



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

DÉBORA FERREIRA DE BARROS

**APLICAÇÃO DE REDES NEURAS CONVOLUCIONAIS NA
DETECÇÃO DE ASSENTAMENTOS PRECÁRIOS EM JOÃO
PESSOA**

CAMPINA GRANDE - PB

2023

DÉBORA FERREIRA DE BARROS

**APLICAÇÃO DE REDES NEURAIS CONVOLUCIONAIS NA
DETECÇÃO DE ASSENTAMENTOS PRECÁRIOS EM JOÃO
PESSOA**

**Trabalho de Conclusão de Curso
apresentado ao Curso Bacharelado em
Ciência da Computação do Centro de
Engenharia Elétrica e Informática da
Universidade Federal de Campina
Grande, como requisito parcial para
obtenção do título de Bacharela em
Ciência da Computação.**

Orientador: Professor Dr. Eanes Torres Pereira.

CAMPINA GRANDE - PB

2023

DÉBORA FERREIRA DE BARROS

**APLICAÇÃO DE REDES NEURAIS CONVOLUCIONAIS NA
DETECCÃO DE ASSENTAMENTOS PRECÁRIOS EM JOÃO
PESSOA**

**Trabalho de Conclusão de Curso
apresentado ao Curso Bacharelado em
Ciência da Computação do Centro de
Engenharia Elétrica e Informática da
Universidade Federal de Campina
Grande, como requisito parcial para
obtenção do título de Bacharela em
Ciência da Computação.**

BANCA EXAMINADORA:

**Professor Dr. Eanes Torres Pereira
Orientador – UASC/CEEI/UFCG**

**Professor Dr. Wilkerson de Lucena Andrade
Examinador – UASC/CEEI/UFCG**

**Professor Tiago Lima Massoni
Professor da Disciplina TCC – UASC/CEEI/UFCG**

Trabalho aprovado em: 14 de FEVEREIRO de 2023.

CAMPINA GRANDE - PB

ABSTRACT

Due to population growth and the expansion of informal settlements, it is necessary to monitor and map these places so that public policies can be developed aimed at solving the precarious nature present in these spaces. Some of the current solutions involve image classification based on machine learning algorithms, however, those present in the state of the art require the extraction of many features, which takes a lot of time and generates a large amount of parameters that need to be processed by the algorithms. This work presents the use of a Convolutional Neural Network, a U-Net with Inception ResNet-V2, as a solution for the feature extraction automation and parameter reduction in satellite images, focusing on the city of João Pessoa, in Paraíba, together with the segmentation of the images aiming at the detection and classification of precarious settlements in urban spaces. The model was evaluated using the Jaccard and Dice coefficients, which presented respectively 53% and 69%, in the test dataset.

Aplicação de Redes Neurais Convolucionais na Detecção de Assentamentos Precários em João Pessoa

Débora Ferreira de Barros

Unidade Acadêmica de Sistemas e Computação
Universidade Federal de Campina Grande
Campina Grande, Paraíba, Brasil
debora.barros@ccc.ufcg.edu.br

Eanes Torres Pereira

Unidade Acadêmica de Sistemas e Computação
Universidade Federal de Campina Grande
Campina Grande, Paraíba, Brasil
eanes@computacao.ufcg.edu.br

RESUMO

Devido ao crescimento populacional e a expansão dos assentamentos informais faz-se necessário o monitoramento e mapeamento desses locais para que possam ser desenvolvidas políticas públicas visando a solução da precariedade característica presente nesses espaços. Algumas das soluções atuais envolvem classificação de imagem baseada em algoritmos de aprendizagem de máquina, entretanto, as presentes no estado da arte necessitam da extração de muitas características, o que demanda muito tempo e gera uma grande quantidade de parâmetros que precisam ser processados pelos algoritmos. Este trabalho apresenta o uso de uma rede neural convolucional, a U-Net com Inception ResNet-V2, como solução para a automação de extração de características e redução de parâmetros em imagens de satélite, com foco na cidade de João Pessoa, na Paraíba, junto com a segmentação das imagens visando a detecção e classificação de assentamentos precários nos espaços urbanos. O modelo foi avaliado utilizando os coeficientes Jaccard e Dice, que apresentaram respectivamente 53% e 69%, nos dados de teste¹.

Palavras-chave

Redes Neurais Convolucionais, Aprendizagem profunda, Segmentação de imagens; Assentamentos precários

¹ Os autores retêm os direitos, ao abrigo de uma licença Creative Commons Atribuição CC BY, sobre todo o conteúdo deste artigo (incluindo todos os elementos que possam conter, tais como figuras, desenhos, tabelas), bem como sobre todos os materiais produzidos pelos autores que estejam relacionados ao trabalho relatado e que estejam referenciados no artigo (tais como códigos fonte e bases de dados). Essa licença permite que outros distribuam, adaptem e evoluam seu trabalho, mesmo comercialmente, desde que os autores sejam creditados pela criação original.

1. INTRODUÇÃO

De acordo com o Programa das Nações Unidas para os Assentamentos Humanos, UN-Habitat, os indicadores de informalidade podem ser definidos pela falta de acesso à água potável, ao saneamento básico, à espaços adequados para vivência, à durabilidade e qualidade das construções e ao direito de posse [10]. Estima-se que no mundo cerca de 1 bilhão de pessoas viviam nestas condições no ano de 2018 e que no Brasil cerca de 27 milhões de pessoas moravam em assentamentos precários e moradias inadequadas, essas condições sendo causadas principalmente pelo rápido crescimento urbano não planejado [17][18]. Com o crescimento populacional e dos assentamentos precários, e com a falta de informações sobre essas áreas, o monitoramento e a elaboração de planos de melhorias pelas autoridades responsáveis são dificultados, surgindo a necessidade de mapeamento e detecção de tais locais para auxiliar na busca por soluções para esses espaços [6].

Uma das soluções para o mapeamento de áreas urbanas é a análise de imagens geradas por satélites, pois estes conseguem capturar com alta precisão os elementos terrestres [5]. Os métodos presentes no estado da arte para classificação de imagens com o intuito de identificar assentamentos precários, em muitos casos, se baseiam na utilização de algoritmos de aprendizagem de máquina como árvores de decisão, análises de estruturas espaciais, estruturais e contextuais como lacunaridade, histograma de gradientes, matriz de co-ocorrência de níveis de cinza (*Gray Level Co-occurrence Matrix - GLCM*) e detecção de linhas [4][7][13]. Entretanto, devido ao grande volume de parâmetros e ao tempo gasto na extração de características importantes das imagens, se faz necessário o uso de algoritmos de aprendizagem de máquina que consigam detectar características das imagens automaticamente e com alta eficácia [1].

A Rede Neural Convolucional (*Convolutional Neural Network - CNN*) é um dos algoritmos de aprendizagem profunda que vem sendo utilizado na classificação de imagens, pois consegue lidar com grande volume de dados, reduzindo parâmetros de entrada e extraindo características importantes independentemente da localização espacial dos elementos [2].

Este trabalho apresenta a utilização de uma CNN, a U-Net com uma Inception ResNet-V2 integrada, para treinamento de um modelo de classificação de imagens de satélite baseado em pixels como solução para a detecção automática de assentamentos precários na cidade de João Pessoa, na Paraíba, com a utilização de técnicas de pré-processamento para melhoria da qualidade das imagens e geração de novos dados sintéticos devido às limitações de quantidade e resolução das imagens originais.

2. TRABALHOS RELACIONADOS

Devido à quantidade de dados extraídos na análise de imagens de alta resolução, Bergado [3] e Mboga [9] sugeriram o uso da CNN como solução para a extração de características de imagens de satélite devido a sua automação, alta performance e capacidade de aprendizagem. Bergado [3] comparou a CNN com outros três métodos: (i) os canais vermelho, verde e azul (RGB) extraídos das imagens em conjunto com um modelo de superfície digital, (ii) a GLCM, e (iii) padrões binários locais com variação do tamanho do conjunto de treinamento. Em todas as variações de tamanho a CNN obteve um melhor desempenho quando comparada aos outros métodos.

Ulmas e Liiv [16] apresentaram um estudo do uso da CNN na classificação de mapeamento de cobertura de terra a nível de pixel em imagens de satélite. Um modelo ResNet-50 foi usado na tarefa de classificação, em seguida o modelo pré-treinado foi utilizado como codificador de um modelo U-Net na tarefa de segmentação. Patil [12], por sua vez, também apresentou uma combinação de dois modelos, utilizando um codificador Inception ResNet-V2 na estrutura de um modelo U-Net, criando blocos integrados das duas arquiteturas.

Os artigos observados foram utilizados como base para a construção da arquitetura utilizada neste trabalho devido ao potencial da CNN para a detecção de padrões em imagens de satélite de alta resolução, entretanto, utilizando-se neste, imagens de menor resolução no treinamento da rede em conjunto com técnicas de aumento de dados.

3. FUNDAMENTAÇÃO TEÓRICA

Esta seção apresenta os conceitos de algoritmos de aprendizagem de máquina necessários para o entendimento deste trabalho.

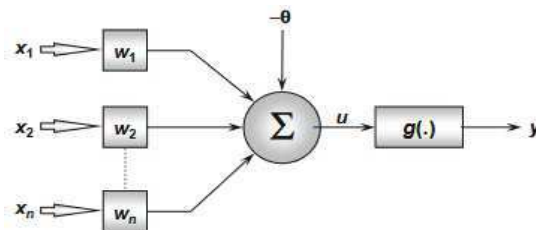
3.1 Rede Neural Artificial

Segundo Silva [15] a Rede Neural Artificial (*Artificial Neural Network* - ANN) é um modelo computacional inspirado no sistema nervoso dos seres vivos, com a capacidade de adquirir e armazenar conhecimento e que pode ser definido como um conjunto de unidades de processamento formadas por neurônios artificiais interligados por conexões, definidas como sinapses artificiais, representadas por vetores e matrizes de pesos sinápticos.

Os neurônios utilizados em ANNs, representados pela Figura 1, são não-lineares e recebem dados de entrada, representados pelo

conjunto $\{x_1, x_2, \dots, x_n\}$, que são multiplicados pelos pesos $\{w_1, w_2, \dots, w_n\}$, o resultado da multiplicação indica quais dados de entrada são mais relevantes. Em seguida o resultado da multiplicação é somado e um limiar de ativação é utilizado (θ), este que é uma variável que indica qual limiar o resultado produzido pela soma deve ter para gerar um valor de disparo para a saída do neurônio. O resultado (u) então é passado para a função de ativação $g(\cdot)$ cuja função é limitar a saída do neurônio dentro de um conjunto de valores, com (y) sendo o valor de saída do neurônio.

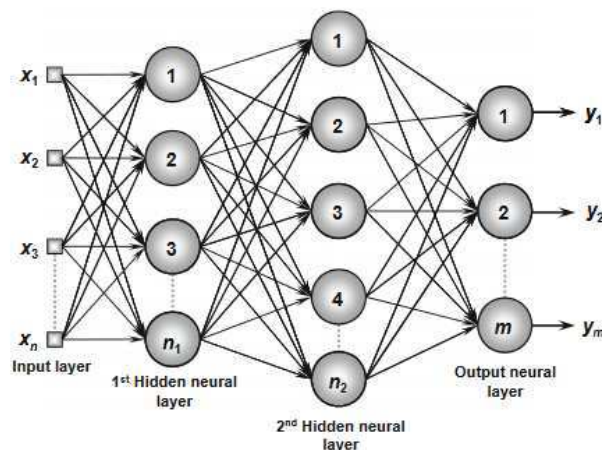
Figura 1. Representação do neurônio artificial



Fonte: Artificial Neural Network: A Practical Course (2017)

A arquitetura de uma ANN é constituída por três elementos: a camada de entrada, que é responsável por receber os dados, a camada intermediária, que é responsável pelo processamento dos dados e detecção de padrões, e a camada de saída, que é responsável pela apresentação dos dados gerados pela camada intermediária. A profundidade da rede neural é definida pela quantidade de camadas intermediárias. A Figura 2 representa uma ANN com duas camadas intermediárias.

Figura 2. Representação de uma rede neural artificial com duas camadas intermediárias



Fonte: Artificial Neural Network: A Practical Course (2017)

3.2 Rede Neural Convolutiva

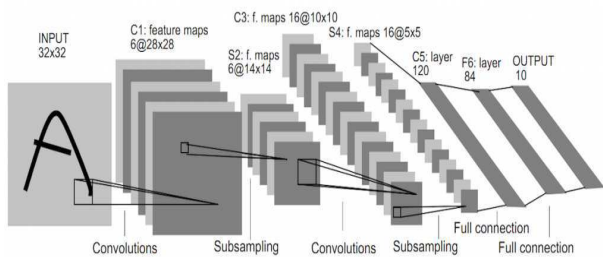
A CNN é um dos algoritmos de aprendizagem de máquina que fazem parte das Redes Neurais Profundas, estas que são ANNs de múltiplas camadas utilizadas principalmente para resolver tarefas de reconhecimento de padrões em imagem, cuja função é minimizar a quantidade de parâmetros de entrada da rede neural [11]. Segundo Albawi [2] a CNN possui os seguintes

elementos que a compõem: (i) a camada de convolução, que funciona como um filtro de extração de características em forma de matriz. Vários filtros podem ser aplicados a mesma camada ou a diferentes camadas de uma imagem, realizando a extração de várias características distintas. A camada de convolução possui dois principais parâmetros: o tamanho do passo, que tem como função delimitar a quantidade de passos que o filtro de convolução efetua na matriz, e o preenchimento, que tem como função controlar o tamanho da matriz de saída quando aplicado o filtro; (ii) a camada de não-linearização, que pode ser utilizada para ajustar ou remover o resultado gerado pelo passo de preenchimento da camada de convolução. Uma das funções de ativação mais utilizadas nesta camada é a função ReLU, que dada a entrada retorna o próprio valor de entrada se este for positivo, caso contrário retorna zero; (iii) a camada de agrupamento, cuja principal função é a redução de parâmetros que serão utilizados nas próximas camadas. O agrupamento máximo é o tipo mais comum utilizado, este particiona a imagem em sub-matrizes no qual o maior valor de cada uma é escolhido para a próxima camada. A posição da informação não é mantida com o agrupamento, sendo essa camada usada apenas para detectar as informações essenciais; e (iv) a camada completamente conectada, que é similar ao arranjo dos neurônios em uma ANN comum. As informações da última camada de agrupamento são passadas para a camada completamente conectada para serem treinadas em forma de um vetor gerado a partir da compressão das matrizes.

Uma das CNNs mais conhecidas é a LeNet-5 apresentada por Yann LeCun [8], representada na Figura 3.

Figura 3. Representação da Rede Neural

Convencional LeNet-5



Fonte: Gradient-Based Learning Applied to Document Recognition (1998)

3.3 U-Net e Inception ResNet-V2

A U-Net, um tipo de CNN proposta por Ronneberger [14], foi inicialmente designada para análise de imagens médicas. Consiste em uma rede que possui um caminho de contração, que possibilita a extração de características mais profundas das imagens, e outro de expansão, que possibilita a propagação de informações contextuais das imagens para camadas de maior resolução.

A Inception ResNet-V2, por sua vez, é uma CNN pré-treinada em uma base de dados chamada ImageNet, que tem como base uma arquitetura Inception com conexões residuais.

3.4 Segmentação de Imagens

A segmentação de imagens é uma técnica de visão computacional para classificação de imagens a nível de pixel, de forma que pequenas frações das imagens são analisadas individualmente e mapeadas para suas determinadas classes [12].

4. METODOLOGIA

Esta seção apresenta detalhes sobre a obtenção dos dados e suas delimitações, o tratamento dos dados, tecnologias e os algoritmos de aprendizagem de máquina utilizados.

4.1 Obtenção dos Dados

Para o treinamento e avaliação dos modelos foram utilizados dois conjuntos de imagens, o primeiro foi obtido através do site do Instituto Nacional de Pesquisa Espacial (INPE), que dispõe de imagens da cidade de João Pessoa, na Paraíba, capturadas pelo satélite CBERS-04A. Foi escolhida a imagem que possui cobertura de nuvem de 10%, a menor disponível no catálogo.

A delimitação dos dados, caracterizados como assentamentos precários, foi determinada pelas demarcações das Zonas Especiais de Interesse Social (ZEIS) disponibilizadas pela prefeitura de João Pessoa, e pelas demarcações dos Setores Especiais de Aglomerados Subnormais (SEAS) disponibilizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE). A Figura 4 representa as marcações de ZEIS, em vermelho, e as de SEAS, em roxo, na imagem original.

Figura 4. Demarcação de ZEIS e SEAS

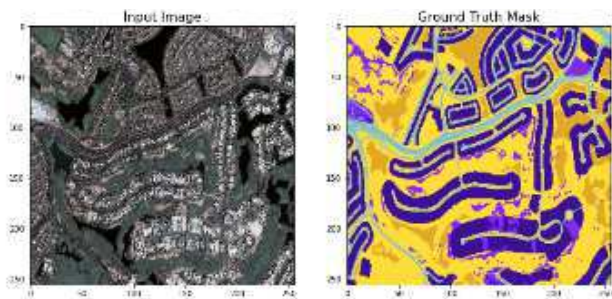


Fonte: Autor (2023)

O segundo conjunto de dados foi obtido através do satélite MBRSC da cidade de Dubai, disponibilizado publicamente através do site *Humans in the Loop*. Foram utilizadas as imagens e seus respectivos rótulos, que consistem em máscaras geradas por segmentação distribuídas em 6 classes, sendo elas

classificadas como: construções, representadas pela cor roxo-escuro, terrenos em roxo-claro, estradas em azul-claro, vegetação em amarelo-claro, água em amarelo-escuro, e áreas sem rótulo em cinza, como ilustradas na Figura 5.

Figura 5. Amostra dos dados de Dubai



Fonte: Humans in the Loop. Semantic Segmentation of Aerial Imagery (2020)

4.2 Pré-processamento

Para o conjunto de imagens de João Pessoa foram coletadas imagens da banda Pan, com resolução de 2 metros e da banda RGB com resolução de 8 metros. Utilizando o *software QGIS* foi feita a mesclagem das bandas para gerar uma nova imagem de composição colorida de melhor resolução.

Em seguida foram feitos recortes de tamanho 256 x 256 pixels na imagem original e na imagem gerada com as demarcações das ZEIS e SEAS. Para a marcação dos rótulos nas imagens originais foi utilizado a ferramenta de anotação *LabelMe*. Os rótulos em questão foram marcados considerando o seguinte critério: entre as ZEIS e SEAS que tiveram interseção, foram escolhidas as que abrangeram uma maior quantidade de área nas imagens, no caso das que não tiveram interseção tanto ZEIS quanto SEAS foram marcadas.

Para a criação dos rótulos nas máscaras, estas foram segmentadas em três tipos de classes: ZEIS em vermelho, SEAS em roxo, e áreas sem rótulo em preto, como mostra a Figura 6. No total 306 imagens com rótulos originais foram geradas, sendo divididas em 244 para treino e 62 para teste.

Figura 6. Amostra dos dados de João Pessoa



Fonte: Autor (2023)

Em seguida, foram utilizadas técnicas de aumento de dados nas imagens para a criação de novos dados sintéticos, com o objetivo de aumentar o conjunto de dados.

Rotações de 45, 90, 135, 180 e 270 graus, giros horizontais e verticais foram aplicados nas imagens e nas máscaras originais. E filtros de alteração de cor e de suavização foram aplicados nas imagens, gerando um total de 5856 de imagens de treinamento e 1488 de teste.

Para o conjunto de dados de Dubai, foram divididas 58 imagens para treino e 14 para teste. Foram aplicadas rotações de 90, 180 e 270 graus, giros horizontais e verticais nas imagens e nas máscaras. E filtros de alteração de cor, de suavização, de brilho e de contraste foram aplicados nas imagens, gerando um total de 1334 imagens para treino e 112 para teste. Todas as imagens geradas nesta etapa, incluindo as originais, também possuem tamanho padronizado de 256 x 256 pixels e canais de cores RGB. Um resumo dos dados pode ser visto na Tabela 1.

4.3 Treinamento da Rede

Os códigos para os experimentos foram implementados utilizando bibliotecas de aprendizagem de máquina e visão computacional como *Keras*, *Tensorflow* e *OpenCV*, e linguagem de programação *Python*. O *Google Colab* foi utilizado para a execução dos códigos.

Para o treinamento foram utilizados dois tipos de CNN. O primeiro é uma U-Net com caminhos de contração e expansão contendo camadas de convolução com função de ativação ReLU, com agrupamento máximo e camadas de *dropout*, que funcionam como regularizador da rede para evitar seu sobre-ajuste. O segundo tipo de CNN utilizado é também uma U-Net, porém que utiliza a Inception-ResNet-V2, com pesos pré-treinados da ImageNet, como camada de codificação. O modelo possui uma camada de decodificação e outra de convolução que usa funções de ativação ReLU, com função de ativação Softmax, esta que converte números reais em probabilidades, na camada de saída.

Nos dois modelos foram utilizados um otimizador Adam e função de perda de entropia cruzada. Como métricas de avaliação foram utilizados os coeficientes Jaccard (4) e Dice (5), estes que são utilizados para medir a similaridade entre os pixels dos conjuntos de máscaras geradas pela predição do modelo nos dados de teste e seus rótulos verdadeiros.

$$\text{Coeficiente Jaccard}(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (4)$$

$$\text{Coeficiente Dice}(A,B) = 2 * \frac{|A \cap B|}{|A| + |B|} \quad (5)$$

Os conjuntos A e B de (4) e (5) representam os pixels das máscaras verdadeiras e das geradas pelo modelo, classificados entre ZEIS, SEAS e áreas sem rótulo. A interseção dos conjuntos representa a sobreposição das áreas dos pixels, enquanto a união representa o número total de pixels dos dois conjuntos.

Os modelos foram treinados em lotes de tamanho 4 por 200 épocas, com condição de parada baseada na perda da validação. Os dados de treino foram divididos 80% para treinamento e 20% para validação do modelo.

5. RESULTADOS

As avaliações dos modelos treinados, apresentadas na Tabela 2, foram comparadas com as da U-Net modificada sugerida por Patil [12]. A U-Net treinada com os dados de João Pessoa obteve um coeficiente Jaccard de 47,43% e Dice de 64,34%, enquanto a U-Net treinada com os dados de Dubai apresentou um coeficiente Jaccard de 17,73% e Dice de 30,12%. A U-Net com Inception ResNet-V2 treinada com os dados de João Pessoa apresentou um coeficiente Jaccard de 53,41% e Dice de 69,63%, e a treinada com os dados de Dubai apresentou um coeficiente Jaccard de 55,98% e Dice de 71,78%. As figuras 7 e 8 apresentam os resultados das predições dos conjuntos de teste gerados pela U-Net com Inception ResNet-V2.

Utilizando os dados de Dubai e a U-Net modificada [12] como parâmetros de avaliação da sanidade da rede apresentada, os resultados das métricas de avaliação obtidos pelos modelos mostram que a U-Net com Inception ResNet-V2 obteve melhores resultados tanto com os dados das imagens de Dubai quanto com os de João Pessoa, se comparados com os resultados obtidos pela rede U-Net, o que mostra que a combinação das duas redes foi mais eficaz na extração e na aprendizagem de características de ZEIS e SEAS, mesmo em imagens que não possuem alta resolução.

Tabela 1. Conjunto de dados após pré-processamento

Base de dados	Imagens originais	Imagens geradas	Imagens para treino	Imagens para teste	Tamanho das imagens
João Pessoa	306	7344	5856	1488	256x256
Dubai	72	1446	1344	112	256x256

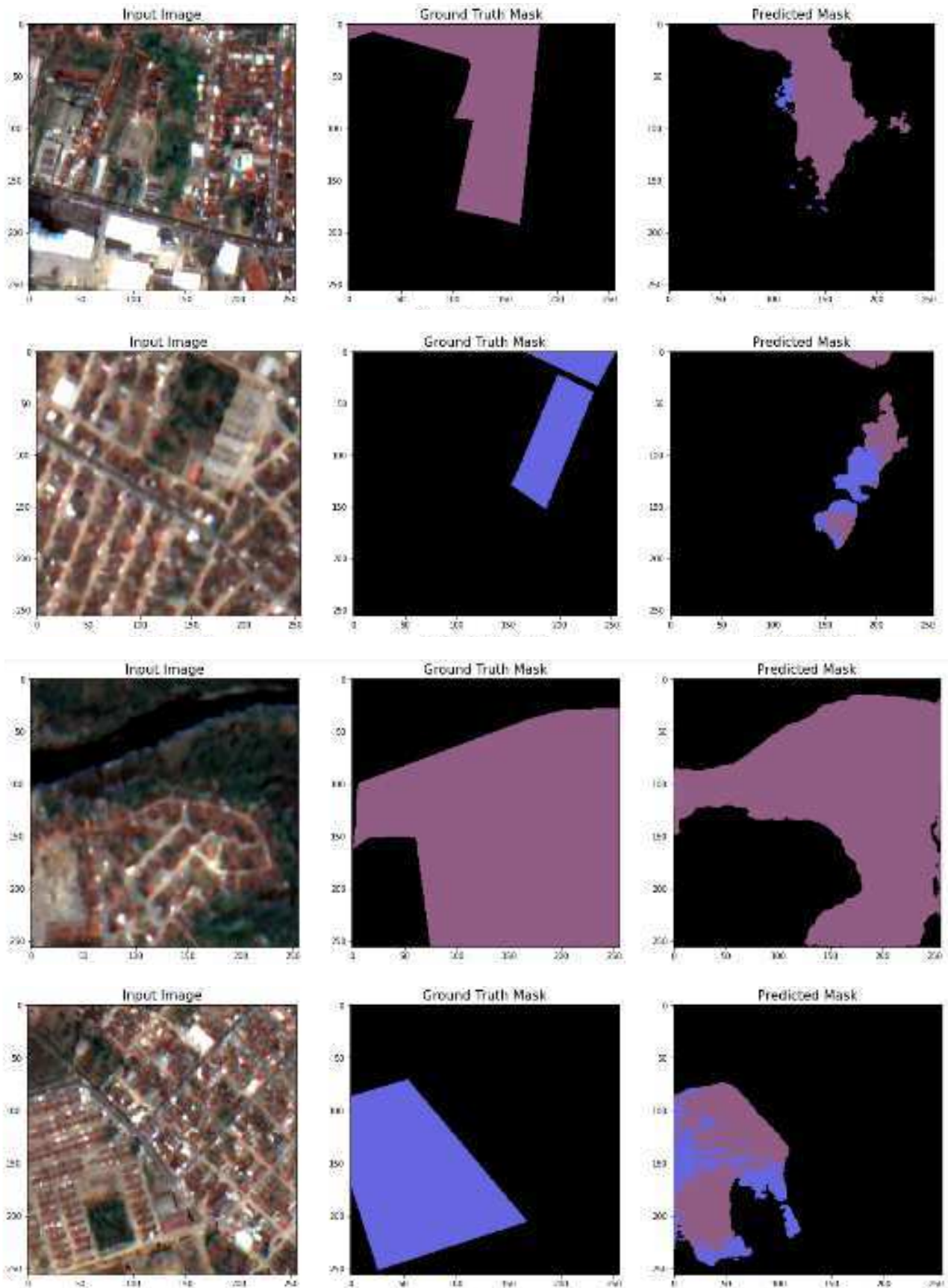
Fonte: Autor (2023)

Tabela 2. Resultados do treinamento e teste das redes

Modelo	Jaccard Treino	Jaccard Teste	Dice Treino	Dice Teste	Acurácia Treino	Acurácia Validação	Época de parada
U-Net - João Pessoa	51,81%	47,43%	67,84%	64,34%	80,87%	81,08%	65
U-Net - Dubai	15,97%	17,73%	27,51%	30,12%	42,31%	43,13%	69
U-Net com Inception ResNet-V2 - João Pessoa	78,22%	53,41%	87,61%	69,63%	92,32%	74,71%	21
U-Net com Inception ResNet-V2 - Dubai	76,36%	55,98%	86,47%	71,78%	91,58%	88,20%	11
U-Net modificada - Dubai (Patil [12])	-	-	≅ 87%	≅ 82%	≅ 92%	≅ 87%	42

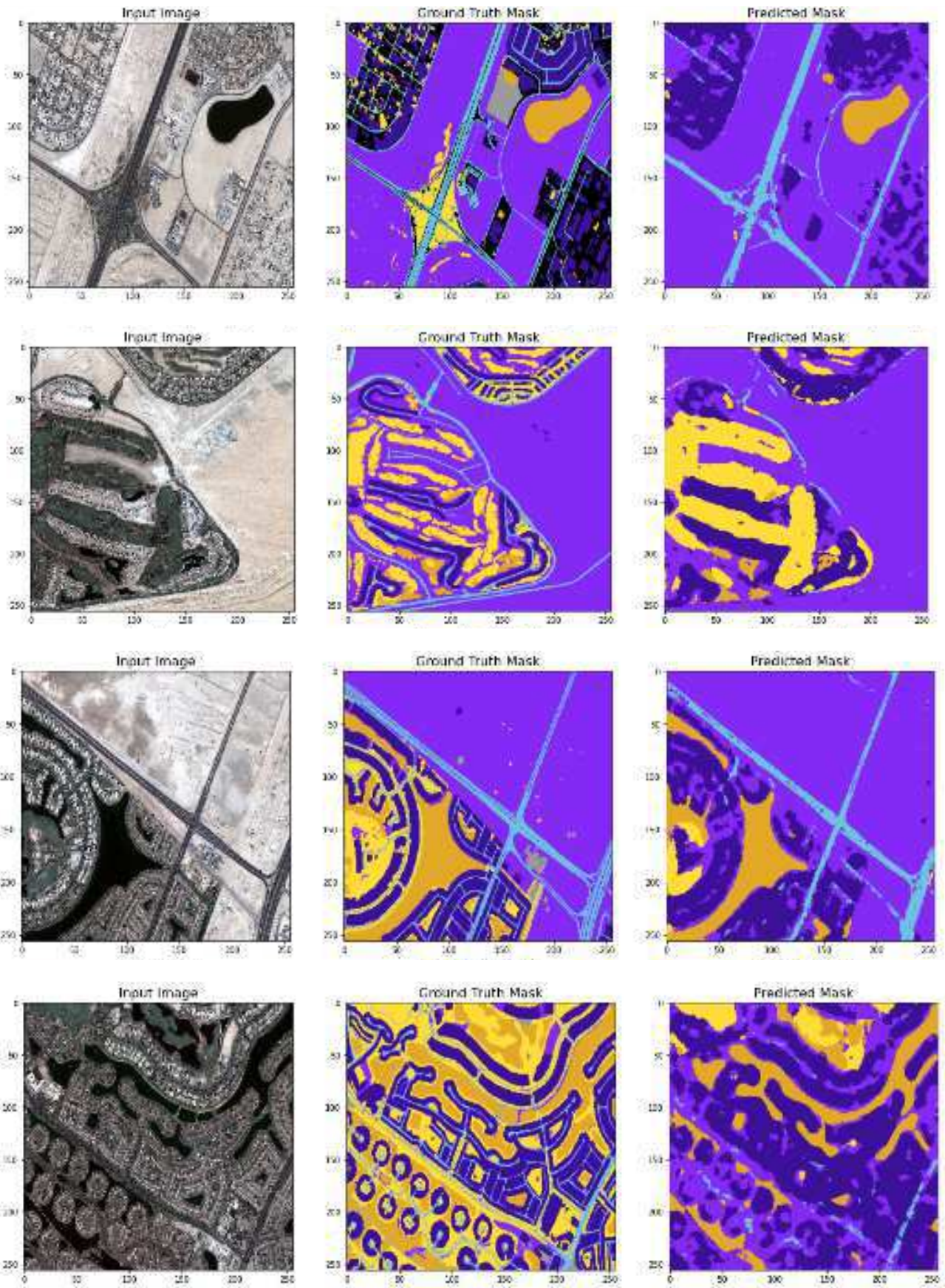
Fonte: Autor (2023)

Figura 7. Predições da U-Net com Inception ResNet-V2 com os dados de João Pessoa



Fonte: Autor (2023)

Figura 8. Predições da U-Net com Inception ResNet-V2 com os dados de Dubai



Fonte: Autor (2023)

6. CONCLUSÃO

Este artigo apresentou as Redes Neurais Convolucionais como alternativa de solução para o problema de eficácia na extração de características e detecção de padrões em imagens de satélite, com o intuito de mapear assentamentos precários nos meios urbanos. Foram utilizados dois tipos de CNN, um modelo U-Net e um modelo U-Net com Inception ResNet-V2, e a partir da avaliação das métricas e dos resultados das predições dos modelos, viu-se que o segundo modelo apresentou melhor desempenho nas predições tanto com os dados de João Pessoa quanto com os de Dubai. Os resultados indicam que as Redes Neurais Convolucionais são promissoras quanto à detecção de padrões e na segmentação de imagens de satélite e podem ser úteis em diversas áreas de interesse urbano, podendo auxiliar no planejamento de políticas públicas das cidades.

6.2 Limitações e Trabalhos Futuros

Algumas limitações foram encontradas com a construção dos experimentos. As imagens coletadas de João Pessoa e suas demarcações de ZEIS e SEAS dificultaram a aprendizagem do modelo de melhor desempenho por apresentarem uma menor resolução e demarcações de classes não muito bem definidas, como a mistura de vegetação e construções em uma única classe, se comparadas com os dados de Dubai.

Como sugestão para trabalhos futuros, espera-se que utilizando uma base de dados com melhor qualidade e mais divisões de classes possa-se obter um melhor desempenho nas predições das imagens de modo que seja possível utilizá-las como forma de auxiliar no mapeamento de novos assentamentos precários.

7. REFERÊNCIAS

- [1] AJAMI, Alireza; KUFFER, Monika; PERSELLO, Claudio; PFEFFER, Karin. Identifying a Slums' Degree of Deprivation from VHR Images Using Convolutional Neural Networks. *Remote Sensing*, vol. 11, ed. 11:1282, 2019.
- [2] ALBAWI, Saad; MOHAMMED, Tareq Abed; AL-ZAWI, Saad. Understanding of a convolutional neural network. 2017 International Conference on Engineering and Technology (ICET), p. 1-6, 2017.
- [3] BERGADO, John Ray; PERSELLO, Claudio; GEVAERT, Caroline. A deep learning approach to the classification of sub-decimeter resolution aerial images. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), p. 1516-1519, 2016.
- [4] GRAESSER, Jordan; CHERIVADAT, Anil; VATSAVAI, Ranga Raju; CHANDOLA, Varun; LONG, Jordan; BRIGHT, Eddie. Image Based Characterization of Formal and Informal Neighborhoods in an Urban Landscape. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, ed. 4, p. 1164-1176, 2012.
- [5] GUEGUEN, Lionel. Classifying Compound Structures in Satellite Images: A Compressed Representation for Fast Queries. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, ed. 4, p. 1803-1818, 2015.
- [6] HOFMANN, Peter; STROBL, Josef; BLASCHKE, Thomas; KUX, Hermann. Detecting informal settlements from QuickBird data in Rio de Janeiro using an object-based approach. *Lecture Notes in Geoinformation and Cartography*. ed. Springer, Berlin, p. 531-553, 2008.
- [7] KIT, Oleksandr; LÜDEKE, Matthias; RECKIEN, Diana. Texture-based identification of urban slums in Hyderabad, India using remote sensing data. *Applied Geography*, vol. 32, issue 2, p. 660-667, 2012.
- [8] LECUN, Yann *et al.* Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, vol. 86, ed. 11, p. 2278-2324, 1998.
- [9] MBOGA, Nicholus *et al.* Detection of Informal Settlements from VHR Satellite Images Using Convolutional Neural Networks. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), p. 5169-5172, 2017.
- [10] MONITORING SDG Indicator 11.1.1. The Urban SDG Monitoring Series, Issue 1, 2019. Disponível em: https://unhabitat.org/sites/default/files/2020/06/the_urban_sdg_monitoring_series_monitoring_sdg_indicator_11.1.1.pdf.
- [11] O'SHEA, Keiron; NASH, Ryan. An Introduction to Convolutional Neural Networks. ArXiv, 2015. Disponível em: <https://arxiv.org/abs/1511.08458>.
- [12] PATIL, Dhanishtha *et al.* Semantic Segmentation of Satellite Images using Modified U-Net, 2022 IEEE Region 10 Symposium (TENSYMP), p. 1-6, 2022.
- [13] RIBEIRO, Bárbara Maria Giacom. Mapping informal settlements using WorldView-2 imagery and C4.5 decision tree classifier. 2015 Joint Urban Remote Sensing Event (JURSE), p. 1-4, 2015.
- [14] RONNEBERGER, Olaf; FISCHER, Phillip; BROX, Thomas. U-Net: Convolutional Networks for Biomedical Image Segmentation, ArXiv, 2015. Disponível em: <https://arxiv.org/abs/1505.04597>
- [15] SILVA, Ivan Nunes da *et al.* Artificial Neural Networks: A Practical Course. ed. Springer, cap. 1-2, 2017.
- [16] ULMAS, Priit. LIIV, Innar. Segmentation of Satellite Imagery using U-Net Models for Land Cover Classification. ArXiv, 2020. Disponível em: <https://arxiv.org/abs/2003.02899>
- [17] UN-HABITAT: For a better urban future. UN-HABITAT, [201-]. Disponível em: <https://unhabitat.org/brazil>.
- [18] URBAN Population Living in Slums by Country or Area, 1990-2018. Urban indicators database. UN-HABITAT, 2021. Disponível em: <https://data.unhabitat.org/datasets/urban-population-living-in-slums-by-country-or-area-1990-2018-thousands/explore>.

Sobre os autores:

Débora Ferreira de Barros. Graduada em Ciência da Computação. Atualmente faz parte do Laboratório de Mineração de Dados da UFCG, trabalhando como pesquisadora e desenvolvedora em projeto com foco em Processamento de Linguagem Natural.

Eanes Torres Pereira. Professor Dr. na UFCG. Orientador.