XVIII SIMPÓSIO DE RECURSOS HÍDRICOS DO NORDESTE
XVIII ENCONTRO DE RECURSOS HÍDRICOS EM SERGIPE

MAPEAMEANTO AUTOMÁTICO DE PEQUENOS RESERVATÓRIOS UTILIZANDO REDES NEURAIS CONVOLUCIONAIS E IMAGENS SENTINEL 1

Germano Gondim Ribeiro Neto [1], Bruno Berkamp Mota Brandão[2], Cinthya Santos da Silva[3] & Adriano Rolim da Paz [4]

RESUMO – Este estudo avaliou a viabilidade do desenvolvimento de uma Rede Neural Convolucional (CNN) para o mapeamento de pequenos reservatórios na região semiárida brasileira, utilizando a bacia hidrográfica do Rio Sangue, localizada no Estado do Ceará, como estudo de caso. Essa bacia hidrográfica apresenta uma alta densidade de pequenos reservatórios e conta com um mapeamento manual detalhado, tornando-a uma área ideal para esta análise. Foram testadas diferentes configurações de dados de entrada, e o melhor desempenho foi alcançado utilizando os valores médios anuais das bandas VV e VH do Sentinel-1 combinados com um Modelo Digital de Elevação. Essa configuração atingiu uma acurácia geral de 80% e permitiu a identificação de aproximadamente 1.000 reservatórios anteriormente não mapeados. Apesar de sua alta precisão, a CNN apresentou limitações na detecção completa de reservatórios afetados pela eutrofização. Nesses casos, áreas cobertas por vegetação flutuante eram frequentemente classificadas como terra em vez de água. Embora preliminares, os resultados demonstram que técnicas de aprendizado de máquina combinadas com dados de sensoriamento remoto representam uma abordagem promissora para o mapeamento e monitoramento de pequenos reservatórios na região semiárida brasileira.

ABSTRACT – This study evaluated the feasibility of developing a Convolutional Neural Network (CNN) for mapping small reservoirs in the Brazilian semiarid region, using the Riacho do Sangue watershed, located in Ceará State, as a case study. This watershed contains a high density of small reservoirs and benefits from a detailed manual inventory, making it an ideal area for this analysis. Different input data configurations were tested, and the best performance was achieved using the annual mean values of the Sentinel-1 VV and VH bands combined with a Digital Elevation Model (DEM). This configuration reached an overall accuracy of 82% and enabled the identification of approximately 1,000 previously unmapped reservoirs. Despite its high accuracy, the CNN showed limitations in fully detecting reservoirs affected by eutrophication. In these cases, areas covered by floating vegetation were often classified as land rather than water. Therefore, future studies should include a larger number of eutrophic reservoirs in the training dataset to improve the network's ability to recognize such patterns and increase mapping accuracy. Although preliminary, the results demonstrate that machine learning techniques combined with remote sensing data represent a promising approach for mapping and monitoring small reservoirs in the Brazilian semiarid region.

[1]) Universidade Federal da Paraíba – Cidade Universitária, João Pessoa/PB, fone: 83 32167200, e-mail: germano.ribeiro@ct.ufpb.br
[2]) Universidade Federal da Paraíba – Cidade Universitária, João Pessoa/PB, fone: 83 32167200, e-mail: contatobrunoberkamp@gmail.com
[3]) Instituto Federal da Paraíba – campus Cajazeiras, Cajazeiras/PB, fone: 8335324100, e-mail: cinthya.santos@ifpb.edu.br
[4]) Universidade Federal da Paraíba – Cidade Universitária, João Pessoa/PB, fone: 83 32167200, e-mail: adrianorpaz@ct.ufpb.br

Palavras-Chave – Mapeamento de açudes, Redes Neurais Convolucionais, Sentinel 1

1 - INTRODUÇÃO

É esperado que o Brasil seja um dos países mais vulneráveis aos impactos das mudanças climáticas, cenário que se intensifica significativamente na região semiárida do Nordeste (Marengo et al., 2020). As projeções indicam que esta área, historicamente caracterizada por baixa disponibilidade hídrica, enfrentará uma redução do período chuvoso e a elevação das perdas evaporativas devido ao aumento das temperaturas médias. Isso implicará em uma maior pressão sobre os recursos hídricos, elevando o risco de secas mais intensas e prolongadas.

Associado a esse contexto, as populações locais do semiárido nordestino dependem intrinsecamente de reservatórios de pequeno e médio porte, também conhecidos como açudes, para suprir suas demandas hídricas básicas (Pereira et al., 2025). Contudo, a maioria dessas estruturas foram construídas informalmente, sem regulação e não são monitoradas por órgãos públicos. A ausência de um mapeamento sistemático e completo desses reservatórios na maioria dos estados da região faz com que o seu real impacto hidrológico permaneça desconhecido (Ribeiro Neto et al., 2022; Ribeiro Net al., 2024). Essa proliferação de pequenos açudes não monitorados pode interferir diretamente no balanço hídrico regional, pois eles retêm o escoamento superficial e afetam a recarga dos reservatórios estratégicos situados a jusante, podendo induzir ou prolongar eventos de seca.

Para mitigar esses riscos e subsidiar uma gestão de recursos hídricos mais holística e eficiente, torna-se estratégico o desenvolvimento de metodologias que permitam o monitoramento operacional dessas redes de reservatórios. Estudos anteriores têm demonstrado que metodologias que combinam aprendizado de máquina (Machine Learning) com imagens de satélite têm se mostrado eficientes para a identificação de padrões espaciais específicos, tais como: identificação de água, mapeamento de áreas irrigadas e mapeamento de culturas agrícolas (Latif et al., 2022; Bem et al., 2021; Wang et al., 2020). Neste sentido, este trabalho tem como objetivo geral apresentar uma metodologia capaz de identificar pequenos açudes na região semiárida do Brasil a partir técnicas de aprendizado de máquinas, redes neurais convolucionais (CNN) e imagens de satélite do tipo SAR (*Synthetic Aperture Radar*). Para viabilizar a aplicação prática da metodologia, foi adotado como estudo de caso a bacia hidrográfica do Riacho do Sangue, localizada no estado do Ceará. A escolha desta bacia justifica-se estrategicamente por se tratar de uma das regiões com maior densidade e concentração de açudes no território cearense e, fundamentalmente, por já dispor de um mapeamento detalhado dessas estruturas, servindo como uma base de calibração e validação ideal para o modelo proposto.

2 - ÁREA DE ESTUDO

Aplicamos nossa análise à bacia hidrográfica do Riacho do Sangue (1.368 km²), que faz parte da bacia hidrográfica do rio Jaguaribe (75.000 km²) e está localizada no estado do Ceará, no Brasil (Figura 2). A precipitação média anual e a evapotranspiração potencial na bacia hidrográfica do Riacho do Sangue são, respectivamente, 740 mm e 2.300 mm (1981–2019). Um complexo cristalino associado a solos rasos limita a ocorrência de aquíferos nessas regiões. Consequentemente, em combinação com as características meteorológicas, todos os rios da região são intermitentes (de Araújo & Bronstert, 2016; Fontenele et al., 2014). A área de estudo possui três reservatórios que são gerenciados e operados por órgãos governamentais: Barragem do Riacho do Sangue, Barragem do Tigre e Barragem do Jenipapeiro. Suas capacidades máximas de armazenamento são, respectivamente, 58,43 hm³, 3,51 hm³ e 14,59 hm³. A Fundação Cearense de Meteorologia e Recursos Hídricos (Funceme) realizou um mapeamento detalhado da pequena açudagem e constatou que existem cerca de 2450 açudes de pequeno porte que não são monitorados (Lima et al., 2025).

3 - METODOLOGIA

Para alcançar o objetivo de desenvolver uma metodologia capaz de conduzir um mapeamento automático de pequenos açudes no semiárido nordestino, foi utilizado uma abordagem baseada em aprendizado profundo (*deep learning*) associado com produtos de sensoriamento remoto. Assim, foi proposto a utilização de uma Rede Neural Convolucional (CNN) capaz de preservar feições espaciais detalhadas associadas com estratégias de modelagem matemática para compensar o desbalanceamento de classes de uso e ocupação do solo típico da paisagem da região.

3.1 - Dados Utilizados

A identificação contínua de corpos d'água no semiárido é frequentemente dificultada pela alta cobertura de nuvens durante a quadra chuvosa que além de potencialmente impedir a visualização de corpos hídrico, produz regiões de sombra que apresentam resposta espectral semelhante a água quando utilizado imagens óticas convencionais (Souza et al., 2022). Para contornar essa limitação, o optou-se por utilizar como dados de entrada imagens do sensor de Radar de Abertura Sintética (SAR) da constelação Sentinel-1. Foram utilizadas as polarizações cruzadas (VH) e co-polarizada (VV), cujos coeficientes de retroespalhamento (*backscatter*) são sensíveis à rugosidade da superfície e à constante dielétrica, permitindo um forte contraste entre os espelhos d'água e o solo ou vegetação adjacente.

Além das bandas SAR, incorporou-se a CNN um Modelo Digital de Elevação (DEM) como a terceira dimensão (canal) dos dados de entrada. Esta variável atua como uma referência espacial para a rede neural, ajudando a eliminar falsos positivos ao contextualizar o relevo. Isso permite, por

exemplo, que o modelo distinga de forma inteligente entre um trecho de rio que se pereniza durante a estação chuvosa de um açude com formato mais retilíneo.

3.2 - Arquitetura Da Rede Neural Convolucional

As Redes Neurais Convolucionais (CNNs) são uma classe de algoritmos de aprendizado profundo projetada primariamente para o processamento de dados multidimensionais, como imagens. Seu funcionamento baseia-se no uso de filtros matemáticos (convoluções) que varrem a matriz de entrada para extrair automaticamente características espaciais, tais como bordas, texturas e formas complexas. Neste estudo, adotou-se uma arquitetura do tipo *Encoder-Decoder* (Codificador-Decodificador). A codificação reduz as dimensões da imagem para abstrair a semântica e o contexto, ou seja, a CNN aprende o que está sendo mostrado na imagem. Já decodificação simétrica expande e reconstrói a resolução espacial para determinar onde os alvos a serem classificados estão localizados. Isso permite uma classificação precisa, pixel a pixel. Por exemplo, se você alimenta no treinamento dessas redes como dados de entrada imagens Sentinel 1 e como saída de treinamento um mapa de uso e ocupação do solo, a rede primeiro aprende (semântica e contexto) as características de cada classe e depois passa a ter a capacidade (decodificação simétrica) de identificar essas classes em outras imagens.

Como os pequenos açudes não ocupam grande espaço nas imagens Sentinel 1, algumas modificações foram feitas para evitar o seu desaparecimento nos mapas de características muito profundos da rede. A etapa de contração (*Encoder*) foi modificada para realizar a redução de dimensionalidade espacial apenas uma vez. A arquitetura foi estruturada da seguinte forma:

• Bloco Codificador 1: Duas camadas convolucionais sequenciais de 3x3 com 64 filtros, cada uma seguida por Normalização em Lote e função de ativação ReLU. O bloco é finalizado com uma camada de *Max Pooling* de 2x2, reduzindo a resolução espacial pela metade.

• Bloco Codificador 2: Composto por duas convoluções de 3x3 com 128 filtros (*Batch Norm* e ReLU). Diferente de arquiteturas tradicionais, a camada de *pooling* foi omitida neste bloco para preservar o possível número reduzido de pixels dos pequenos corpos d'água.

• Gargalo (*Bottleneck*): Duas camadas convolucionais de 3x3 com 256 filtros, intercaladas por uma camada de *Dropout* com taxa de retenção de 70% (abandono de 30%), inserida para atuar como regularizador estrutural e prevenir o sobreajuste (overfitting).

• Bloco Decodificador: A recuperação da resolução espacial é iniciada com uma camada de convolução transposta. Como o Bloco 2 não aplicou redução espacial, a primeira transposição foi ajustada com passo (stride) 1. A reconstrução final para as dimensões originais de entrada ocorre no

segundo decodificador por meio de uma convolução transposta de 2x2. Ambas as etapas de transposição são seguidas por camadas convolucionais de refinamento.

• Camada de Saída: Uma convolução de 1x1 mapeia os 64 canais resultantes para o número total de classes. A classificação final por pixel é computada por uma camada *Softmax*, que gera um mapa de probabilidades para as classes observadas. A classe final é então escolhida sendo a com maior probabilidade encontrada.

3.2.1 – Pré-Processamento E Normalização Dos Dados

Para garantir a estabilidade numérica e acelerar a convergência do modelo durante o treinamento, os dados de entrada foram submetidos a um processo de padronização estatística (*z-score normalization*). A normalização foi aplicada de forma independente para cada canal físico da imagem. A média global e o desvio padrão de cada banda foram calculados em todo o conjunto de treinamento. Em seguida, cada pixel foi ajustado subtraindo-se a média de sua respectiva banda e dividindo-se pelo desvio padrão. Este procedimento mitigou o impacto de diferentes escalas de amplitude entre os dados de radar e os dados topográficos do DEM, assegurando que todas as variáveis de entrada contribuíssem de forma equilibrada para o cálculo dos gradientes. Os rótulos de treinamento (máscaras de segmentação) foram convertidos para matrizes categóricas, compreendendo três classes principais de interesse para o cenário: terra, rio e água/açudes.

3.2.2 – Compensação Do Desbalanceamento De Classes

A paisagem do semiárido impõe um severo desbalanceamento de classes, onde a caatinga e o solo exposto (Land) dominam majoritariamente a cena, enquanto os rios e açudes representam uma fração mínima dos pixels. Para impedir que a rede fosse enviesada em favor da classe majoritária, adotou-se uma função de perda de Entropia Cruzada Ponderada. Os pesos foram calculados de forma inversamente proporcional à frequência empírica de cada classe no conjunto de dados (aproximadamente 90% para Terra, 4% para rio e 6% para açudes). O vetor de pesos foi normalizado e integrado diretamente na dimensão de canais da função de perda. Desta forma, erros de classificação cometidos sobre pixels de água ou rios foram penalizados com rigor significativamente maior do que falsos negativos em áreas de terra firme.

3.2.3 – Otimização E Hiperparâmetros De Treinamento

O treinamento da rede foi conduzido utilizando o otimizador Adam (*Adaptive Moment Estimation*). A fim de garantir um refinamento suave dos pesos nas épocas finais do treinamento,

implementou-se um escalonamento de taxa de aprendizado decrescente em degraus. A taxa de aprendizado inicial foi de 0,001, com um fator de redução de 0,5 aplicado a cada 15 épocas. O tamanho do mini-lote (*mini-batch size*) foi configurado para 64 amostras, visando reduzir o custo computacional e a introdução de ruído estocástico necessário para evitar a convergência para mínimos locais indesejados. O modelo foi treinado por um total de 50 épocas, com embaralhamento completo dos dados (*shuffle*) a cada iteração, garantindo a generalização geométrica e espectral da rede.

As amostras para o treinamento correspondem a 20% da área de estudo. A avaliação da acurácia da rede foi avaliada considerando os 80% não utilizados para o treino em que verificou-se a rede foi capaz de identificar os açudes mapeados previamente pela FUNCEME. Além disso, analisou-se visualmente se a rede foi capaz de identificar açudes não mapeados por essa instituição.

4 – RESULTADOS E DISCUSSÕES

Foram testadas seis configurações de dados de entrada para o treinamento das redes, variando a utilização combinada das bandas VH e VV da missão Sentinel-1 e do Modelo Digital de Elevação (DEM): VH; VV; VH+VV; VH+DEM; VV+DEM; e VH+VV+DEM. No presente trabalho, serão apresentados apenas os melhores resultados obtidos, correspondentes à rede que utilizou ambas as bandas do Sentinel-1 associadas ao DEM. A Figura 1 apresenta um histograma com a distribuição do número de açudes identificados por intervalo de área superficial, bem como a acurácia na identificação dessas estruturas. Neste estudo, a acurácia foi definida como a razão entre o número de açudes identificados pela CNN e o número de açudes identificados pela FUNCEME em cada classe de área.

Na Figura 1a, observa-se que cerca de 68% dos açudes possuem área superficial inferior a 1 ha. Como esses reservatórios são de pequeno porte, espera-se que permaneçam secos durante grande parte do ano, o que dificulta sua identificação por meio de produtos de sensoriamento remoto. Ainda assim, a CNN foi capaz de identificar aproximadamente 70% dessas estruturas (Figura 1b). Para as demais classes de açudes, observa-se uma acurácia média em torno de 80%, indicando que o uso de técnicas de aprendizado de máquina associadas a produtos de sensoriamento remoto de alta resolução, como as imagens Sentinel-1, apresenta potencial para constituir uma solução viável para o mapeamento desse tipo de estrutura na região semiárida do Nordeste brasileiro. A acurácia global da CNN que utilizou as duas bandas do Sentinel-1 em conjunto com informações do DEM foi de 82%, enquanto a mesma arquitetura, sem a inclusão de informações topográficas, alcançou uma acurácia de 75%. Esse resultado indica que a incorporação do DEM contribui positivamente para o aprendizado da rede, principalmente na diferenciação entre trechos de rios e açudes.

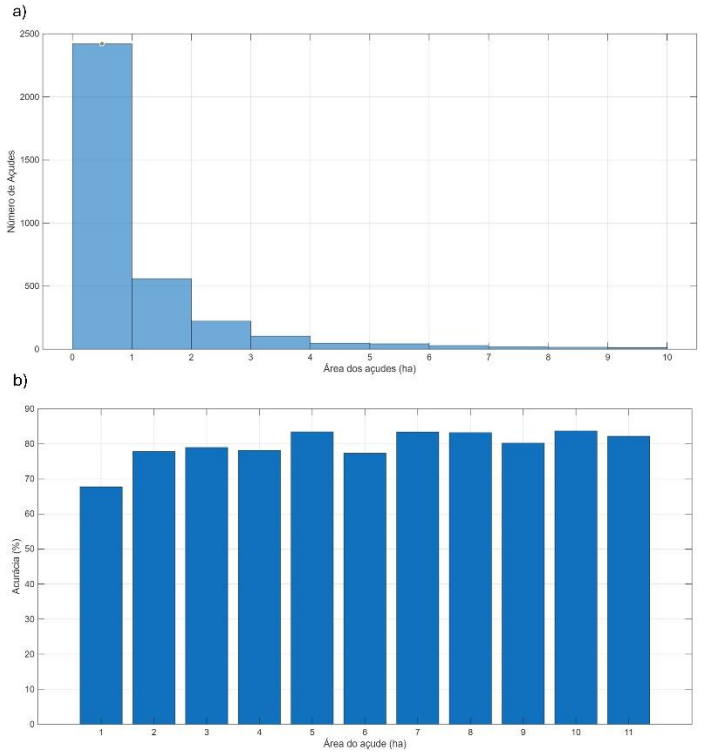


Figura 1 – Histograma da distribuição das áreas dos açudes mapeados pela melhor CNN (a) e acurácia da melhor CNN no mapeamento de açudes.

A Figura 2 apresenta exemplos do nível de detalhamento do mapeamento obtido pela rede que apresentou o melhor desempenho. Observa-se que a CNN foi capaz de delimitar adequadamente as zonas de transição entre solo úmido e água. Isso é evidenciado pela identificação individual de uma sequência de açudes em cascata, que não foram agrupados como um único reservatório, apesar da proximidade espacial entre eles. A precisão do detalhamento do mapeamento gerado pela CNN também pode ser observada na sua capacidade de identificar novos açudes (polígonos rosa com tracejado). Além disso, a Figura 2 evidencia uma limitação da versão atual da CNN, que tende a não identificar adequadamente áreas de água cobertas por vegetação flutuante, fenômeno comum na região em decorrência de processos de eutrofização (setas verdes). Para futuros trabalhos será necessário testar a introdução nos arquivos de treinamento mais casos específicos de açudes eutrofizados para que a CNN aprenda esse padrão espectral e espacial. Espera-se, assim, que a rede passe reconhecer áreas cobertas por vegetação flutuante como parte do corpo hídrico, aumentando a precisão do mapeamento desses reservatórios.

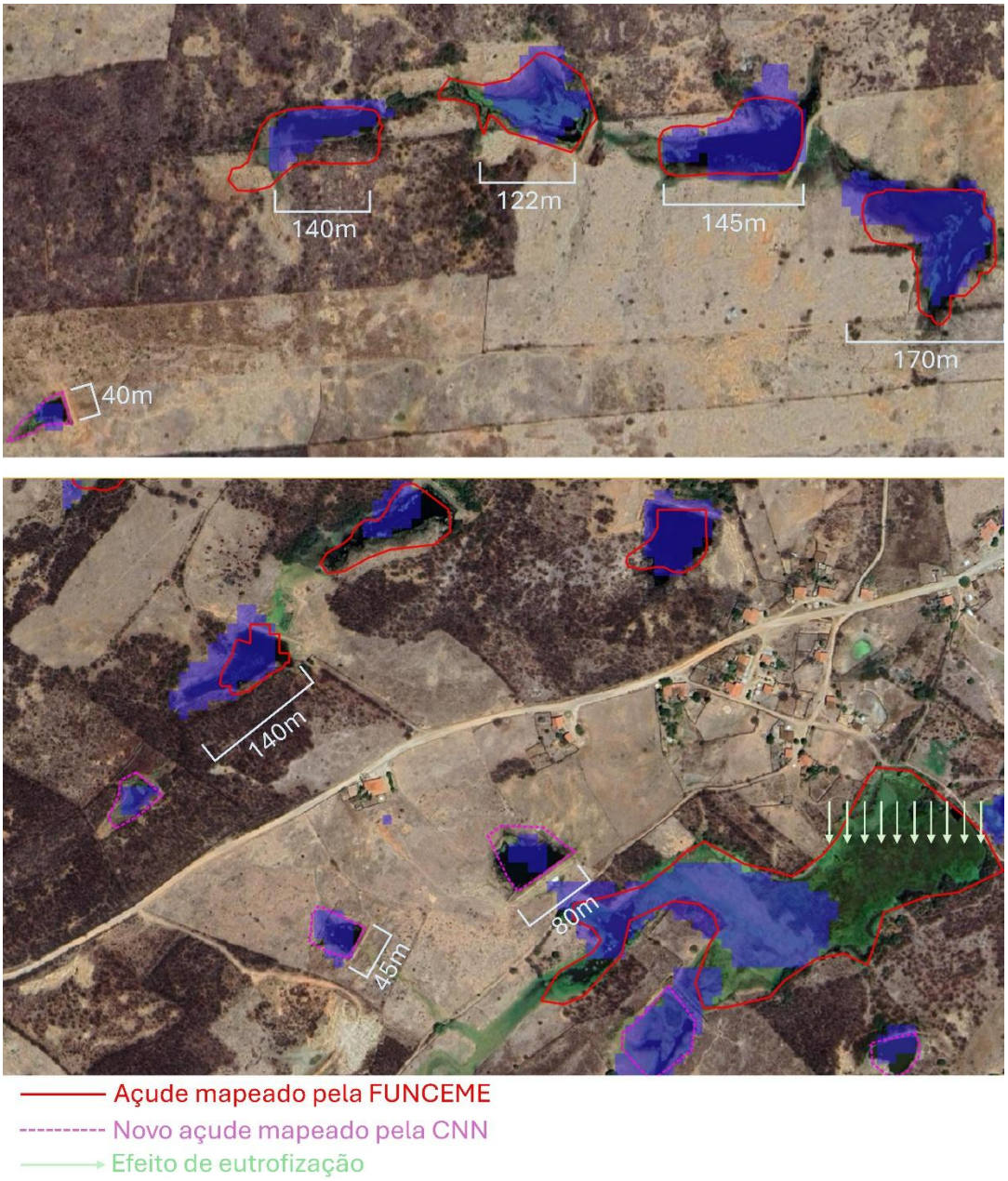


Figura 2 – Amostra do desempenho da melhor CNN no mapeamento de açudes.

5 - CONCLUSÃO

Nesse estudo foi avaliado a viabilidade do desenvolvimento e uma rede neural convolucional (CNN) voltada para o mapeamento de pequenos açudes para região semiárida do Brasil com base em dados da missão espacial Sentinel 1 informações topográficas. Considerou-se como estudo de caso a bacia hidrográfica do riacho do Sangue, localizada no Ceará. Essa bacia apresenta uma alta concentração de pequenos açudes e dispõe de um mapeamento manual detalhado dessas estruturas, sendo assim uma área de estudo ideal para conduzir tal análise.

Foram testadas diferentes configurações de dados de entrada para o treinamento da CNN e conclui-se que o melhor desempenho foi obtido para a rede que utilizou como variáveis de entrada os valores médios anuais das bandas VV e VH do Sentinel 1 e o modelo digital de elevação da região. Essa configuração alcançou uma acurácia geral de 82% e possibilitou a identificação de aproximadamente 1000 novos açudes não contemplados no mapeamento original. Ainda que a acurácia tenha sido alta, a rede teve dificuldade de mapear por completo os açudes com alta eutrofização. Nesses casos, áreas cobertas por vegetação flutuante foram frequentemente classificadas como terra em vez de água. Para futuros trabalhos é indicado que seja modificada a metodologia de treinamento das redes para incluir mais açudes com essas características a fim de garantir que a CNN aprenderá a identificar com mais precisão padrões de açudes eutrofizados. Dessa forma, recomenda-se que trabalhos futuros incorporem um maior número de exemplos de açudes eutrofizados nos conjuntos de treinamento, de modo a aprimorar a capacidade da CNN de reconhecer esse padrão e aumentar a precisão do mapeamento dessas estruturas. Os resultados obtidos são preliminares, mas indicam que a utilização de técnicas de aprendizado de máquina associadas a dados de sensoriamento remoto é uma abordagem promissora para o mapeamento e o monitoramento de pequenos açudes na região semiárida do Brasil.

AGRADECIMENTOS – Os autores agradecem ao CNPq pelo apoio e fomento a pesquisa através da bolsa 314827/2025-0 e projeto 446157/2024-4.

REFERÊNCIAS

BEM, P.P.D.; DE CARVALHO JÚNIOR, O.A.; CARVALHO, O.L.F.; GOMES, R.A.T.; GUIMARÃES, R.F.; PIMENTEL, C.M.M. (2021). “Irrigated rice crop identification in Southern Brazil using convolutional neural networks and Sentinel-1 time series”. Remote Sensing Applications: Society and Environment 24, 100627.

DE ARAÚJO, J.C.; BRONSTERT, A. (2016). “A method to assess hydrological drought in semi-arid environments and its application to the Jaguaribe River basin, Brazil”. Water International 41(2), pp. 213–230.

DE BRITO FONTENELE, S.; MENDONÇA, L.A.R.; DE ARAÚJO, J.C.; SANTIAGO, M.M.F.; DE BRITO GONÇALVES, J.Y. (2014). “Relationship between hydrogeological parameters for data-scarce regions: the case of the Araripe sedimentary basin, Brazil”. Environmental Earth Sciences 71(2), pp. 885–894.

LATIF, G.; ABDELHAMID, S.E.; MALLOUHY, R.E.; ALGHAZO, J.; KAZIMI, Z.A. (2022). “Deep Learning Utilization in Agriculture: Detection of Rice Plant Diseases Using an Improved CNN Model”. Plants 11, 2230.

LIMA, D.M.; DA PAZ, A.R.; XUAN, Y.; PICCILLI, D.G.A. (2024). “Incorporating spatial variability in surface runoff modeling with new DEM-based distributed approaches”. Computational Geosciences 28(6), pp. 1331–1348.

MARENGO, J.A.; CUNHA, A.P.M.; NOBRE, C.A.; RIBEIRO NETO, G.G.; MAGALHÃES, A.R.; TORRES, R.R.; ÁLVALA, R.C. (2020). “Assessing drought in the drylands of northeast Brazil under regional warming exceeding 4°C”. Natural Hazards 103(2), pp. 2589–2611.

PEREIRA, B.S.; UCHÔA, J.G.S.; FREITAS, G.S.; MEIRA NETO, A.A.; ANACHE, J.A.; WENDLAND, E.C.; MEDEIROS, P.H. (2025). “Hydrological heritage: A historical exploration of human–water dynamics in northeast Brazil”. Hydrological Sciences Journal 70(4), pp. 546–558.

RIBEIRO NETO, G.G.; MELSEN, L.A.; MARTINS, E.S.; WALKER, D.W.; VAN OEL, P.R. (2022). “Drought cycle analysis to evaluate the influence of a dense network of small reservoirs on drought evolution”. Water Resources Research 58(1), e2021WR030799.

RIBEIRO NETO, G.G.; MELSEN, L.A.; COSTA, A.C.; WALKER, D.W.; CAVALCANTE, L.; KCHOUK, S.; VAN OEL, P.R. (2024). “Clash of drought narratives: A study on the role of small reservoirs in the emergence of drought impacts”. Earth's Future 12(7), e2023EF004311.

SOUZA, W.D.O.; REIS, L.G.D.M.; RUIZ-ARMENTEROS, A.M.; VELEDA, D.; RIBEIRO NETO, A.; FRAGOSO JR., C.R.; MONTENEGRO, S.M.G.L. (2022). “Analysis of environmental and atmospheric influences in the use of SAR and optical imagery from Sentinel-1, Landsat-8, and Sentinel-2 in the operational monitoring of reservoir water level”. *Remote Sensing* 14(9), 2218.

WANG, Y.; LI, Z.; ZENG, C.; XIA, G.-S.; SHEN, H. (2020). “An Urban Water Extraction Method Combining Deep Learning and Google Earth Engine”. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13, pp. 768–781.