

Detecção de presença de dragas de garimpo em rios da Amazônia via classificação de imagens SAR com *Few-Shot Learning*

Caio Henrique De Felice¹, Elcio Hideiti Shiguemori², Rafael Lemos Paes³, Daniel Ferreira Martins⁴

¹Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos/SP - Brasil

²Instituto de Estudos Avançados (IEAv), São José dos Campos/SP - Brasil

³Estado Maior da Aeronáutica (EMAER), Brasília/DF - Brasil

⁴Universidade Federal de São Paulo (UNIFESP), São Paulo/SP - Brasil

Resumo—A mineração ilegal de ouro em regiões remotas da Amazônia brasileira representa grandes ameaças ambientais e de segurança. O desenvolvimento de sistemas para detectar dragas de mineração é dificultado pela escassez de dados rotulados, seja pelo isolamento da região ou por condições meteorológicas. Apesar disso, poucos estudos exploraram abordagens automatizadas usando imagens SAR sob regimes de poucos dados para encontrar dragas de garimpo. Este trabalho apresenta um algoritmo automatizado para classificar imagens contendo dragas em imagens SAR da constelação ICEYE. O método inicia recortando imagens GRD de cena completa em recortes menores, que são classificados por uma Rede de Protótipos treinada via *Few-Shot Learning*. Embora o método não determine a posição dos alvos, reduz em aproximadamente 98% o número de imagens a serem analisadas, mantendo um *recall* com perdas máximas de 13% das imagens com alvos no pior cenário. Apesar da incapacidade atual em distinguir dragas de outras pequenas embarcações, o método pode ser utilizado para reduzir o escopo de busca para monitoramento em larga escala de áreas vulneráveis à mineração ilegal.

Palavras-Chave—Imagens SAR, Inteligência Artificial, Garimpo Ilegal

I. INTRODUÇÃO

A mineração de ouro no Brasil é uma atividade antiga [1] e já fez o país se destacar como um grande produtor de ouro no cenário mundial [2]. No entanto, a mineração ilegal na região amazônica, especialmente próxima ou até mesmo dentro dos rios, resulta em consequências catastróficas, como a destruição da biodiversidade e a disseminação de doenças entre povos indígenas [3], [4].

A mineração em rios é geralmente realizada por meio de dragas — plataformas flutuantes que extraem sedimentos do leito do rio para recuperar ouro ou outros metais preciosos, descartando o material restante. Substâncias tóxicas, como o mercúrio (utilizado na amalgamação), são frequentemente utilizadas nesse processo [5], [1], [6], [7].

Embarcações como as dragas, (denominadas como "alvos" neste trabalho) tendem a ter tamanho reduzido e grande mobilidade devido ao nível de informalidade e precariedade da atividade, além de não utilizarem dispositivos de localização em tempo real, como sistemas de navegação por satélite (GNSS) ou o Sistema de Identificação Automática

(AIS), dificultando sua detecção por sensoriamento remoto e consequente monitoramento e controle de suas atividades [8], [5], [1].

Ainda, a própria região amazônica apresenta cobertura de nuvens densa e frequente [1], fazendo com que a obtenção de imagens ópticas sem nuvens seja desafiadora em regiões como o estado de Roraima e o norte do Amazonas [8]. Nesse contexto, o radar de abertura sintética (SAR) surge como uma alternativa viável, por não depender das condições meteorológicas [1], [6], [7], [9], [10], [11].

Conforme já apontado por [1], [6], há poucos exemplos de uso de imagens SAR para aplicações relacionadas às dragas fluviais. [9], [7] utilizaram metodologias baseadas no contraste de retroespalhamento entre superfícies de água lisa e as estruturas metálicas das dragas; [9] empregou uma análise temporal para localizar os alvos no rio Ayeyarwady (Mianmar). De forma semelhante, o sistema implementado por [7] possibilitou a detecção quase em tempo real de balsas de mineração de areia com precisão satisfatória no rio Kadéi (Sangha), na República Centro-Africana. [6] avaliou três abordagens semi-automatizadas para detecção de alvos no rio Madeira, no Brasil (Search for Unidentified Maritime Objects - SUMO - baseado no algoritmo de Alarme Falso Constante - CFAR, técnica de limiarização global e método local usando filtro passa-alta e agrupamento de regiões). Em seu estudo, [1] aplicou um algoritmo baseado em CFAR combinado com avaliação geométrica no rio Madeira. Todos os estudos referenciados utilizaram imagens SAR de banda C do Sentinel-1 e observaram que o canal VV apresenta maior retroespalhamento em comparação ao canal VH; [1] também relatou que essa polarização é mais propensa a falsos positivos (FP). [10] utilizou imagens do ENVISAT ASAR para detectar dragas de areia no rio Yangtzé com base no contraste de retroespalhamento entre a água e as embarcações. O método combinava limiarização baseada em histograma, crescimento de regiões e filtragem morfológica. Até onde se sabe, ainda não foram reportadas abordagens baseadas em redes neurais convolucionais (CNN) para detecção de dragas em imagens SAR.

Contrariamente, a utilização de imagens SAR pode ser observada na detecção de navios em ambientes marítimos, com métodos baseados em aprendizado profundo (DL) apresentando resultados significativamente superiores aos das técnicas baseadas em CFAR [12]. Aplicações como [13], [14], [15]

Caio Henrique De Felice, caiochf@ita.br; Elcio Hideiti Shiguemori, elcio@lac.inpe.br; Rafael Lemos Paes, rlpaes@gmail.com; Daniel Ferreira Martins, danielfermartins@gmail.com.

também reforçam o potencial das redes neurais convolucionais na detecção de navios por DL. No entanto, o desempenho desses métodos baseados em DL depende da quantidade de dados rotulados disponíveis [12], [1] e, como a coleta de grandes conjuntos de dados anotados costuma ser inviável, abordagens alternativas como o aprendizado com poucos exemplos (Few-Shot Learning - FSL) têm sido exploradas [16].

Este trabalho propõe a classificação de imagens SAR da constelação ICEYE de regiões da floresta amazônica brasileira suscetíveis ao garimpo ilegal utilizando um regime de escassez de dados baseado em FSL utilizando uma CNN. Na segunda seção são apresentados os dados, sua interação com o sensor SAR e a arquitetura de rede neural utilizada.

Os resultados são apresentados na terceira seção e as discussões acerca deles na seção IV. Por fim, são expostas conclusões e sugestões de trabalhos futuros.

II. MATERIAIS E MÉTODOS

A. Áreas de Estudo e Conjunto de Dados

Devido à indisponibilidade de banco de dados específico para a tarefa aqui proposta, foi necessária a criação de um banco de dados próprio, assim sendo, o conjunto de dados consiste de imagens provenientes dos satélites da constelação ICEYE. Esses sensores operam na banda X e utilizam antenas *phased-array* [17]. A utilização dessa banda possibilita emprego de imagens com melhores resoluções, facilitando distinção de alvos menores [7], [18], entretanto, apresenta pouca penetração em regiões de vegetação, ficando mais suscetível ao seu efeito [18].

A resolução espacial das imagens pode variar entre 15 metros e 0,25 metros, dependendo da geração do satélite, da estratégia de aquisição (*Spotlight*, *Stripmap* ou *ScanSAR*) e do modo de imageamento. Apenas a polarização VV está disponível nas imagens da ICEYE [17], [7]. Essa seleção de sensor é justificada pelos seguintes fatores: o formato e o tamanho das embarcações são mais nítidos nos dados ICEYE [7], a polarização VV oferece maior retroespalhamento, o que favorece a detecção [1], além da disponibilidade das imagens e de seus rótulos por si só.

No total foram obtidas 21 imagens de cena completa com resolução espacial de 3 metros e 2 imagens com 1 metro. Os valores de pixel representam valores de amplitude e todas são do formato *Ground Range Detected* (GRD). As áreas de estudo, o Vale do Javari e a Estação Ecológica Juami-Japurá, estão ambas localizadas ao sul da região conhecida como “Cabeça de Cachorro”, respectivamente no noroeste e oeste do estado do Amazonas. Além de serem regiões de fronteira isoladas, ambas possuem histórico de atividades de mineração [2] e continuam sendo alvo de interesse devido ao elevado número de permissões de lavra garimpeira (PLG) [1], [19]. As imagens foram adquiridas entre 6 de abril e 14 de setembro de 2024, sem um intervalo regular de aquisição. Nenhuma das imagens contempla áreas urbanas e cinco rios principais com atividade garimpeira foram imageados: Japurá, Puruê, Javari, Boia e Juami.

Um total de 220 instâncias rotuladas (dragas previamente detectadas) estavam disponíveis, e os rótulos foram obtidos a partir do seguinte processo: as imagens GRD foram coletadas, um modelo YOLOv4 baseado no trabalho de [20] foi utilizado

para detecções preliminares, e a validação das detecções foi feita por um especialista em foto inteligência, gerando os rótulos finais. Um exemplo de detecção validada por esse processo é exibido na Fig. 1, que foi o único exemplo cuja correspondência óptica foi possível devido à indisponibilidade de imagens para as demais regiões de ou, quando disponíveis, apresentavam uma cobertura de nuvens que nunca foi inferior a 80%.

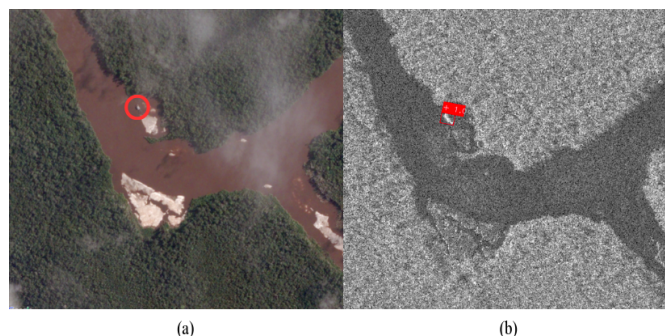


Fig. 1: Imagem óptica (a) e sua correspondência em SAR (b). O alvo é destacado na imagem (a) e sua detecção através do modelo YOLOv4 é exibida em (b).

B. Interação entre Dragas e Radar de Abertura Sintética (SAR)

As propriedades dielétricas do meio afetam sua interação com micro-ondas [18]. Por exemplo, embarcações são objetos artificiais com ângulos retos, capazes de gerar sinais de retroespalhamento intensos, conhecidos como (*double bounce*) [7], [6], [1]. O ângulo de incidência, ou ângulo de interação, também influencia a intensidade do retroespalhamento [7], [6], [18]. Todos esses fatores resultam em diferentes representações de valor de pixel e quando vistos em contexto, geram representações específicas. No caso da região de estudo desta pesquisa, quatro principais elementos (água, floresta ou terreno, pequenas ilhas ou pilhas de sedimentos e dragas) são observados e estão representados na Fig. 2. Os mesmos elementos também se destacaram em trabalhos como [6], [7], [1], [9].

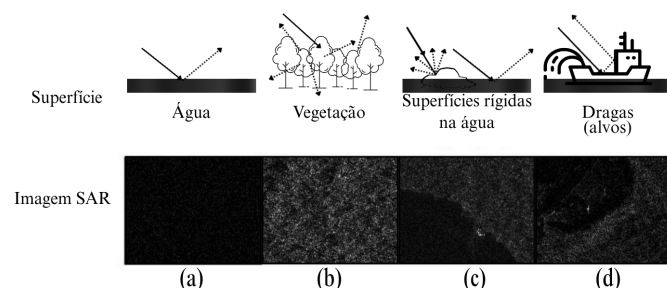


Fig. 2: Diferenças de retroespalhamento para cada alvo principal da área de estudo.

A Fig. 2 apresenta as interações das micro-ondas com: (a) superfícies planas ocasionando retroespalhamento especular, (b) e (c) superfícies rígidas com retroespalhamento difuso e (d) objetos artificiais como dragas com efeito *double bounce*. A representação desses fenômenos em imagens (pixels) é exibida abaixo de cada um deles.

C. Aprendizado com Poucos Exemplos e Redes Prototípicas

O aprendizado por meta-aprendizado (Meta-Learning, MTL), frequentemente descrito como "aprender a aprender", difere do aprendizado de máquina tradicional por seu nível de adaptação [21]. Um dos principais desafios do MTL é treinar modelos de aprendizado profundo utilizando apenas poucos exemplos [22]. Um exemplo típico é a tarefa de classificação "*N*-way *K*-shot", na qual um classificador é adaptado para "*N*" classes nunca vistas com apenas "*K*" exemplos por classe, com base em sua experiência prévia em tarefas similares com grandes volumes de dados [23], [16], [22]. O FSL baseado em MTL pode ser classificado em três grupos: baseado em otimização, baseado em modelo e baseado em métrica [16], [21].

O aprendizado baseado em métrica depende da comparação de vetores de características (*feature vectors*) extraídos de imagens base para determinar sua similaridade, o que permite um processamento de dados rápido e eficaz [21], [16]. Um exemplo de rede baseada nessa abordagem é a Rede de Protótipos (Prototypical Network - PN) [16], [23].

Redes dessa natureza representam uma abordagem simples para FSL [23]: aprendem um espaço métrico no qual a classificação é realizada com base em representações prototípicas de cada classe (médias dos vetores de *embedding*), utilizando uma métrica de distância. Em resumo, após extrair as características de uma imagem com uma CNN e representá-las em um vetor, por exemplo, compara-se a distância entre ele e os vetores de suporte obtidos através de outras imagens, a menor distância entre eles determinará a classe daquela imagem [23], [22], [16].

D. EfficientNetV2

A família de redes neurais convolucionais EfficientNetV2 foi projetada com foco em eficiência de treinamento e desempenho escalável em tarefas de visão computacional. Diferentemente da versão anterior (EfficientNet), que apresentava limitações como uso ineficiente de imagens grandes e convoluções profundas em estágios iniciais, a EfficientNetV2 propõe otimizações estruturais que reduzem significativamente o tempo de treinamento [24], [25]. Primeiramente, essa arquitetura faz uso de imagens menores que sua versão anterior, o que resulta em um acréscimo de velocidade de treinamento e de acurácia [25]. Além disso, destaca-se a introdução das camadas *Fused-MBConv*, que fazem uma única convolução 3×3 padrão, melhor aproveitando os recursos computacionais, trazendo assim ganhos especialmente nos estágios iniciais da rede. Por fim, a estratégia de *progressive training* adotada permite iniciar o treinamento com imagens menores e aumentos de dados simples, aumentando gradualmente a complexidade combinada com a adição de regularização, o que contribui para uma melhor generalização, acurácia e aceleração do processo de aprendizado.

E. Fluxo do Algoritmo

A Fig. 3 ilustra o fluxo de processamento adotado. Basicamente, as imagens GRD de cena completa são recortadas em blocos menores, seguindo a abordagem utilizada em [13]. Esses blocos são classificados entre quatro classes principais conforme Fig. 2 utilizando-se de uma PN refinada especificamente para essa tarefa.

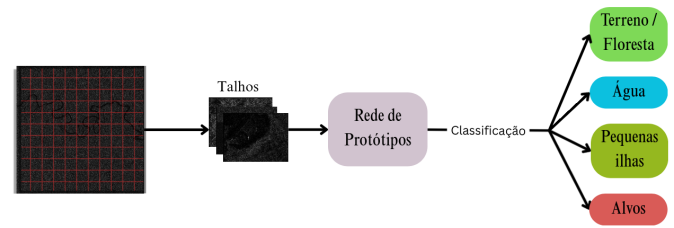


Fig. 3: Fluxo de processamento.

F. Configuração Experimental

Experimentos foram conduzidos utilizando PN para avaliar o desempenho da classificação sob diferentes configurações. As imagens GRD foram recortadas em talhos menores de 200×200 pixels, e os experimentos realizados com diferentes valores de *K*-shot (5 e 10, representando respectivamente menor e maior suporte de dados) e configurações *N*-way (3 e 4). Na abordagem com *N*-way= 3, as classes "água" e "pequenas ilhas" foram mescladas, gerando apenas uma classe "água" genérica.

Para cada experimento, o conjunto de suporte (*K* imagens por classe) foi selecionado aleatoriamente. Todas as amostras restantes do conjunto de dados foram classificadas e as métricas correspondentes foram computadas. As distâncias utilizadas incluíram aquelas propostas em [23]: Euclidiana, Euclidiana ao quadrado e similaridade de cossenos.

Como cobrir todos os episódios possíveis era inviável, cada novo episódio utilizou um conjunto de suporte diferente, selecionado aleatoriamente (com reposição), configurando um treinamento e validação por episódios.

Durante a validação foram obtidas as métricas de precisão (1), *recall* (2) - ambas apenas para a classe de alvos, foco deste trabalho- [26] e o desempenho geral baseado em acurácia.

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

Tais que *TP* (verdadeiro positivo) representa a classificação correta (alvo classificado como alvo), *FP* (falso positivo) classificação de outra classe como alvo e *FN* (falso negativo) alvo não classificado como alvo.

Todos os modelos foram avaliados em 500 episódios, as métricas obtidas e, para eleger o de melhor desempenho, foi realizada uma análise de variância. Os resultados são apresentados na seção seguinte.

III. RESULTADOS

Todos os desempenhos foram obtidos utilizando blocos de imagem com resolução de 200×200 pixels, configuração que resultou em 124 imagens por classe. Em média, havia aproximadamente 2 alvos por imagem rotulada como "alvo". Os dados foram divididos em 70% para treinamento e 30% para teste.

A divisão com apenas três classes incluiu apenas "terra", "água" e "alvos". Notavelmente, as feições inicialmente rotuladas como "superfícies rígidas" foram incorporadas à classe "água" com o objetivo de melhorar a coerência semântica. Um

processo de reequilíbrio também foi realizado para igualar a quantidade de amostras das outras duas classes.

O EfficientNetV2 [25] foi escolhido como extrator de feições. Sua escolha se deu por dois principais motivos: estava presente na configuração vencedora de [13] e sua estrutura compacta e eficiente pode levar a bom desempenho bem como evitar *overfitting* devido à escassez de dados [27]. Todos os modelos foram ajustados para o novo conjunto de dados utilizando refinamento *fine-tuning* do modelo já treinado (*transfer learning*) e uma taxa de aprendizado cíclica para melhorar o desempenho sem a necessidade de ajustes manuais [28].

Na Tabela I, todos os resultados são apresentados, os modelos com os melhores resultados gerais foram destacados.

A Fig.4 exibe o resultado do emprego do algoritmo em uma região nunca vista no processo de treinamento ou validação. A imagem representa uma porção de uma cena completa.

IV. DISCUSSÕES

Esta seção foca em três aspectos principais deste trabalho: conjunto de dados, configuração e resultados de detecção.

A. Tamanho da imagem e do alvo

Em cenários nos quais apenas uma pequena porção da imagem de entrada é relevante para o rótulo, a razão sinal-ruído torna-se baixa, dificultando a classificação. Essa razão, também denominada "Objeto-para-Imagem (O2I)" é definida como a proporção entre a área do alvo (A_{object}) e a área total da imagem (A_{image}) [29]:

$$O2I = \frac{A_{object}}{A_{image}} \quad (3)$$

Para um tamanho de alvo fixo (A_{object}), o aumento da dimensão da imagem (A_{image}) reduz a razão O2I, o que pode degradar o desempenho de generalização CNNs. Esse desafio é ainda mais acentuado em regimes com poucos dados, como discutido em [29]. Considerando que os alvos deste trabalho variam de 12 a 60 metros — o que equivale a aproximadamente 4 a 20 pixels em uma resolução espacial de 3 metros —, em imagens de dimensões 200x200 pixels, ainda podem ser considerados pequenos.

No entanto, diferentes tamanhos de recortes (talhos) introduzem *trade-offs*. Para talhos menores, o contexto espacial reduzido e a necessidade de janelas sobrepostas para evitar o corte de alvos devido ao processo adotado aumentam tanto a complexidade quanto o custo computacional do treinamento ou da validação e futuro emprego. Por exemplo, uma única imagem de cena completa pode gerar cerca de 3.000 talhos de 200x200 pixels, mas cerca de 16.000 talhos de 100x100 pixels — resultando em um aumento de 33% no total de pixels processados por imagem. Para chips maiores, há mais contexto disponível e menos pixels processados, mas a razão O2I diminui.

B. Desempenho dos modelos

Conforme descrito na Seção II, foi realizado um processo de fusão e reorganização das classes de água e superfícies rugosas. Essa fusão levou a uma melhoria de desempenho de maneira geral, quando todas as métricas são consideradas.

Tal resultado era esperado, dado que reduzir o número de classes ("*N-way*") pode levar a melhorias de desempenho, como observado em [30].

Alguns pontos principais podem ser observados na Tabela I. Primeiro, a maioria dos desvios padrão é considerável e, quando combinados com os valores médios, revelam uma sobreposição significativa entre algumas métricas computadas, principalmente considerando os modelos aparentemente melhores (destacados). Essa sobreposição dificulta a identificação intuitiva da melhor configuração geral.

A variabilidade observada (desvio padrão) pode ser atribuída às imagens de suporte utilizadas aleatoriamente durante a avaliação. Uma análise dos episódios individuais de avaliação revelou que certas imagens proporcionaram melhor representação de suas respectivas classes, resultando em melhor desempenho nesses casos, conforme já mencionado em [27]. Isso destaca a importância da seleção cuidadosa das imagens de suporte em aplicações reais.

Para determinar o modelo com melhor desempenho, foi adotado o teste de Kruskal–Wallis [31] aos dados brutos das três métricas — ou seja, os valores das métricas de cada um dos 500 episódios de avaliação usados para calcular as médias e desvios padrão apresentados. Esse teste não paramétrico foi escolhido devido ao não atendimento das premissas da Análise de Variância (ANOVA), como normalidade dos resíduos e homogeneidade de variâncias. A aplicação permitiu rejeitar a hipótese nula para todos os modelos e todas as métricas, indicando que há diferenças estatisticamente significativas entre pelo menos dois resultados para cada métrica, considerando $\alpha = 0,05$.

As diferenças múltiplas par-a-par foram então avaliadas através do teste de Dunn [32]. As seguintes equivalências estatísticas foram encontradas por métrica, para os modelos destacados:

- **Recall:** 3-way 10-shot com Euclidiana = 3-way 10-shot com Similaridade de Cossenos = 3-way 5-shot com Similaridade de Cossenos;
- **Precisão:** 3-way 10-shot com Euclidiana = 3-way 10-shot com Similaridade de Cossenos;
- **Acurácia Geral:** 3-way 10-shot com Euclidiana = 3-way 10-shot com Similaridade de Cossenos = 3-way 5-shot com Similaridade de Cossenos;

Como todos os demais são iguais estatisticamente de acordo com as outras métricas, o modelo 3-way 5-shot com Similaridade de Cossenos pode ser elencado como melhor devido à sua precisão pois, além de não se igualar aos demais estatisticamente, possui maior média e menor DP para essa métrica.

Curiosamente, esse resultado diverge da proposta original das PNs [23], onde a distância Euclidiana ao quadrado era recomendada por sua interpretação como uma divergência de Bregman e seu embasamento probabilístico. A similaridade de cossenos, por outro lado, não é uma divergência de Bregman. Ainda assim, na prática, ela pode superar a distância Euclidiana, especialmente quando usada com *backbones* modernos como o EfficientNetV2, que podem produzir *embeddings* com direção consistente, mesmo com magnitudes variadas. Isso sugere que, sob certas condições, relações angulares entre características podem oferecer uma separação de classes mais robusta do que distâncias absolutas no espaço de *embedding*.

TABELA I: Resultados experimentais

N-way	K-shot	Distância	Recall (alvos)	Precisão (alvos)	Acurácia Geral
3	5	Euclidiana	$0,89 \pm 0,11$	$0,88 \pm 0,10$	$0,85 \pm 0,06$
3	5	Similaridade de Cossenos	$0,94 \pm 0,07$	$0,94 \pm 0,07$	$0,92 \pm 0,05$
3	5	Euclidiana Quadrática	$0,89 \pm 0,11$	$0,88 \pm 0,10$	$0,85 \pm 0,06$
3	10	Euclidiana	$0,93 \pm 0,09$	$0,91 \pm 0,09$	$0,91 \pm 0,05$
3	10	Similaridade de Cossenos	$0,92 \pm 0,09$	$0,90 \pm 0,08$	$0,91 \pm 0,05$
3	10	Euclidiana Quadrática	$0,90 \pm 0,09$	$0,89 \pm 0,09$	$0,88 \pm 0,06$
4	5	Euclidiana	$0,90 \pm 0,14$	$0,85 \pm 0,13$	$0,81 \pm 0,06$
4	5	Similaridade de Cossenos	$0,94 \pm 0,08$	$0,84 \pm 0,09$	$0,85 \pm 0,05$
4	5	Euclidiana Quadrática	$0,82 \pm 0,15$	$0,78 \pm 0,10$	$0,77 \pm 0,07$
4	10	Euclidiana	$0,90 \pm 0,14$	$0,84 \pm 0,13$	$0,85 \pm 0,06$
4	10	Similaridade de Cossenos	$0,93 \pm 0,08$	$0,85 \pm 0,09$	$0,85 \pm 0,05$
4	10	Euclidiana Quadrática	$0,89 \pm 0,11$	$0,82 \pm 0,10$	$0,81 \pm 0,06$

^a Todos os resultados representam valores médios e seus respectivos desvios padrão para 500 episódios de classificação de imagens em dados não vistos durante o treinamento.

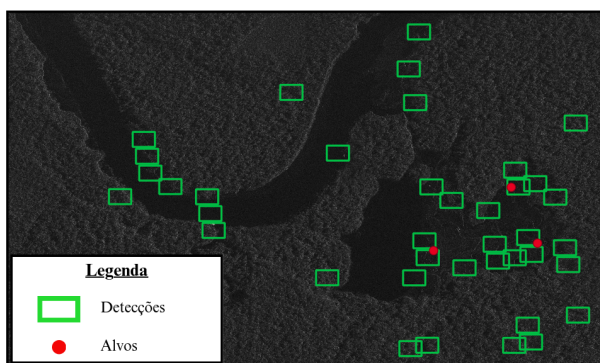


Fig. 4: Teste do algoritmo e gabarito

O fato de o melhor desempenho ter sido alcançado pelo modelo com apenas 5 amostras de suporte por classe sugere que o método de obtenção dos protótipos — via média dos *embeddings* — pode não ser tão eficaz nesta aplicação quando há maior quantidade de exemplos. Isso pode ocorrer porque a média tende a suavizar ou diluir características discriminativas relevantes, especialmente em cenários com alta variabilidade intra-classe, fazendo com que informações importantes sejam ofuscadas ao aumentar o número de amostras de suporte.

Ademais, realizar comparações de desempenho com aplicações voltadas à tarefa de detecção dos mesmos alvos ainda não se mostra viável pois implicaria numa comparação de quantidade de imagens com alvos contra detecções (posição do alvo) realizadas. Entretanto, o presente método de classificação permite reduzir significativamente o escopo de busca para posteriores análises visuais ou emprego de algoritmos de visão computacional. Seguindo os parâmetros antes mencionados, uma imagem de cena completa que originalmente geraria 3000 recortes de 200x200 pixels foi reduzida para apenas 60 recortes relevantes para análise em etapas posteriores, representando uma redução aproximada de 98% na área a ser inspecionada sem prejuízos maiores que 13% no pior cenário para a métrica de *recall* do melhor modelo.

Através da Fig. 4 também pode-se notar que em certas regiões, todos os alvos presentes são corretamente localizados pela classificação dos talhos da classe "alvos", entretanto, alguns falsos positivos ainda se mostram presentes, principalmente em regiões com superfícies rígidas na água nos leitos. Tal fato já era esperado devido ao formato e a valores de pixels semelhantes com os dos alvos conforme já observado

em [7], [9].

Vale destacar, contudo, que o presente método não é capaz de distinguir entre dragas de garimpo e outras pequenas embarcações que podem estar presentes nas imagens SAR. Entretanto, o contexto geográfico da aplicação (uma reserva indígena remota próxima a uma fronteira nacional, conhecida por diversas atividades ilegais como tráfico de drogas e armas) justifica o monitoramento de qualquer embarcação de pequeno porte detectada na região. Dessa forma, o método cumpre seu objetivo principal: auxiliar na detecção de dragas em áreas onde a vigilância manual é inviável.

V. CONCLUSÃO

Este estudo apresentou um algoritmo em múltiplas etapas projetado para classificar imagens contendo dragas de mineração em rios localizados em regiões remotas e fronteiriças na Amazônia brasileira. Para tal, utilizando imagens SAR provenientes da constelação de satélites ICEYE.

O algoritmo proposto inicia com uma imagem GRD de cena completa, que é recortada em talhos menores (200x200 pixels), os quais são classificados utilizando uma Rede de Protótipos treinada em uma estrutura de *few-shot learning*, uma abordagem motivada pela disponibilidade limitada de dados rotulados e pelas dificuldades associadas à sua obtenção.

Embora não seja possível realizar uma comparação direta em termos de quantidade de alvos detectados com outros trabalhos, o método proposto resulta numa redução de aproximadamente 98% na quantidade de imagens a serem analisadas com uma perda de, no máximo, 13% dos alvos considerando o pior cenário de acordo com o desvio padrão.

Uma limitação notável do método é sua incapacidade de distinguir dragas de outras pequenas embarcações que operam no mesmo ambiente. Entretanto, dado o contexto da região, marcado por atividades ilegais como mineração ilegal, tráfico de drogas e contrabando de armas, qualquer detecção de pequenas embarcações é relevante para fins de monitoramento.

Ainda, vale ressaltar que o presente método apenas indica que há alvos numa determinada cena, mas não determina onde estão. Nesse contexto, trabalhos futuros podem explorar métodos de detecção de objetos para regimes de poucos dados com técnicas de *machine learning* de visão computacional baseados em *thresholding* conforme feito por [6], entretanto, já utilizando-se da redução de escopo de busca trazido pela presente pesquisa.

REFERÊNCIAS

- [1] D. H. C. Pereira, "Detection and monitoring of boat-based gold mining activities in the amazon using remote sensing and artificial intelligence, in portuguese: Detecção e monitoramento de atividades de garimpo de ouro embarcado na amazônia com sensoriamento remoto e inteligência artificial," Ph.D. dissertation, Universidade de Brasília, 2024.
- [2] M. A. da Costa and F. J. Rios, "The gold mining industry in brazil: A historical overview," *Ore Geology Reviews*, vol. 148, p. 105005, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0169136822003134>
- [3] MapBiomas, "Collection 8 of the Annual Series of Land Cover and Use Maps of Brazil: Proximity of mining, rivers, and lakes in the Amazon, in Portuguese: Coleção 8 da Série Anual de Mapas de Cobertura e Uso da Terra do Brasil: Proximidade de garimpo, rios e lagos na Amazônia," <https://brasil.mapbiomas.org/2024/04/19/77-dogarimpo-na-amazonia-esta-a-menos-de-500-metros-da-agua/>, 2024, accessed on April 21, 2024.
- [4] C. H. D. Felice, A. L. B. Rodrigues, G. Stabile, N. C. R. Bergiante, M. C. N. Belderrain, R. Paes, and E. H. Shiguemori, "Analyzing the Impact of Illegal Gold Mining on the Yanomami Indigenous Land in Roraima, Brazil: An Approach with Soft Systems Methodology (SSM) in Portuguese: Analisando o Impacto do Garimpo Ilegal de Ouro na Terra Indígena Yanomami em Roraima, Brasil: Uma Abordagem com Soft Systems Methodology (SSM)," in *Proceedings of the XLIX Brazilian Symposium on Operational Research (SBPO)*, 2024. [Online]. Available: https://proceedings.science/proceedings/100475/_papers/193576
- [5] D. H. C. Pereira, R. A. T. Gomes, O. A. de Carvalho Júnior, and R. F. Guimarães, "Identification of boat-based gold mining through classification algorithms on sentinel-2 imagery," *Society & Nature*, vol. 36, no. 1, Nov. 2024, original title in Portuguese: "Identificação de Garimpos de Ouro Embarcados por meio de Algoritmos de Classificação em Imagens Sentinel-2".
- [6] M. A. Alessi, P. G. Chirico, and M. Millones, "Artisanal mining river dredge detection using sar: A method comparison," *Remote Sensing*, vol. 15, 12 2023.
- [7] M. A. Alessi, P. G. Chirico, S. Sunder, and K. L. O'Pry, "Detection and monitoring of small-scale diamond and gold mining dredges using synthetic aperture radar on the kadé (sangha) river, central african republic," *Remote Sensing*, vol. 15, 2 2023.
- [8] F. d. L. Lobo, P. W. M. Souza-Filho, E. M. L. d. M. Novo, F. M. Carlos, and C. C. F. Barbosa, "Mapping mining areas in the Brazilian amazon using MSI/Sentinel-2 imagery (2017)," *Remote Sensing*, vol. 10, no. 8, 2018.
- [9] C. R. Gruel and E. M. Latrubesse, "A monitoring system of sand mining in large rivers and its application to the ayeyarwady (irrawaddy) river, myanmar," *Water (Switzerland)*, vol. 13, 9 2021.
- [10] X. Cheng, Y. Wan, and B. Cao, "Identification of sand dredges in yangtze river based on asar remote sensing data," in *2009 17th International Conference on Geoinformatics*, 2009, pp. 1–5.
- [11] J. Li, L. Tian, X. Chen, X. Li, J. Huang, J. Lu, and L. Feng, "Remote-sensing monitoring for spatio-temporal dynamics of sand dredging activities at poyang lake in china," *International Journal of Remote Sensing*, vol. 35, pp. 6004–6022, 8 2014.
- [12] J. Li, C. Xu, H. Su, L. Gao, and T. Wang, "Deep learning for sar ship detection: Past, present and future," *Remote Sensing*, vol. 14, no. 11, June 2022.
- [13] F. Paolo, T. ting Tim Lin, R. Gupta, B. Goodman, N. Patel, D. Kuster, D. Kroodsmas, and J. Dunnmon, "xview3-sar: Detecting dark fishing activity using synthetic aperture radar imagery," 2022. [Online]. Available: <https://arxiv.org/abs/2206.00897>
- [14] R. Yang, Z. Pan, X. Jia, L. Zhang, and Y. Deng, "A Novel CNN-Based Detector for Ship Detection Based on Rotatable Bounding Box in SAR Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 1938–1958, 2021.
- [15] M. Sun, Y. Li, X. Chen, Y. Zhou, J. Niu, and J. Zhu, "A Fast and Accurate Small Target Detection Algorithm Based on Feature Fusion and Cross-Layer Connection Network for the SAR Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 8969–8981, 2023.
- [16] J. Vanschoren, "Meta-learning," in *Automated Machine Learning: Methods, Systems, Challenges*, F. Hutter, L. Kotthoff, and J. Vanschoren, Eds. Springer, 2019, pp. 35–61. [Online]. Available: https://doi.org/10.1007/978-3-030-05318-5_2
- [17] ICEYE, *ICEYE Data Product Specification*, Feb. 2025, proprietary document. [Online]. Available: <https://www.iceye.com/>
- [18] A. I. Flores-Anderson, K. E. Herndon, E. Cherrington, and R. B. Thapa, Eds., *SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation*. Huntsville, AL, USA: NASA, 2019, accessed: 2025-05-20. [Online]. Available: https://gis1.servirglobal.net/TrainingMaterials/SAR/SARHB_FullRes.pdf
- [19] SIGIMINE - Mining Geographic Information System, "Sigimine. mining geographic information system," <https://geo.anm.gov.br>, 2022, accessed: 2025-05-19.
- [20] H. C. T. de Oliveira, "Automatic detection of small vessels in sar images from the iceye constellation using a convolutional neural network, in portuguese: Detecção automática de pequenas embarcações em imagens sar da constelação iceye com uso de uma rede neural convolucional," Course Completion Work (Lato Sensu), Instituto Tecnológico de Aeronáutica, São José dos Campos, Brazil, 2021, advisors: Prof. Dr. Elcio Hideiti Shiguemori and M.Sc. Tahisa Neitzel Kuck.
- [21] F. G. Mohammadi, M. H. Amini, and H. R. Arabnia, "An introduction to advanced machine learning: Meta-learning algorithms, applications, and promises," in *Advances in Intelligent Systems and Computing*. Springer, 2020, vol. 1123, pp. 129–144.
- [22] J. Vanschoren, "Meta-learning: A survey," *CoRR*, vol. abs/1810.03548, 2018. [Online]. Available: <http://arxiv.org/abs/1810.03548>
- [23] J. Snell, K. Swersky, and R. Zemel, "Prototypical networks for few-shot learning," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 4080–4090.
- [24] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97. PMLR, Jun 2019, pp. 6105–6114. [Online]. Available: <https://proceedings.mlr.press/v97/tan19a.html>
- [25] —, "EfficientNetV2: Smaller models and faster training," in *Proceedings of the 38th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, M. Meila and T. Zhang, Eds., vol. 139. PMLR, Jul 2021, pp. 10 096–10 106. [Online]. Available: <https://proceedings.mlr.press/v139/tan21a.html>
- [26] A. Geron, *Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow*, 2nd ed. O'Reilly Media, 2019, pp. 90–96.
- [27] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016, ch. 1 and 20, pp. 1–4, 428–429, <http://www.deeplearningbook.org>.
- [28] L. N. Smith, "No more pesky learning rate guessing games," *CoRR*, vol. abs/1506.01186, 2015. [Online]. Available: <http://arxiv.org/abs/1506.01186>
- [29] N. Pawlowski, S. Bhooshan, N. Ballas, F. Ciompi, B. Glocker, and M. Drozdal, "Needles in haystacks: On classifying tiny objects in large images," *CoRR*, vol. abs/1908.06037, 2019. [Online]. Available: <http://arxiv.org/abs/1908.06037>
- [30] X. Luo, H. Wu, J. Zhang, L. Gao, J. Xu, and J. Song, "A closer look at few-shot classification again," 2023. [Online]. Available: <https://arxiv.org/abs/2301.12246>
- [31] D. C. Montgomery, *Design and Analysis of Experiments*, 9th ed. Hoboken, NJ: John Wiley & Sons, Inc., 2017, pp. 78–86, 123–125, library of Congress Control Number: 2017002355. ISBN (ebook): 9781119299363. [Online]. Available: <https://lccn.loc.gov/2017002355>
- [32] O. J. Dunn, "Multiple comparisons using rank sums," *Technometrics*, vol. 6, no. 3, pp. 241–252, 1964.