

A Deep Transfer Learning Strategy for Identifying Photovoltaic Panels Mounted on Building Rooftops from Orthophotos

Uma Estratégia de Transferência de Aprendizado Profundo para Identificação de Painéis Fotovoltaicos Instalados em Telhados de Edificações a partir de Ortofotos

Marco A. M. Freitas Junior^{1*}, Ricardo A. S. Fernandes¹

Resumo: This work presents an analysis of orthophotos, which were used to train a predictive model based on a Convolutional Neural Network with the objective of identifying the presence of photovoltaic panels on building rooftops. To this end, the dataset named as Kortviseren, from Denmark, was used, which contains images of rooftops with and without photovoltaic panels. From this dataset, a convolutional neural network ResNet18 was employed, whose synaptic weights were adjusted through transfer learning. During the validation process, it was possible to observe a f1-score of approximately 98,7%. The results demonstrated the feasibility of the proposed approach for identifying photovoltaic panels mounted on building rooftops and, consequently, supporting power distribution utilities with the inspection of these consumer units that have the capacity to produce electricity.

Keywords: photovoltaic panels — identification — orthophoto — convolutional neural network

Resumo: Este trabalho apresenta uma análise de ortofotos, as quais foram usadas para treinar um modelo preditivo baseado em uma Rede Neural Convolucional com o objetivo de identificar a presença de painéis fotovoltaicos em telhados de edificações. Empregou-se o conjunto de dados denominado *Kortviseren*, da Dinamarca, o qual contém imagens de telhados com e sem a presença de painéis fotovoltaicos. A partir desse conjunto de dados, empregou-se a rede neural convolucional ResNet18, em que seus pesos sinápticos foram ajustados por meio de transferência de aprendizado. Durante o processo de validação, foi possível observar uma taxa de *f1-score* de aproximadamente 98,7%. Os resultados demonstraram a viabilidade da abordagem proposta para a identificação dos painéis fotovoltaicos instalados nos telhados das edificações e, consequentemente, auxiliar as concessionárias de distribuição de energia elétrica na inspeção dessas unidades consumidoras que têm a capacidade de produzir energia elétrica.

Palavras-Chave: painéis fotovoltaicos — identificação — ortofoto — redes neurais convolucionais

¹ Departamento de Engenharia Elétrica e de Computação, Universidade de São Paulo (USP), Brasil

*Corresponding author: marco.defreitas@usp.br

DOI: <http://dx.doi.org/10.22456/2175-2745.150976> • Received: 15/10/2025 • Accepted: 03/11/2025

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introdução

A geração solar fotovoltaica, segundo dados disponibilizados pela Empresa de Pesquisa Energética (EPE), foi a fonte que mais cresceu no Brasil nos últimos quatro anos [1], com destaque especial para as modalidades de mini e micro geração distribuída. Segundo dados da Agência Internacional de Energia, as instalações de fontes renováveis mais que dobraram em dois anos, partindo de menos de 400 GWp em 2022, para mais de 800 GWp em 2024, sendo as plantas solares responsáveis por 60% da energia gerada pelas novas instalações [2]. Dados da EPE mostram que o Brasil reflete essa tendência, saindo de 30 TWh para 70 TWh de geração solar fotovoltaica anual

no mesmo período [3]. Portanto, pode-se inferir que a quantidade de unidades consumidoras com instalações de painéis fotovoltaicos em telhados cresceu nos últimos anos no Brasil, trazendo consigo dificuldades para as concessionárias de distribuição de energia elétrica em relação à inspeção em termos da regularidade dessas instalações.

Portanto, o processamento de ortofotos mostra-se conveniente no âmbito dessas inspeções, como ferramenta de suporte à decisão. Nesse sentido, vale mencionar que o processamento de imagens geradas por satélite não é uma tarefa trivial. Diversas áreas do conhecimento têm se beneficiado das técnicas de processamento de imagens de satélite. Entretanto, nos últimos anos, grande parte desses trabalhos têm focado no uso de Re-

des Neurais Convolucionais (CNNs, do inglês *convolutional neural network*).

Em [4], os autores compararam o desempenho de seis tipos de classificadores no mapeamento de zonas urbanas, dentre as quais a CNN apresentou o melhor desempenho. Em [5], diversas CNNs foram treinadas para identificar as várias categorias de intensidade de ciclones tropicais, conseguindo uma acurácia de 86,0% com seu modelo mais eficiente, que possuía 14 camadas. O sistema de categorização apresentado em [6] utilizou uma CNN com 12 camadas para identificar formações ciclônicas, obtendo um *f1-score* de 92,3%. Já o trabalho de [7] utilizou uma combinação de dois tipos de CNNs, sendo uma delas uma rede residual, para estimar a altura da copa da vegetação em zonas remotas da tundra russa. Todos os trabalhos listados comprovam a versatilidade e a robustez das CNNs no âmbito do processamento de imagens de satélite.

No âmbito da energia solar, os autores de [8] realizaram uma revisão aprofundada sobre as soluções disponíveis de visão computacional para a detecção de falhas em painéis fotovoltaicos, destacando os avanços de técnicas e comparando o desempenho de diversos algoritmos. Em [9], foi realizada a comparação de desempenho de cinco diferentes tipos de modelos YOLO, com a finalidade de detectar pontos quentes em painéis fotovoltaicos. Modelos YOLO também foram aplicados em [10], em que foi proposta a criação de uma arquitetura executada em nuvem para a detecção de falhas em tempo real. Dentre as falhas detectadas, destacam-se trincas, problemas de soldagem e danos a células fotovoltaicas.

Considerando o desempenho das CNNs e a necessidade de ferramentas de suporte à decisão por parte das concessionárias de distribuição de energia elétrica, esse artigo propõe uma abordagem que visa identificar painéis de mini e micro geração distribuída em ambientes urbanos, especificamente em telhados de prédios comerciais e residenciais utilizando ortofotos geradas por satélite. Para esse fim, foi empregado o conjunto de dados Kortviseren[11], da Dinamarca, o qual fora submetido a um processo de limpeza e organização. Na sequência, houve uma etapa de pré-processamento das imagens. Assim, com o conjunto de dados devidamente condicionado, o modelo preditivo baseado em CNN ResNet18 foi treinado e validado utilizando uma estratégia de transferência de aprendizado.

O restante do artigo está dividido de forma que a Seção 2 apresenta a fundamentação teórica sobre redes neurais convolucionais, com destaque para o modelo residual. Na Seção 3 são apresentados os aspectos da metodologia proposta. Os resultados são discutidos na Seção 4. Por fim, as conclusões e os possíveis desdobramentos futuros deste trabalho são apresentados na Seção 5.

2. Redes Neurais Convolucionais

Essa seção traz uma fundamentação geral sobre as CNNs e uma explanação sobre as *Residual Networks* (ResNet), arquitetura de CNN que foi aplicada na classificação de ortofotos

no contexto desse trabalho.

2.1 Fundamentação

Desde os primeiros trabalhos de reconhecimento de escrita, aos sistemas modernos de detecção de tumores, as CNNs seguem evoluindo ao redor dos mesmos componentes fundamentais [12, 13, 14, 15]:

1. Camada Convolucional;
2. Camada de *Pooling*;
3. Camada Totalmente Conectada.

As camadas convolucionais são responsáveis por aplicar filtros, também chamados de *kernels*, que são matrizes de pequenas dimensões capazes de realizar a varredura da imagem, aplicando transformações numéricas. Essas camadas se especializam em destacar características da imagem, como contornos, formas geométricas ou padrões de cores. A Figura 1 mostra um exemplo de imagem do conjunto de dados utilizado, sem nenhum tipo de processamento, enquanto a Figura 2 mostra o resultado de várias etapas de convolução.



Figura 1. Imagem extraída do conjunto de dados Kortviseren.

As camadas de *pooling* são responsáveis por priorizar informação. Elas são treinadas para identificar quais saídas das camadas de convolução são mais importantes no problema de classificação proposto, o que permite realizar a redução espacial do modelo, ou seja, reduzir o número de parâmetros que precisam ser ajustados e, consequentemente, acelerar o processo de treinamento sem sacrificar desempenho [16]. As camadas totalmente conectadas, são camadas de um classificador tradicional, como um *Perceptron feed-forward*, que recebem os dados processados e priorizados das camadas anteriores [12, 13, 15]. É nesse nível que a CNN classifica a imagem a ser processada.

2.2 ResNet

As *Residual Networks* correspondem a uma arquitetura de CNN que utiliza um método de treinamento denominado aprendizagem residual. Esse método foi desenvolvido para

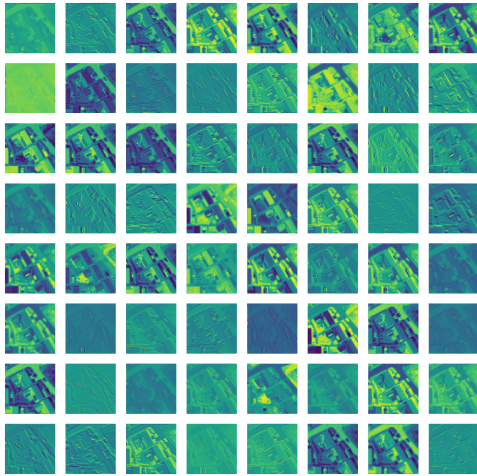


Figura 2. Demonstração das saídas de uma camada de convolução ao processar a imagem do conjunto de dados Kortviseren.

endereçar o problema de degradação de função de ativação que afeta redes profundas, conhecido com *gradient vanishing* [17]. Esse problema limitava a ampliação da CNN, pois, a partir de um certo nível, cada camada adicional degradava o desempenho do classificador. Para tanto, a ResNet utiliza *skip connection*, que são conexões sinápticas que saltam camadas, como representado na Figura 3. As *skip connections* permitem que neurônios chave para o treinamento da rede levem informação para um ponto numa camada posterior do modelo, saltando a hierarquia tradicionalmente imposta pelas camadas de neurônios.

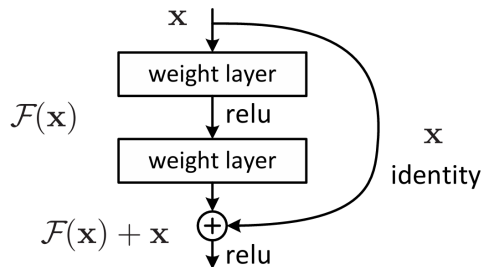


Figura 3. Representação do bloco fundamental do processo de aprendizagem residual [18].

O modelo ResNet18 é considerado o mais simples dessa arquitetura, contando com apenas 18 camadas de processamento, sendo 17 camadas convolucionais, seguida de uma camada totalmente conectada [18]. As camadas convolucionais são projetadas com matrizes de convolução de 3×3 , progredindo de 64 filtros nas primeiras camadas, para 512 nas últimas. A Figura 4 representa a estrutura desse modelo. Como o desempenho de redes profundas também está atrelado ao tamanho do conjunto de dados utilizado [19, 20], esse modelo é indicado para aplicações com conjuntos de dados menores, por ser o mais simples de treinar. Vale ressaltar que, mesmo que se trate do modelo mais simples da arquitetura,

o ResNet18 ainda é um modelo de aprendizagem profunda, com mais de 11 milhões de parâmetros treináveis.

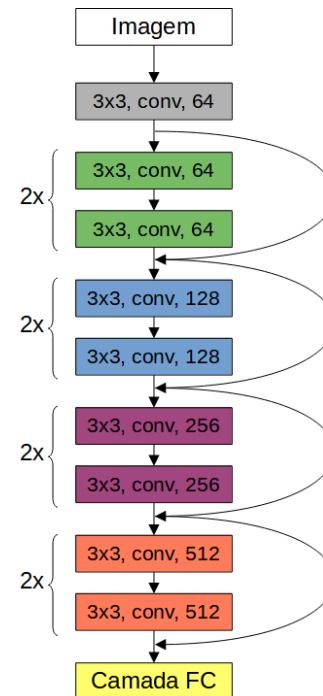


Figura 4. Representação da estrutura da CNN ResNet18.

Vale mencionar que a ResNet18 utilizada nesse trabalho foi previamente treinada a partir da base de dados ImageNet [21]. Essa base de dados contém mais de 1 milhão de imagens de treinamento, em 1000 categorias distintas, a qual se tornou popular para estratégias de transferência de aprendizado. Portanto, é importante ressaltar que a ResNet18 empregada tem seus pesos sinápticos pré-ajustados para a ImageNet, os quais foram então apenas treinados para o domínio da aplicação, ou seja, para a identificação dos painéis fotovoltaicos.

3. Metodologia

A Figura 5 apresenta uma visão geral da metodologia proposta nesse artigo. As subseções que seguem explanam de forma detalhada cada uma das etapas que compõem a metodologia.

3.1 Conjunto de Dados

O conjunto de dados utilizado nesse artigo foi gerado manualmente a partir do Sistema de Informações Geográficas (SIG) Kortviseren, da Dinamarca, que apresenta ortofotos atualizadas de todo o país em domínio público. Foram extraídas 3025 fotos de satélite, das quais 1000 possuem edificações com painéis fotovoltaicos e as 2025 restantes não possuem. Tal conjunto de dados é composto por ortofotos de prédios com e sem a presença de painéis fotovoltaicos em seus telhados. Visto que o modelo preditivo necessita de uma ampla gama de dados de ambas as categorias para treiná-lo e validá-lo, houve a necessidade de realizar a preparação do conjunto de dados.

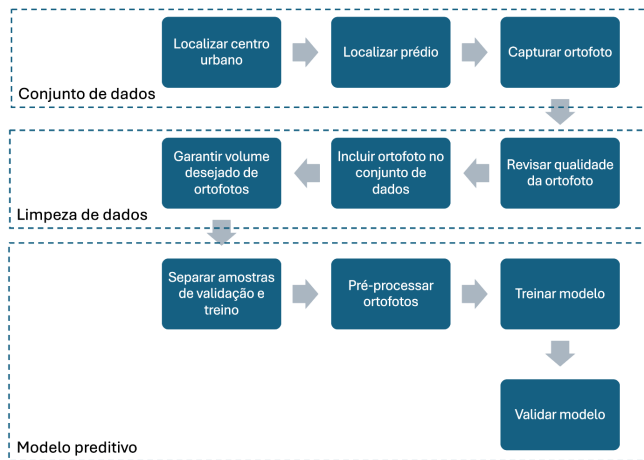


Figura 5. Visão geral da metodologia proposta.

A partir dessa organização do conjunto de dados, na sequência, realizou-se um estágio de limpeza dos dados.

3.2 Limpeza dos Dados

A princípio, foram utilizados os mesmos critérios de resolução e escala de [22], especificamente para garantir que todas as imagens tivessem uma resolução de 400×400 pixels, e que as fotos estivessem numa escala de representação de 0.1m/pixel, de forma que cada imagem representa uma área de 40×40 metros. Esse processo foi realizado de forma manual.

Em seguida, devido aos telhados dos prédios dinamarqueses apresentar uma diversidade de cores, como mostrado na Figura 6, buscou-se diversificar ao máximo a variedade de cores na captura das imagens, para que o conjunto de dados resultante fosse o mais abrangente possível.



Figura 6. Exemplos de edificações com telhados com cores variadas.

Especialmente para o conjunto de imagens sem a presença de painéis fotovoltaicos, foi importante garantir que nenhuma das edificações ao redor possuisse algum painel fotovoltaico

que estivesse na cena, ainda que parcialmente, ou seja, na borda da ortofoto. Por outro lado, para as imagens com a presença de painéis fotovoltaicos, houve a necessidade de garantir que todo o arranjo de painéis fotovoltaicos estivesse presente na imagem, bem como a edificação estivesse afastada da borda, conforme previamente ilustrado na Figura 1.

Vale destacar um exemplo de imagem que foi rejeitada nesse processo de limpeza de dados devido ao arranjo de painéis fotovoltaicos estar parcialmente exposto e na borda superior esquerda (Figura 7).



Figura 7. Exemplo de ortofoto rejeitada durante o processo de limpeza de dados.

Devido ao conjunto de dados não apresentar informações de coordenadas geográficas por questões óbvias de privacidade das unidades consumidoras, foi necessário garantir que não havia duplicidade para determinada edificação. Para tanto, empregou-se um método de varredura em um único sentido. Foram identificados centros urbanos, partindo do norte e indo rumo ao sul do país. E cada centro urbano era varrido partindo do extremo noroeste, rumo ao sudeste, como representado na Figura 8. Ao alcançar o fim do centro urbano, a coleta de dados seguia em um novo centro, em uma latitude mais ao sul. Os centros urbanos não foram revisitados. Esse processo foi repetido em sua integridade para os conjuntos com e sem painéis fotovoltaicos.

3.3 Modelo Preditivo

Conforme previamente mencionado, o modelo preditivo foi baseado em uma CNN com arquitetura ResNet18 pré-treinada sobre o conjunto de dados ImageNet. Sua implementação foi realizada em Python, por meio da biblioteca Torchvision, utilizando a estratégia de transferência de aprendizado [23].

Devido à arquitetura ResNet18, todas as imagens foram redimensionadas para 224×224 pixels, as quais são posteriormente convertidas em tensores e normalizadas utilizando os valores padrão recomendados pela documentação da biblioteca Torchvision (mean = [0.485, 0.456, 0.406], std = [0.229, 0.224, 0.225]).



Figura 8. Varredura das imagens visando reduzir duplicidade de edificações.

As imagens das duas classes rotuladas, “com painel” e “sem painel”, aleatoriamente distribuídas como amostras de treinamento e validação, considerando uma razão de 80% e 20%, respectivamente.

Os dados foram submetidos a 30 épocas de treinamento. Em cada época, o conjunto de dados de treinamento foi dividido em lotes de 32 imagens, e o modelo foi ajustado iterativamente através do processo de *forward pass*, cálculo da perda, retropropagação dos gradientes (*backpropagation*) e atualização dos parâmetros [24].

O modelo foi treinado utilizando a função de perda *Cross Entropy Loss* [25]. Para a atualização dos pesos foi empregado o otimizador Adam [26], com taxa de aprendizado inicial definida em 0,001, devido à sua boa estabilidade e convergência em problemas de visão computacional.

Foram registradas as métricas de perda média (*loss*) e acurácia a cada uma das 30 épocas, analisando tanto o conjunto de treinamento quanto o de validação. Essas métricas permitem monitorar a convergência do modelo, e podem dar indícios de que o modelo sofre de sobreajuste (*overfitting*).

4. Resultados e discussões

Os resultados obtidos demonstram a efetividade do classificador desenvolvido. A curva de perda e acurácia, apresentada na Figura 9 indica estabilização do desempenho a partir da vigésima época, evidenciando que o modelo alcançou convergência de forma consistente. A proximidade entre as curvas de treinamento e validação sugere ainda que não houve ocorrência significativa de sobreajuste, reforçando a capacidade de generalização do modelo.

O desempenho pode ser observado nas Tabelas 1 e 2, que apresentam, respectivamente, a matriz de confusão e o relatório de classificação do modelo sobre o conjunto de validação. Os resultados mostram um *f1-score* de 98,7%, derivado de valores elevados de precisão (*precision*) e revocação

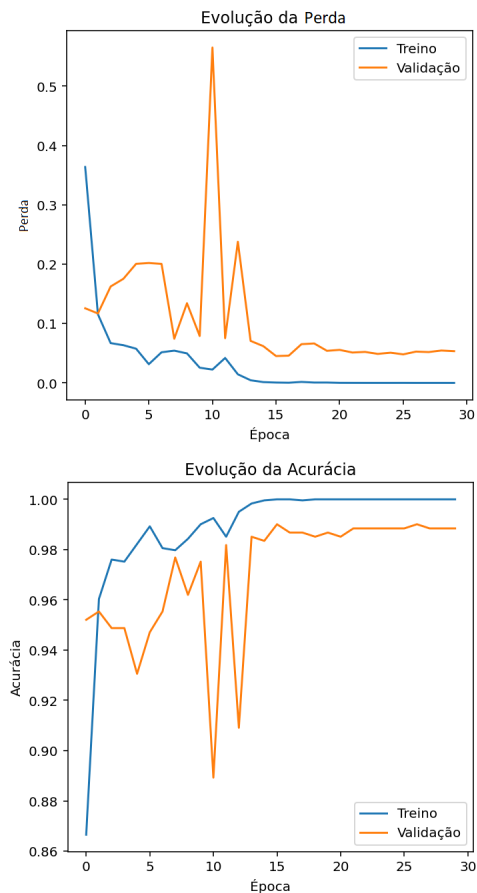


Figura 9. Curvas de perda e acurácia do treinamento e validação da CNN ResNet18 ao longo de 30 épocas.

(*recall*) para ambas as classes. Esse equilíbrio indica que o modelo não possui um viés para priorizar apenas uma das categorias, conseguindo identificar de forma confiável tanto telhados com painéis fotovoltaicos quanto os sem painéis.

Tabela 1. Matriz de confusão da validação da CNN ResNet18.

	Class. com placa	Class. sem placa
Imagem com placa	195	5
Imagem sem placa	2	403

A análise da matriz de confusão revela que os erros residuais se concentram em poucos casos. Dado o objetivo de identificar painéis fotovoltaicos, é importante analisar os falsos negativos gerados pelo modelo. Todos os cinco casos indicados na Figura 10 são de edificações com telhados na cor cinza, que apresenta pouco contraste com os painéis. É possível que a baixa resolução das imagens tenha prejudicado a capacidade do modelo de detectar a fronteira entre eles e o telhado.

Embora os resultados tenham demonstrado alto desempenho do modelo preditivo, é importante destacar as limitações relacionadas à origem do conjunto de dados. As ortofotos



Figura 10. Edificações com painéis fotovoltaicos que não foram identificadas pela CNN ResNet18.

Tabela 2. Relatório de classificação da validação da CNN ResNet18

	<i>Precision</i>	<i>Recall</i>	<i>f1-score</i>
com placa	0.9898	0.9750	0.9824
sem placa	0.9877	0.9951	0.9914
Acurácia	0.9884		

utilizadas foram obtidas a partir do cliente SIG Kortviseren, que fornece imagens da Dinamarca. Essa escolha se deu pela ausência de um ortomosaico de livre acesso que apresente todo o território nacional e pela disponibilidade pública do *Kortviseren*, publicado sob licença *Creative Commons*. Entretanto, deve-se considerar que existem diferenças relevantes entre telhados dinamarqueses e brasileiros no que diz respeito a arquitetura e urbanismo, como materiais de cobertura, inclinação das superfícies, presença de claraboias, jardins e vegetação específicos, além das próprias dimensões das edificações. Essas diferenças podem impactar a capacidade de generalização do modelo quando aplicado a imagens locais.

Adicionalmente, destaca-se que o modelo atual foi treinado para categorização de ortofotos representando áreas de 40×40 metros, centradas em uma edificação. Ou seja, ele não foi projetado para varrer territórios de grandes dimensões, como cidades, bairros ou mesmo bairros completos. A metodologia proposta também não faz a identificação do polígono das edificações, ele apenas identifica a presença dos painéis fotovoltaicos nas edificações representadas na

imagem.

Outra limitação relevante é a seleção e limpeza de imagens, que foi realizada manualmente neste trabalho. A utilização desse modelo em aplicações práticas exigirá a automatização dessas tarefas, possivelmente pela integração com uma plataforma de SIG.

Finalmente, a sintonia fina de hiperparâmetros foi feita de forma a otimizar o *f1-score*, sem priorizar nenhuma das categorias. Esses parâmetros devem ser revisitados caso haja um critério adicional na aplicação do modelo, como a eliminação de falsos positivos ou falsos negativos.

5. Conclusões

Este trabalho demonstrou a viabilidade de uma metodologia baseada em redes neurais convolucionais, mais especificamente da arquitetura ResNet18 no âmbito da identificação de painéis fotovoltaicos em imagens de satélite. A utilização de uma estratégia de transferência de aprendizado, a partir de modelos pré-treinados no conjunto de dados ImageNet, mostrou-se eficaz e alcançou um desempenho elevado. Os resultados obtidos na validação indicaram que a abordagem proposta apresenta potencial para auxiliar em iniciativas de mapeamento de unidades consumidoras que tenham painéis fotovoltaicos instalados em seus telhados, atuando como uma ferramenta de suporte à decisão de concessionárias de distribuição de energia elétrica. Entretanto, como o modelo foi treinado apenas com imagens da Dinamarca, sua capacidade de generalização para outras regiões ainda demanda

investigação, configurando uma limitação do presente estudo.

Como perspectivas de trabalhos futuros, destacam-se a adaptação do modelo preditivo para análises de imagens em escala ampliada, abrangendo bairros ou cidades inteiras; o treinamento do modelo com ortofotos obtidas de diferentes regiões, como áreas urbanas brasileiras, de forma a avaliar e ampliar a capacidade de generalização; o refinamento de hiperparâmetros ou mesmo a utilização de arquiteturas diferentes de CNNs; e a exploração de estratégias distintas de transferência de aprendizado.

Agradecimentos

Esse trabalho foi apoiado pela Pró-Reitoria de Pesquisa e Inovação (PRPI) [22.1.09345.01.2] - Universidade de São Paulo e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) [305215/2022-1].

Contribuição dos Autores

Marco A. M. Freitas Junior contributed with the creation of the dataset, the implementation and training of the CNN models and processing the results. Ricardo A. S. Fernandes contributed with the definition of work steps, the definition of the performance validation criteria and the revision and proofreading of the article.

Referências

- [1] MICRO e Minigeração Distribuída Baterias Atrás do Medidor. https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-894/Caderno_MGD_Baterias_2035_2025.07.31.pdf, 2025.
- [2] STRATEGIC PV Analysis and Outreach. <https://iea-pvps.org/snapshot-reports/snapshot-2025>, 2025. 26 p.
- [3] ANUÁRIO Estatístico de Energia Elétrica. <https://dashboard.epe.gov.br/apps/anuario-livro/>, 2025.
- [4] MOHAMMADI, M.; SHARIFI, A. Evaluation of convolutional neural networks for urban mapping using satellite images. *Journal of the Indian Society of Remote Sensing*, Springer, v. 49, n. 9, p. 2125–2131, 2021.
- [5] WANG, C. et al. Tropical cyclone intensity estimation from geostationary satellite imagery using deep convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, v. 60, p. 1–16, 2022.
- [6] JENA, K. K. et al. Deep convolutional network based machine intelligence model for satellite cloud image classification. *Big Data Mining and Analytics*, v. 6, n. 1, p. 32–43, 2023.
- [7] ILLARIONOVA, S. et al. Estimation of the canopy height model from multispectral satellite imagery with convolutional neural networks. *IEEE Access*, v. 10, p. 34116–34132, 2022.
- [8] POLYMERPOULOS, I. et al. Enhancing solar plant efficiency: A review of vision-based monitoring and fault detection techniques. *Technologies*, MDPI, v. 12, n. 10, p. 175, 2024.
- [9] VOURKOS, E. G. et al. Improving photovoltaics hotspot and dirt accumulation detection with computer vision: A systematic yolo model comparison based on field aerial images. *Sustainable Energy Technologies and Assessments*, Elsevier, v. 78, p. 104355, 2025.
- [10] WANG, H. et al. Novel cloud-edge collaborative detection technique for detecting defects in pv components, based on transfer learning. *Energies*, MDPI, v. 15, n. 21, p. 7924, 2022.
- [11] HØJSTED, P. *Kortviseren: a Danish map portal for topography, culture and history*. 2017. <https://kortviseren.dk/>. Front-end GIS client based on OpenLayers, displaying WMS services and historical maps.
- [12] GU, J. et al. Recent advances in convolutional neural networks. *Pattern Recognition*, v. 77, p. 354–377, 2018. ISSN 0031-3203. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0031320317304120>.
- [13] TAYE, M. M. Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. *Computation*, v. 11, n. 3, 2023. ISSN 2079-3197. Disponível em: <https://www.mdpi.com/2079-3197/11/3/52>.
- [14] SHARMA, P. et al. A survey on cancer detection via convolutional neural networks: Current challenges and future directions. *Neural Networks*, v. 169, p. 637–659, 2024. ISSN 0893-6080. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0893608023006287>.
- [15] KRICHEN, M. Convolutional neural networks: A survey. *Computers*, v. 12, n. 8, 2023. ISSN 2073-431X. Disponível em: <https://www.mdpi.com/2073-431X/12/8/151>.
- [16] GALANIS, N.-I. et al. Convolutional neural networks: A roundup and benchmark of their pooling layer variants. *Algorithms*, v. 15, n. 11, 2022. ISSN 1999-4893. Disponível em: <https://www.mdpi.com/1999-4893/15/11/391>.
- [17] HU, Z.; ZHANG, J.; GE, Y. Handling vanishing gradient problem using artificial derivative. *IEEE Access*, v. 9, p. 22371–22377, 2021.
- [18] HE, K. et al. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 770–778.
- [19] FENG, Q. et al. When do gans replicate? on the choice of dataset size. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. [S.l.: s.n.], 2021. p. 6701–6710.

- [20] BARBEDO, J. G. A. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Computers and Electronics in Agriculture*, v. 153, p. 46–53, 2018. ISSN 0168-1699. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168169918304617>.
- [21] DENG, J. et al. ImageNet: A Large-Scale Hierarchical Image Database. In: *CVPR09*. [S.l.: s.n.], 2009.
- [22] KASMI, G. et al. A crowdsourced dataset of aerial images with annotated solar photovoltaic arrays and installation metadata. v. 10, n. 1, p. 59, 2023. Disponível em: <https://doi.org/10.1038/s41597>.
- [23] DESAI, C. Image classification using transfer learning and deep learning. *International Journal of Engineering and Computer Science*, Valley International, v. 10, n. 9, p. 25394–25398, 2021.
- [24] LECUN, Y. A. et al. Efficient backprop. In: _____. *Neural Networks: Tricks of the Trade: Second Edition*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 9–48. ISBN 978-3-642-35289-8. Disponível em: https://doi.org/10.1007/978-3-642-35289-8_3.
- [25] LI, L.; DOROSLOVAČKI, M.; LOEW, M. H. Approximating the gradient of cross-entropy loss function. *IEEE Access*, v. 8, p. 111626–111635, 2020.
- [26] KINGMA, D. P.; BA, J. *Adam: A Method for Stochastic Optimization*. 2017. Disponível em: <https://arxiv.org/abs/1412.6980>.