DREWS-JR, P. et al. Underwater image segmentation in the wild using deep learning. *Journal of the Brazilian Computer Society*, Springer, v. 27, p. 1–14, 2021.
- [52] MACQUEEN, J. et al. Some methods for classification and analysis of multivariate observations. In: OAKLAND, CA, USA. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. [S.l.], 1967. v. 1, n. 14, p. 281–297.
- [53] PENA, J. M.; LOZANO, J. A.; LARRANAGA, P. An empirical comparison of four initialization methods for the k-means algorithm. *Pattern recognition letters*, Elsevier, v. 20, n. 10, p. 1027–1040, 1999.