DOI: 10.48011/asba.v2i1.1417

Visão Computacional e Redes Neurais Convolucionais aplicadas à Detecção de Vazamentos de Óleo

Marcel M. Mendes * **Levi W. de Resende Filho** *,**
Lucas V. Vargas * **Alessandro Jacoud Peixoto** *

* *Programa de Eng. Elétrica, COPPE, Universidade Federal do Rio de Janeiro (marcel.mendes@coppe.ufrj.br, vares.lucas@gmail.com, jacoud@poli.ufrj.br)*
** *Programa de Pós-Graduação em Instrumentação, Controle e Automação de Processos de Mineração, PROFICAM, Universidade Federal de Ouro Preto e Instituto Tecnológico Vale (levi.filho@aluno.ufop.edu.br)*

Abstract: In this work, two algorithms are developed for the automatic detection of oil spills using information from local sensors, such as images from visible cameras. The first method uses Computer Vision and Image Processing, based on unsupervised segmentation operations and color thresholds. The second uses Convolutional Neural Networks, in a supervised learning approach. Numerical simulations illustrate the performance of both detection methodologies and a comparison between them is presented.

Resumo: Neste trabalho são desenvolvidos dois algoritmos para a detecção automática de derramamentos de óleo usando informações de sensores locais, como imagens de câmeras visíveis. O primeiro método usa Visão Computacional e Processamento de Imagens, com base em operações de segmentação não supervisionadas e limiares de cores. O segundo usa Redes Neurais Convolucionais, em uma abordagem de aprendizado supervisionado. Simulações numéricas ilustram o desempenho de ambas as metodologias de detecção e é apresentada uma comparação entre eles.

Keywords: oil spill, computer vision, image processing, convolutional neural network, detection algorithms.

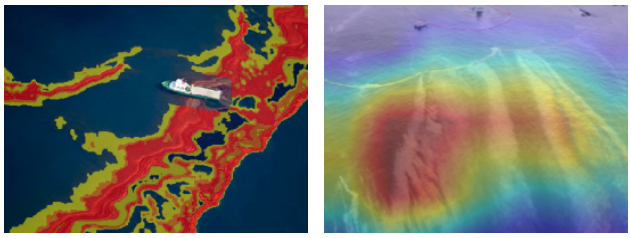
Palavras-chave: vazamento de óleo; visão computacional; processamento de imagem; redes neurais convolucionais; algoritmos de detecção.

1. INTRODUÇÃO

A exploração de petróleo no mar teve início por volta dos anos de 1930 e 1950, na Venezuela e no Golfo do México. No Brasil, o primeiro poço *offshore* foi descoberto em 1968, no Campo de Guaricema (SE), seguido do Campo Garoupa, na Bacia de Campos. Atualmente, o Brasil é pioneiro em exploração de petróleo em águas profundas, tendo recebido diversas premiações (Petrobras, 2020). Dentre os riscos presentes na exploração de petróleo em alto mar, pode-se destacar acidentes relacionados a vazamentos de óleo, sendo diversas suas causas, tais como: rompimentos ou fraturas em dutos, naufrágios, perfuração ou sobre-pressão nos poços.

O principal efeito do derramamento de óleo é o seu impacto negativo na natureza, incluindo a destruição de ambientes e ecossistemas oceânicos. Também é importante destacar seus riscos à saúde humana (Pérez-Cadahía et al., 2007) e seu grande impacto econômico devido ao custo das operações de contenção, recuperação e limpeza.

A grande dificuldade em detectar derramamentos de óleo está no fato de que o óleo não reage fundamentalmente diferente da água com o espectro eletromagnético. Além disso, a aparência do óleo sobre a água pode variar condicionada por fatores ambientais (e.g. agitação da água), pela qualidade do óleo e tempo de exposição (Lammoglia, 2010).



(a) (b)

Figura 1. Resultados dos algoritmos de detecção de óleo. Em (a) resultado obtidos por técnicas de Visão Computacional e Processamento de Imagens, em (b) a partir de Redes Neurais Convolucionais.

O monitoramento remoto de vazamentos constitui um componente importante da gestão de desastres em derramamentos de óleo. Reduzir o risco de desastres é essencial

para proteger o meio ambiente e reduzir as perdas econômicas. Dessa forma, diversos estudos vem sendo realizados. Em (Fingas and Brown, 1997), os autores apresentam algumas tecnologias atualmente presentes na detecção de óleo, exemplificando o uso de sensores utilizados neste processo, classificados como passivos e ativos. Os sensores passivos, em que a fonte de informação é uma consequência da luz solar, detectam a radiação emitida ou refletida dentro de um espectro, como infravermelho, ultravioleta e visível. Nos sensores ativos, existe uma interação com o ambiente em que este transmite um sinal de onda, geralmente em micro-ondas ou radiofrequência, e o sinal refletido é medido. Além disso, são apresentadas algumas vantagens e desvantagens do uso de câmeras visíveis, câmeras térmicas, imagens de satélites e o uso de sensores do tipo *fluorsensor*.

Em particular, no contexto de sensores ativos, muitos dos trabalhos desenvolvidos nesta área envolvem imagens de satélites ou aeronaves que utilizam radar de abertura sintética (*Synthetic Aperture Radar* – SAR), que consiste de sensores de micro-ondas ativos que capturam imagens bidimensionais. O brilho da imagem é um reflexo das propriedades de retroespalhamento de micro-ondas da superfície. As manchas de óleo amortecem as ondas de Bragg (comprimento de onda de alguns centímetros) na superfície do oceano e reduzem o coeficiente de retroespalhamento do radar (SAR), resultando em regiões escuras ou manchas nas imagens obtidas de um satélite (SAR). Essas imagens são capazes de detectar camadas finas de óleo a uma velocidade de até 3 m/s e óleo emulsificado de até 12 m/s (Brekke and Solberg, 2005).

Por outro lado, no contexto de sensores passivos, destaca-se o uso de imagens de câmeras hiperespectrais. Além disso, vêm sendo propostos esquemas de sensoriamento híbrido que combinam informações de radar ou radiômetro de micro-ondas (sensor ativo) e câmera térmica (sensor passivo) ou câmera visível (Lammoglia, 2010).

Esse arsenal de sensoriamento híbrido (ou não) combinado com algoritmos de detecção automática permite desenvolver sistemas automáticos de detecção de óleo menos suscetíveis a falhas. Os algoritmos de detecção automática de derramamento de óleo são normalmente divididos em três etapas: detecção, extração e classificação das manchas escuras.

Kubat et al. (1998) e Solberg et al. (1999) descrevem algoritmos de detecção de vazamentos de óleo utilizando imagens de satélites (SAR). Em (Solberg et al., 1999), o algoritmo detecta manchas escuras na imagem e, utilizando um modelo, classifica a imagem com a presença ou não de óleo. Este modelo foi testado em 84 imagens e apresentou um resultado de 94% de acerto para detecção de imagens com a presença de óleo e 99% para imagens que não apresentavam óleo. Outros exemplos nesta linha podem ser encontrados em (Del Frate et al., 2000) e (Fiscella et al., 2000). Em (Kubat et al., 1998), os autores destacam uma série de questões durante o desenvolvimento de um componente de aprendizado de máquina para um sistema de detecção de derramamento de óleo. A primeira questão é a escassez de dados, pois a maioria das imagens SAR de satélite adquiridas não contém derramamentos de óleo. O segundo é o conjunto de treinamento desequilibrado. A detecção de derramamento de óleo é uma aplicação em que o classificador deve detectar um evento raro, mas importante. Os derramamentos de óleo aparecem em lotes, onde exemplos retirados da mesma imagem constituem um único lote. Pode haver uma semelhança entre lotes e uma dissimilaridade entre lotes, o que pode influenciar o treinamento do classificador. E, finalmente, o desempenho do classificador está relacionado à possibilidade do usuário de diminuir o número de alarmes à custa de vazamentos de óleo não detectados (especificidade).

O uso de imagens de satélites *remotos* (SAR) para identificação e inspeção de vazamento é muito útil por sofrer pouca influência do clima, por poder ser utilizado em monitoramentos 24/7 (limitado pelo número de satélites e a frequência de varredura) e por ter uma ampla cobertura (Brekke and Solberg, 2005). Porém, o maior desafio é a discriminação precisa entre derramamentos de óleo e eventos visualmente similares, como proliferação de algas.

Neste artigo, visando contornar o desafio da baixa resolução das técnicas de sensoriamento *remoto* via imagens de satélites (SAR) e visando aumentar a frequência de varredura do monitoramento, propõe-se utilizar sensores passivos e *locais*, instalados de forma apropriada para a obtenção de imagens da mancha de óleo, com dois algoritmos de detecção automática, como ilustra a Figura 1. O primeiro algoritmo busca avaliar até que ponto técnicas convencionais de Processamento de Imagens e Visão Computacional são suficientes para prover uma detecção robusta utilizando informação de câmeras visíveis. O segundo algoritmo utiliza aprendizado de máquina baseado em Redes Neurais Convolucionais para a classificação das imagens. Por fim, são comparados os resultados obtidos com as duas abordagens, utilizando um conjunto de dados (*dataset*) com 228 imagens de câmeras visíveis.

1.1 Construção do dataset

O *dataset* utilizado neste trabalho foi construído a partir de imagens de sites de busca, devido a dificuldade de se reproduzir uma condição ideal para análise. Desta forma, com as imagens obtidas, é possível obter uma hipótese inicial mais acurada, isto é, se há ou não a presença de óleo naquela região. Ele contém 228 imagens, sendo categorizado em duas classes balanceadas: *clean* e *oil spill*, os quais representam o conjunto de imagens sem vazamento de óleo e com vazamento, respectivamente. No grupo *oil spill* foram selecionadas imagens em três diferentes estados de contaminação: filme, cru e emulsificado (Lammoglia, 2010).

2. VISÃO COMPUTACIONAL

Nesta seção, será tratado o problema de detecção de manchas e derramamentos de óleo de uma forma tradicional, utilizando uma abordagem típica por visão computacional. Em um nível abstrato, o objetivo é utilizar os dados de imagem observados para inferir algo sobre o mundo (Prince, 2012) - neste caso, deseja-se constatar algo extremamente indesejado no mundo. Assim, ferramentas convencionais para processamento e tratamento de imagens serão amplamente exploradas, sendo essenciais para a extração de elementos e/ou características marcantes e também para a busca por padrões conhecidos. Ao final

desta seção, deseja-se, portanto, obter um algoritmo capaz de usar fotos e vídeos para identificar e monitorar com precisão satisfatória manchas de óleo e poluentes indesejados no ambiente. Para o desenvolvimento do que se segue, foram utilizados o *MATLAB* e sua *Image Processing Toolbox*.

2.1 Tratamentos Preliminares e Espaços de Cores.

A base do algoritmo de detecção proposto consiste no conhecimento do meio em que está inserido e na busca por elementos que se destaquem e que também apresentem determinadas características aproximadamente conhecidas (e.g. busca por manchas mais avermelhadas para uma detecção de óleo emulsificado). Neste ponto, torna-se natural recorrer a análise de diferentes espaços de cor (normalizados ou não) e também pseudo-cores (combinações lineares de componentes) que amplifiquem determinadas características.

Para a detecção de óleo, uma análise mais básica realizaria uma simples busca em faixas predominantemente vermelhas do espectro de cores - canal H na representação em HSV (*hue-saturation-value*) (Jang and Nam, 2020). A Figura 2 apresenta duas fotos de um ambiente predominantemente marítimo e as distribuições dos valores H observados. Nota-se que na Figura 2(a) uma grande parcela da imagem é referente à costa e sua vegetação, sendo tal característica evidenciada pelo segundo pico de H na faixa verde (ver Figura 2(b)). Nas Figuras 2(c) e 2(d) nota-se a existência de componentes na faixa vermelha, potencial indicador de óleo ao se assumir que tal imagem provém de um ambiente marítimo. Ainda assim, tal análise é insuficiente, pouco confiável como fonte única e bastante generalizada, uma vez que acusa qualquer elemento nessa faixa de cor e também teria sua aplicação restrita aos casos em que o óleo é muito marcante na imagem.

Na análise do que se segue, são utilizados portanto outros dois espaços de cores. São eles: $YCbCr$ e $L^*a^*b^*$. Como usual, as cores são então expressas por três componentes e, de forma geral, a primeira componente em ambos os espaços está ligada à luminosidade. As seguintes estão associadas a existência de tons azuis e vermelhos. Cb é denominada crominância azul e Cr crominância vermelha. Para o espaço $L^*a^*b^*$, $-a/+a$ indica o nível de verde/vermelho e $-b/+b$ indica azul/amarelo. Para um ambiente marítimo em que são esperadas tonalidades azuis e possíveis contaminações - em diferentes níveis - de vermelho, tais escolhas são naturais e serão cruciais para a identificação de segmentos similares e posterior classificação.

Algoritmos podem extrair de imagens algumas características e aspectos marcantes, fornecendo valiosas informações do que se vê. Porém, as etapas preliminares são igualmente importantes e afetam diretamente o desempenho desses métodos por visão. Operações, como *blur, dilate, erode* - e composições destas como *open* e *close* - se mostram essenciais para a homogeneização e redução de pequenos ruídos na imagem. Para o ambiente em questão, torna-se ainda mais crucial para amenizar e assemelhar pequenos pontos brancos existentes por conta de ondulações e da luz incidente.

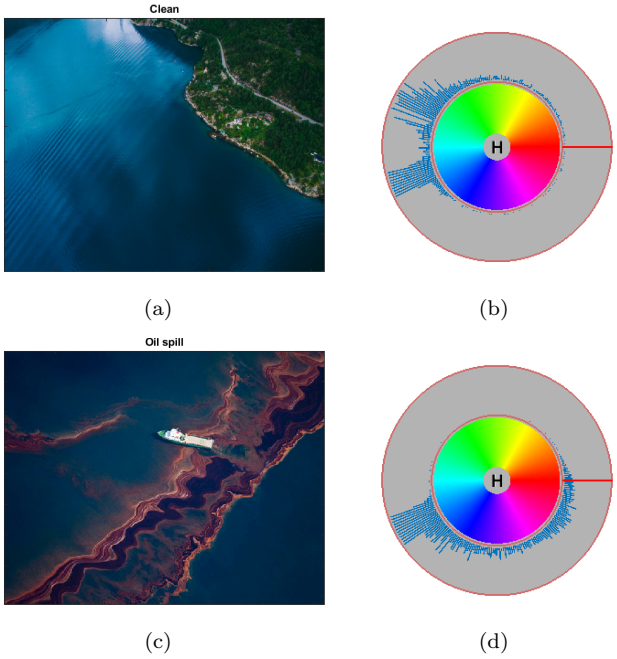


Figura 2. Fotos sem (a) e com (c) derramamento de óleo, e respectivos histogramas de H (b)(d).

Para a aplicação de interesse, é realizado um filtro das componentes altas de V (HSV), associadas ao céu, pontos iluminados na água e remanescentes, e elementos brancos/vívidos (tais como barcos, plataformas, equipamentos e espumas no mar). Note que o conhecimento de parâmetros intrínsecos e extrínsecos da câmera também poderia ser utilizado para identificar a linha do horizonte na imagem e assim remover a parcela referente ao céu da zona de busca por óleo. Porém, dada a característica luminosidade, uma filtragem de valores altos de V, Y e L também têm o efeito desejado. Filtros adicionais podem ser sempre utilizados para filtrar outros componentes conhecidos, tais como um barco amarelo. A principal motivação deste pré-processamento é tornar a imagem mais tratável e própria para a análise, facilitando os passos seguintes de detecção.

No algoritmo proposto, esta detecção consiste de duas etapas. A primeira se baseia na segmentação da imagem, ou seja, na realização de um agrupamento automático dos elementos que possuem alto grau de semelhança e que, uma vez agrupados, estarão associados ao centroide característico de sua classe no espaço de cor de interesse. O segundo passo consiste em, dadas as características gerais dos centroides representativos, classificar os *clusters* como "óleo" ou "mar". Note que esta classificação pode ser ainda mais específica, envolvendo, por exemplo, uma distinção entre: mar caribenho, oceano, vegetação, recife, coral, rochas/areia, óleo emulsificado, óleo filme, etc. Neste caso, a classificação se baseará na pertinência (conceitos de *soft* e *hard-clustering*) dos grupos segmentados numa imagem em análise a esses centroides representativos de classe previamente calibrados/identificados. A Figura 3 exemplifica no plano $Cb \times Cr$ algumas dessas possíveis distribuições de cores características, suas dispersões e seus centroides. A classificação em alguns grupos se mostra viável dessa forma, porém, por se basear apenas em cores médias dos segmentos, confusões típicas podem ocorrer - por exemplo, rocha e areia. Para uma classificação mais

robusta, torna-se interessante a adição de outras análises por visão, como textura.

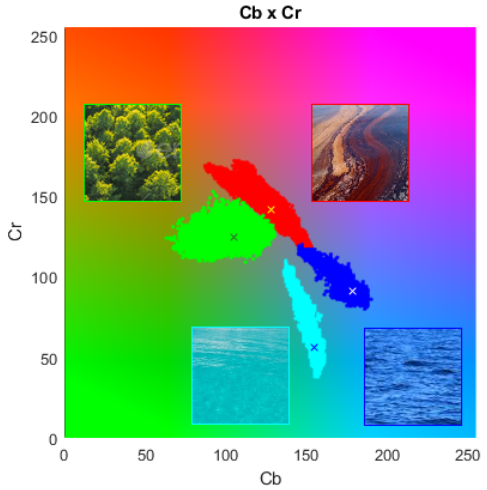


Figura 3. Plano $Cb \times Cr$, distribuições de cores e *clusters* característicos.

2.2 Segmentação e Classificação.

O primeiro passo do algoritmo envolve a segmentação da imagem em grupos/*clusters* de elementos semelhantes, isto é, dadas as ocorrências no espaço de cores considerado (planos $Cb \times Cr$ e $a^* \times b^*$), esta etapa fornecerá a forma de particionar tais dados, num esquema de aprendizado não-supervisionado. Portanto, há um ajuste automático de k gaussianas aos dados e estes passam a ser modelados pelos centroides de seus respectivos grupos. Este algoritmo, denominado *k-means*, é uma forma iterativa/alternada de se encontrar o melhor grupamento (Bishop, 2006). A Figura 4(a) apresenta uma imagem previamente tratada com *blur* e com áreas de rejeição já definidas (e.g. altos valores de H). Nas Figuras 4(c), 4(b) e 4(d) são apresentados o histograma em H, as ocorrências no plano $Cb \times Cr$ e as ocorrências no plano $a^* \times b^*$, respectivamente - já com gaussianas ajustadas. Note que seus centroides definem a cor média do segmento.

A segmentação por *k-means* ajusta a melhor - ou quase ótima - forma de se dividir uma imagem em k grupos, porém ainda é necessário escolher o valor de k. Para que essa tarefa seja realizada de forma adequada e automática, foi utilizado o método heurístico do cotovelo, do inglês *elbow method*. Tal método se baseia em múltiplos testes para diferentes valores de k e escolha de um valor em que o aumento do número de *clusters* começa a ter pouco impacto na variância intra-grupo, isto é, quando não há grande benefício relativo em se particionar em mais grupos. Tal ponto é evidenciado como um "cotovelo" no gráfico e pode ser obtido traçando-se uma reta do ponto inicial ao final e checando-se o ponto de maior distância à reta. Por ser um método heurístico, apresenta suas limitações e, por vezes, se mostra ambíguo.

Para o algoritmo proposto, o método do cotovelo é avaliado para cinco diferentes pontos iniciais, compreendidos entre o ponto inicial original e um ponto acima do projetado com a reta definida pelos segundo e terceiro pontos (25% do intervalo acima). Há outras formas de se chegar ao valor

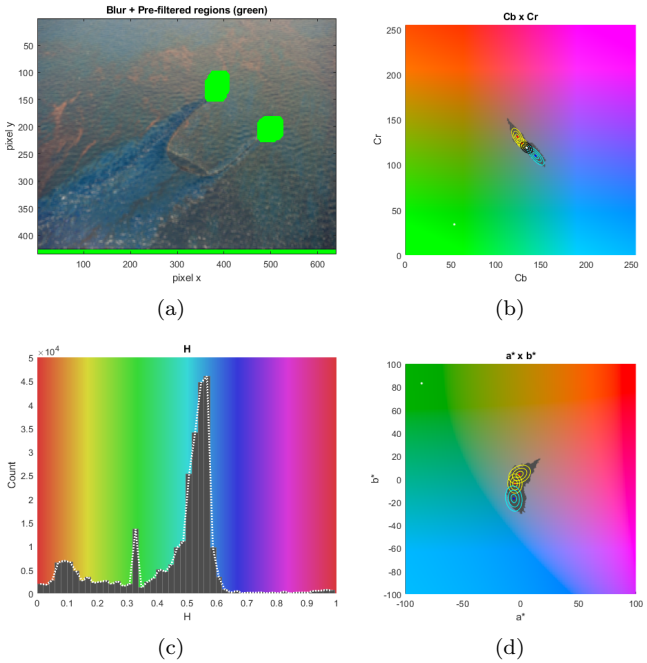
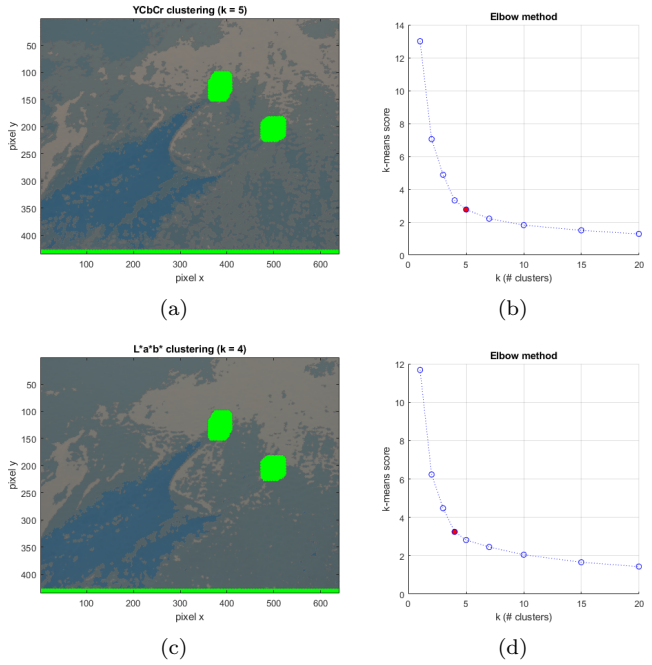


Figura 4. Foto tratada (a), plano $Cb \times Cr$ (b), histograma de H (c) e plano $a^* \times b^*$ (d).

de k, como o método *silhouette*, em que se avalia métricas tanto de coerência intra-*cluster* como de distanciamento inter-*cluster*. Tais métodos não são excludentes e podem ser usados em conjunto. As Figuras 5(a) e 5(c) apresentam os resultados da segmentação para os dois espaços de interesse, sendo os segmentos representados pela média de cor, isto é, pela cor do centroide da classe. As Figuras 5(b) e 5(d) evidenciam a escolha de k por *elbow*. Note que o número de segmentos ajustados foi diferente para os dois distintos planos. Em $Cb \times Cr$, quatro segmentos na região de busca e, para $a^* \times b^*$, três segmentos válidos.

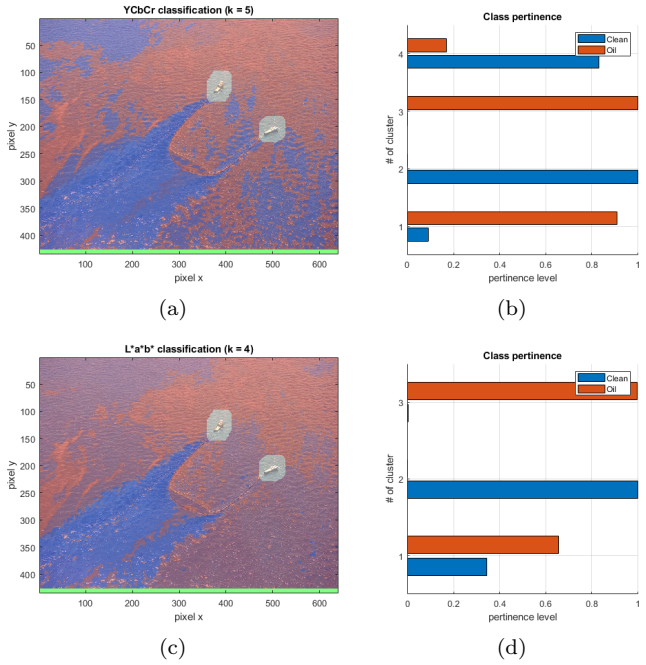
O próximo passo do algoritmo consiste na etapa de classificação, sendo, portanto, o ponto mais crucial do método proposto. Neste ponto, assume-se uma segmentação adequada proveniente da camada anterior (número adequado de grupos e um bom ajuste de gaussianas) e cada classe/segmento é submetido a um filtro convencional de cor, sendo definidos limiares específicos para a aplicação. Assim, filtram-se possíveis segmentos e tornam-se candidatos a "óleo" aqueles cujos centroides pertencem ao envelope de cor definido, aproximadamente conhecido e adequado ao que se busca (e.g. óleo emulsificado, óleo filme, etc). A definição deste envelope (parâmetro do algoritmo) é igualmente crucial, podendo ser realizada de forma absoluta em um espaço de cor tradicional mas podendo também utilizar análises relativas - por exemplo, em uma análise de tonalidade dominante, normalizam-se as componentes e se estabelece um limiar e/ou limiares condicionais para estas variáveis normalizadas.

Tendo em vista o problema de detecção de óleo, buscam-se segmentos mais avermelhados e/ou acinzentados/prateados num ambiente marítimo. Assim, sempre é esperada uma tonalidade azul mas com presença ou até predominância de tons vermelhos quando há óleo. Por outro lado, segmentos em que tonalidades verdes se mostram muito dominantes estão comumente associados a vegetação e re-



(a) (b) (c) (d)

Figura 5. Segmentação (a)(c) e gráficos para escolha de k ótimo (b)(d).



(a) (b) (c) (d)

Figura 6. Classificação (a)(c) e gráficos de pertinência às classes (b)(d).

cifes, podendo ser geralmente descartados como candidatos a "óleo". Nesta etapa é escolhida também, dentre todas as gaussianas ajustadas, aquela que será referência para cor do "mar"/"oceano", escolhendo-se naturalmente o segmento cujo centroide está mais próximo da cor azul - nos planos $Cb \times Cr$ e $a^* \times b^*$, são usados os vértices extremos. Nota-se que, caso haja apenas um segmento identificado na parte válida da imagem, avalia-se a predominância de azul ou vermelho, e caracteriza-se como "mar" ou "óleo", adequadamente.

Uma vez definidos todos os candidatos a "óleo" e identificado um segmento de referência/representativo do "oceano", definem-se suas pertinências de classe. Dentre os possíveis candidatos, o mais distante da referência pré-estabelecida como "mar" é instantaneamente classificado como "óleo". Todos os segmentos intermediários (e possíveis candidatos) têm sua pertinência de classe avaliada, num esquema clássico de *soft-clustering*. A equação 1 evidencia o cálculo geral da pertinência dessas gaussianas intermediárias às gaussianas de referência (Theodoridis and Koutroumbas, 2010):

$$P(C_i|x_k) = \frac{e^{-d_M(x_k,\mu_i)}}{\sum_j e^{-d_M(x_k,\mu_j)}} \tag{1}$$

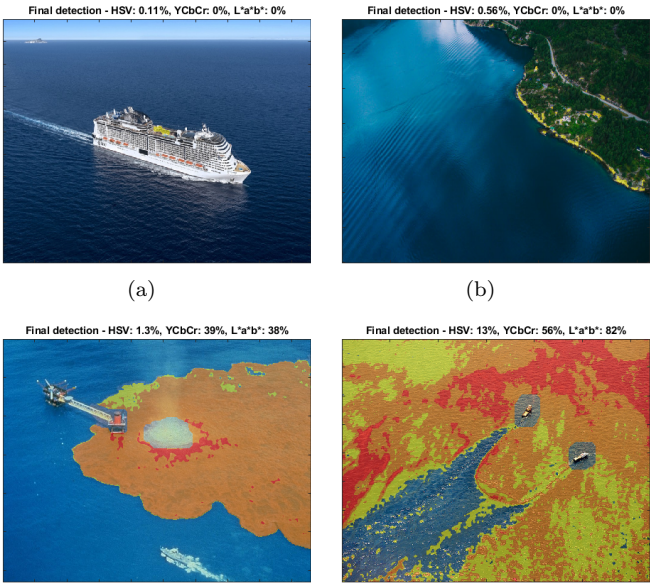
em que x_k são os centroides dos segmentos que se deseja avaliar, μ_i é o centroide da classe C_i (referência) e $d_M(x_k, \mu_i)$ é a distância de *Mahalanobis* de x_k ao centroide μ_i, com covariância Σ_i. Cada gaussiana é finalmente classificada como pertencente ao grupo que melhor a explica, isto é, ao qual sua probabilidade de pertencer é maior (classificador linear *softmax*). A Figura 6 apresenta visualizações das classificações obtidas com as análises em $YCbCr$ e $L^*a^*b^*$, assim como os gráficos indicadores das pertinências. Note que a segmentação no plano $a^* \times b^*$ foi feita com um grupamento a menos, acarretando maior dúvida/equilíbrio de pertinência para um dos segmentos em questão - sendo este mesmo segmento divido em dois pela análise em $YCbCr$, um classificado como "mar" e outro classificado como "óleo".

2.3 Algoritmo Final de Detecção

Como visto, o método proposto apresenta um esquema de classificação condicionado/relativo ao segmento mais marcante de óleo identificado na imagem. Além disso, ao se basear em três análises independentes e complementares (limiar de cor em H + segmentação/classificação em $Cb \times Cr$ + segmentação/classificação em $a^* \times b^*$) também torna-se necessário a adoção de algum critério para combinar as detecções finais de cada camada. Portanto, é preciso ponderar para cada aplicação os pesos de falsos positivos e falsos negativos, determinando assim se um sinal de alerta deve ser emitido com a detecção em ao menos um ou na maioria dos canais. Note que, para vazamentos de óleo, falsos negativos têm impactos ambientais grandes e são altamente indesejados. Assim, sugere-se a forma mais sensível, uma vez que uma sensibilidade excessiva pode ser contornada com a persistência de detecções durante uma janela de tempo. O algoritmo proposto é, por fim, avaliado para as 228 fotos do *dataset*, apresentando uma acurácia de 77,19%. A Figura 7 apresenta resultados de detecção para alguns casos de interesse, com e sem derramamento de óleo.

3. REDES NEURAIS CONVOLUCIONAIS

Tarefas de classificação, localização, detecção de objetos e segmentação de instâncias podem ser realizadas por meio da utilização de redes neurais artificiais. No contexto de identificação de padrões em imagens, a utilização de Redes Neurais Convolucionais, do inglês *Convolutional Neural Network* (CNN), são mais aplicadas devido a sua



(a) (b) (c) (d)

Figura 7. Detecções finais: sem derramamento de óleo (a)(b) e diferentes níveis de derramamento (c)(d).

capacidade de extrair características em um *dataset* de imagens (Haykin, 2009). Portanto, supõe-se que as CNN's possuam a capacidade de encontrar um padrão de óleo nas imagens e fazer as suas devidas classificações. Nesta análise, os algoritmos foram desenvolvidos em Python utilizando-se as bibliotecas do Pytorch (Paszke et al., 2019).

Considerando que o *dataset* obtido apresenta limitações, devido a presença de poucas imagens, técnicas de transferência de aprendizado (Torrey and Shavlik, 2010), do inglês *Transfer Learning* (TL), foram utilizadas. Para os testes foram selecionadas as redes que apresentam *Top-5 Error* inferior a 10%: DenseNet, Inception V3, MnasNet, MobileNet V2, ResNet, ResNeXt, VGG e Wide ResNet (PyTorch, 2020).

Antes do treinamento das redes, foi realizado o aumento do conjunto de dados (Goodfellow et al., 2016), do inglês *dataset augmentation*, nas imagens, realizando-se as seguintes operações:

- ajuste de 0,5 na saturação e matiz;
- rotação horizontal com probabilidade de 60% de acontecer;
- rotação vertical com probabilidade de 60% de acontecer;
- distorção de 0,3 na escala com a probabilidade de 30% de acontecer;
- rotação aleatória no eixo de -45° a +45°.

Em seguida, os treinamentos e testes das redes foram realizados, os quais foram divididos em 75% e 25% do *dataset*, respectivamente. Para cada CNN foram executadas 25 épocas e os testes foram repetidos três vezes. A Tabela 1 mostra a média dos resultados obtidos.

Conforme pode ser verificado na Tabela 1 a maioria das redes apresentou precisão acima de 87%, o que demonstra

Tabela 1. Resultado da precisão média para cada modelo.

Modelo	**Média da melhor precisão obtida nos testes [%]**
DenseNet161	$91,37 \pm 3,45$
Inception v3	$87,35 \pm 1,99$
MNasNet1.0	$73,56 \pm 8,15$
MobileNet v2	$92,52 \pm 1,99$
ResNet34	$91,37 \pm 1,73$
ResNext 50 32x4d	$67,24 \pm 21,53$
VGG 16	$90,22 \pm 3,59$
Wide ResNet50.2	$89,08 \pm 3,59$

que a hipótese de utilizar CNN's para detecção de óleo em superfícies de água é extramente válida.

Os modelos MobileNet v2, ResNet34 e DenseNet foram selecionados para as próxima etapa de testes por apresentarem os melhores resultados.

3.1 Validação Cruzada

Em seguida, as redes foram submetidas à validação cruzada através do método conhecido por *K-Fold Cross Validation* (Rodriguez et al., 2010). Os testes foram realizados com $k = 4$, para os quais foram executadas 25 épocas, repetindo-os três vezes para cada *fold* e modelo, totalizando 12 testes por CNN. O objetivo desta análise foi validar a distribuição de imagens dentro do *dataset* e verificar qual das redes apresenta melhor desempenho. A Tabela 2 apresenta os resultados obtidos nesta batelada de testes.

Tabela 2. Resultados médios dos *folds* testados

Modelo	**Média dos k-fold da melhor precisão obtida [%]**
DenseNet161	$92,24 \pm 2,72$
MobileNet v2	$88,91 \pm 6,24$
ResNet34	$91,52 \pm 3,94$

Conforme pode ser verificado na Tabela 2, a rede DenseNet161 apresentou resultado superior às demais, obtendo a maior média e o menor desvio.

3.2 Mapa de Ativação de Classe

Uma vez que o melhor modelo foi selecionado, DenseNet161, foi realizado o Mapa de Ativação de Classes (Zhou et al., 2016), do inglês *Class Activation Mapping* (CAM). O resultado do CAM é um mapa de calor sobre as imagens testadas, o qual demonstra o que foi mais relevante na imagem na classificação. A Figura 8 demonstra duas imagens categorizadas como *clean* e duas categorizadas como *oil spill*.

Analisando a Figura 8 é possível notar que imagens categorizadas como *clean* apresentam maiores áreas quentes, tal fato pode ser explicado pela rede ter que fazer buscas em diferentes regiões da imagem para ter certeza que não existe a presença de óleo, enquanto para as imagens categorizadas como *oil spill*, pelo fato do óleo estar presente, apenas a região que o limita apresenta cores quentes.

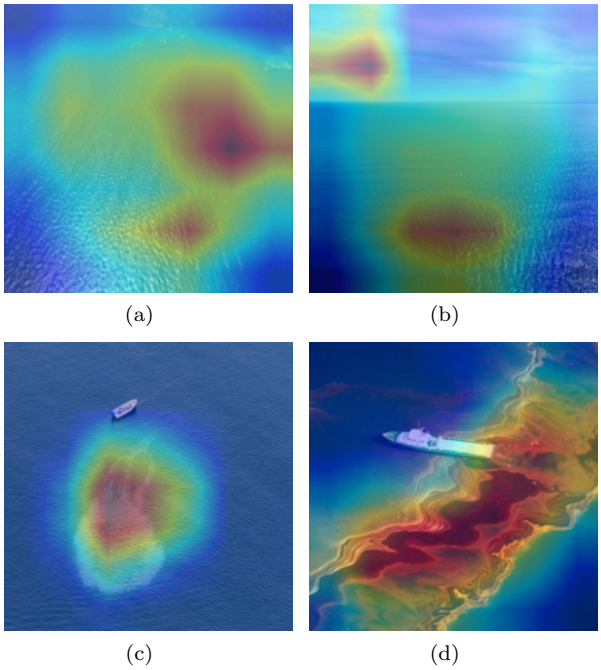


Figura 8. CAM usando o DenseNet161 aplicado às imagens do *dataset*: exemplos sem (a)(b) e com (c)(d) a presença de óleo na superfície da água.

4. COMPARAÇÃO ENTRE MÉTODOS

Após ajustar parâmetros/limiares de cor para o algoritmo por visão e determinar os melhores modelos para detecção de óleo utilizando Redes Neurais Convolucionais, um teste comparativo entre os métodos foi realizado. Neste teste, 75% do *dataset* foi utilizado para treinamento das redes e os 25% restantes, que não foram utilizadas em nenhum momento, são testados. Para cada metodologia deve-se obter a matriz confusão (Japkowicz and Shah, 2011), evidenciando os pares entrada-alvo dos classificadores propostos e facilitando a visualização de indicadores de desempenho, como acurácia, *precision* e *recall*.

4.1 Resultados com Visão Computacional

O algoritmo desenvolvido é avaliado para as 58 fotos do grupo de teste definido, apresentando uma acurácia de 81,04%. A Tabela 3 mostra a matriz confusão obtida.

Tabela 3. Matriz Confusão dos testes realizados com o algoritmo por visão.

		Predito	
	Categorias	Clean	Oil Spill
Real	Clean	23	6
	Oil Spill	5	24

Neste caso, observam-se pouco desbalanço e viés nas classificações. Porém, como discutido ao final da seção 2.3, uma taxa de *recall* de 82,75% pode ser considerada pouco satisfatória para a aplicação de interesse, justificando contínuas melhoria e exploração do algoritmo e de suas assunções.

4.2 Resultados com Redes Neurais Convolucionais

A Tabela 4 apresenta a matriz confusão obtida pela rede DenseNet161 no *dataset* do teste.

Tabela 4. Matriz Confusão dos testes realizados com a rede DenseNet161 treinada.

		Predito	
	Categorias	Clean	Oil Spill
Real	Clean	29	0
	Oil Spill	2	27

Tais resultados demonstram que a rede treinada apresentou desempenho similar aos obtidos durante os treinamentos, com acurácia de 96,55%. Obteve-se *recall* de 93,10% devido à presença de apenas dois falsos negativos.

5. CONCLUSÕES

Neste trabalho, foram propostos dois esquemas para a detecção automática de derramamentos de óleo usando sensores locais (câmeras térmicas ou RGB): um baseado em Visão Computacional e outro em aprendizado de máquina. Ao longo do desenvolvimento deste trabalho, foi observado que uma abordagem a partir de Visão Computacional demonstrou ser analítica pois, a partir dos padrões vistos/percebidos nas imagens, foi desenvolvido um algoritmo que fosse capaz de fazer a classificação. As principais limitações encontradas para a aplicação destes algoritmos recaem na dificuldade de se estabelecer limiares de cores suficientemente gerais para tratar imagens muito diversas, obtidas em diferentes horários do dia, sob distintas condições climáticas e de luminosidade, etc. Paralelamente, foi realizada uma abordagem por Redes Neurais Convolucionais, que podem ser consideradas empíricas, pois a partir de um *dataset* pré-classificado, os modelos CNN's foram capazes de extrair as principais características da imagens, generalizá-las e na sequência efetuar a classificação. As limitações das técnicas de aprendizado de máquina estão relacionadas à dificuldade de obter *datasets* representativos, garantir a capacidade de generalização e, também, evitar *overfitting* dos modelos.

Ainda assim, é importante notar que, uma vez conhecidas as limitações do algoritmos, medidas práticas e alguns cuidados podem ser adotados para garantir um melhor desempenho (e.g. posicionamento adequado de câmera visando reduzir interferências de alguns elementos externos), simulação de um *dataset* adequado, etc.

Comparações de operações de classificação já foram abordadas na literatura (He et al., 2016) e (Krizhevsky et al., 2012) e demonstram que em condições igualitárias, as CNN's têm apresentado desempenhos superiores às metodologias clássicas desde 2012 e até mesmo melhor do que humanos a partir de 2015, o que, não necessariamente, afirmam que as CNN's obterão sempre os melhores resultados. Neste trabalho, a comparação dos resultados foi feita a partir de um *dataset* com 58 imagens, conforme Seção 4, e, como esperado, as Redes Neurais Convolucionais apresentaram o melhor desempenho.

5.1 Trabalhos futuros

Este trabalho apresenta a investigação inicial de um projeto robótico que visa ser capaz de detectar óleo em superfícies de água. Como demonstrado, com o *dataset* atual foi possível obter resultados de precisão na casa dos 90%. Portanto, como próximas etapas sugere-se: i) aplicação dos algoritmos desenvolvidos em testes de campo; ii) retroalimentação do *dataset* a partir de aquisição de imagens (proveniente dos testes); iii) utilização de câmera termal para análise em outro espectro e consequente utilização em paralelo com as imagens RGB; (iv) fusão dos algoritmos propostos que utilizam sensores locais (câmeras térmicas e RGB) com técnicas que utilizam sensores remotos, como imagens de radar (SAR), visando obter um esquema robusto para detectar derramamentos de óleo.

REFERÊNCIAS

Bishop, C.M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Brekke, C. and Solberg, A.H. (2005). Oil spill detection by satellite remote sensing. *Remote sensing of environment*, 95(1), 1–13.

Del Frate, F., Petrocchi, A., Lichtenegger, J., and Calabresi, G. (2000). Neural networks for oil spill detection using ers-sar data. *IEEE Transactions on geoscience and remote sensing*, 38(5), 2282–2287.

Fingas, M.F. and Brown, C.E. (1997). Review of oil spill remote sensing. *Spill Science & Technology Bulletin*, 4(4), 199–208.

Fiscella, B., Giancaspro, A., Nirchio, F., Pavese, P., and Trivero, P. (2000). Oil spill detection using marine sar images. *International Journal of Remote Sensing*, 21(18), 3561–3566.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.

Haykin, S.S. (2009). *Neural networks and learning machines*, volume 3. Pearson Upper Saddle River.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.

Jang, H.J. and Nam, J.H. (2020). Algorithm to Estimate Oil Spill Area Using Digital Properties of Image. *Journal of Ocean Engineering and Technology*, 34(1), 46–54.

Japkowicz, N. and Shah, M. (2011). *Evaluating learning algorithms: a classification perspective*. Cambridge University Press.

Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. In F. Pereira, C.J.C. Burges, L. Bottou, and K.Q. Weinberger (eds.), *Advances in Neural Information Processing Systems 25*, 1097–1105. Curran Associates, Inc. URL `http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf`.

Kubat, M., Holte, R.C., and Matwin, S. (1998). Machine learning for the detection of oil spills in satellite radar images. *Machine learning*, 30(2-3), 195–215.

Lammoglia, T. (2010). *Sensoriamento Remoto Ótico & Exploração Petrolífera Onshore e Offshore*. Ph.D. thesis, Universidade Estadual de Campinas, Instituto de Geociências, Campinas, SP.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems 32*, 8024–8035. Curran Associates, Inc. URL `http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf`.

Pérez-Cadahía, B., Lafuente, A., Cabaleiro, T., Pásaro, E., Méndez, J., and Laffon, B. (2007). Initial study on the effects of prestige oil on human health. *Environment international*, 33(2), 176–185.

Petrobras (2020). Exploração e produção de petróleo e gás: Pré-sal. URL `https://petrobras.com.br/pt/nossas-atividades/areas-de-atuacao/exploracao-e-producao-de-petroleo-e-gas/pre-sal/`.

Prince, S. (2012). *Computer Vision: Models Learning and Inference*. Cambridge University Press.

PyTorch (2020). Torchvision.models. URL `https://pytorch.org/docs/stable/torchvision/models.html`.

Rodriguez, J.D., Perez, A., and Lozano, J.A. (2010). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3), 569–575.

Solberg, A.H.S., Storvik, G., Solberg, R., and Volden, E. (1999). Automatic detection of oil spills in ers sar images. *IEEE Transactions on Geoscience and Remote Sensing*, 37(4), 1916–1924.

Theodoridis, S. and Koutroumbas, K. (2010). *Pattern recognition and Matlab intro*. Academic Press.

Torrey, L. and Shavlik, J. (2010). Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, 242–264. IGI Global.

Zhou, B., Khosla, A., A., L., Oliva, A., and Torralba, A. (2016). Learning Deep Features for Discriminative Localization. *CVPR*.