



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus Presidente Prudente

**DETECÇÃO DE MUDANÇAS DO USO E COBERTURA DO SOLO EM IMAGENS
DE ALTA RESOLUÇÃO USANDO REDE NEURAL CONVOLUCIONAL**

Orientanda: Thamires Gil Godoy
Orientador: Erivaldo Antonio da Silva
Coorientadora: Jussara de Oliveira Ortiz

Presidente Prudente -SP
2024



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

Campus Presidente Prudente

**DETECÇÃO DE MUDANÇAS DO USO E COBERTURA DO SOLO EM IMAGENS
DE ALTA RESOLUÇÃO USANDO REDE NEURAL CONVOLUCIONAL**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciências Cartográficas da Faculdade de Ciências e Tecnologia – UNESP - Universidade Estadual Paulista "Júlio de Mesquita Filho" – Campus de Presidente Prudente - SP, como parte dos requisitos para obtenção do título de Mestre em Ciências Cartográficas.

Orientador: Dr. Erivaldo Antonio da Silva
Coorientadora: Dra. Jussara de Oliveira Ortiz

G589d

GODOY, THAMIRES GIL

Detecção de mudanças do uso e cobertura do solo em imagens de alta resolução usando rede neural convolucionária / THAMIRES GIL

GODOY. -- Presidente Prudente, 2024

115 p.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Tecnologia, Presidente Prudente

Orientador: Erivaldo Antonio da Silva

Coorientadora: Jussara de Oliveira Ortiz

1. Sensoriamento Remoto. 2. Redes Neurais Convolucionais. I.
Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Tecnologia, Presidente Prudente. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Detecção de Mudanças do Uso e Cobertura do Solo em imagens de alta resolução usando Rede Neural Convolucional

AUTORA: THAMIRES GIL GODOY

ORIENTADOR: ERIVALDO ANTONIO DA SILVA

COORIENTADORA: JUSSARA DE OLIVEIRA ORTIZ

Aprovada como parte das exigências para obtenção do Título de Mestra em Ciências Cartográficas, área: Aquisição, Análise e Representação de Informações Espaciais pela Comissão Examinadora:


Prof. Dra. JUSSARA DE OLIVEIRA ORTIZ (Participação Virtual)
Divisão de Observação da Terra e Geoinformática – DIOTG / Instituto Nacional de Pesquisas Espaciais

Prof. Dr. SIDNEI JOÃO SIQUEIRA SANT'ANNA (Participação Virtual)
Divisão de Observação da Terra e Geoinformática – DIOTG / Instituto Nacional de Pesquisas Espaciais

Prof. Dr. GUILHERME HENRIQUE BARROS DE SOUZA (Participação Virtual)
Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente - FCT/Unesp

Presidente Prudente, 29 de abril de 2024

*"O cérebro eletrônico faz tudo
Quase tudo
Quase tudo
Mas ele é mudo"*
Gilberto Gil

AGRADECIMENTOS

Quero expressar meu agradecimento a todos que contribuíram para o sucesso desta dissertação.

Agradeço aos meus orientadores Prof. Dr. Erivaldo Antônio da Silva e Prof.^a Dr.^a Jussara de Oliveira Ortiz, pelo direcionamento e apoio durante todo o desenvolvimento deste trabalho. Sua orientação e a confiança que depositaram no meu processo de aprendizagem desempenharam um papel fundamental em meu progresso.

Também estendo minha gratidão à professora Prof.^a Dr.^a Maria de Lourdes Bueno Trindade Galo, ao Prof. Dr. Sidnei João Siqueira Sant'Anna e ao Prof. Dr. Guilherme Henrique Barros de Souza, por suas contribuições e sugestões que apontaram melhorias na execução deste trabalho e enriqueceram o desenvolvimento deste projeto, permitindo minha evolução científica.

Agradeço à minha família, uma árvore repleta de frutos inspiradores. Em especial, quero expressar minha profunda gratidão aos meus pais, Sandra e Celso, que sempre me incentivaram a buscar novos conhecimentos, independentemente da trajetória acadêmica, e sempre me transmitiram a confiança necessária para perseguir meus objetivos.

Aos meus amigos, acadêmicos ou não, por compartilharem suas experiências e conhecimentos. Essa troca enriqueceu a pesquisa e me motivou na conclusão deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Portanto, agradeço o apoio financeiro e os recursos que tornaram possível a realização deste projeto. Também não posso deixar de agradecer a UNESP e a Geopixel, cujo apoio foi fundamental para a viabilidade deste trabalho

Por fim, desejo expressar minha gratidão a todas as pessoas que compartilham suas competências de forma ampla e gratuita, pois essas entendem que o conhecimento não se constrói de forma isolada. E em particular, agradeço ao Dr. Sreenivas Bhattiprolu pela valiosa contribuição em meus entendimentos sobre redes neurais convolucionais.

Obrigada a todos por fazerem parte deste significativo capítulo da minha jornada acadêmica.

RESUMO

A superfície terrestre está sempre em transformação, devido tanto a fenômenos naturais quanto à intervenção humana. Nesse cenário, o uso de imagens de Sensoriamento Remoto, aéreo ou orbital, capturadas em diferentes períodos de tempo, é importante para detectar e monitorar as mudanças que ocorrem no uso e cobertura do solo. A extração de informações dessas imagens requer métodos de análise para identificar as feições de interesse, para tal, a interpretação visual de imagens, analógicas ou digitais, e a classificação automática de imagens digitais são considerados os principais métodos. Especialmente em áreas urbanas, devido à heterogeneidade dos alvos, a interpretação visual é comumente usada. Esse processo consiste em percorrer visualmente toda a imagem e compará-la com outra de data diferente, entretanto, essa abordagem consome tempo e, conseqüentemente, recursos financeiros. No que diz respeito ao processo de classificação digital, ferramentas de processamento de imagens e algoritmos que incorporam Inteligência Artificial (IA), ou seja, que replicam habilidades humanas como o reconhecimento de padrões, têm sido empregados com o objetivo de aprimorar o desempenho na classificação do uso e cobertura do solo. Dentre as técnicas de IA, destaca-se a Rede Neural Convolucional (RNC), concebida em 1998 pelo francês Yann LeCun, cuja principal finalidade é a classificação de imagens, permitindo a diferenciação de alvos, mesmo diante de mudanças em suas características espaciais ocasionadas, por exemplo, por altura e escala. O principal objetivo desse estudo é propor e avaliar uma metodologia para a detecção de mudanças em imagens de alta resolução espacial, usando imagens do satélite *WorldView – 2* em conjunto com as classificações resultantes da RNC U-Net. Essa abordagem, que identifica as mudanças por meio de duas imagens classificadas pela RNC, visa simplificar a identificação de alterações no uso e cobertura do solo para fins de gestão municipal. Os resultados obtidos no processo de classificação e detecção de mudanças foram avaliados para verificar a aplicabilidade dessa metodologia na melhoria da tomada de decisões em áreas urbanas que carecem desse tipo de informação de maneira ágil e eficaz. Ao adotar a arquitetura U-Net buscou-se o aproveitamento da sua capacidade de lidar com variações de textura e forma, fatores frequentemente encontrados em ambientes urbanos. A rede demonstrou ser ideal para classificar alvos como *construções*, *vegetação arbórea* e *outros*, mesmo diante do desequilíbrio na representação dessas

classes. Na primeira data, a área utilizada no treinamento apresentou 89,64% de acurácia geral e 87,25 % para o índice de concordância Kappa, enquanto na segunda data esses valores foram de 81,15% e 86,02%, respectivamente, indicando uma ótima classificação. A detecção de mudanças foi realizada utilizando as imagens classificadas e um conjunto de regras pré-definido; e a análise de qualidade se baseou em pontos fornecidos pela empresa GeoPixel Inteligência Municipal. Embora a detecção de mudanças tenha alcançado valores excelentes para a métrica *recall*, apontando a capacidade do modelo de identificar as mudanças presentes, a precisão foi notavelmente baixa, inferior a 1%, sugerindo uma alta taxa de falsos positivos. O trabalho evidenciou, portanto, a eficácia da metodologia na identificação de alterações em áreas urbanas, apesar de ser necessário abordar as diferenças de angulação na aquisição das imagens, que resultaram na identificação equivocada de várias áreas sem relevância como mudanças. A crescente necessidade de monitorar e gerenciar áreas urbanas demanda soluções tecnológicas avançadas, sendo que a combinação de imagens de alta resolução com a arquitetura U-Net se destaca como uma abordagem promissora, podendo acelerar a tomada de decisões baseadas em dados atualizados, contribuindo para uma melhor gestão municipal.

Palavras-chave: U-Net, Sensoriamento Remoto, Detecção de Mudanças, Classificação de Imagens.

ABSTRACT

The Earth's surface is constantly undergoing transformation, due to both natural phenomena and human intervention. In this scenario, the use of Remote Sensing imagery, whether aerial or orbital, captured at different time periods, is important for detecting and monitoring changes that occur in land use and land cover. Extracting information from these images requires analysis methods to identify features of interest; for this purpose, visual interpretation of images, whether analog or digital, and automatic classification of digital images are considered the main methods. Especially in urban areas, due to the heterogeneity of targets, visual interpretation is commonly used. This process involves visually scanning the entire image and comparing it with another image from a different date; however, this approach is time-consuming and consequently, financially resource-intensive. Regarding the digital classification process, image processing tools and algorithms incorporating Artificial Intelligence (AI), i.e., replicating human skills such as pattern recognition, have been employed to improve performance in land use and land cover classification. Among AI techniques, Convolutional Neural Networks (CNNs) stand out, conceived in 1998 by the Frenchman Yann LeCun, whose main purpose is image classification, allowing differentiation of targets, even in the face of changes in their spatial characteristics caused, for example, by height and scale. The main objective of this study is to propose and evaluate a methodology for change detection in high spatial resolution images, using images from the *WorldView – 2* satellite in conjunction with the classifications resulting from the U-Net CNN. This approach, which identifies changes through two images classified by the CNN, aims to simplify the identification of changes in land use and land cover for municipal management purposes. The results obtained in the classification and change detection process were evaluated to verify the applicability of this methodology in improving decision-making in urban areas lacking this type of information in an agile and effective manner. By adopting the U-Net architecture, we sought to leverage its ability to handle variations in texture and shape, factors often encountered in urban environments. The network proved to be ideal for classifying targets such as buildings, tree vegetation, and others, even in the face of imbalance in the representation of these classes. On the first date, the training area showed 89.64% overall accuracy and 87.25% for the Kappa

agreement index, while on the second date these values were 81.15% and 86.02%, respectively, indicating excellent classification. Change detection was performed using the classified images and a predefined set of rules; and the quality analysis was based on points provided by the company GeoPixel Municipal Intelligence. Although change detection achieved excellent values for the recall metric, indicating the model's ability to identify present changes, precision was notably low, less than 1%, suggesting a high false positive rate. The work therefore demonstrated the effectiveness of the methodology in identifying changes in urban areas, although it is necessary to address differences in the angle of image acquisition, which resulted in the mistaken identification of several areas without relevance as changes. The growing need to monitor and manage urban areas demands advanced technological solutions, with the combination of high-resolution images with the U-Net architecture standing out as a promising approach, which can accelerate decision-making based on updated data, contributing to better municipal management.

Keywords: U-Net, Remote Sensing, Change Detection, Image Classification.

LISTA DE FIGURAS

Figura 1 - Classificação de Métodos de Detecção de Mudança.....	24
Figura 2 - Aumento de publicações da IA em SR.....	26
Figura 3 - Neurônio Artificial.....	28
Figura 4 - Exemplo de uma Rede Neural.....	28
Figura 5 - Exemplo de um maxpooling 2x2.....	29
Figura 6 - Exemplo simplificado de uma Rede Neural Convolucional.....	30
Figura 7 - Gráfico ReLU (Rectified Linear Units).....	32
Figura 8 - Função Softmax.....	33
Figura 9 - Exemplo de uma Rede Neural Convolucional.....	35
Figura 10 - Exemplo de correspondência entre uma Fully Connected Layer e uma Rede Neural Totalmente Convolucional.....	37
Figura 11 - Arquitetura da rede U-Net.....	38
Figura 12 - Matriz de Confusão 3x3 com indicadores (VP, FP e FN).....	40
Figura 13 - A: Matriz de confusão 3x3 B: Matriz de confusão normalizada 3x3.....	41
Figura 14 - Matriz de Confusão Binária (Detecção de Mudanças).....	44
Figura 15 - Localização da área de estudo.....	47
Figura 16 - Evolução populacional em Guaratinguetá (1920-2022).....	48
Figura 17 - Pontos identificados como mudanças pela Geopixel na área de estudo.....	50
Figura 18 - Fluxograma da metodologia.....	52
Figura 19 - Categorização da metodologia de detecção de mudanças.....	53
Figura 20 – Exemplos das classes adotadas na classificação.....	55
Figura 21 - Conjunto de regras para detecção de mudanças.....	59
Figura 22 - Exemplo de uma matriz de transição.....	60
Figura 23 - Exemplo da interpretação visual para análise de qualidade da detecção de mudanças.....	62
Figura 24 - Distribuição dos pontos utilizados no registro das imagens.....	63
Figura 25 - Delimitação das áreas 1 e área 2.....	64
Figura 26 - Comportamento da função perda e da métrica IoU ao longo das épocas na imagem de 2021.....	65
Figura 27 - Comportamento da função perda e da métrica IoU ao longo das épocas na imagem de 2022.....	66

Figura 28 - Matrizes de confusão normalizadas.....	66
Figura 29 - Exemplo que pode acarretar a confusão entre pixels de construções e outros.	67
Figura 30 – Exemplo que pode acarretar a confusão entre pixels de vegetação arbórea e outros.	68
Figura 31 - Resultado da classificação da área 1 em 2021.....	70
Figura 32 – da classificação da área 1 em 2022.	71
Figura 33 - Região de ampliação A: resultado de classificação da área 1 nas duas datas.	72
Figura 34 - Região de ampliação B: resultado de classificação da área 1 nas duas datas.	73
Figura 35 - Região de ampliação C: resultado de classificação da área 1 nas duas datas.	74
Figura 36 - Resultado da classificação da área 2 em 2021.....	75
Figura 37 - Resultado da classificação da área 2 em 2022.....	76
Figura 38 - Região de aproximação A: resultado de classificação da área 2 nas duas datas.	77
Figura 39 - Região de aproximação B: resultado de classificação da área 2 nas duas datas.	78
Figura 40 - Região de aproximação C: resultado de classificação da área 2 nas duas datas.	79
Figura 41 - Resultado da detecção de mudanças na área 1.....	80
Figura 42 - Região de ampliação A: resultado da detecção de mudanças da área 1 nas duas datas.	81
Figura 43 - Região de aproximação B: resultado da detecção de mudanças da área 1 nas duas datas.	82
Figura 44 - Região de aproximação C: resultado da detecção de mudanças da área 1 nas duas datas.	83
Figura 45 - Resultado da detecção de mudanças na área 2.....	84
Figura 46 - Região de aproximação A: resultado da detecção de mudanças da área 2 nas duas datas.	85
Figura 47 - Região de aproximação B: resultado da detecção de mudanças da área 2 nas duas datas.	86

Figura 48 - Região de aproximação C: resultado da detecção de mudanças da área 2 nas duas datas.	87
Figura 49 - Matriz de transição da área 1.....	88
Figura 50 - Matriz de transição da área 2.....	89
Figura 51 - Mudanças detectadas pela metodologia na área 1 e os pontos de mudanças de referência.	91
Figura 52 - Mudanças detectadas pela metodologia na área 2 e os pontos de mudanças de referência.	92
Figura 53 -Mudanças detectadas pela metodologia na área 1 e a não correspondência com os pontos de mudanças de referência.	93

LISTA DE TABELAS

Tabela 1 - A qualidade da classificação associada aos valores do índice Kappa.	43
Tabela 2 - Resumo das métricas utilizadas.....	46
Tabela 3 - Características dos sensores WorldView – 2.	49
Tabela 4 - Dados das imagens utilizadas.....	50
Tabela 5 - Correspondência entre as classes e os identificadores numéricos.....	55
Tabela 6 - Métricas avaliativas das classificações usando a U-Net.	69
Tabela 7 - VP, FP e FN da detecção de mudanças.	90
Tabela 8 - Análise de qualidade da detecção de mudanças com as métricas, precisão, recall e f1-score.	93

LISTA DE ABREVIATURAS E SIGLAS

<i>Adam</i>	<i>Adaptive Moment estimation</i>
AG	Acurácia Global
AC	Acurácia do Consumidor
<i>FCL</i>	Fully Conected Layer
FN	Falso Negativo
FP	Falso Positivo
IA	Inteligência Artificial
<i>IoU</i>	<i>Intersecction over Union</i>
<i>ReLU</i>	<i>Rectified Linear Unit</i>
<i>RMSE</i>	Raiz Quadrada do Erro Médio (<i>Root Mean Squared Error</i>)
RNC	Rede Neural Convolucional
RNTC	Rede Neural Totalmente Convolucional
SR	Sensoriamento Remoto
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

SUMÁRIO

1. INTRODUÇÃO.....	17
1.1 JUSTIFICATIVA E RELEVÂNCIA DO TEMA.....	19
1.2 OBJETIVOS.....	20
1.3 ORGANIZAÇÃO DO DOCUMENTO.....	21
2. FUNDAMENTAÇÃO TEÓRICA	22
2.1 DETECÇÃO DE MUDANÇAS.....	22
2.2 CLASSIFICAÇÃO DE IMAGENS PARA DETECÇÃO DE MUDANÇAS	24
2.3 REDES NEURAIS CONVOLUCIONAIS	27
2.3.1 HIPERPARÂMETROS PARA ESTRUTURAÇÃO DA REDE	31
2.3.2 HIPERPARÂMETROS PARA O TREINAMENTO DA REDE	33
2.4 ARQUITETURA DE REDE: U-NET.....	36
2.5 MÉTRICAS DE AVALIAÇÃO	38
3. MATERIAL E MÉTODOS	47
3.1 ÁREA DE ESTUDO	47
3.2 MATERIAL	49
3.3 METODOLOGIA	51
3.3.1 CRIAÇÃO DO CONJUNTO DE DADOS.....	53
3.3.2 IMPLEMENTAÇÃO DA U-NET	56
3.3.2.1 TREINAMENTO, TESTE E ANÁLISE DE QUALIDADE DA U-NET	57
3.3.3 DETECÇÃO DE MUDANÇAS	57
3.3.3.1 ANÁLISE DE QUALIDADE DA DETECÇÃO DE MUDANÇAS	60
4. RESULTADOS E DISCUSSÃO	63
5. CONSIDERAÇÕES FINAIS.....	95
REFERÊNCIAS.....	97
APÊNDICE A.....	111
APÊNDICE B.....	112

1. INTRODUÇÃO

A cartografia é fundamental para o planejamento e tomadas de decisões e dada à rapidez com que ocorrem mudanças no uso e cobertura do solo, faz-se necessário o contínuo monitoramento dessas alterações. O Sensoriamento Remoto (SR), área afim da cartografia, é uma ciência que busca detectar e analisar essas mudanças. Esse processo envolve a comparação de imagens da mesma região geográfica em datas distintas, permitindo a identificação de variações no estado dos objetos, como textura, forma e resposta espectral. Conforme destacado por Shafique et al. (2022), a detecção de mudanças é uma técnica de reconhecimento das discrepâncias entre imagens adquiridas em momentos diferentes e é aplicada em diversas esferas, como no desenvolvimento de áreas urbanas, no gerenciamento de desastres ambientais e na identificação de objetos na cobertura do solo.

Nas regiões urbanas, principalmente nas grandes cidades, a identificação de mudanças no uso e cobertura do solo tem se tornado um desafio para as autoridades públicas, devido à crescente extensão territorial e à rápida ocorrência de eventos que podem acometer problemas ambientais, como por exemplo, as ocupações irregulares. O Instituto Brasileiro de Geografia e Estatística (IBGE) apresenta estimativas indicando que existem mais de cinco milhões de moradias em aglomerados subnormais¹ (IBGE, 2020) e muitas vezes em áreas de risco. Outra ocorrência importante a ser identificada é o desmatamento em áreas urbanas, onde Mota (2003) ressalta a preocupação com esse evento, ao considerá-lo como um fator significativo às mudanças climáticas, pois, resulta na exposição do solo, assoreamento de recursos hídricos, inundações e diminuição na capacidade de infiltração das águas pluviais. Além disso, o descarte irregular de entulho é mais uma preocupação dos órgãos públicos, onde conforme relatado pela Associação Brasileira de Empresas de Limpeza Pública e Resíduos Especiais (ABELPRE), somente no ano de 2017 e considerando apenas os resíduos descartados ou abandonados em espaços públicos, os municípios brasileiros coletaram aproximadamente 45,1 milhões de toneladas de resíduos de entulho (ABRELPE, 2017).

O SR é considerado uma ferramenta importante para a detecção de

¹ Aglomerado Subnormal é uma forma de ocupação irregular de terrenos de propriedade alheia para fins de habitação em áreas urbanas (IBGE, 2020)

mudanças espaciais e temporais na classificação do uso e cobertura do solo, especialmente quando se utiliza imagens provenientes de sensores de alta resolução espacial (Maxwell et al., 2017; Chen et al., 2021). Além disso, essas imagens possibilitam a análise de uma cena em diversas regiões do espectro eletromagnético, enriquecendo a quantidade de dados disponíveis e aprimorando a qualidade da informação obtida.

Os sensores orbitais de alta resolução espacial têm a capacidade de detectar objetos com dimensões reduzidas, identificando detalhes com tamanho inferior a 1,0 metro na banda pancromática e aproximadamente de 4,0 a 5,0 metros nos sensores multiespectrais. O uso desses sensores oferece vantagens significativas, incluindo a obtenção de informações altamente detalhadas e uma melhor resolução temporal, quando comparada às imagens aéreas (Panizza e Fonseca, 2011), isso ocorre porque esses sensores geralmente revisitam um determinado local com maior frequência. O Instituto Nacional de Pesquisas Espaciais - INPE (2022) destaca a ampla utilização desses sensores na cartografia, especialmente no mapeamento do uso e cobertura do solo, com foco no planejamento urbano e na atualização de bases cartográficas. Em contraste, satélites equipados com sensores de baixa e média resolução espacial, como o Landsat, MODIS e AVHRR, enfrentam limitações na diferenciação de diversos alvos. Por exemplo, o Landsat/TM abrange uma área de pixel de 900 metros quadrados, o que supera as dimensões de muitos lotes urbanos (Zhou e Troy, 2008).

Existem várias abordagens para a detecção de mudanças no uso e cobertura do solo utilizando imagens orbitais, e esses métodos podem diferir em diversos aspectos, como a unidade de análise (pixel ou objeto), as operações e as características dos dados de entrada (Si Salah et al., 2019). Dessa forma, esse trabalho, inserido no contexto do SR e da Cartografia, propõe e avalia uma metodologia para a detecção de mudanças em áreas urbanas. Essa abordagem visa identificar possíveis alterações nas classes de *vegetação arbórea* e *construções*, que podem indicar a ocorrência de fenômenos como aumento ou supressão de área construída, desmatamento, descarte irregular de entulho e ocupação irregular. A metodologia envolve o uso de uma RNC para classificar imagens de alta resolução espacial do *WorldView – 2* em duas datas distintas e, em seguida, identificar mudanças no uso e cobertura do solo, por meio de um conjunto de regras entre elas, contribuindo assim para a melhoria dos procedimentos de detecção de mudanças no

uso e na cobertura do solo em áreas urbanas.

1.1 JUSTIFICATIVA E RELEVÂNCIA DO TEMA

A Ciência Cartográfica tem buscado soluções para otimizar o processo de mapeamento em áreas urbanas em todo o mundo. Por exemplo, França et al. (2013) discutem a utilização de Sistemas de Informações Geográficas para a detecção de edificações irregulares ou expansões em áreas urbanas. Miler et al. (2010) apresentam uma solução eficaz baseada em tecnologias de código aberto para o planejamento e a gestão urbana por meio de uma plataforma web-GIS. Cantoni e Vassena (2019) exploram a aplicação de sistemas de mapeamento para obter dados geoespaciais de forma rápida, como escaneamento por laser e mapeamento móvel.

Um dos procedimentos pertinentes ao mapeamento compreende a detecção de mudanças no uso e cobertura do solo, que consiste na análise de uma determinada área geográfica em datas distintas para identificar variações no estado do objeto, como mudanças na textura, forma e resposta espectral. Essa identificação de modificações pode englobar diversas transformações em ambientes urbanos, como por exemplo a expansão de áreas construídas, a substituição de áreas verdes por edificações, alterações na infraestrutura viária, a evolução de áreas industriais e mudanças em zonas comerciais. No entanto, nesse trabalho, o foco está voltado para o aumento ou supressão de construções, desmatamento, descarte irregular de entulho e ocupação irregular. É importante ressaltar que essa detecção não ocorre de maneira direta, uma vez que exigiria uma análise abrangente considerando diversos fatores, como sociais e ambientais. Com base nessa premissa, optou-se por utilizar a *vegetação arbórea* e *construções* como classes que sugerem a ocorrência dessas modificações.

Devido ao crescente uso de Veículos Aéreos Não Tripulados (VANTs), algumas cidades têm realizado, ao menos a cada quatro anos, o mapeamento aéreo para fins de cadastro (Antunes e Hollatz, 2015; Bezerra et al., 2019). No entanto, é relevante observar que as imagens obtidas por VANTs costumam apresentar menor resolução temporal em comparação com as imagens de satélite. Essa limitação é relevante, uma vez que as transformações urbanas são contínuas e o intervalo entre esses levantamentos retarda medidas públicas que podem afetar o planejamento dos municípios. E como argumentam Amorim et al. (2018) os métodos envolvidos em atualização cadastral devem ser constantes, e com informações públicas, permitindo

o questionamento e controle da atividade governamental.

Desse modo, a utilização de imagens orbitais de alta resolução oferece um equilíbrio adequado entre a resolução espacial e temporal, permitindo uma análise periódica das áreas urbanas. Como Kux et al. (2006) também reforçam ao mostrarem o potencial das imagens de satélite de alta resolução para gerar informações nas áreas urbanas no Brasil, onde muitas vezes faltam informações atualizadas nos bancos de dados municipais, estaduais e federais.

Muitos processos de detecção de mudanças em áreas urbanas acabam dependendo da capacidade de interpretação visual do usuário. Além disso, quando o processo abrange extensas áreas e utiliza imagens de alta resolução, o aumento no volume de dados e o tempo necessário tendem a ser significativos (Cândido et al., 2015). Assim, é pertinente promover pesquisas que desenvolvam metodologias para aprimorar a detecção de mudanças no uso e cobertura do solo em áreas urbanas, promovendo uma gestão mais eficiente dessas áreas.

Nesse sentido, os estudos referentes às RNCs têm ganhado espaço na área do SR visando à segmentação e classificação de imagens multiespectrais, como mostra o trabalho de Ghanbari et al. (2021), que indica um aumento na frequência anual e cumulativa de trabalhos de SR utilizando RNCs. Ainda de acordo com Ghanbari et al. (2021), em 2016 havia menos de 50 publicações dessa natureza, nas onze revistas analisadas, ocorrendo um salto em 2019 para quase 400 publicações, de modo que 63% utilizavam as RNCs para classificação do uso e cobertura do solo e detecção de objetos. O uso de RNC, especialmente a U-Net, tem indicado aspectos positivos com relação a outros procedimentos como *Random Forest* e *GEographic Object-Based Image Analysis (GEOBIA)* (Guirado et al., 2017; Pan et al., 2020).

Portanto, o aumento da utilização de RNCs na classificação do uso e cobertura do solo em imagens multiespectrais destaca a eficácia desse método. Todavia, é importante levantar uma questão que vai nortear este trabalho: qual é o comportamento deste tipo de classificador, para detecção de mudanças em áreas urbanas, usando imagem de satélite de alta resolução espacial?

1.2 OBJETIVOS

Avaliar a eficácia da metodologia de detecção de mudanças no uso e cobertura do solo em áreas urbanas, visando simplificar a identificação de alterações para fins de gestão municipal. A abordagem emprega a classificação pela RNC U-Net

e imagens de alta resolução espacial, para em seguida, identificar áreas onde ocorreram alterações nas classes de *construções* e *vegetação arbórea*. Com base nesse objetivo geral, são delineados os seguintes objetivos específicos:

- Configurar os parâmetros da U-Net para ambas as datas visando obter uma classificação precisa das imagens.
- Avaliar a classificação da área usada no treinamento nas duas datas usando métricas como matriz de confusão normalizada, Acurácia do Consumidor (AC), Acurácia Global (AG), índice Kappa e *Intersection over Union (IoU)*.
- Analisar qualitativamente a classificação da área não utilizada no treinamento da U-Net com relação à utilizada no treinamento.
- Avaliar a detecção de mudanças de toda a área de estudo usando métricas como *recall*, precisão e *f1-score*.
- Investigar e identificar as possíveis causas de erros na detecção de mudanças.

1.3 ORGANIZAÇÃO DO DOCUMENTO

Este documento está organizado em 5 principais tópicos.

- Capítulo de Introdução: contextualiza o problema em análise, justificando sua importância para gestão municipal. Também delinea os objetivos e as questões a serem abordadas.
- Capítulo de Fundamentação Teórica: estabelece uma base conceitual para facilitar a compreensão. Os temas explorados abrangem a detecção de mudanças, a classificação de imagens em áreas urbanas, as RNCs, a arquitetura da U-Net e métricas avaliativas que são usadas nesse trabalho.
- Capítulo de Material e Métodos: descreve os materiais para a execução do estudo. Além disso, a metodologia proporciona uma compreensão do processo de pesquisa. Estas seções incluem a criação do conjunto de dados, a implementação da U-Net, o treinamento, teste e análise de qualidade da U-Net, bem como a detecção de mudanças e a análise de qualidade resultante dessa detecção.
- Capítulo de Resultados: apresenta os resultados obtidos a partir da aplicação da metodologia descrita no capítulo anterior. Imagens, tabelas e interpretações são fornecidas para esclarecer os estudos.
- Capítulo de Considerações Finais: o último capítulo consiste em uma conclusão abrangente que resume os principais resultados discutidos no estudo.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 DETECÇÃO DE MUDANÇAS

As aplicações de detecção de mudanças fornecem um meio para estudar transformações no uso e cobertura do solo, biodiversidade, processo de urbanização, identificação de desastres e outras modificações (Bai et al., 2021). A detecção de mudanças no SR tem sido objeto de pesquisa desde que as imagens orbitais se tornaram disponíveis, e na literatura são encontradas várias definições deste conceito. Em Firouzabadi e Ghanavati (2007), é fundamentado que a detecção de mudanças é o processo de analisar a mesma região geográfica e identificar áreas que não são semelhantes em duas ou mais imagens de datas distintas. Singh (2010) definiu a detecção de mudanças como o processo de identificação de variações no estado de um objeto ou fenômeno observando-o em momentos diferentes. E Deng et al. (2009) ainda complementam, mencionando que, essas variações são dinâmicas e resultam de uma complexa interação entre fatores físicos, biológicos e sociais. Assim, a detecção de mudanças, ocasionada por fatores físicos, biológicos e sociais, pode ser compreendida como a análise de uma determinada área geográfica em datas diferentes, com o propósito de identificar variações no estado do objeto, tais como textura, forma e resposta espectral.

Imagens ópticas de diferentes datas geralmente possuem discrepâncias no valor dos pixels devido à irradiância solar no topo da atmosfera, distância Sol-Terra, geometria de iluminação, geometria de aquisição, além de variações atmosféricas (O'Connell, 2013; Dave et al., 2015). Dessa forma, antes de se realizar a detecção de mudanças, é necessária a aplicação de alguns procedimentos, tais como a normalização radiométrica e o registro das imagens (Mateos et al., 2010; Tondewad e Dale, 2020). A normalização eficaz da imagem desempenha um papel crucial na redução da influência radiométrica provocada por fatores não relacionados à superfície (Hong e Zhang, 2005). O registro de imagens, em essência, é responsável por alinhar geometricamente as duas imagens (Tondewad e Dale, 2020).

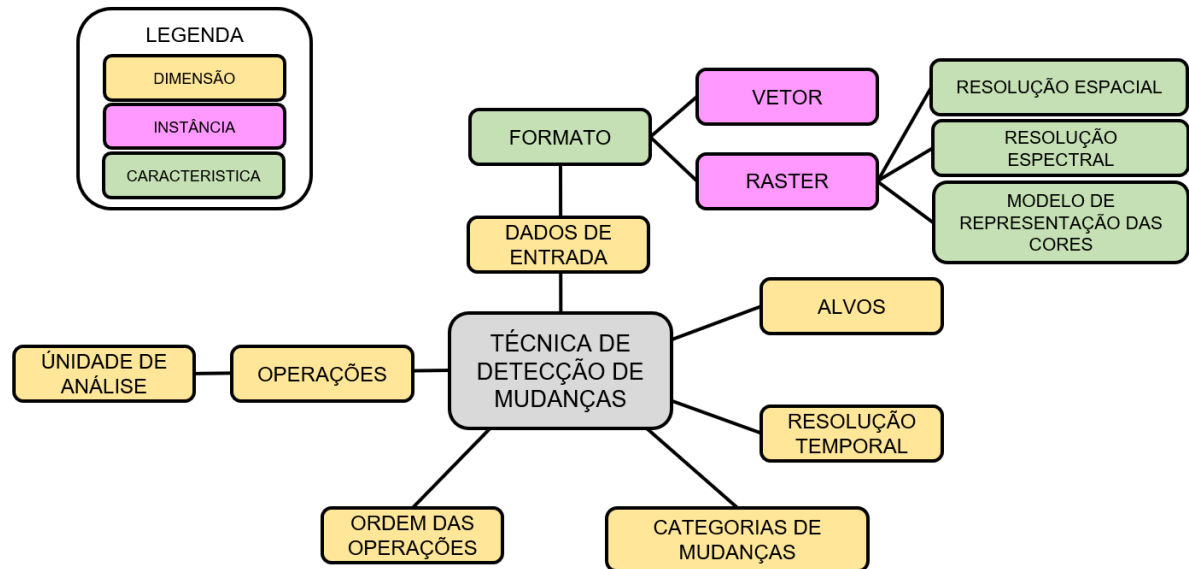
Após concluir o pré-processamento para a correção radiométrica e geométrica em nas imagens, é possível realizar a detecção de mudanças aplicando uma variedade de métodos disponíveis. Em Si Salah et al. (2019) é proposto um meio de classificar as metodologias de detecção de mudanças, onde eles se baseiam em

outros trabalhos consolidados. Nesse trabalho, os autores, partem de três conceitos: dimensão, instância e característica. A dimensão tem o propósito de descrever como a metodologia funciona, permitindo discerní-la de outras técnicas. Isso abrange determinar se a detecção é de natureza bi-temporal ou multi-temporal, se é aplicada pixel a pixel ou ao nível de objetos, entre outros exemplos. Na instância, são especificados os tipos de dados de entrada, como dados *raster* ou vetoriais, bem como os parâmetros utilizados. Por fim, as características abrangem aspectos como resolução espacial, resolução espectral e o modelo de representação das cores dos dados de entrada.

A Figura 1 ilustra a metodologia proposta por Si Salah et al. (2019) para a categorização de metodologias de detecção de mudanças em dados de SR. A dimensão *unidade de análise* determina se a detecção ocorre ao nível pixel a pixel ou ao nível de objetos. Com base nessa dimensão, é estabelecida a dimensão *operações* de forma apropriada, ou seja, em casos de detecção ao nível de objetos, por exemplo, poderão ser empregadas operações que se concentram na extração de características de objetos para a identificação de mudanças. Além disso, o trabalho determina que a ordem em que as operações (*ordem das operações*) são executadas também é uma dimensão à parte.

Ainda, de acordo com os autores acima, as *categorias de mudanças* se referem à classificação das alterações de interesse. Existem três tipos principais de mudanças: mudança binária, que envolve determinar se há mudança ou não; múltiplas categorias de mudança, que classificam as mudanças em várias classes, como objeto demolido, novo objeto, objeto modificado, etc.; e mudança direcional, que descreve a transição de uma classe para outra, como de vegetação para construção, por exemplo. A *resolução temporal* dita se a detecção de mudanças é bi-temporal ou multi-temporal. E a dimensão *alvos* refere-se a quais objetos ou alvos são importantes no processo de detecção de alterações, por exemplo, em uma análise temporal de desmatamento a vegetação arbórea seria um alvo dessa detecção de mudanças. Os *dados de entrada* são uma dimensão que informa o tipo de dados que serão processados, podendo ter várias fontes, como fotografias aéreas e imagens de satélite óptico ou SAR, dados vetoriais híbridos que integram informações de mapas, e nuvens de pontos geradas por sensores LIDAR.

Figura 1 - Classificação de Métodos de Detecção de Mudança.



Fonte: Adaptado de Si Salah et al. (2019).

Nos últimos anos, a tecnologia integrada a IA tornou-se um foco de pesquisa para desenvolver novas técnicas de detecção de mudanças (Shi et al., 2020), permitindo diversificações em todas as dimensões propostas por Si Salah et al. (2019).

Janga et al. (2023) fornecem uma revisão abrangente das aplicações de IA em SR, incluindo classificação de imagens, mapeamento da cobertura do solo e detecção de alterações. Loyola et al. (2019) discutem a integração da IA e do SR no ensino e na pesquisa, com foco na coleta e análise de dados utilizando sistemas aéreos não tripulados. Pour et al. (2018) avaliam a eficácia dos algoritmos de IA no mapeamento geológico, demonstrando sua capacidade em distinguir alvos de pixel e subpixel em dados de satélite.

2.2 CLASSIFICAÇÃO DE IMAGENS PARA DETECÇÃO DE MUDANÇAS

A classificação de imagens pode ser entendida como uma operação no processo de detecção de mudanças (Si Salah et al., 2019). Esse processo é de particular importância nas regiões urbanas, uma vez que fornece suporte essencial para pesquisas em diversas áreas, incluindo urbanismo, hidrologia, questões ambientais e energia, conforme destacado por Ohata e Quintanilha (2005).

A classificação de imagens de satélite de áreas urbanas pode ser encontrada em diversos trabalhos. Em Passo et al. (2013) é avaliada a potencialidade e a acurácia

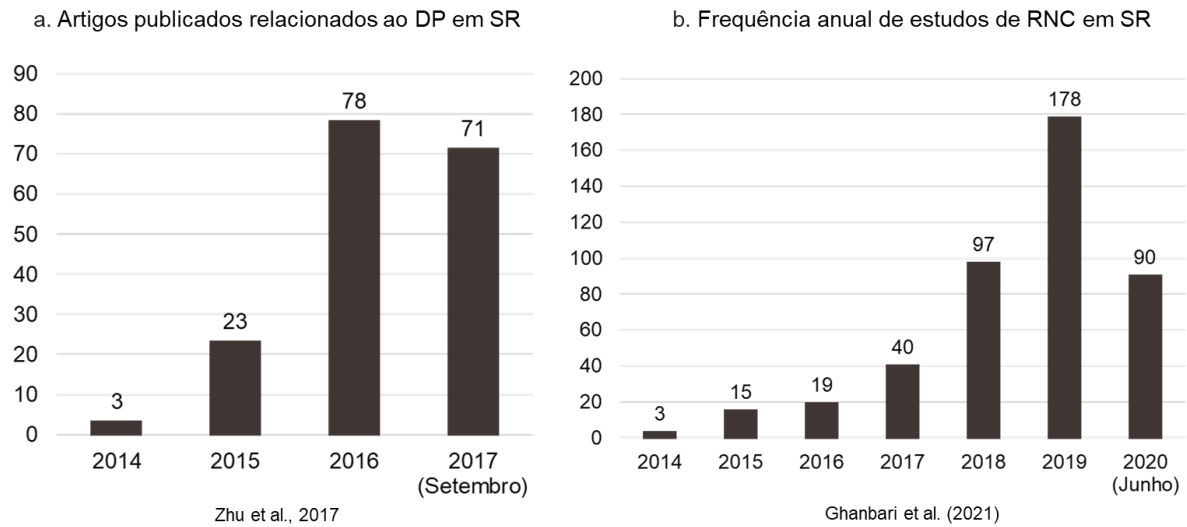
da classificação baseada em objeto em uma imagem do satélite *WorldView* – 2 sobre uma área urbana. Nesse trabalho o índice Kappa (Cohen, 1960) variou de 0,6 a 1 entre as classes definidas, indicativo de classificações muito boas à excelente, de acordo com Landis e Koch (1977).

Ribeiro et al. (2011) propõem o uso do sistema de classificação de imagens InterImage (disponível em: <http://www.lvc.ele.puc-rio.br/projects/interimage/pt-br/>), baseada em análise orientada a objeto, para o mapeamento da cobertura do solo urbano usando imagens *WorldView* – 2. Esse método resultou em um índice de classificação Kappa (Cohen, 1960) superior a 0,75, indicando classificação muito boa, como mostra a Tabela 1 no Capítulo 2.5. Pan et al. (2020) apresentam a classificação de vilas urbanas usando RNC, que atingiu um nível de precisão geral de 87,2% na identificação das vilas urbanas.

A demanda por metodologias para a classificação urbana aumentou nas últimas décadas, como relatam Chen et al. (2021), que discutem e revisam os avanços das pesquisas e metodologias de mapeamento e classificação do uso do solo urbano. Blaschke et al. (2011), por exemplo, enfatizaram, à época, que a classificação por pixel em imagens de alta resolução não apresentou bom desempenho e, também, houve a dificuldade em extrair características urbanas de forma satisfatória. Conforme os mesmos autores, na data dessa publicação, uma das abordagens mais eficaz seria a GEOBIA, que operava em nível de objeto.

A partir de 2014, as RNCs ganharam destaque no SR, como evidenciado na Figura 2, que apresenta dois gráficos ilustrativos. O primeiro gráfico a esquerda, baseado na pesquisa de Zhu et al. (2017), ilustra a frequência de artigos relacionados ao *Deep Learning* no SR. Observa-se claramente um aumento exponencial no número de publicações desse tipo ao longo do tempo. O segundo gráfico, resultado do estudo de Ghanbari et al. (2021), retrata a frequência anual de pesquisas relacionadas as RNCs no SR, revelando um crescimento igualmente exponencial.

Figura 2 - Número de publicações usando IA em SR.



Fonte: Adaptado de Zhu et al. (2017) e Ghanbari et al. (2021).

Ao comparar as RNCs com a GEOBIA, percebe-se que ambas realizam extração de atributos das imagens e os utilizam para treinar classificadores. No entanto, as abordagens diferem na forma como executam esse processo. No caso da GEOBIA, ocorre a extração de textura, resposta espectral, características geométricas e contextuais. Por outro lado, as RNCs realizam a extração de características em uma hierarquia com base em imagens de entrada, geralmente usando kernels quadrados. Além disso, essa hierarquia não é pré-definida ou guiada por intervenção humana, mas sim orientada pelos dados (LeCun et al., 1998; Kucharczyk et al., 2020).

Essa evolução na abordagem de análise de imagens tem implicações significativas para a detecção de mudanças em ambientes urbanos. Em um estudo conduzido por Fu et al. (2018), é comparada uma RNC com a técnica GEOBIA para a classificação do uso e cobertura do solo em imagens de altíssima resolução espacial. A RNC demonstrou uma melhora de cerca de 10% na precisão geral ao classificar oito classes distintas. De maneira similar, Zhang et al. (2017) concluíram que as RNCs alcançaram uma precisão maior em aproximadamente 6% na classificação da cobertura do solo em cinco classes utilizando imagens de altíssima resolução. Em outro estudo conduzido por Pan et al. (2020), uma imagem *WorldView - 2* de uma área urbana é classificada em construções e não construções usando ambas as técnicas. Nesse caso, a RNC obteve uma precisão geral de 87,2%, enquanto a GEOBIA alcançou 58,1%.

Dessa forma, os autores mostraram a viabilidade das RNCs para classificar imagens de SR de alta resolução. Além disso, como comentam Shi et al. (2020),

diversos estudos provaram que a IA atinge ótimos resultados na classificação do uso e cobertura do solo e que esses resultados podem ser usados para detecção de mudanças.

2.3 REDES NEURAIS CONVOLUCIONAIS

Em 2006, Lambin e Linderman (2006) enfatizavam previamente a importância das imagens de satélite no monitoramento global, especialmente no estudo das mudanças de uso e cobertura do solo. Em 2020, Gong et al. (2020) destacaram a crescente relevância de pesquisas em SR para áreas urbanas, apontando para o acesso a dados de alta resolução e o surgimento de novos sensores, tornando a observação de ambientes urbanos cada vez mais viável.

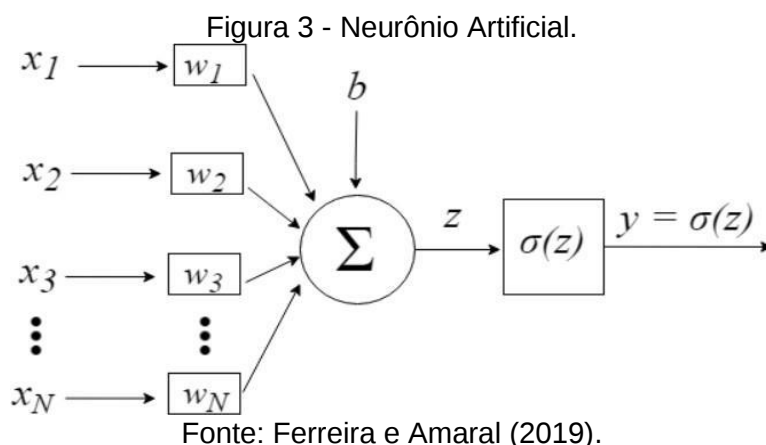
No contexto de imagens de alta resolução espacial em ambientes urbanos, as metodologias de aprendizado de máquina estão ganhando destaque, como afirmado por Maxwell et al. (2017) ao compararem o aprendizado de máquinas com classificadores paramétricos:

Uma ampla gama de estudos descobriu que esses métodos tendem a produzir maior precisão em comparação com classificadores paramétricos tradicionais, especialmente para dados complexos com uma alta dimensão, com muitas variáveis preditoras.

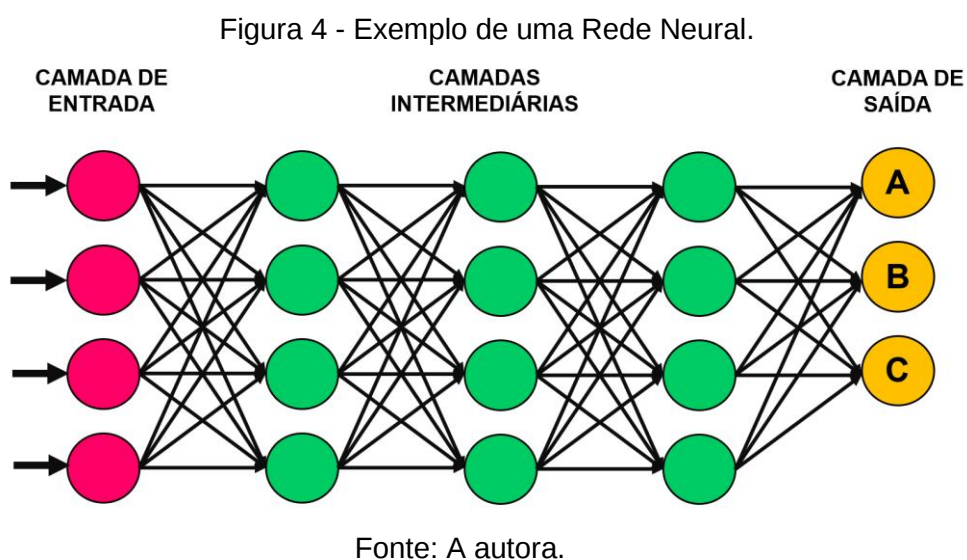
O aprendizado de máquina, também conhecido por *machine learning*, é um subcampo da IA, responsável por analisar dados que automatizam a construção de modelos analíticos, dando a possibilidade de sistemas computacionais aprenderem padrões e características. Dentre as técnicas de aprendizado de máquinas mais utilizadas em SR estão: as: Redes Neurais, *Support Vector Machine*, Árvores de Decisão e *Random Forest* (Shao e Lunetta, 2012). Ainda dentro do processo de aprendizado de máquina, há o subcampo do Aprendizado Profundo, conhecido como *Deep Learning*, e que diz respeito às redes neurais.

As redes neurais são baseadas no funcionamento dos neurônios biológicos, que são compostos por dendritos, axônio e um corpo celular; os dendritos são responsáveis pela recepção de sinais elétricos de outros neurônios, o corpo celular processa o sinal e o axônio transmite o fluxo para realizar determinadas ações (Goodfellow et al., 2016). De forma mais detalhada, uma rede neural artificial possui diversos neurônios artificiais (Figura 3), que consistem basicamente em entradas ($x_i; i \in \{1, 2, \dots, n\}$) (como sinapses) multiplicadas por pesos ($w_i; i \in \{1, 2, \dots, n\}$) e

somadas com um viés (b). Esses pesos representam o fluxo de informações computados por uma função matemática (Z) que determina a ativação do neurônio. No final há outra (ou a mesma) função ($\sigma(z)$) que calcula a saída (y) do neurônio artificial (Gupta, 2013).



De maneira simplificada, na Figura 4, apresenta-se um exemplo de uma rede neural artificial com vários neurônios representados por círculos. A camada de entrada recebe os dados juntamente com seus rótulos correspondentes. Posteriormente, nas camadas intermediárias, ocorre o processamento da rede, no qual os pesos e viés são ajustados para determinar a relação ótima entre os dados e seus rótulos. Por fim, a camada de saída classifica os dados em uma das classes A, B ou C, com base na relação estabelecida.



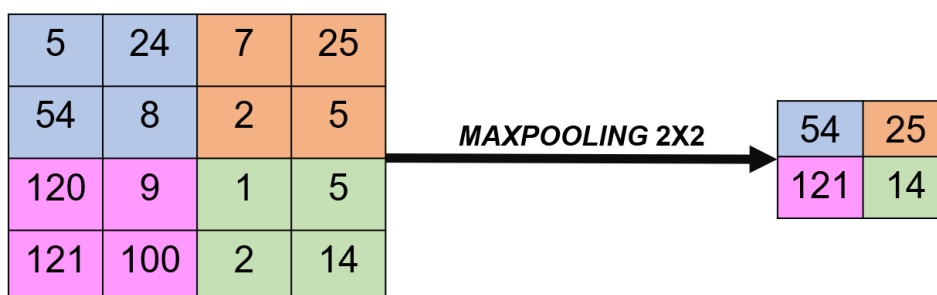
Em 1943 as redes neurais começaram a ser desenvolvidas pelo Walter Pitts e o Warren McCulloch (McCulloch e Pitts, 1943), mas só em 1998 as RNCs surgiram

com o trabalho de LeCun et al. (1998); essa metodologia tem como dados de entrada imagens que percorrem camadas de neurônios e aprendem pesos para vários objetos da cena, assim, a RNC é capaz de diferenciar os tantos alvos presentes.

A estrutura de uma RNC é dividida em três camadas: camada convolucional, camada de *pooling* e pela camada totalmente conectada, conhecida como *Fully Connected Layer (FCL)*. A camada convolucional é responsável por extrair as características da entrada, como um exemplo dado por Kucharczyk et al. (2020): a primeira camada convolucional de uma RNC extrai recursos de baixo nível, como linhas retas, linhas curvas e cantos, a camada seguinte pode combinar certos recursos de baixo nível para extrair recursos de nível médio, como objetos de telhado (por exemplo, chaminés, painéis solares) ou vegetação sobre telhados. A última camada convolucional pode combinar recursos de nível médio para extrair recursos de alto nível, por exemplo, telhados inteiros.

A camada *pooling* é responsável por simplificar os mapas de características gerados pelo processo de convolução, tal que tem como objetivos: diminuir o número de parâmetros e consequentemente o custo computacional, além de controlar o *overfitting* da rede (Gholamalinezhad e Khosravi, 2020). Dentre os métodos de *pooling* conhecidos existe o *average pooling*, que simplifica um conjunto de pixel pela média deles e o *maxpooling*, que é o mais utilizado e atribui o valor de saída como o máximo do conjunto analisado, como explicam Krizhevsky et al. (2017). Na Figura 5 há a ilustração do *maxpooling*, de tamanho 2x2, onde na saída se atribui o valor máximo.

Figura 5 - Exemplo de um *maxpooling* 2x2.

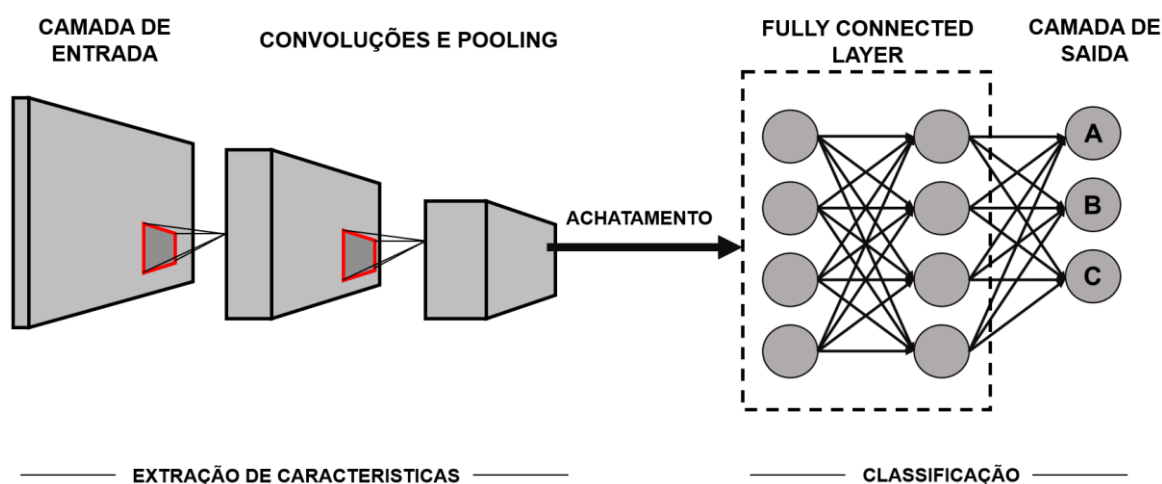


Fonte: A autora.

A *FCL* é a camada que classifica os dados e, portanto, formam as últimas camadas da rede. A entrada para a *FCL* pode ser a saída de uma camada convolucional ou de uma *pooling*, que então é achatada para diminuir sua dimensão para um vetor. Dentro da *FCL* há neurônios que são totalmente conectados com as

camadas vizinhas e seu procedimento é semelhante ao da rede neural exemplificada na Figura 4, ou seja, pesos e vieses são ajustados para encontrar a melhor relação entre as imagens e as classes. Na Figura 6 há um exemplo de uma RNC, onde a primeira função da rede é aprender as características para depois classificar.

Figura 6 - Exemplo simplificado de uma Rede Neural Convolucional.



Fonte: A autora.

Um das principais vantagens das RNCs com relação às demais técnicas de classificação é a invariância das RNCs aos fatores de transformação, ou seja, uma rede bem treinada é capaz de identificar um objeto mesmo esse tendo sido rotacionado, transladado ou mudado de escala, como descrevem Ronneberger et al. (2015) ao incluir deformações no conjunto de dados. Essa característica se deve principalmente ao compartilhamento do mesmo peso e viés entre os neurônios de uma camada, que acaba por definir o filtro de convolução.

Portanto, as RNCs são redes de aprendizagem supervisionada que extraem propriedades a partir dos dados de entrada e utilizam a convolução em pelo menos uma de suas camadas (Lima et al., 2020). Um tipo de RNC é a rede U-Net, proposta por Ronneberger et al. (2015), ela é um procedimento que tem bons resultados com poucos exemplos de treinamento e produz uma segmentação semântica precisa, como indicam Pan et al. (2020).

As RNCs se diferenciam por sua arquitetura e pelos hiperparâmetros que apresentam. De acordo com Aszemi e Selvam (2019) “as variáveis na configuração podem ser chamadas de hiperparâmetros e seus valores não podem ser estimados a partir de dados”; além disso, os autores os dividem em: hiperparâmetros que

determinam a estrutura da rede e hiperparâmetros que determinam o treinamento da rede.

2.3.1 HIPERPARÂMETROS PARA ESTRUTURAÇÃO DA REDE

A estrutura de uma RNC depende dos fatores: tamanho do filtro convolucional e preenchimento das bordas, camadas ocultas e função de ativação. “A largura e a altura dos filtros de convolução nas RNCs determinam o tamanho do campo receptivo no qual extraem as informações” (Ahmed e Karim, 2020). O preenchimento das bordas é o meio pelo qual se garante que o filtro passe pelas extremidades da imagem, onde normalmente são adicionados valores nulos (Aszemi e Selvam, 2019).

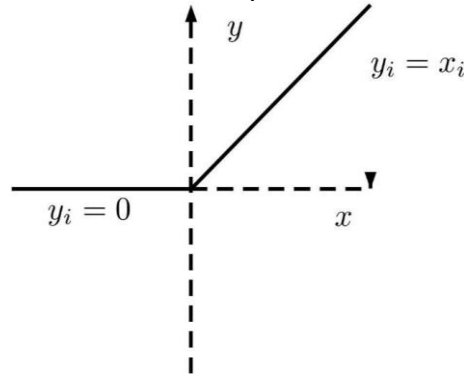
Outro hiperparâmetro para a estruturação da rede é a quantidade de camadas ocultas, pois são elas que vão gerar os recursos necessários em uma complexidade crescente para a classificação de imagens; começando com orientações de linhas simples e terminando com formas generalizadas completas (Josephine et al., 2021 e Kucharczyk et al., 2020).

O processo que descreve uma rede se propõe a encontrar pesos e vieses que se adequem ao valor de entrada de forma que sua saída tenha a melhor correspondência com os valores de treinamento (Goodfellow et al. 2016). No entanto, por ser essa estrutura linear, não permite formular problemas complexos.

Para modelar problemas mais complexos, como solução, as arquiteturas de redes incluem funções de ativação, o que permitem que mudanças nos pesos ou vieses não alterem drasticamente o resultado de saída. Como definem Krizhevsky et al. (2017) as “funções de ativação permitem que o modelo aprenda limites de previsão não lineares”.

Em redes neurais há diversas funções de ativação, como a tangente hiperbólica e a sigmoide, mas em RNC a *ReLU* (*Rectified Linear Unit*) é amplamente difundida, pois como explicam Costa et al. (2018) “a escolha da *Rectified Linear Unit* ocorreu pelo fato de ser, atualmente, o estado da arte das funções de ativação para RNC”. O comportamento da *ReLU* permite o aprendizado mais rápido, se comparado com a função logística, além de não saturar e evitar problemas de desaparecimento do gradiente, o que permite RNC com um grande número de camadas intermediárias (He et al., 2015; Nair e Hinton, 2010). A Figura 7 mostra o gráfico da *ReLU* dado pela formulação $f(x) = \max(0, x)$, onde é visível que não há a ativação de todos os neurônios ao mesmo tempo, pois é atribuído nulo como saída para valores negativos.

Figura 7 - Gráfico *ReLU* (*Rectified Linear Units*).



Fonte: A autora.

As funções de ativação também são aplicadas ao final das RNCs para classificar os dados de saída. Dentre estas funções, destaca-se a *softmax* (descrita pela Equação 1 e apresentada graficamente na Figura 8), colocada após uma *FCL* para atribuir probabilidades de um objeto pertencer à determinada classe (Yuan et al., 2017). Independente das entradas, a *softmax* tem como contradomínio valores entre 0 e 1 e; além disso, devido à função exponencial, ela resulta em valores pequenos para dados negativos e valores próximos à 1 para grandes valores.

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (1)$$

Onde:

z : Vetor com valores de saída da *FCL*

z_i : Elemento de z

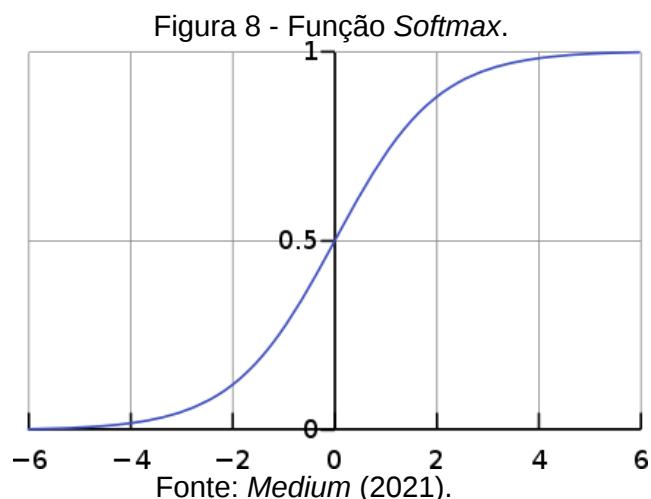
e^{z_i} : Função exponencial aplicada a cada elemento de z

$\sum_{j=1}^K e^{z_j}$: Esse termo permite normalizar o resultado para o intervalo de 0 a 1

K : Número de classes

Fonte: Goodfellow et al. (2016).

Na Figura 8 há o comportamento dessa função.



As RNCs podem ainda ter um processo de *dropout* ao longo das camadas. Essa técnica tem como característica eliminar aleatoriamente e temporariamente alguns neurônios, para em seguida retropropagar o resultado e atualizar pesos e vieses. Esse processo é repetido para outros conjuntos de neurônios para reduzir a possibilidade de coadaptações entre eles e, conseqüentemente, reduzir a possibilidade de *overfitting* no sistema (Khan et al., 2019), ou seja, a possibilidade do modelo se ajustar demasiadamente aos dados.

2.3.2 HIPERPARÂMETROS PARA O TREINAMENTO DA REDE

O treinamento de uma RNC depende dos seguintes hiperparâmetros: taxa de aprendizagem, época e tamanho do lote (*batch size*). A determinação dessas medidas deve ser antecessora ao treinamento e seu valor estipulado tende a ditar a qualidade da classificação (Aszemi e Selvam, 2019). Como dito, hiperparâmetros são designados para melhorar a precisão da classificação; no entanto, o objetivo da taxa de aprendizagem é monitorar o custo computacional do treinamento, ou seja, esse parâmetro controla o grau de ajuste das sinapses ao decorrer das épocas de treinamento (Costa et al., 2018). Com isso, uma taxa de aprendizagem baixa causa um aprendizado lento da rede, enquanto valores maiores podem impedir a convergência do processo de aprendizado (Faceli et al. 2011; Goodfellow et al. 2016; Deng e Yu, 2014).

De acordo com Aszemi e Selvam (2019), o número de épocas determina a quantidade de iterações do treinamento de todo conjunto de dados. Esse valor, empírico, depende do tamanho do modelo e da variação do conjunto, sendo que poucas épocas podem não ser suficientes para treinar o modelo e muitas épocas

podem ocasionar um *overfitting* e um alto custo computacional.

Por fim, o tamanho do lote, também conhecido como *batch size*, é um hiperparâmetro que define o número de exemplos de treinamento usado em cada época (Krizhevsky et al., 2017), quanto maior o seu valor, mais rápido o processamento poderá ser executado, no entanto, grandes valores podem ocasionar uma generalização.

Apesar de não ser um hiperparâmetro de treinamento, por não ditar a qualidade da classificação, outra informação importante a se estipular antes de treinar a rede é a função perda. Esta função desempenha o papel de comparar as saídas da rede com os valores de referência, ou seja, as classes reais (Zhao et al., 2018). Existem diversas funções disponíveis para calcular essa discrepância, como o erro logarítmico, o erro médio quadrático e a entropia cruzada, que avalia a diferença entre as probabilidades previstas e as probabilidades reais de classificação em cada classe. Além disso, a entropia cruzada é especialmente projetada para estimar estatísticas relacionadas a eventos de probabilidade muito baixa e pode ser empregada em problemas de otimização (Fyfe et al., 2012).

Para uma melhor compreensão da entropia cruzada, considere uma rede neural com três classes (Equação 2): A, B e C. O primeiro passo envolve o cálculo das probabilidades ($P(A)$, $P(B)$ e $P(C)$) de um determinado pixel (X) pertencer a cada uma das classes A, B e C. Em seguida, cada uma dessas probabilidades é comparada com os rótulos verdadeiros correspondentes ($V(A)$, $V(B)$ e $V(C)$). Além disso, é aplicado o logaritmo natural às probabilidades, o que enfatiza os erros de previsão e atenua o impacto das previsões corretas. Posteriormente, as multiplicações entre os valores verdadeiros e os logaritmos das probabilidades são somadas; e para garantir um valor não negativo, o resultado é multiplicado por -1. Quanto mais próximas às previsões estiverem dos rótulos reais, menor será o valor da entropia cruzada.

$$H(x) = -[V(A) * \log(P(A)) + V(B) * \log(P(B)) + V(C) * \log(P(C))]$$
 (2)

Fonte: Boltzmann (1878).

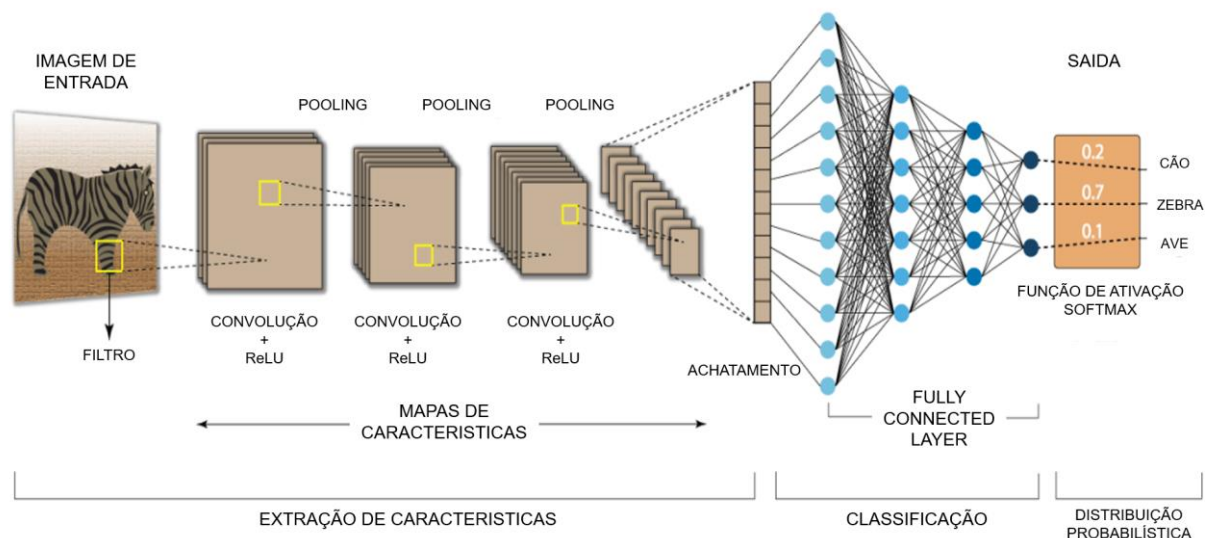
Além da função perda, a rede normalmente conta com um otimizador para acelerar o processamento. Como explicam Tato e Nkambou (2018), o objetivo de um otimizador é minimizar uma função perda, e isso consiste em encontrar o conjunto de parâmetros da arquitetura que forneça os melhores resultados nas tarefas visadas

como classificação, previsão ou agrupamento. Existem diversos algoritmos com esse objetivo, como o *back-propagation*, gradiente descendente e o gradiente adaptativo, mas como ressaltam Goodfellow et al. (2016), Isola et al. (2017) e Xu et al. (2015), o *adam* (*Adaptive Moment Estimation*) é um dos algoritmos mais utilizados na literatura. A atualização do *adam* ocorre pelo primeiro momento normalizado pelo segundo momento e sua metodologia acaba sendo eficiente para trabalhar com grandes problemas envolvendo muitos dados ou parâmetros (Tato e Nkambou, 2018).

A Figura 9 mostra um exemplo em que a RNC recebe dados de entrada do tipo *raster*, que podem conter qualquer quantidade de bandas. Em seguida, ocorre o processo de convolução, a função de ativação (*ReLU*) é aplicada e depois o conjunto de dados é reduzido pelo processo de *pooling*. Esta parte inicial (extração de características) pode ter quantas camadas ocultas forem necessárias para extrair as características pertinentes.

O último mapa de características é achatado, isso é, transformado em um vetor unidimensional, para que cada neurônio da camada de entrada receba a informação do pixel referente. Dentro da *FCL* há um processo similar à rede neural convencional e, desta forma, os pesos e vieses são ajustados para encontrar a melhor relação entre as imagens e as classes. Por fim, na última camada é aplicada uma função de ativação que atua como classificador e mapeia a saída não normalizada de uma rede para uma distribuição de probabilidade.

Figura 9 - Exemplo de uma Rede Neural Convolutacional.



Fonte: Adaptado de Jinadasa et al. (2022).

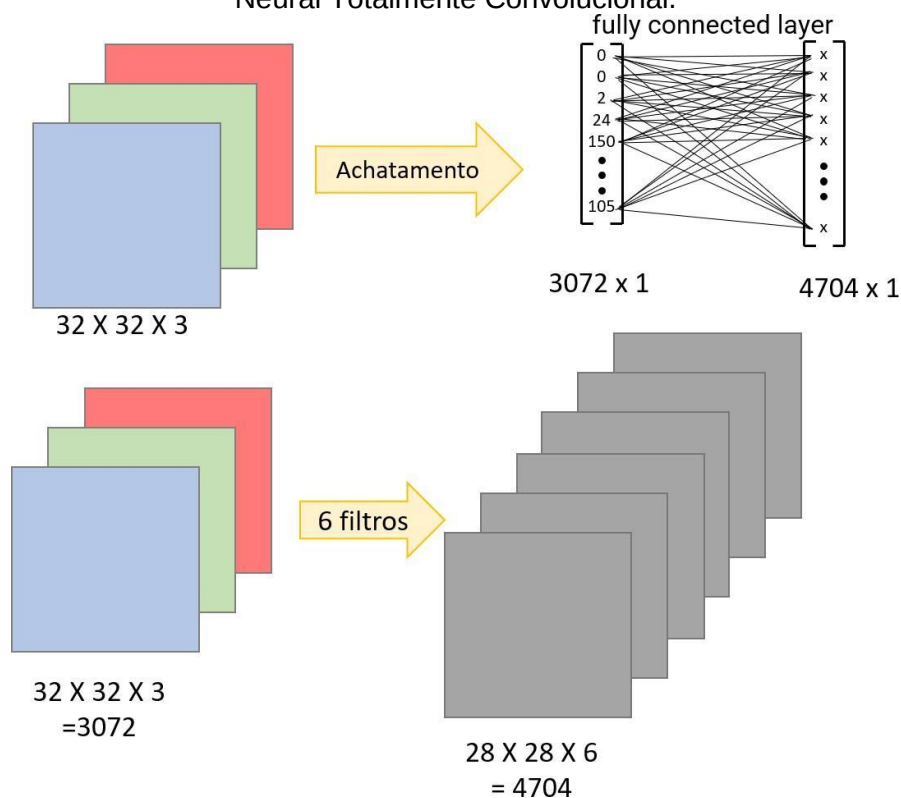
2.4 ARQUITETURA DE REDE: U-NET

A arquitetura de rede U-Net é uma técnica de segmentação semântica e sua nomeação é devido a sua estrutura em forma de U, como mostra a Figura 11 no final desse capítulo. Desenvolvida em 2015 por Ronneberger et al. (2015), a sua principal aplicação é a classificação de imagens biomédicas, mas apesar disso, ela tem ganhado espaço na área de processamento de imagens em SR (Chen et al., 2021; Ren et al., 2020; Wu et al., 2021).

Em 2015, ano da publicação da arquitetura U-Net, Long et al. (2015) anteciparam as futuras direções das Redes Neurais Convolucionais (RNCs), notando que, além da detecção de objetos por meio de caixas delimitadoras e correspondência de locais, essas redes seriam capazes de prever objetos em cada pixel de uma imagem, tal como realizado pela U-Net. Diferentemente de outras RNCs, a U-Net se distingue pela sua arquitetura de Rede Neural Totalmente Convolucional (RNTC), na qual em vez de depender de camadas totalmente conectadas (*FCL*), a U-Net realiza o processo de classificação utilizando camadas convolucionais totalmente conectadas e uma função de ativação no final, que determina a classe mais provável. Os criadores desta rede, Ronneberger et al. (2015), a projetaram de forma que ela não tenha uma *FCL*, focando apenas nas partes válidas de cada operação de convolução, mantendo o contexto completo disponível na imagem de entrada.

As *FCL* de RNCs típicas, como a LeNet e a AlexNet (Zhang, 2021) recebem entradas de tamanho fixo e produzem saídas não espacializadas. No entanto, essas *FCL* podem ser vistas como convoluções, com filtros que cobrem toda a sua região de entrada (Long et al. 2015). Na Figura 10 há um exemplo de como a *FCL* e as RNTC se equivalem. Na parte de cima da Figura 10 há o achatamento de uma imagem de dimensão $32 \times 32 \times 3$ para entrar em uma *FCL* e classificar a imagem de entrada, onde a camada escondida possui 4704 neurônios, resultando em $3072 \times 4704 = 14450688$ conexões. Na parte de baixo, da Figura 10, utilizando uma RNTC, o processo é simplificado ao contar com 6 filtros e gerar a mesma quantidade de neurônios. Ainda de acordo com Long et al. (2015), quando ocorre essa equivalência, como a apresentada na Figura 10, o cálculo é altamente amortizado, de maneira que a RNTC chega a ser de 5 vezes mais rápida.

Figura 10 - Exemplo de correspondência entre uma *Fully Connected Layer* e uma Rede Neural Totalmente Convolucional.



Fonte: A autora.

Além de ser uma RNTC, a U-Net é caracterizada por um processo de contração e outro de expansão. Conforme descrito por Ronneberger et al. (2015), o caminho de contração segue um padrão típico de RNC, consistindo em repetidas aplicações de duas convoluções 3×3 , seguidas por uma camada *ReLU* e uma operação de *maxpooling* 2×2 com passo 2.

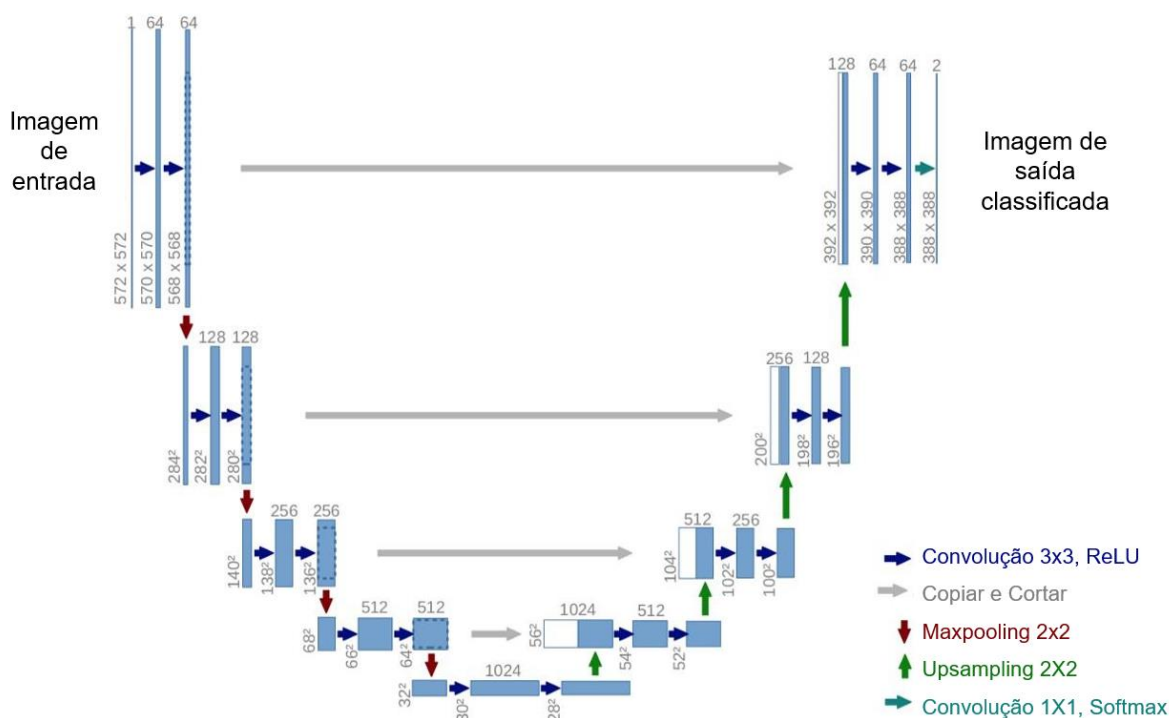
Na fase de contração, conhecida como a "descida do U" (conforme ilustrado na Figura 11), a rede recebe a imagem de entrada e inicia a extração de características por meio de operações de convolução. Por exemplo, considere uma imagem de entrada com dimensões de $572 \times 572 \times 1$. Conforme essa imagem passa por 64 filtros de tamanho 3×3 , seguidos pela aplicação da função de ativação *ReLU*, sua dimensão muda para $570 \times 570 \times 64$. Após um segundo ciclo de convolução semelhante, a imagem é reduzida para $568 \times 568 \times 64$. Em seguida, uma operação de *maxpooling* é aplicada, reduzindo pela metade as dimensões espaciais da imagem. Esse processo se repete quatro vezes até alcançar a "base do U," onde ocorrem duas convoluções, substituindo o *maxpooling* pela *upsampling*.

O recurso de *upsampling* é um dos principais aspectos da "subida do U" (Figura 11), que constitui o caminho de expansão. O *upsampling* aumenta a resolução

das saídas geradas pela fase de contração. Após o *upsampling*, segue-se a concatenação com o mapa de características gerado no processo de contração, e duas convoluções seguidas por uma camada *ReLU*. O objetivo desta etapa é combinar informações das camadas anteriores para obter previsões mais precisas.

A rede possui 23 camadas convolucionais, distribuídas da seguinte forma: 10 camadas no caminho de contração, 8 camadas no caminho de expansão, 4 camadas nos processos de concatenação (copiar e cortar) e uma última camada que aplica uma convolução 1x1 com a função *softmax* para mapear os recursos no número desejado de classes. Ronneberger et al. (2015) também ressaltam a importância de selecionar o tamanho das imagens de entrada de modo que todas as operações de *maxpooling* 2x2 sejam aplicadas a uma camada com dimensões uniforme.

Figura 11 - Arquitetura da rede U-Net.



Fonte: Adaptado de Ronneberger et al. (2015).

2.5 MÉTRICAS DE AVALIAÇÃO

A finalidade desta seção é apresentar as métricas de avaliação que podem ser utilizadas para avaliar os resultados obtidos com a classificação e a detecção de mudanças.

Antes da detecção de mudanças é necessário garantir que as imagens utilizadas estejam no mesmo referencial de coordenadas, e para isso uma imagem é

georreferenciada em relação à uma outra de referência. Nessa etapa, localizar com precisão os pontos de controle não é tão simples, e isso acaba gerando ruídos (Hong e Zhang, 2005), ou seja, diferenças nas coordenadas pontuadas com relação àquelas calculadas pelo processo de registro.

O erro médio quadrático (*RMSE*) (Equação 3) é comumente adotado para medir a diferença entre os valores previstos e os reais observados. Em essência, o *RMSE* representa a raiz quadrada da média do quadrado da diferença entre os valores previstos e os valores reais. E quanto mais próximo o valor do *RMSE* estiver de zero, mais precisas serão as previsões em relação aos valores reais, sendo que o resultado é expresso na mesma unidade que os dados originais.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_i^R - y_i^P)^2}{n}} \quad (3)$$

Onde:

y_i^R : Valor real;

y_i^P : Valor previsto;

n : Quantidade de pontos.

Fonte: Chai e Draxler (2014).

Ao considerar a classificação obtida por meio da interpretação visual como referência, a avaliação do desempenho da classificação utilizando a U-Net pode ser realizada, por exemplo, por meio da matriz de confusão. Esta matriz possibilita o cálculo de métricas estatísticas para a avaliação de mapeamentos temáticos, como a Acurácia Global (AG) (Black e Dwyer, 1990), que fornece uma estimativa da proporção de acerto global dos classificadores; o índice Kappa (Cohen, 1960), que indica a concordância entre as classificações observadas e as esperadas; e o coeficiente de similaridade de Jaccard ou *Intersection over Union (IoU)* (Jaccard, 1894), que mostra a sobreposição entre as áreas classificadas e as áreas de referência.

Provost et al. (1998) declaram que a matriz de confusão é responsável por apresentar informações sobre a diferença real e a prevista feita por um modelo de classificação e fornece uma visualização desse desempenho. As linhas geralmente representam a distribuição classe a classe dos dados resultantes, enquanto as colunas apresentam os dados de referência. É importante notar que a posição dos dados de referência e classificação, seja nas linhas ou colunas, é totalmente arbitrária

e depende das preferências do usuário (Diniz, 2021). Na computação, é comum dispor as referências em linhas, enquanto no SR, é mais frequente dispor as referências nas colunas.

A matriz de confusão também pode ser interpretada usando importantes indicadores de desempenho, como Verdadeiro Positivo (VP), Falso Positivo (FP) e Falso Negativo (FN). O FP ocorre quando um pixel é erroneamente adicionado a uma classe, embora não deva pertencer a ela. Por outro lado, um FN ocorre quando um pixel que deveria pertencer a uma determinada classe é classificado incorretamente, e os VP são os pixels que são classificados de forma correspondente ao dado de referência.

A Figura 12 ilustra duas disposições de uma Matriz de Confusão. Ao lado esquerdo, na matriz A, as classificações são representadas nas linhas, identificadas por rótulos em verde, enquanto os dados de referência são dispostos nas colunas, identificadas por rótulos em vermelho. A diagonal principal, destacada em cinza, representa os valores de VP para cada classe. Acima da diagonal, em azul, estão os FP, enquanto abaixo da diagonal, em bege, estão os FN. Na matriz B, observa-se que a inversão do posicionamento das referências e classificações, em linhas ou colunas, resulta na inversão da posição de FN e FP. No caso do SR, a matriz mais usual é a A, e é a que se utilizará esse trabalho. Nessa matriz, o somatório da coluna representa a verdade terrestre (referência), enquanto o total da linha representa o resultado da classificação da imagem.

Figura 12 - Matriz de Confusão 3x3 com indicadores (VP, FP e FN).

A		Dado de Referência		
		A	B	C
Dado Classificado	A	VP_A	FP	FP
	B	FN	VP_B	FP
	C	FN	FN	VP_C

B		Dado Classificado		
		A	B	C
Dado de Referência	A	VP_A	FN	FN
	B	FP	VP_B	FN
	C	FP	FP	VP_C

Fonte: Adaptado de Diniz (2021).

Para uma interpretação mais precisa, recomenda-se normalizar a matriz de confusão, proporcionando precisões uniformes para cada categoria de classificação

(Zhuang et al., 1995). No contexto de imagens classificadas digitalmente, a matriz de confusão oferece informações sobre a contagem de pixels classificados em uma classe específica quando deveriam estar em outra classe. Por sua vez, a matriz de confusão normalizada expressa esses valores como proporções em relação ao total de instâncias reais na classe correspondente. Na Figura 13A, à esquerda, é apresentada uma matriz de confusão abrangendo três classes, A, B e C, enquanto à direita, na Figura 13B, encontra-se a versão normalizada.

Figura 13 - A: Matriz de confusão 3x3 B: Matriz de confusão normalizada 3x3.

A		Dado de Referência		
		A	B	C
Dado Classificado	A	VP_A	FP	FP
	B	FN	VP_B	FP
	C	FN	FN	VP_C
		T_A	T_B	T_C



B		Dado de Referência		
		A	B	C
Dado Classificado	A	$\frac{VP_A}{T_A}$	$\frac{FP}{T_B}$	$\frac{FP}{T_C}$
	B	$\frac{FN}{T_A}$	$\frac{VP_B}{T_B}$	$\frac{FP}{T_C}$
	C	$\frac{FN}{T_A}$	$\frac{FN}{T_B}$	$\frac{VP_C}{T_C}$
		$\frac{T_A}{T_A}$	$\frac{T_B}{T_B}$	$\frac{T_C}{T_C}$

Fonte: A autora.

A normalização da matriz de confusão destaca uma métrica essencial em seus elementos: a AC (Equação 4) para cada classe ao longo da diagonal principal. Essa medida avalia a proporção de elementos classificados corretamente em relação ao total de elementos de interesse (Utsumi et al., 2023). Sua escala varia de 0% a 100%, onde 100% indica que, para aquela classe específica, os pixels classificados correspondem exatamente aos de referência. Desta forma, a AC indica a probabilidade de uma categoria classificada representar a mesma categoria de referência.

$$AC_j = \frac{VP_j}{T_j} \quad (4)$$

Onde:

AC_j : é a acurácia do consumidor para a classe j ;

VP_j : é o valor do Verdadeiro Positivo para a classe j ;

T_j : é o total da coluna (referência) da classe j .

Fonte: Adaptado de Diniz (2021).

Além da análise da acurácia para cada classe, os indicadores de desempenho (VP, FP e FN) também possibilitam o cálculo da AG. Essa métrica indica a proporção geral de acertos do classificador e pode ser calculada conforme demonstrado na Equação 5, determinando quão precisamente o modelo é capaz de classificar (Silva, 2018). O resultado varia de 0% a 100%, sendo 100% indicativo de uma alta acurácia do modelo, enquanto valores mais baixos sugerem uma acurácia média ou baixa, indicando que o modelo pode não estar tão preciso na classificação das amostras.

$$AG = \frac{\sum VP}{\sum T} \quad (5)$$

Onde:

AG : *Acurácia Global*;

$\sum VP$: *Somatório dos valores corretamente classificados*;

$\sum T$: *Somatório das amostras disponíveis*.

Fonte: Adaptado de Diniz (2021).

Além da AG, o coeficiente Kappa proposto por Cohen (1960) também pode ser calculado. Essa métrica estatística é utilizada para avaliar o nível de concordância ou reprodutibilidade entre dois conjuntos de dados. De acordo com Rennó (2004), este coeficiente supõe que parte do acerto total poderia ser devido ao acaso e, portanto, o coeficiente Kappa (Cohen, 1960) estima a proporção de acerto que efetivamente ocorre. A Equação 6 mostra a formulação do coeficiente Kappa (Cohen, 1960) considerando os elementos da matriz de confusão, de tal forma que, quanto mais próximo de 1, melhor a classificação.

$$\kappa = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})} \quad (6)$$

Onde:

K: Coeficiente Kappa;

N: Número de observações;

r: Número de linhas da matriz de confusão

x_{ii}: Observação na linha *i* e coluna *i*;

x_{i+}: Total da linha *i*;

x_{+i}: Total da coluna *i*;

Fonte: Bishop et al. (1975).

Os valores do índice Kappa (Cohen, 1960) (Tabela 1) conduzem à interpretação da qualidade da classificação como propõe Landis e Koch (1977).

Tabela 1 - A qualidade da classificação associada aos valores do índice Kappa.

Valor de Kappa	Interpretação
Menor do que 0	Péssima
Entre 0,00 e 0,20	Ruim
Entre 0,21 e 0,40	Razoável
Entre 0,41 e 0,60	Boa
Entre 0,61 e 0,80	Muito boa
Entre 0,81 e 1,00	Excelente

Fonte: Landis e Koch (1977).

Outra medida de avaliação da classificação é o *IoU*, também conhecido como índice Jaccard (Jaccard, 1894). Esta métrica é frequentemente utilizada para comparar semelhanças entre duas formas: uma considerada verdade absoluta e outra prevista pelo modelo (Rezatofighi et al., 2019 e Pashaei et al., 2020). A resolução da métrica é obtida pela divisão da interseção pela união para cada classe, seguida pela média desses valores para avaliar o desempenho do modelo. Isso significa que, para cada classe, calcula-se a proporção da área onde a previsão do modelo e a verdade absoluta se sobrepõem (interseção) em relação à área total coberta por ambas (união). Valores próximos a 1 indicam alta semelhança entre as feições, valores próximos a 0 indicam baixa semelhança e valores intermediários indicam uma similaridade parcial entre as classes. Este índice é amplamente utilizado na avaliação

dos resultados da classificação de imagens devido às suas qualidades perceptivas e à invariância de escala, que atribui relevância apropriada a objetos pequenos (Berman et al., 2018). A Equação 7 mostra o cálculo da métrica *IoU* (Jaccard, 1894) utilizando os indicadores VP, FP e FN.

$$IoU = \frac{VP}{FP+VP+FN} \quad (7)$$

Fonte: Adaptado de Jha et al. (2020).

Para avaliar o desempenho da detecção de mudanças, uma matriz de confusão binária pode ser utilizada, conforme ilustrado na Figura 14, onde VN é entendido como Verdadeiro Negativo. Além disso, as métricas analisadas abrangem a precisão, o *recall* e o *f1-score*; a precisão mede a proporção de VP em relação a todas as amostras positivas previstas, enquanto o *recall* indica a proporção de VP em relação a todas as amostras positivas reais e o *f1-score* é uma média harmônica entre precisão e *recall*, fornecendo um equilíbrio entre essas métricas.

Figura 14 - Matriz de Confusão Binária (Detecção de Mudanças).

		REFERÊNCIA	
		MUDANÇA	NÃO MUDANÇA
PREVISTO	MUDANÇA	VP	FP
	NÃO MUDANÇA	FN	VN

Fonte: A autora.

A precisão, conforme definida na Equação 8, é obtida como a razão entre o número de VP e o total de detecções, e como ressalta Mao et al. (2019), é empregada para mitigar FP. A precisão varia no intervalo de 0 a 1, onde o valor 1 indica que o modelo não gerou FP.

$$Precisão = \frac{VP}{VP+FP} \quad (8)$$

Fonte: Lipton et al. (2014).

Para identificar a capacidade do modelo em evitar FN, pode-se utilizar a métrica conhecida como *recall* (Equação 9). Como observam Buckland e Gey (1994), essa métrica avalia quão eficaz é o processo de detectar os itens aos quais a metodologia se propõe. Assim, como a precisão, o *recall* varia de 0 a 1, e o valor 1 indica que a metodologia não cometeu nenhum FN.

$$Recall = \frac{VP}{VP+FN} \quad (9)$$

Fonte: Lipton et al. (2014).

No entanto, quando o conjunto de dados é desequilibrado, ou seja, quando há uma distribuição desigual entre classes, a precisão não pode mais ser considerada uma medida confiável (Chicco e Jurman, 2020). Dessa forma a métrica conhecida como *f1-score* (Equação 10) pode ser calculada, levando em consideração tanto a precisão quanto o *recall*, proporcionando uma medida equilibrada que considera, tanto os VP quanto a capacidade de recuperar corretamente os exemplos positivos, mesmo em situações de desequilíbrio de classes.

$$F1 - Score = \frac{2*Precisão*Recall}{Precisão+Recall} \quad (10)$$

Fonte: Lipton et al. (2014)

A Tabela 2 sintetiza as métricas de avaliação a serem empregadas nesse estudo e detalhadas nesse capítulo. Na primeira coluna, é indicado o procedimento que cada métrica analisará. Em seguida, são apresentadas a métrica utilizada, sua fórmula e o intervalo de valores possíveis, com uma progressão que reflete a melhoria: o limite esquerdo representa o pior valor esperado, enquanto o limite direito representa o melhor valor esperado.

Tabela 2 - Resumo das métricas utilizadas

Método Avaliado	Métrica Utilizada	Fórmula	Intervalo
Registro	<i>RMSE</i>	$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_i^R - y_i^P)^2}{n}}$	$] - \infty, +\infty[$
Classificação	Acurácia do Consumidor	$AC_j = \frac{VP_j}{T_j}$	$[0\%, 100\%]$
Classificação	Acurácia Geral	$AG = \frac{\sum VP}{\sum T}$	$[0\%, 100\%]$
Classificação	Kappa	$\kappa = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})}$	$[-1, 1]$
Classificação	<i>IoU</i>	$IoU = \frac{VP}{FP + VP + FN}$	$[0, 1]$
Detecção de Mudanças	Precisão	$Precisão = \frac{VP}{VP + FP}$	$[0, 1]$
Detecção de Mudanças	<i>Recall</i>	$Recall = \frac{VP}{VP + FN}$	$[0, 1]$
Detecção de Mudanças	<i>F1-Score</i>	$F1 - Score = \frac{2 * Precisão * Recall}{Precisão + Recall}$	$[0, 1]$

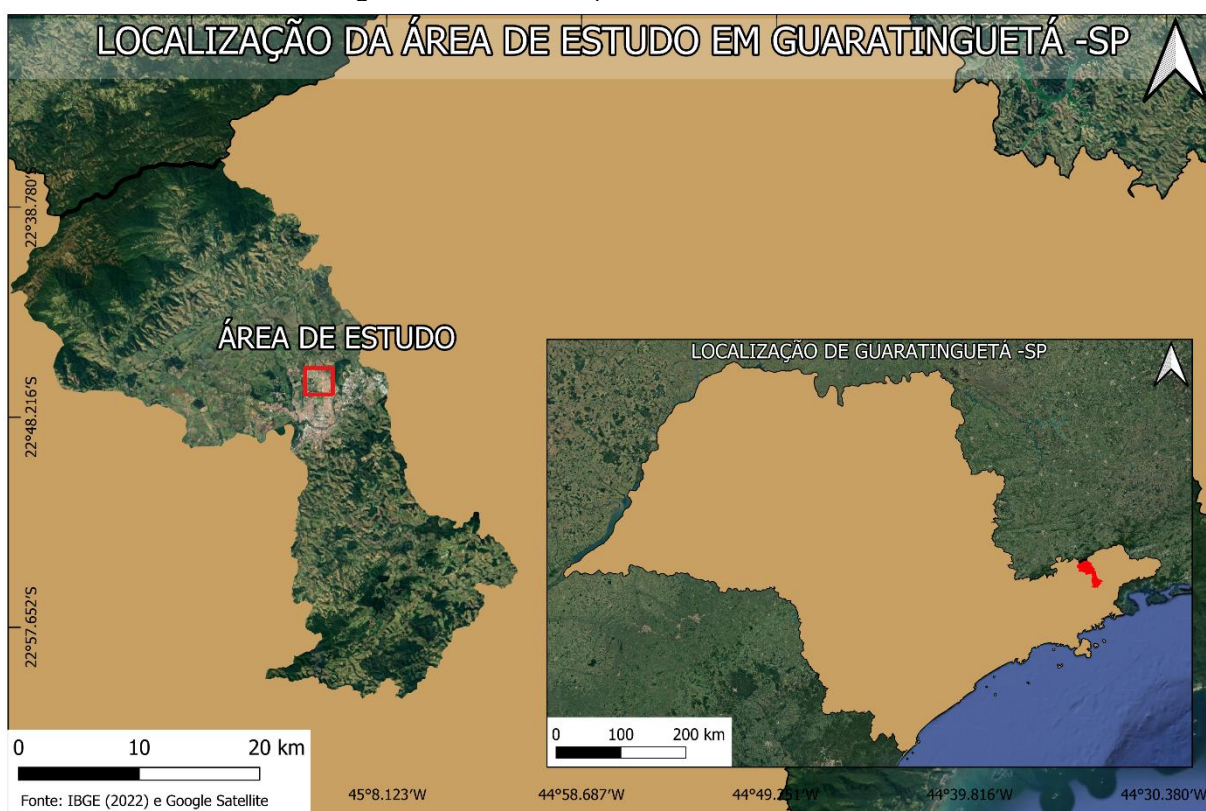
Fonte: A autora.

3. MATERIAL E MÉTODOS

3.1 ÁREA DE ESTUDO

A área de estudo (Figura 15) está situada no município de Guaratinguetá, no estado de São Paulo, delimitada pelas seguintes coordenadas geográficas: 22°46'02"S 45°12'06"W e 22°47'14"S 45°10'45"W. Este município está localizado na mesorregião do Vale do Paraíba Paulista, na parte leste do estado e segundo dados do Censo (IBGE, 2022), sua população é de cerca de 118.044 habitantes.

Figura 15 - Localização da área de estudo.



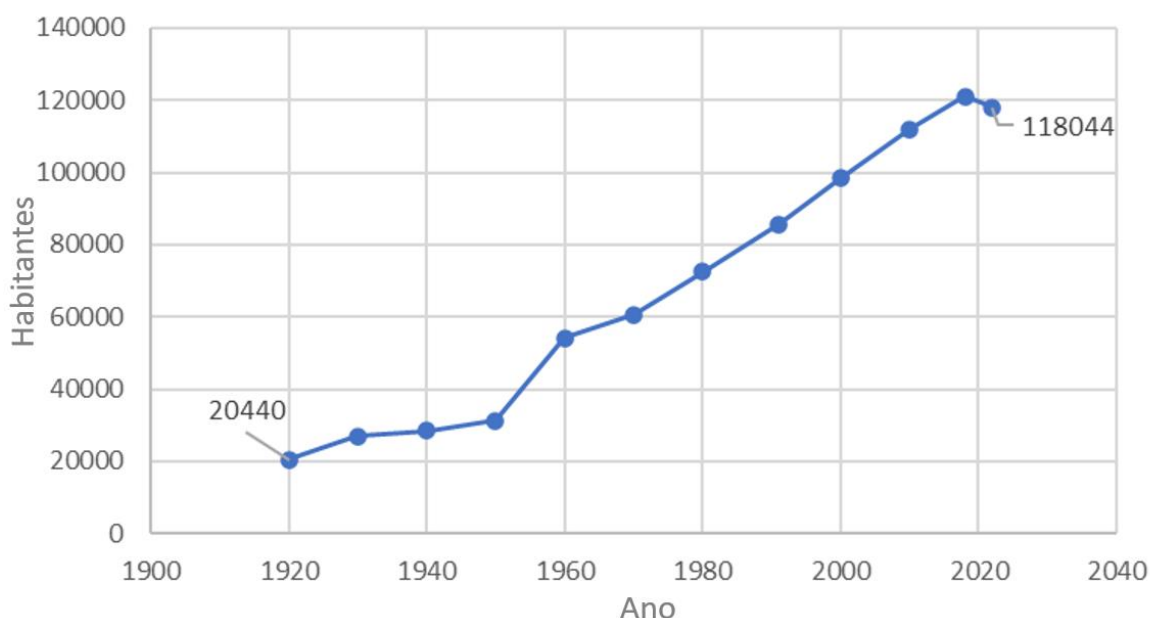
Fonte: A autora.

Conforme a classificação de Nimer (1989), a região onde o município está inserido possui um regime de chuvas tipicamente tropical. As maiores precipitações ocorrem na parte norte, próxima à Serra da Mantiqueira, com uma média anual de 1700 mm e na porção sul, os índices de precipitação são mais baixos, em torno de 1400 mm por ano. Segundo a Embrapa (2012), os meses de maio a setembro registram os menores índices de precipitação, geralmente inferiores a 50 mm por mês, enquanto durante dezembro e fevereiro são observadas as maiores quantidades, frequentemente superiores a 200 mm mensais.

A variação altimétrica do município é de aproximadamente 1500 metros. As elevações mais altas estão localizadas ao norte, ultrapassando os 2000 metros na Serra da Mantiqueira, enquanto na região sul as altitudes giram em torno de 1000 metros, associadas à presença da Serra do Mar. As altitudes mais baixas, em torno de 510 metros, são encontradas na área central, próxima ao Rio Paraíba do Sul.

Fundada em 1630, Guaratinguetá viu seu crescimento concentrado principalmente às margens esquerdas do rio Paraíba do Sul. Sendo que, em cerca de 100 anos, a população da cidade sextuplicou, como revela o gráfico da Figura 16, adaptado a partir de dados do IBGE (2022).

Figura 16 - Evolução populacional em Guaratinguetá (1920-2022).



Fonte: A autora.

O gráfico ilustra a evolução populacional de Guaratinguetá ao longo de várias décadas, desde 1920 até 2023. Durante esse período, é notável um crescimento contínuo da população, com aumentos significativos em determinadas décadas. Por exemplo, entre 1920 e 1960, a população quase triplicou, passando de 20.440 para 54.200 habitantes. Esse crescimento manteve-se gradual nas décadas seguintes, com um incremento expressivo entre 1970 e 1980, quando a população ultrapassou a marca dos 70.000 habitantes. Essa tendência de crescimento continuou até 2018, quando a população atingiu seu pico de 121.073 habitantes. No entanto, em 2023, houve uma leve queda em relação ao ano anterior, totalizando 118.044 habitantes.

Esses dados sugerem um padrão consistente de crescimento populacional ao

longo do tempo, influenciado por diversos fatores, tais como migração, natalidade e desenvolvimento econômico. É importante ressaltar que esse crescimento pode ter consequências, como a expansão urbana desordenada, mencionada no Plano Diretor de Guaratinguetá de 2019 (Líder Engenharia e Gestão de Cidades, 2019). O documento destaca a existência de 42 loteamentos irregulares, bem como questões ambientais, como a ausência de mata ciliar e o descarte irregular de entulho, enfatizando a necessidade de uma metodologia que auxilie a gestão municipal na detecção desses eventos.

3.2 MATERIAL

O mapeamento de uso cobertura do solo para detecção de mudanças em áreas urbanas tem como dados imagens do satélite *WorldView – 2*, cujas características estão apresentadas na Tabela 3.

Tabela 3 - Características dos sensores *WorldView – 2*.

Sensor	Banda Espectral	Resolução Espectral (nm)	Resolução Espacial (m)	Resolução Temporal	Resolução Radiométrica
Pancromático	Pancromática	450 – 800	0,46 (nadir) 0,52 (20º off-nadir)	1,1dias no nadir 3,7 dias a 20º off-nadir	11 bits
Multiespectral	Costal	400 – 450	1,85 (nadir) 2,08 (20º off-nadir)		
	Azul	450 – 510			
	Verde	510 – 580			
	Amarelo	585 – 625			
	Vermelho	630 – 690			
	Red Edge	705 – 745			
	Infravermelho Próximo -1	770–895			
	Infravermelho Próximo -2	860 – 1040			

Fonte: Adaptado de Embrapa (2018).

Nesse estudo, utilizam-se imagens ópticas disponibilizadas pela empresa GeoPixel Geotecnologias Consultoria e Serviços. Essas imagens foram fornecidas exclusivamente para este trabalho, com o acesso restrito a quatro bandas espectrais específicas. As características detalhadas das imagens utilizadas estão na Tabela 4.

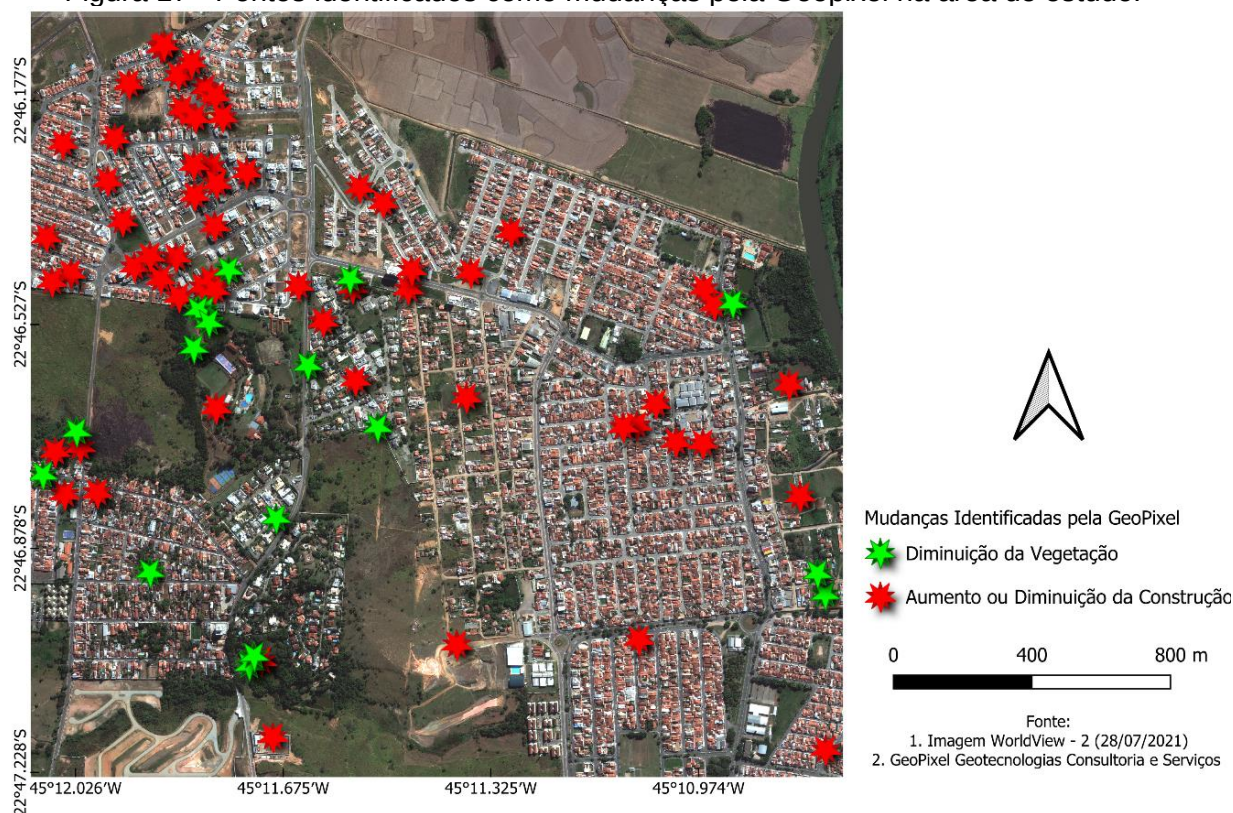
Tabela 4 - Dados das imagens utilizadas.

Data da Imagem	Bandas Utilizadas	Ortorretificação	Correção Atmosférica
28/07/2021	Azul, Verde, Vermelho e Infravermelho Próximo-1	Sim	Sim
24/07/2022	Azul, Verde, Vermelho e Infravermelho Próximo-1	Sim	Sim

Fonte: A autora.

Além das imagens, pontos de ocorrência de mudanças, também fornecidos pela empresa GeoPixel, são empregados como referência para a identificação de áreas em que ocorrem mudanças. Esses pontos indicam locais onde há supressão da *vegetação arbórea*, bem como o aumento ou a diminuição das *construções*. A distribuição desses pontos está representada na Figura 17.

Figura 17 - Pontos identificados como mudanças pela Geopixel na área de estudo.



Fonte: A autora.

Para implementação da metodologia de detecção de mudanças, a linguagem *Python* é utilizada devido às diversas bibliotecas que possui para aprendizado de máquina e processamento de imagens, além de ser uma linguagem bem consolidada na área de SR.

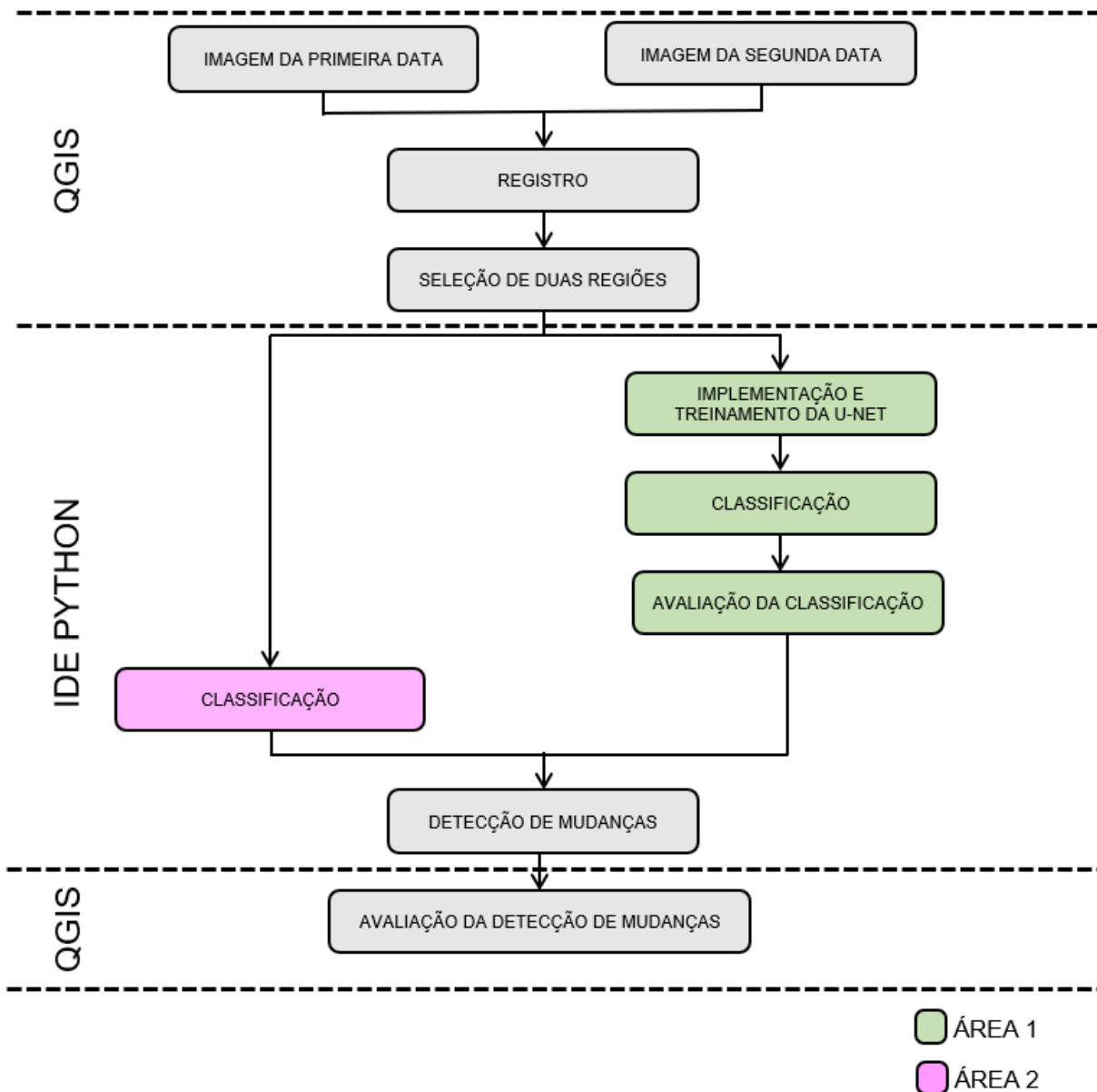
Para realização dos processamentos das imagens e implementação da rede U-Net são utilizados os seguintes aplicativos:

- IDE *Google Colab*, ambiente de desenvolvimento integrado da plataforma *Google Colaboratory*, para armazenamento em nuvem de notebooks voltados à criação e execução de códigos em *Python*. Disponível em “<https://colab.research.google.com/>”;
- QGIS 3.14, Sistema de Informação Geográfica disponível em “<http://qgis.osgeo.org>”;
- Bibliotecas Numpy (<https://numpy.org/>), Keras (<https://keras.io/>), TensorFlow (<https://www.tensorflow.org/?hl=pt-br>), Rasterio (<https://rasterio.readthedocs.io/en/stable/>);
- Notebook Lenovo com GPU GeForce MX150 4GB e 8GB de memória;

3.3 METODOLOGIA

A metodologia adotada no estudo é visualizada no fluxograma apresentado na Figura 18. Na representação gráfica, os procedimentos executados na *área 1* (usada no treinamento da rede) na primeira e segunda datas são destacados em verde. Da mesma forma, os procedimentos aplicados na *área 2* (não usada no treinamento da rede) são indicados em rosa, correspondendo às duas datas de análise. Os retângulos em cinza representam processos que são realizados independentemente da divisão em áreas. Além disso, os tracejados indicam as ferramentas utilizadas nas etapas.

Figura 18 - Fluxograma da metodologia.

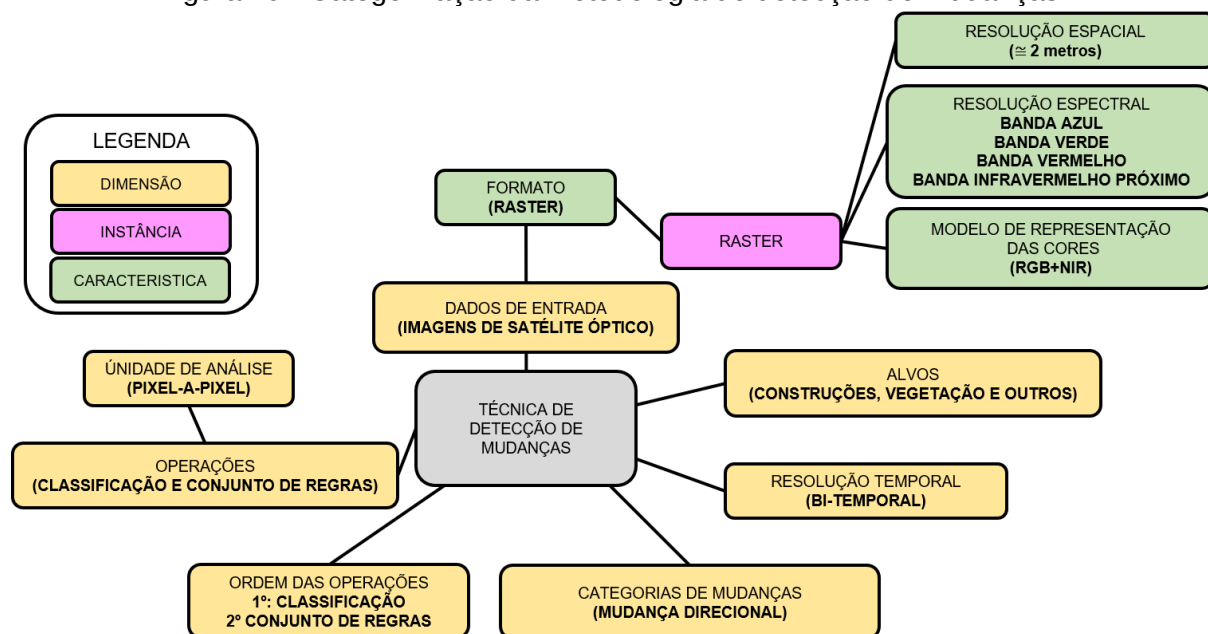


Fonte: A autora.

Para uma compreensão mais detalhada da metodologia, adotou-se a proposta de Si Salah et al. (2019), explicada no capítulo 2, para categorizar o processo de detecção de mudanças. Na Figura 19, em **negrito**, há a descrição de cada característica adotada na metodologia utilizada nesse trabalho para detecção de mudanças do uso e cobertura do solo. A começar, a *unidade de análise* é ao nível de pixel, tanto na operação de classificação, quanto no conjunto de regras. A *ordem de operações* utilizadas se dá primeiramente com a classificação das imagens, para depois a aplicação do conjunto de regras. A *categoria de mudanças* é dita como mudança direcional, pois é analisada a transição de uma classe para outra, e como

só há duas imagens na metodologia, a *resolução temporal* admitida é a bi-temporal. Para esse trabalho, os *alvos* analisados são *construções* e *vegetação arbórea*, no entanto, para a etapa de classificação também é considerado o alvo *outros*, que inclui todas as demais categorias que não se enquadram especificamente como *construções* ou *vegetação arbórea*, como áreas pavimentadas, corpos d'água, e solo exposto. Essa inclusão permite uma classificação mais abrangente e precisa dos diferentes tipos de cobertura do solo presentes na área de estudo. Os dados de entrada, como melhor mencionados na seção anterior, são imagens *WorldView – 2* com as três bandas da região do visível e uma do infravermelho próximo.

Figura 19 - Categorização da metodologia de detecção de mudanças.



Fonte: A autora.

A seguir, os procedimentos metodológicos apresentados no fluxograma da Figura 18 são detalhados.

3.3.1 CRIAÇÃO DO CONJUNTO DE DADOS

Para executar a metodologia proposta, na detecção de mudanças é essencial garantir a compatibilidade de referência entre as duas imagens por meio da operação de registro ou georreferenciamento. Essa operação é realizada no software QGIS, utilizando uma função polinomial de primeiro grau e 20 pontos de controle, identificados em ambas as imagens. Esses pontos, distribuídos de maneira uniforme em toda a área de estudo, são cruciais para alinhar adequadamente as imagens e

assegurar uma análise precisa das mudanças.

A escolha do polinômio de grau 1 baseia-se no trabalho de Mohammed e Elhaj (2013), onde as equações de primeira, segunda e terceira ordem são analisadas no georreferenciamento de imagens de satélites e os autores demonstram que as três obtêm aproximadamente o mesmo resultado, sendo indicado o uso da função polinomial de ordem 1 devido a menor complexidade computacional. Para analisar a qualidade do georreferenciamento, o *RMSE* é calculado, admitindo como valor aceitável, nesse trabalho, um erro menor que meio pixel (Almeida e Oliveira, 2010), isso é, menor que 1 metro.

Para validar a metodologia de detecção de mudanças, a área de estudo é dividida em duas regiões, identificadas como *área 1* e *área 2*, sendo que apenas a *área 1* é utilizada no treinamento da U-Net. A seleção dos tamanhos das áreas leva em consideração as limitações de tempo para a coleta manual de amostras, o que resulta na geração de máscaras provenientes da interpretação visual apenas na *área 1*, isso é, imagens multiclassadas que representam a presença ou ausência das classes de interesse em cada pixel. Essa abordagem possibilita a avaliação da qualidade da detecção de mudanças na região definida pela *área 2*, sem que ocorra influência do processo de treinamento da rede.

Na implementação e treinamento da U-Net, consideram-se três classes: *vegetação arbórea*, *construções* e *outros*. No entanto, apenas duas dessas classes são pertinentes para a detecção de mudanças: *vegetação arbórea* e *construções*. Variações nessas classes podem indicar desmatamento, moradias irregulares, descarte irregular de entulho, supressão ou aumento de construções. Por exemplo, uma diminuição na cobertura vegetal pode indicar desmatamento, enquanto uma redução na área construída pode apontar para a remoção de construções. Por outro lado, um aumento na classe *construções* pode sugerir descarte irregular de entulho, expansão da área construída ou moradias irregulares devido a resposta espectral. Porém, ressalta-se que este trabalho não trata da classificação da detecção de mudanças, mas a identificação de mudanças nessas classes pode implicar em modificações relevantes.

A Figura 20 mostra um exemplo de cada classe considerada nesse trabalho, nas imagens do sensor *WorldView – 2*, em composição verdadeira, onde alvos não considerados *vegetação arbórea* ou *construções* são incluídos na classe *outros*. A

classe *construção* tem como elementos principais edificações e piscinas, enquanto a *vegetação arbórea* é composta por árvores.

Figura 20 – Exemplos das classes adotadas na classificação.



Fonte: A autora.

A etapa seguinte requer a interpretação visual das amostras (*área 1*), de modo que a rede possa ser implementada para as duas datas. Como as máscaras oriundas da interpretação visual são geradas para cada data, não é necessário realizar a correção radiométrica entre as duas imagens, eliminando a necessidade de ajustar os valores de radiância ou reflectância para compará-las.

O processo de interpretação visual consiste em converter cada pixel da área de estudo em pontos rotulados com números inteiros variando de 0 a 2, dos quais cada classe equivale a um valor (Tabela 5). Posteriormente, cada camada vetorial é convertida para formato *raster*, sendo que cada ponto corresponde a um pixel, o que resulta em duas imagens para cada data considerada: a original e a interpretação visual.

Tabela 5 - Correspondência entre as classes e os identificadores numéricos.

Identificador Numérico	Classe
0	Outros
1	Construções
2	Vegetação Arbórea

Fonte: A autora.

Para cada data, tanto a imagem original quanto a imagem resultante da interpretação visual são divididas em um total de 725 imagens. As divisões das imagens originais resultam em uma dimensão de 200x200 pixels e 4 bandas, enquanto as máscaras têm uma dimensão de 200x200 pixels e 1 banda. Essa subdivisão é realizada por meio de uma grade, na qual cada célula tem uma dimensão de 100x100 metros e uma sobreposição de 30 metros. A divisão é determinada com base na análise visual para verificar se as feições são observáveis, ou seja, em uma imagem de 200x200 pixels, é possível identificar feições do tipo *construções* e *vegetação arbórea*, além disso, as imagens são sobrepostas para aumentar o conjunto de dados.

O total de 725 imagens, resultantes do recorte das imagens originais de 2021 e 725 imagens correspondentes a 2022, juntamente com as 725 imagens, previamente rotuladas (máscaras) para cada uma das datas, são armazenadas no ambiente de nuvem do *Google*.

3.3.2 IMPLEMENTAÇÃO DA U-NET

A implementação e parametrização da U-Net são efetuadas utilizando a linguagem de programação *Python*, no ambiente do *Google Colaboratory* e os procedimentos para implementação da U-Net são equivalentes para as duas datas.

O primeiro passo consiste em adequar os dados para que a rede possa lê-los; assim, as imagens resultantes da original passam a ocupar uma posição no tensor² X , resultando uma dimensão de (724, 200, 200, 4), e as máscaras ocupam o tensor y , de dimensão (724, 200, 200, 1). Em aprendizado de máquinas, um conjunto amostral maior tende a proporcionar um melhor aprendizado. No entanto, como o conjunto estudado não é denso, após testes, opta-se por manter 80% dos dados para treinamento e reservar 20% para teste.

O próximo passo é a definição da função perda, que quantifica o que a arquitetura deve minimizar durante o treinamento. Para isso utiliza-se a entropia cruzada, pois mostra uma convergência melhor do modelo, quando comparada a outras funções como *Focal Loss* (Lin et al., 2020) e *Dice Loss* (Sudre et al., 2017).

Em seguida, a U-Net é construída na forma de uma função com os seguintes parâmetros: largura e altura de cada imagem, número de bandas e número de classes.

² Tensor: um conjunto de valores formatados dentro de um vetor de qualquer dimensão. (TensorFlow, 2022)

A arquitetura se manteve a mesma da Figura 11, mas *dropouts* são incluídos após convoluções para reduzir a probabilidade de *overfitting* dos dados.

3.3.2.1 TREINAMENTO, TESTE E ANÁLISE DE QUALIDADE DA U-NET

Nessa etapa, a rede neural é treinada em duas datas distintas, utilizando o otimizador adam. Além disso, a função de perda entropia cruzada, juntamente com a métrica *IoU* (Jaccard, 1894), são aplicadas em cada iteração para análise de convergência. Adicionalmente, de forma empírica, são determinados valores para o *batch size*, taxa de aprendizagem e quantidade de épocas.

Durante o treinamento, espera-se uma queda no resultado da função perda dos dados de treinamento e teste, visto que a cada iteração, as classes determinadas pela rede devem se aproximar das classes rotuladas inicialmente. Por outro lado, se presume que o índice *IoU* (Jaccard, 1894) aumente ao longo do treinamento, indicando uma maior similaridade entre o resultado obtido e o esperado. Portanto, para certificar esse comportamento entre as métricas, um gráfico é gerado, mostrando essa variação ao longo das épocas, tanto para os dados de treinamento quanto de teste.

Após o treinamento, para avaliar o grau de discernimento entre as classes, uma matriz de confusão normalizada é gerada para cada data e as métricas AG, *IoU* (Jaccard, 1894) e o índice Kappa (Cohen, 1960) são calculadas. Com a U-Net treinada, as imagens da *área 1* podem ser classificadas e unidas novamente para resultar em apenas uma imagem classificada para cada data.

Antes de proceder com a detecção de mudanças e para estabelecer uma área de validação adicional, após a conclusão da análise da *área 1*, a *área de estudo 2* é delimitada e classificada. Nesse contexto, a *área 2* é dividida em 372 imagens, tanto as originais quanto as classificadas visualmente, resultando de uma subdivisão de 200x200 pixels, sem sobreposição. Utilizando a U-Net previamente treinada, a *área 2* é então submetida ao processo de classificação e em seguida, as imagens resultantes são posteriormente integradas em uma única imagem consolidada. Dessa forma, o resultado da *área 1* e da *área 2* são comparados visualmente.

3.3.3 DETECÇÃO DE MUDANÇAS

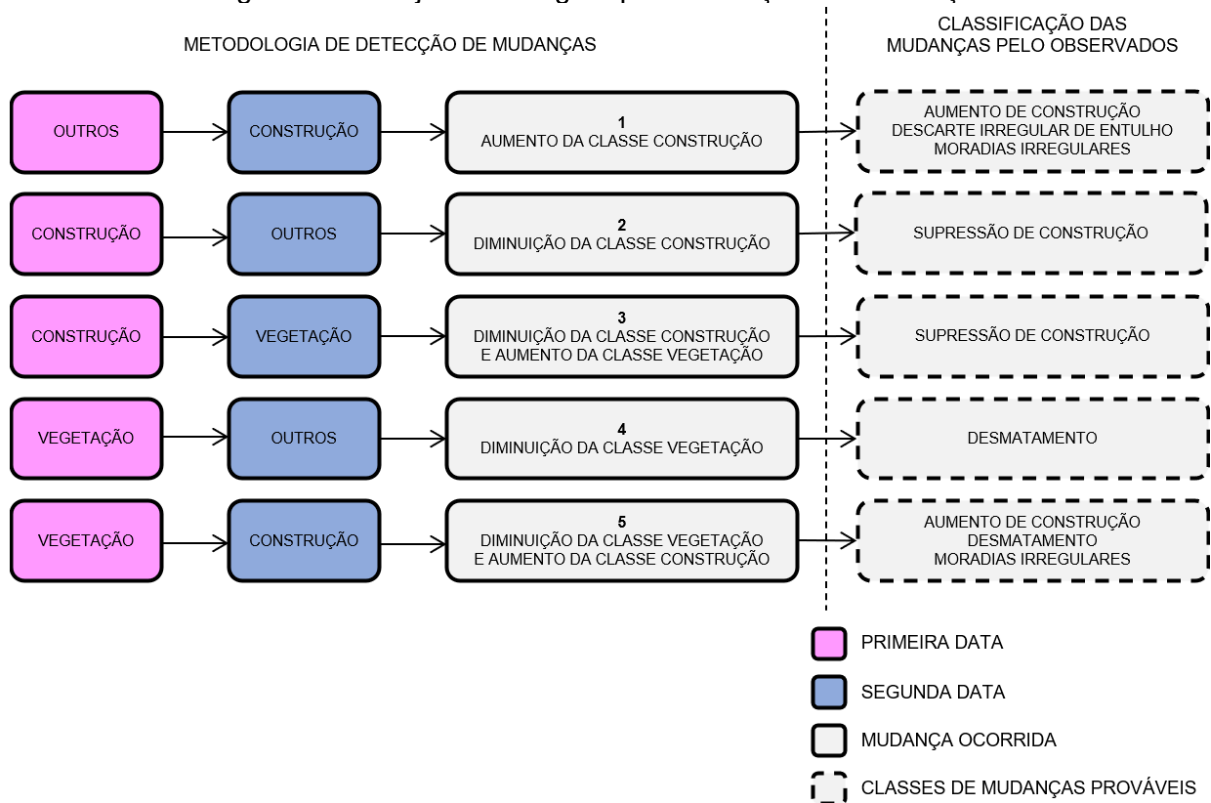
Com o objetivo de otimizar a identificação de áreas onde ocorrem aumento ou supressão de construções, desmatamento, descarte irregular de entulho e ocupação

irregular, este estudo busca detectar variações nas classes de *construções* e *vegetação arbórea*. A detecção de mudanças nessas duas classes serve como indicativo de uma possível modificação, que posteriormente pode ser classificada por um observador.

Após a classificação das quatro imagens, ou seja, das duas áreas em duas datas distintas, a etapa seguinte envolve a criação de um conjunto de regras para avaliar o tipo de alteração ocorrida nas classes de *construções* e *vegetação arbórea*. Essas regras são representadas na Figura 21. Posteriormente, uma imagem para cada área de estudo é gerada para rotular o tipo de modificação, seguindo critérios: 1, para indicar apenas o aumento da classe *construções*; 2, para representar apenas a diminuição da classe *construções*; 3, para denotar o aumento da *vegetação arbórea* e diminuição da classe *construções*; 4, para mostrar a diminuição da classe de *vegetação arbórea* e 5 para indicar a diminuição da classe *vegetação arbórea* e o aumento da classe *construções*.

Com base no resultado da rotulação, uma interpretação visual, levando em conta fatores social e ambiental, pode ser realizada nas áreas sinalizadas, sem a necessidade de percorrer toda a imagem. Dessa forma, pretende-se otimizar a detecção das ocorrências de desmatamento, aumento e supressão de construções, descarte irregular de entulho e construções irregulares. A última coluna da Figura 21 não será realizada nesse trabalho, mas indica as prováveis classes de mudanças que um observador pode classificar.

Figura 21 - Conjunto de regras para detecção de mudanças.



Fonte: A autora.

Para identificar e quantificar as modificações que ocorrem entre as classes, nos resultados da detecção de mudanças são incorporadas as matrizes de transição para as duas áreas de estudo. De acordo com Batty e Cheshire (2011), as matrizes de transição fornecem uma compreensão das mudanças entre as classes e mostra a dinâmica das alterações no uso e ocupação do solo urbano. Campos e Queiroz Filho (2017) complementam informando que a matriz de transição é obtida por meio da análise das proporções das diferentes classes em ambas as datas. Na Figura 22, é apresentada uma ilustração de uma matriz de transição que compreende duas datas e três classes (A, B e C). Cada elemento na matriz representa a probabilidade de transição de uma classe para outra, com a diagonal principal indicando a persistência das classes. A Figura 22 destaca, também, como são calculados os valores de ganho e perda nas transições entre cada classe.

Figura 22 - Exemplo de uma matriz de transição.

		Primeira Data				
		A	B	C	Total	Perda
Segunda Data	A	$P_{A,A}$	$P_{A,B}$	$P_{A,C}$	P_{A+}	$P_{A+} - P_{A,A}$
	B	$P_{B,A}$	$P_{B,B}$	$P_{B,C}$	P_{B+}	$P_{B+} - P_{B,B}$
	C	$P_{C,A}$	$P_{C,B}$	$P_{C,C}$	P_{C+}	$P_{C+} - P_{C,C}$
	Total	P_{+A}	P_{+B}	P_{+C}	1	
	Ganho	$P_{+A} - P_{A,A}$	$P_{+B} - P_{B,B}$	$P_{+C} - P_{C,C}$		

Fonte: A autora.

A perda para uma determinada classe i é calculada como a diferença entre o total de transições para essa classe P_{i+} e a persistência dessa classe P_{ii} . Com relação ao ganho, ele é calculado como a diferença entre o total de transições de todas as classes para essa classe P_{+i} e a persistência dessa classe P_{ii} .

3.3.3.1 ANÁLISE DE QUALIDADE DA DETECÇÃO DE MUDANÇAS

A análise de qualidade da detecção de mudanças é feita tanto para a *área 1*, utilizada no treinamento, quanto para a *área 2*, que não é utilizada no treinamento. Além disso, as classes de interesse da detecção de mudanças são avaliadas pelas métricas precisão, *recall* e *f1-score*, separadamente.

Para validar a detecção e mudanças, como referência, utiliza-se um arquivo de pontos interpretados pela GeoPixel que marcam locais onde as mudanças nas classes *construções* e *vegetação arbórea* ocorreram. Além disso, as imagens resultantes da metodologia de detecção de mudança passam por um processo de vetorização automática para contabilizar os polígonos de mudanças.

Através de uma análise visual, são contabilizados quantos pontos interpretados pela Geopixel aparecem contidos nos polígonos de mudanças detectadas pela metodologia proposta, quantos estão fora e quantos polígonos detectados como mudanças não contêm nenhum ponto da referência. O resultado desta contabilização é atribuído a valores que correspondem VP, FP e FN, definidos da seguinte forma:

- Verdadeiro Positivo (VP): Total de polígonos com pontos.

- Falso Positivo (FP): Total de polígonos sem pontos.
- Falso Negativo (FN): Pontos sem polígonos correspondentes.

Para uma melhor compreensão da interpretação visual destes resultados, a Figura 23 mostra um exemplo em que ocorrem a identificação das mudanças tipo 1 e 4, isto é, a classe *outros* passou para *construções* e a classe *vegetação arbórea* passou para *outros*. Ao realizar a análise visual verifica-se o pertencimento dos pontos de mudança das classes de *vegetação arbórea* e *construções* nos polígonos de mudança, indicando interpretação coerente. Dessa maneira, os VP e FN podem ser quantificados, representando respectivamente: os polígonos com pontos, ou seja, as mudanças detectadas corretamente; e os pontos sem polígonos, que indicam mudanças que ocorreram, mas não são identificadas pela metodologia. Os FP resultam da quantidade total de polígonos, presente na camada vetorial, menos os VP contabilizados anteriormente, evidenciando polígonos de mudanças detectados pela metodologia, mas que não são mudanças de acordo com a referência.

Figura 23 - Exemplo da análise visual para avaliação da qualidade da detecção de mudanças.



Fonte: A autora.

Em seguida, com base nas identificações dos indicadores de desempenho VP, FN e FP, três métricas são calculadas e avaliadas: a precisão, que quantifica a proporção de detecções corretas em relação ao total de detecções; o *recall*, que demonstra a proporção de detecções corretas em relação ao total real de mudanças; e o *f1-score*, uma métrica que combina precisão e *recall* em uma única medida para avaliar o equilíbrio entre esses dois aspectos da detecção. Essas métricas fornecem uma visão abrangente do desempenho da detecção de mudanças e permitem uma avaliação precisa da metodologia utilizada.

Na seção seguinte serão apresentados e discutidos os resultados obtidos com a metodologia proposta.

4. RESULTADOS E DISCUSSÃO

Conforme apresentado na metodologia, o georreferenciamento das imagens (2021 e 2022) é efetuado e a Figura 24 mostra a distribuição dos 20 pontos utilizados nessa operação sobrepostos na imagem em composição colorida verdadeira de 2021.

Figura 24 - Distribuição dos pontos utilizados no registro das imagens.



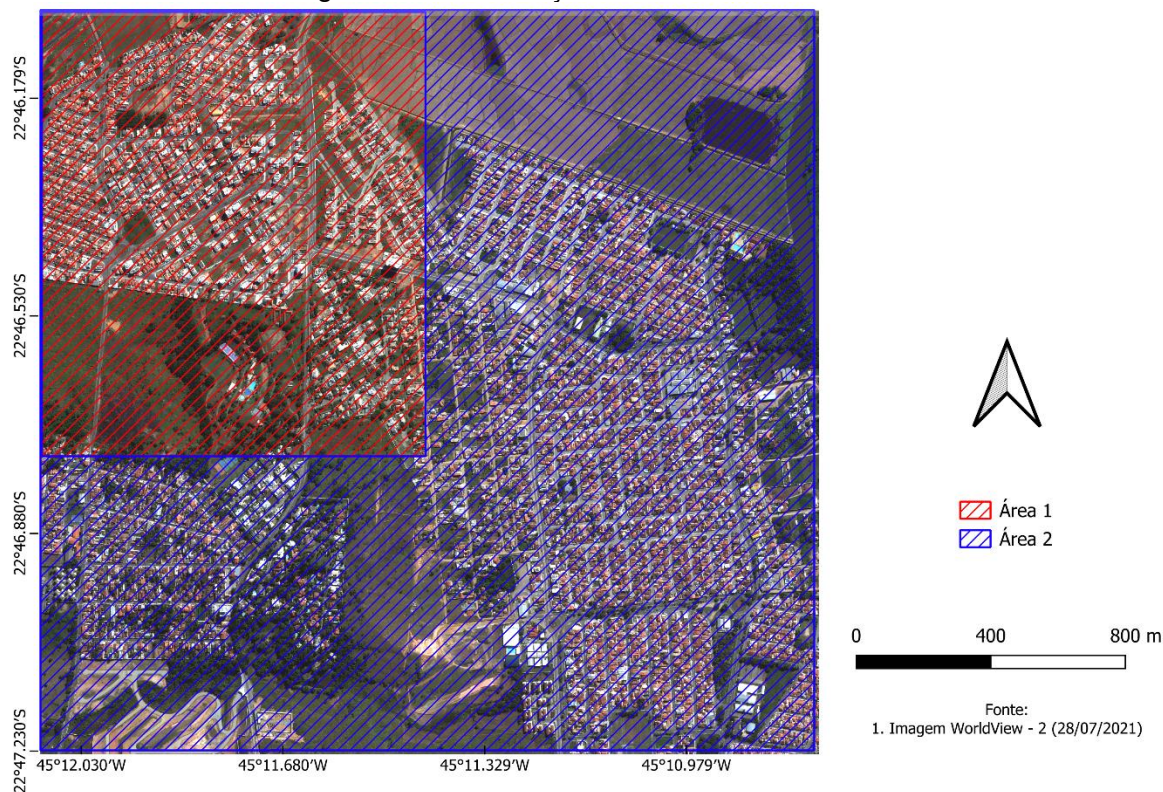
Fonte: A autora.

Por meio dos resíduos gerados em cada ponto, após o georreferenciamento, o *RMSE* é calculado obtendo um valor de 0,37 metros, ou seja, menor do que meio pixel (1 metro), indicando um bom georreferenciamento. No Apêndice A, nas quatro primeiras colunas, estão listadas as coordenadas obtidas nas duas imagens antes do processo de registro. As colunas ΔE e ΔN representam as variações nas coordenadas, no sistema de projeção cartográfica UTM, após o registro, enquanto o resíduo indica a distância euclidiana entre as coordenadas de 2021 e 2022, após o registro.

Em seguida, como mostra a Figura 25, a área de estudo é dividida em duas áreas, uma para treinamento, denominada *área 1*, com retângulo envolvente entre as coordenadas geográficas 22°46'02"S 45°12'06"W e 22°46'45.4"S 45°11'25.9"W e outra para validação da metodologia, denominada *área 2*, no restante da área de estudo. A seleção dos tamanhos das áreas levou em consideração as limitações de

tempo para a coleta manual de amostras, resultando na aplicação das máscaras geradas a partir do processo de interpretação visual apenas na *área 1*.

Figura 25 - Delimitação das áreas 1 e área 2.



Fonte: A autora.

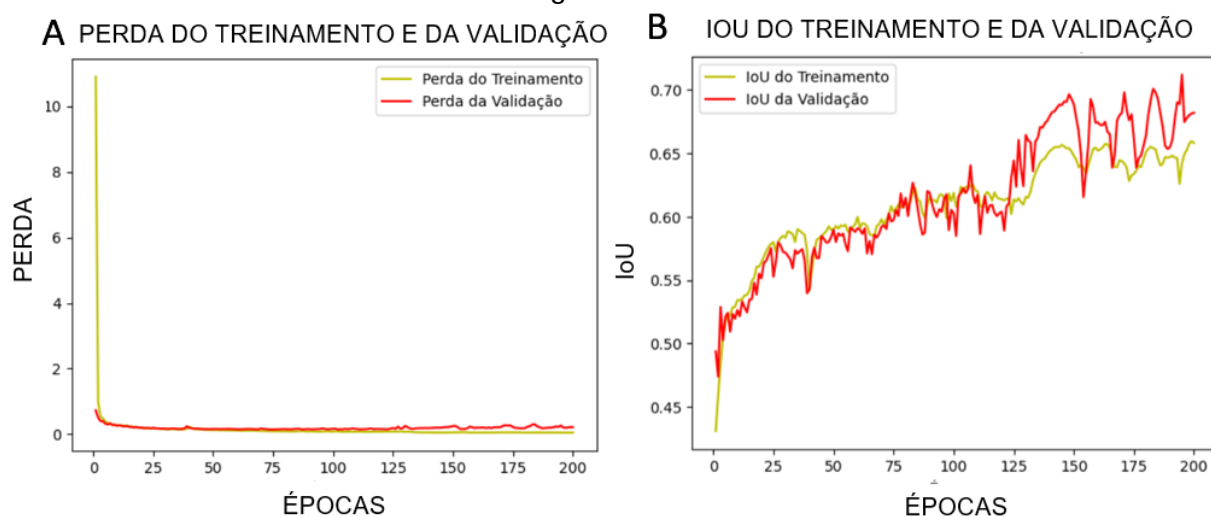
No processo de treinamento, a *área 1* é subdividida em 724 imagens de 4 bandas para ambas as datas, juntamente com 724 imagens classificadas por interpretação visual, ou seja, com uma única banda. Após a criação do conjunto de dados, avança-se com a implementação da U-Net, conforme demonstrado no Apêndice B.

Posteriormente, se conduzem testes com o objetivo de otimizar os hiperparâmetros de treinamento que resultam em uma avaliação satisfatória da qualidade da classificação das imagens. Como resultado de análise empírica, se estabelece uma taxa de aprendizagem de 0,001, 200 épocas de treinamento e um *batch size* de 40 imagens para a parametrização em ambas as datas distintas.

Para cada data, geram-se dois gráficos a fim de avaliar a convergência durante o processo de treinamento. Na Figura 26A, que corresponde à data de 2021, observa-se a diminuição dos valores da função de perda, aproximando-se de zero, tanto nos dados de treinamento quanto nos de teste. O segundo gráfico (Figura 26B) ilustra o aumento da métrica *IoU* (Jaccard, 1894) ao longo das épocas, estabilizando-

se na faixa de 0,6 a 0,7 em relação aos conjuntos de treinamento e validação. É importante notar que, ao considerar um número de épocas superior a 200, é observado um afastamento entre os resultados do treinamento e da validação, sugerindo a ocorrência de *overfitting*.

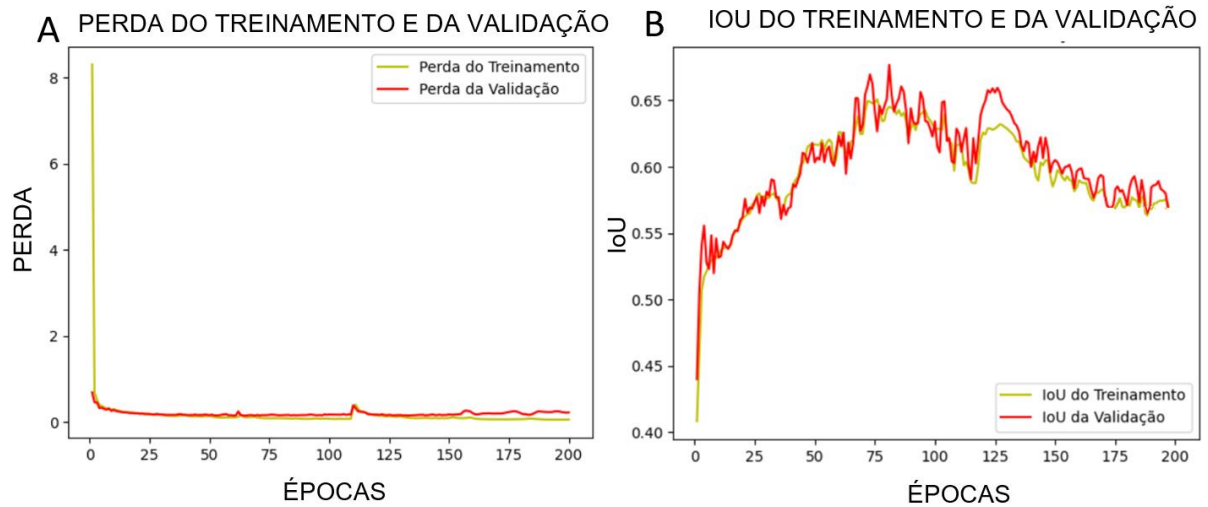
Figura 26 - Comportamento da função perda e da métrica *IoU* ao longo das épocas na imagem de 2021.



Fonte: A autora.

Os dados de treinamento da segunda data também são submetidos à análise utilizando gráficos da função de perda e métrica *IoU* (Jaccard, 1894). Na Figura 27A, observa-se a redução dos valores da função de perda, aproximando-se de zero, de forma semelhante ao que ocorreu na imagem de 2021. Em relação à métrica *IoU* (Figura 27B), é evidenciada uma estabilização em torno de 0,55 e 0,60. Assim como em 2021, o aumento no número de épocas resultou em maior ocorrência de *overfitting*.

Figura 27 - Comportamento da função perda e da métrica *IoU* ao longo das épocas na imagem de 2022.



Após a implementação e treinamento da U-Net, a *área 1* é classificada em ambas as datas e as matrizes de confusão normalizadas (Figura 28) são geradas.

Figura 28 - Matrizes de confusão normalizadas.

2021		Dado de Referência		
		<i>outros</i>	<i>construção</i>	<i>vegetação</i>
Dado Classificado	<i>outros</i>	96%	9,2%	15%
	<i>construção</i>	2,4%	91%	0,04%
	<i>vegetação</i>	1,9%	0,01%	85%

2022		Dado de Referência		
		<i>outros</i>	<i>construção</i>	<i>vegetação</i>
Dado Classificado	<i>outros</i>	98%	11%	19%
	<i>construção</i>	1,4%	89%	0,05%
	<i>vegetação</i>	0,65%	0,07%	81%

Fonte: A autora.

O resultado da matriz de confusão de 2021 mostra que a classe *outros* tem uma alta taxa de VP, ou seja, 96% dos pixels referentes a classe *outros* são corretamente classificados como *outros*. Uma pequena porcentagem, cerca de 2,4% dessa mesma classe é classificada como *construções* e menos ainda (1,9%) erroneamente classificadas como *vegetação arbórea*.

Com relação a classe *construções*, observa-se uma taxa de VP em cerca de 91% e uma proporção menor de pixels, aproximadamente 9,2%, são classificados erroneamente como *outros*. Esse valor de 9,2% se deve a resposta espectral semelhante entre feições consideradas *construções* e *outros*, como exemplifica a Figura 29. Nessa figura, o polígono em vermelho representa um quintal cimentado e

possui semelhança espectral com alvos da classe *construções*, no entanto, pela interpretação visual, apenas o polígono em azul é considerado *construções*, pois esse se trata de uma edificação. Além disso, há um valor extremamente baixo, de apenas 0,01% de pixels que confundiram *construções* com *vegetação arbórea*.

Figura 29 - Exemplo que pode acarretar a confusão entre pixels de *construções* e *outros*.



Fonte: A autora.

Com relação a classe de *vegetação arbórea*, em 2021, 85% dos pixels são classificados corretamente e uma quantidade muito pequena, de 0,04% se confundiu com *construções*. No entanto, uma porcentagem significativa de 15% da classe *vegetação arbórea* se confundiu com *outros*, isso se deve principalmente a delimitação da *vegetação arbórea* ser imprecisa. Na Figura 30, por exemplo, é possível observar que um pixel que delimita a copa de uma árvore pode ser confundido com vegetação rasteira ou sombra, além disso, dentro da própria copa há pixels de sombras.

Figura 30 – Exemplo que pode acarretar a confusão entre pixels de *vegetação arbórea* e *outros*.



Fonte: A autora.

A mesma interpretação da matriz de confusão normalizada de 2021 pode ser aplicada a 2022. O modelo de 2022 demonstra bons valores, sem a ocorrência de VP inferiores a 81%. Entretanto, as classes que mais apresentam confusões são *vegetação arbórea* com *outros*, com uma taxa de 19%, e *construções* com *outros*, com uma taxa de 11%. Essas confusões podem ser atribuídas aos mesmos problemas destacados nas Figuras 29 e 30.

Além disso, pode-se observar a AC para cada classe nos dois anos. Em 2021, a classe *outros* alcançou uma AC de 96%, enquanto *construções* e *vegetação arbórea* obtiveram 91% e 85%, respectivamente. Já em 2022, a AC para as classes *outros*, *construções* e *vegetação arbórea* é de 98%, 89% e 81%, respectivamente.

Após a classificação das duas imagens, são calculadas as métricas IoU (Jaccard, 1894), AG e Índice Kappa (Cohen, 1960), conforme apresentado na Tabela 7. O Índice de Jaccard (*IoU*) (Jaccard, 1894) é determinado para as imagens de 2021 e 2022, revelando valores de 67,80% e 61,26%, respectivamente. Esses números indicam uma considerável sobreposição entre as classificações produzidas pela U-Net e as classes de referência com relação à união das duas áreas. No entanto, é importante notar que a métrica IoU (Jaccard, 1894) é sensível à variação espacial, o que significa que mesmo pequenas variações na localização ou forma das áreas

classificadas podem ter um impacto significativo em seu valor. Isso ocorre porque a IoU (Jaccard, 1894) calcula a sobreposição entre duas áreas classificadas, e qualquer discrepância na localização ou forma das áreas pode resultar em diferenças na sobreposição, afetando assim o valor da métrica IoU (Jaccard, 1894). Devido a essa sensibilidade, entre as três métricas avaliadas (IoU, AG e Kappa), a IoU (Jaccard, 1894) é a que obteve um valor mais baixo.

A AG supera os 81% para ambas as datas, o que sugere que a U-Net apresenta um bom desempenho nas classificações, sendo capaz de classificar corretamente mais de 81% dos pixels em ambas as imagens. Além disso, ao analisar o coeficiente Kappa (Cohen, 1960), conforme proposto por Landis e Koch (1977), os resultados da classificação são considerados excelentes para ambas as datas, uma vez que estão dentro do intervalo de 80% a 100%.

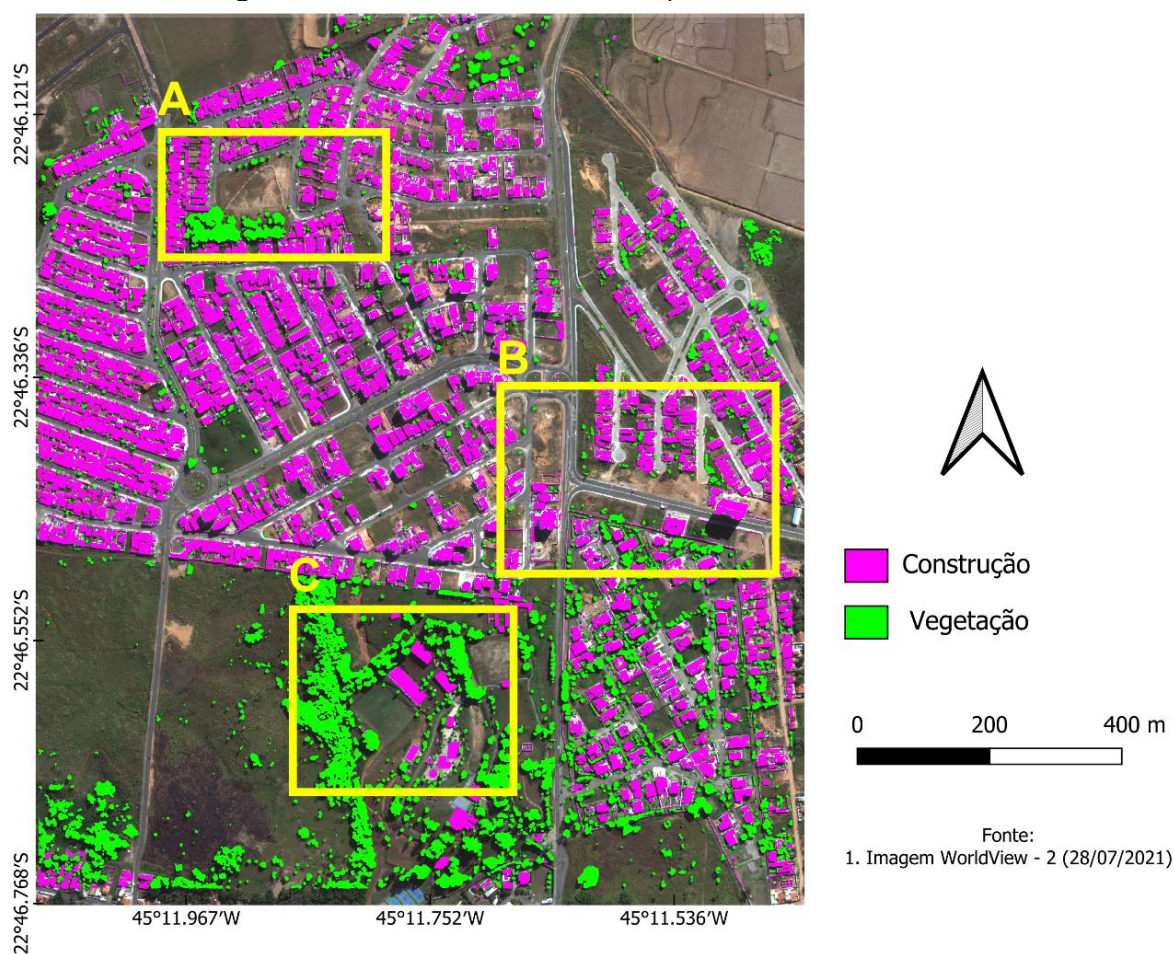
Tabela 6 - Métricas avaliativas das classificações usando a U-Net.

	2021	2022
<i>Intersection over Union (IoU) (%)</i>	67,80	61,26
Acurácia Global (%)	89,64	81,15
Kappa (%)	87,25	86,02

Fonte: A autora.

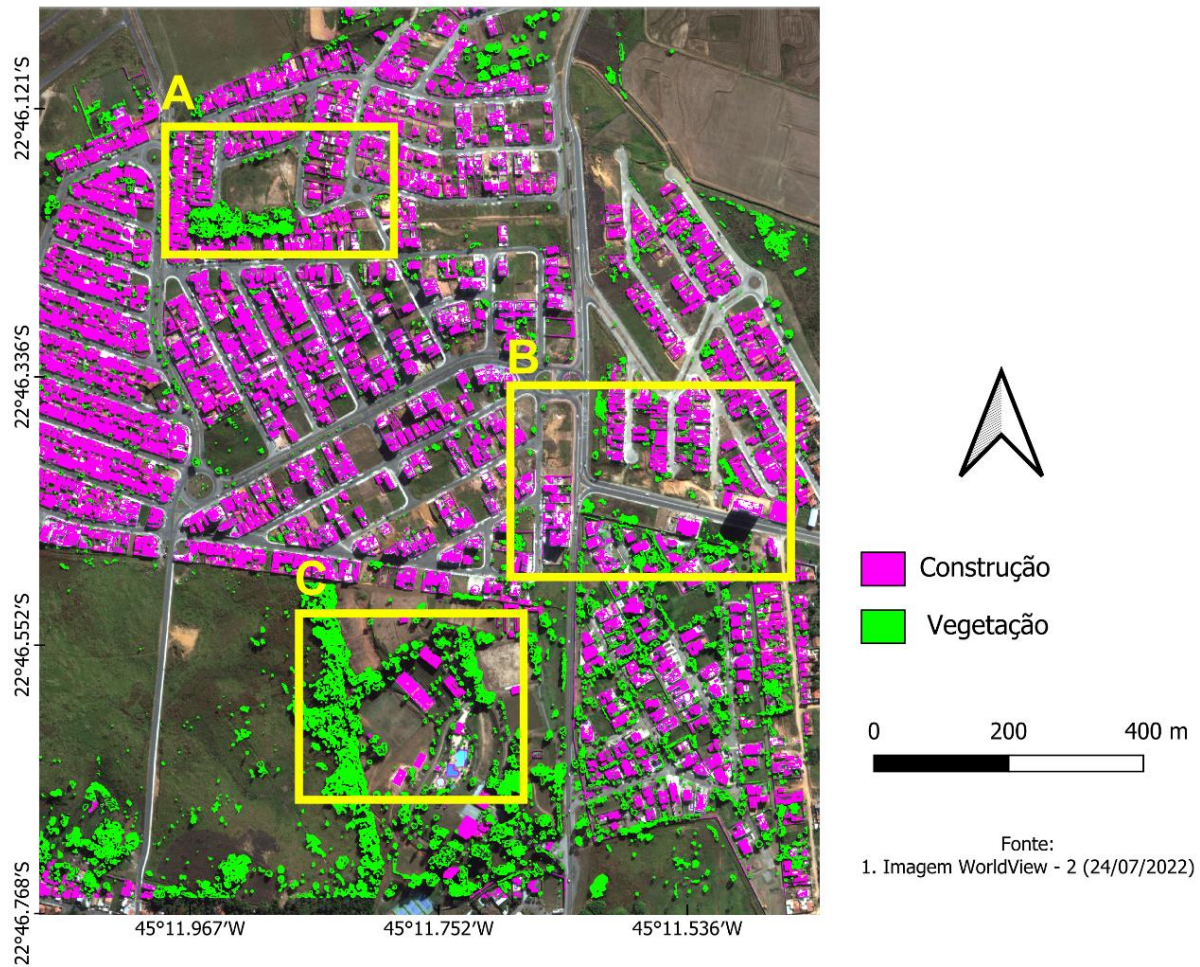
Após o cálculo das métricas, é conduzida uma análise visual dos resultados. De maneira abrangente, o resultado da *área 1*, em 2021, é mostrado na Figura 31, enquanto o resultado da classificação da *área 1*, em 2022, é apresentado na Figura 32. Nessas figuras, os resultados das classificações exibem as *construções* em rosa e a *vegetação arbórea* em verde. A classe *outros* está em transparente para que seja possível a visualização de que as demais classes realmente sobrepõem as feições desejadas da imagem do *WorldView – 2*, apresentada em composição colorida verdadeira. Nas Figura 31 e Figura 32, é possível verificar uma boa correspondência entre o classificado e a imagem de referência reiterando as métricas obtidas (Tabela 7). Entretanto, para uma análise mais detalhada, selecionam-se três regiões (A, B e C) permitindo uma observação aproximada e uma avaliação qualitativa dos resultados (Figura 33, Figura 34 e Figura 35).

Figura 31 - Resultado da classificação da área 1 em 2021.



Fonte: A autora.

Figura 32 – Resultado da classificação da área 1 em 2022.



Fonte: A autora.

Na Figura 33, ao analisar a região A em escala ampliada, constata-se que, em ambas as datas avaliadas, as *construções* e a *vegetação arbórea* são corretamente identificadas. No entanto, a classificação não é absoluta, sendo possível evidenciar algumas confusões desses alvos com a classe *outros*. Observa-se, por exemplo, segmentações da *vegetação arbórea* em 2022, destacada pela seta A. Além disso, a seta B indica casos em que alvos de *construções* são erroneamente classificados como *outros*, corroborando com as métricas apresentadas na Tabela 7 e a matriz de confusão da Figura 28.

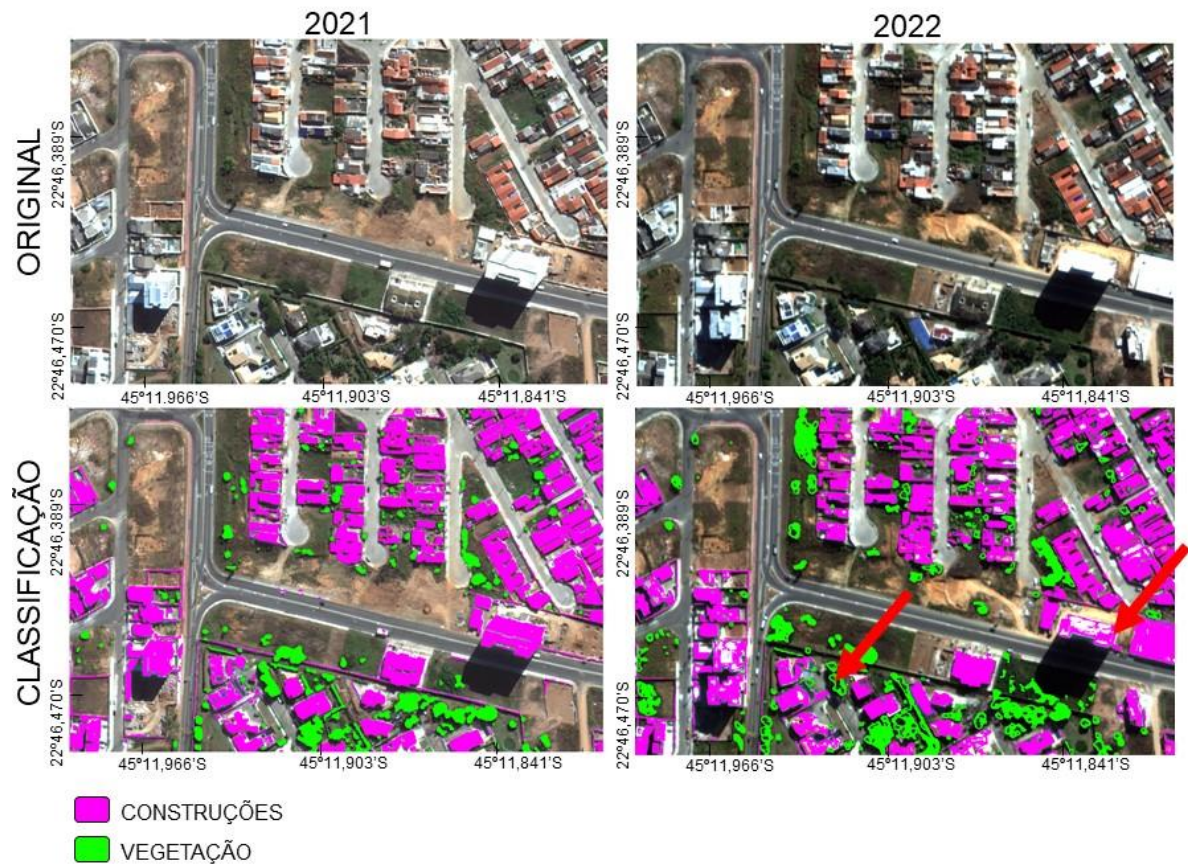
Figura 33 - Região de ampliação A: resultado de classificação da área 1 nas duas datas.



Fonte: A autora.

Na Figura 34, a região *B*, outra ampliação na imagem, mostra que boa parte dos alvos de interesse são classificados corretamente. No entanto, observam-se algumas *construções* e *vegetação arbórea* que são classificadas como *outros* na data de 2022 (setas), reforçando o apresentado na Tabela 7 e a matriz de confusão da Figura 28.

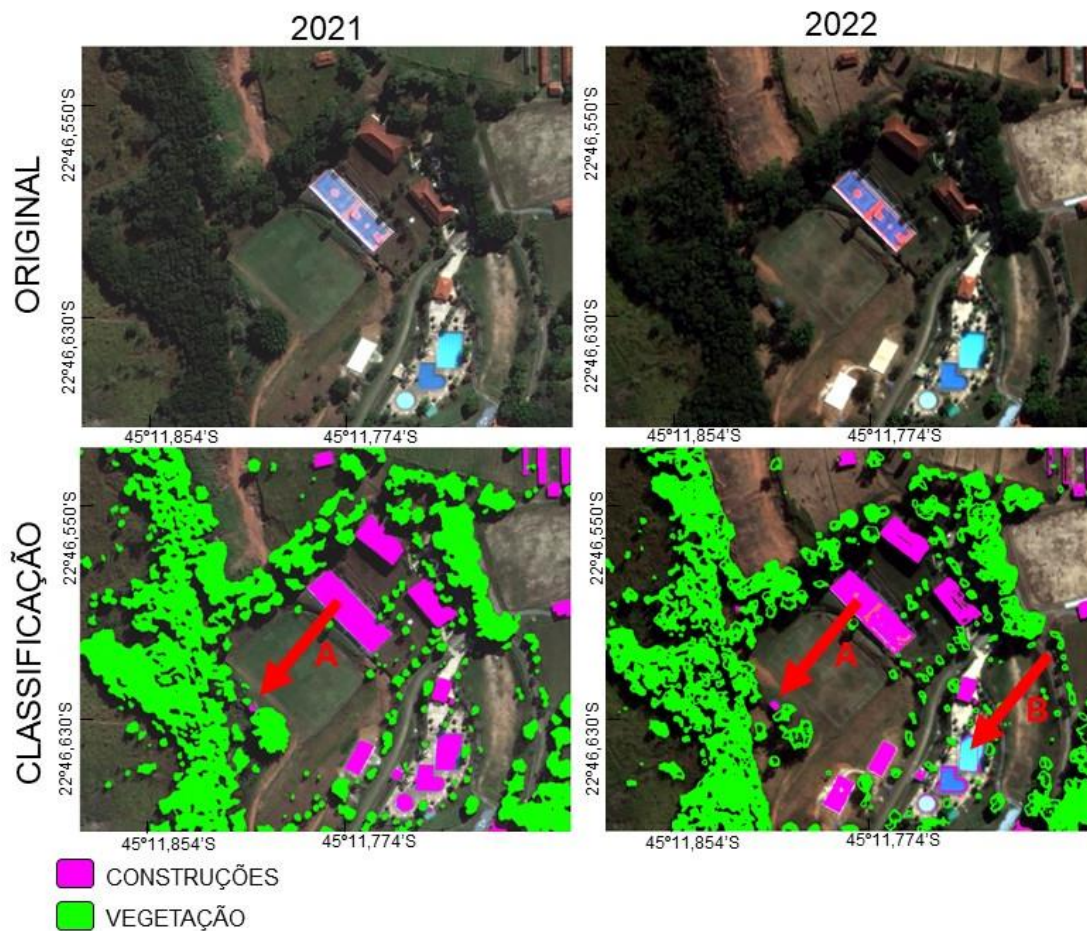
Figura 34 - Região de ampliação B: resultado de classificação da área 1 nas duas datas.



Fonte: A autora.

Na Figura 35, é evidente que diversas estruturas são classificadas corretamente como edificações, árvores, piscinas, telhados e quadras poliesportivas para a data de 2021 independentemente do tamanho dessa feição, como aponta a seta A. No entanto, em 2022, as piscinas não são classificadas corretamente como *construções* e se confundem com a classe *outros* (seta B), indicando mais uma vez o apresentado pelas métricas, isso é, uma classificação melhor para a data de 2021. Além disso, é possível ver, assim como na área de ampliação A, que a copa das árvores em 2022 está mais fragmentada, dessa forma a *IoU* (Jaccard, 1894), que já é sensível ao deslocamento de pixels, diminui mais para a data de 2022, justificando os valores de 67,80% em 2021 para 61,26% em 2022.

Figura 35 - Região de ampliação C: resultado de classificação da área 1 nas duas datas.

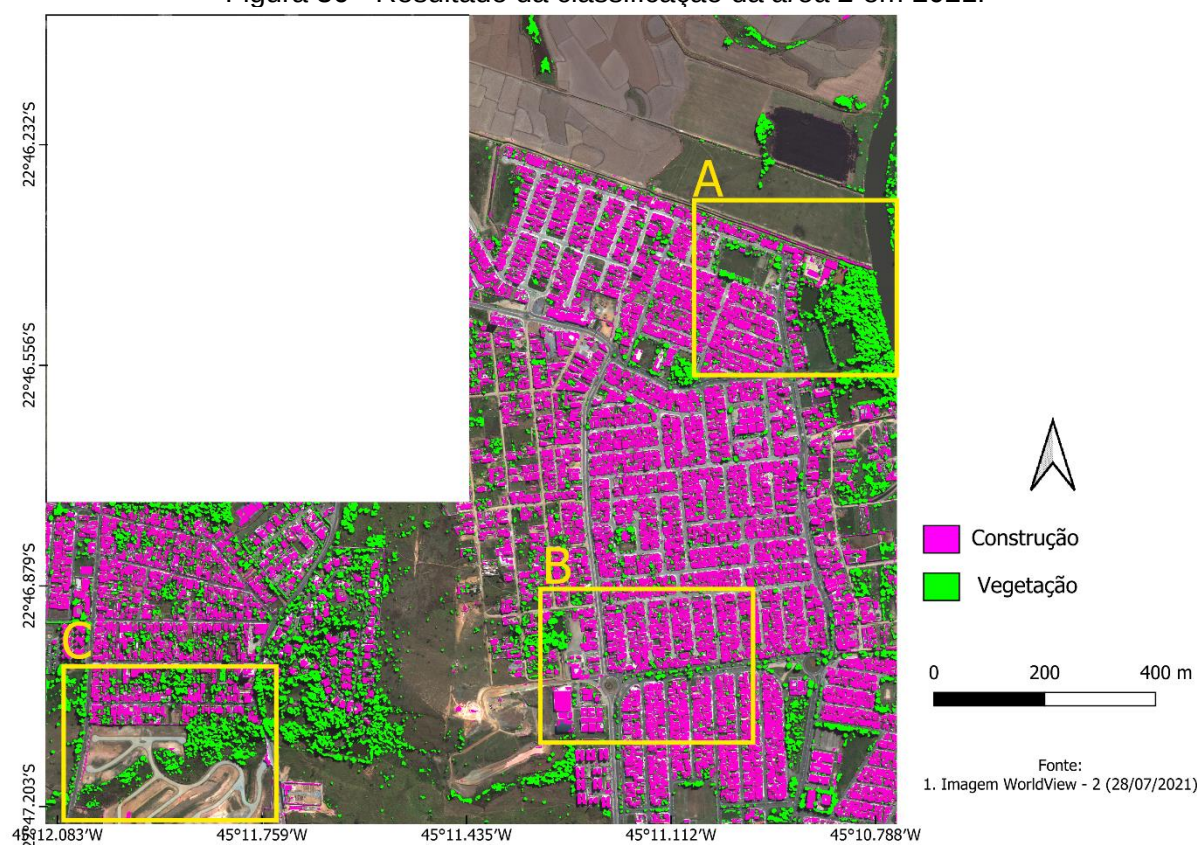


Fonte: A autora.

Para verificar a eficácia da parametrização da rede de forma qualitativa, a *área 2*, que não é utilizada no treinamento da U-Net, é classificada usando a rede parametrizada na *área 1*. Ou seja, a imagem da *área 2*, em 2021, é classificada com base na U-Net parametrizada com dados de 2021 e o mesmo procedimento é realizado para 2022.

A Figura 36 e 37 exibem os resultados da classificação da *área 2* em 2021 e 2022, respectivamente. Nas duas imagens, as *construções* são destacadas em rosa, enquanto a *vegetação arbórea* é representada em verde. A classe *outros* é apresentada em transparente para verificar se há sobreposição entre o resultado dessas duas classes e a imagem original. Em uma análise ampla e visual, ambas as imagens demonstram uma classificação satisfatória. Da mesma forma, como na *área 1*, são ampliadas três regiões da *área 2* (A, B e C) para uma visualização mais detalhada.

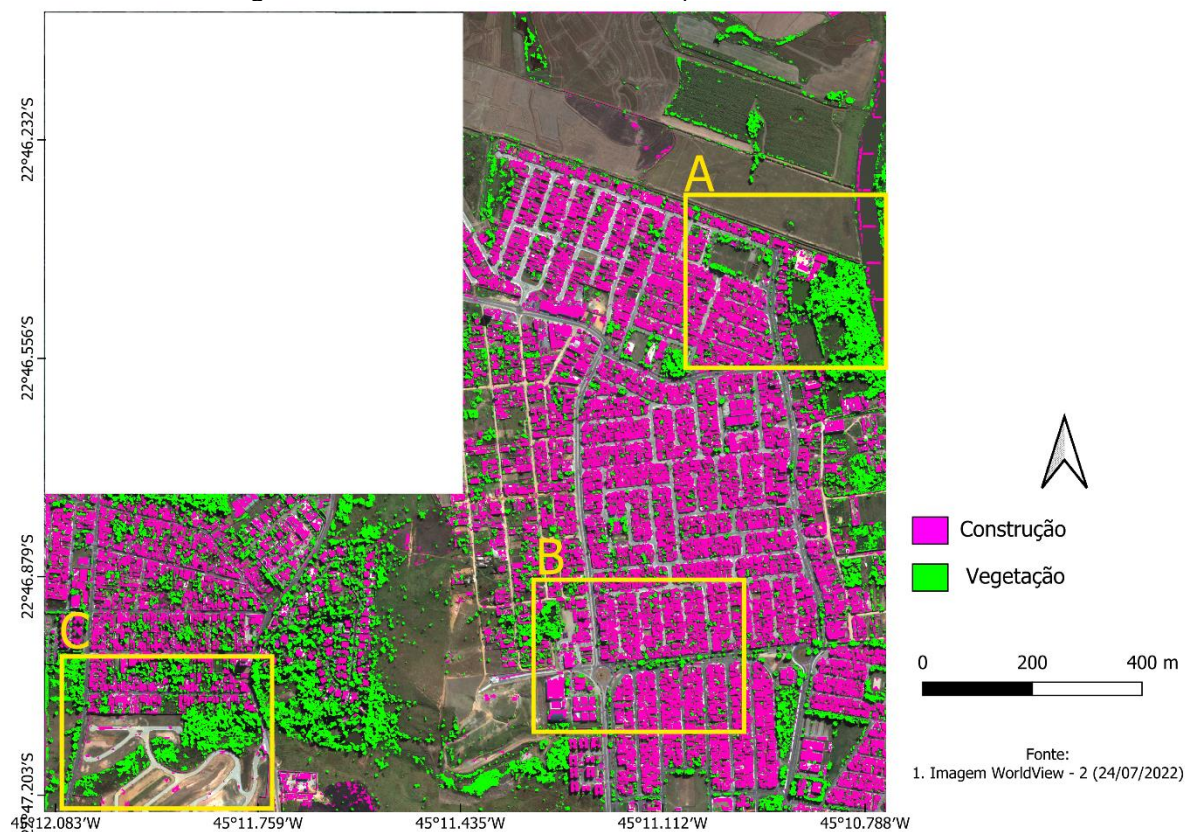
Figura 36 - Resultado da classificação da área 2 em 2021.



Fonte: A autora.

Na Figura 37 abaixo, pode-se ver o resultado da classificação da área 2, em 2022.

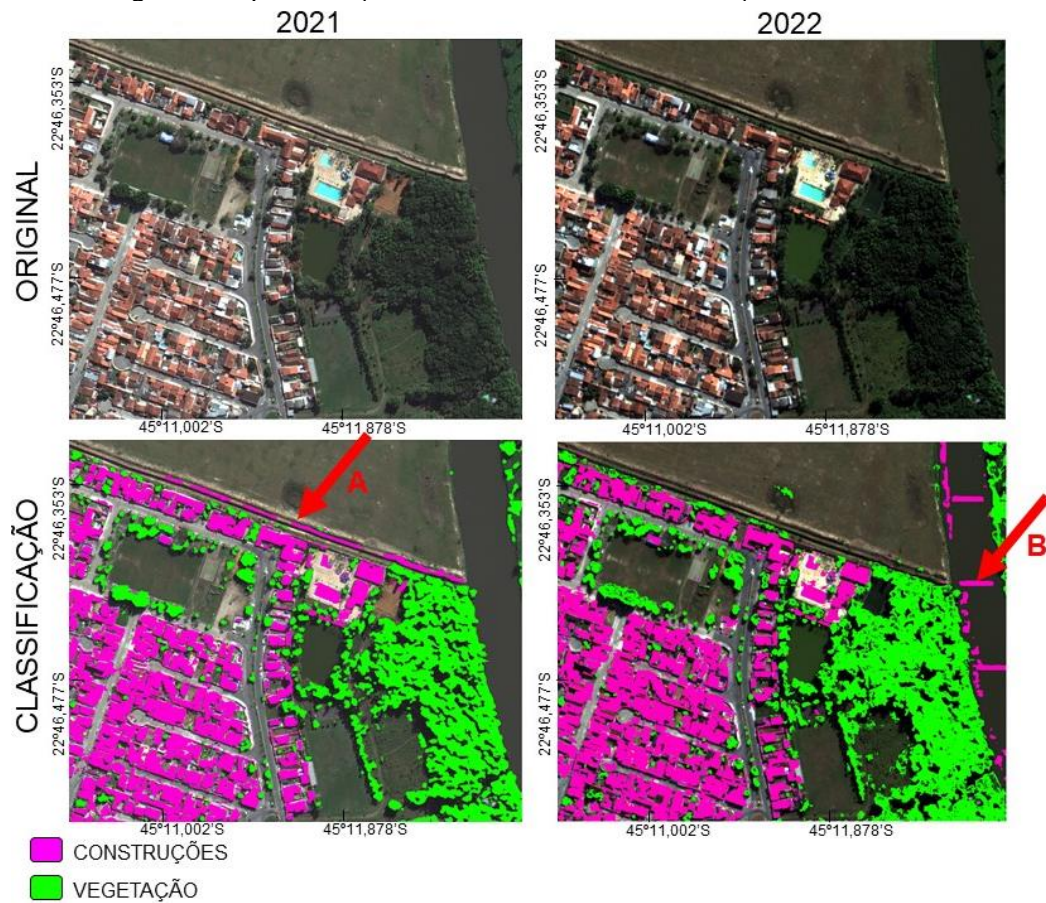
Figura 37 - Resultado da classificação da área 2 em 2022.



Fonte: A autora.

Na região aproximada A, da área 2 (Figura 38), observa-se uma classificação satisfatória, em termos qualitativos, das classes *construções* e *vegetação arbórea*. No entanto, em 2021, é perceptível que segmentos de uma rua são erroneamente classificados como *construções* (indicado pela seta A), enquanto em 2022, um trecho de corpo d'água recebe essa classificação equivocada (indicado pela seta B). Apesar da área 2 não ter sido analisada quantitativamente por não possuir dados de referência, esses resultados corroboram as conclusões da matriz de confusão da área 1, indicando uma tendência de confusão entre a classe *outros* e *construções*.

Figura 38 - Região de aproximação A: resultado de classificação da área 2 nas duas datas.



Fonte: A autora.

Na região *B*, conforme ilustrado na Figura 39, é possível observar que a maioria das *construções* e áreas de *vegetação arbórea* são identificadas corretamente. No entanto, algumas *construções* não são classificadas integralmente; isto é, alguns pixels da classe *construções* são erroneamente agrupados com a classe *outros*, como indicado pelas setas A e B nos anos 2021 e 2022.

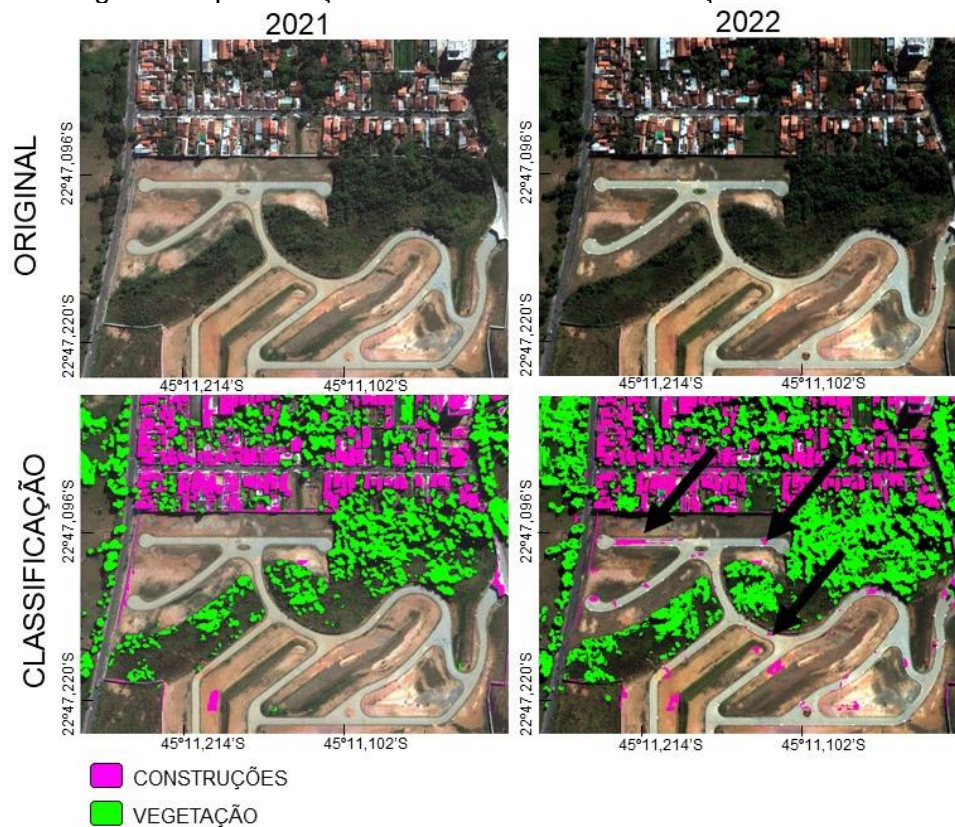
Figura 39 - Região de aproximação *B*: resultado de classificação da área 2 nas duas datas.



Fonte: A autora.

A Figura 40 mostra a região de aproximação *C* da área 2, onde principalmente, na data de 2022, há uma confusão entre a classe *construções* com a classe *outros*, como algumas setas pretas indicam.

Figura 40 - Região de aproximação C: resultado de classificação da área 2 nas duas datas.



Fonte: A autora.

A classificação da *área 2* é realizada com base na parametrização da rede utilizada para a *área 1* e na análise visual das imagens constata-se que a qualidade da classificação não é afetada pela utilização dos dados de treinamento. Isso sugere que os hiperparâmetros admitidos são eficazes para o propósito estabelecido.

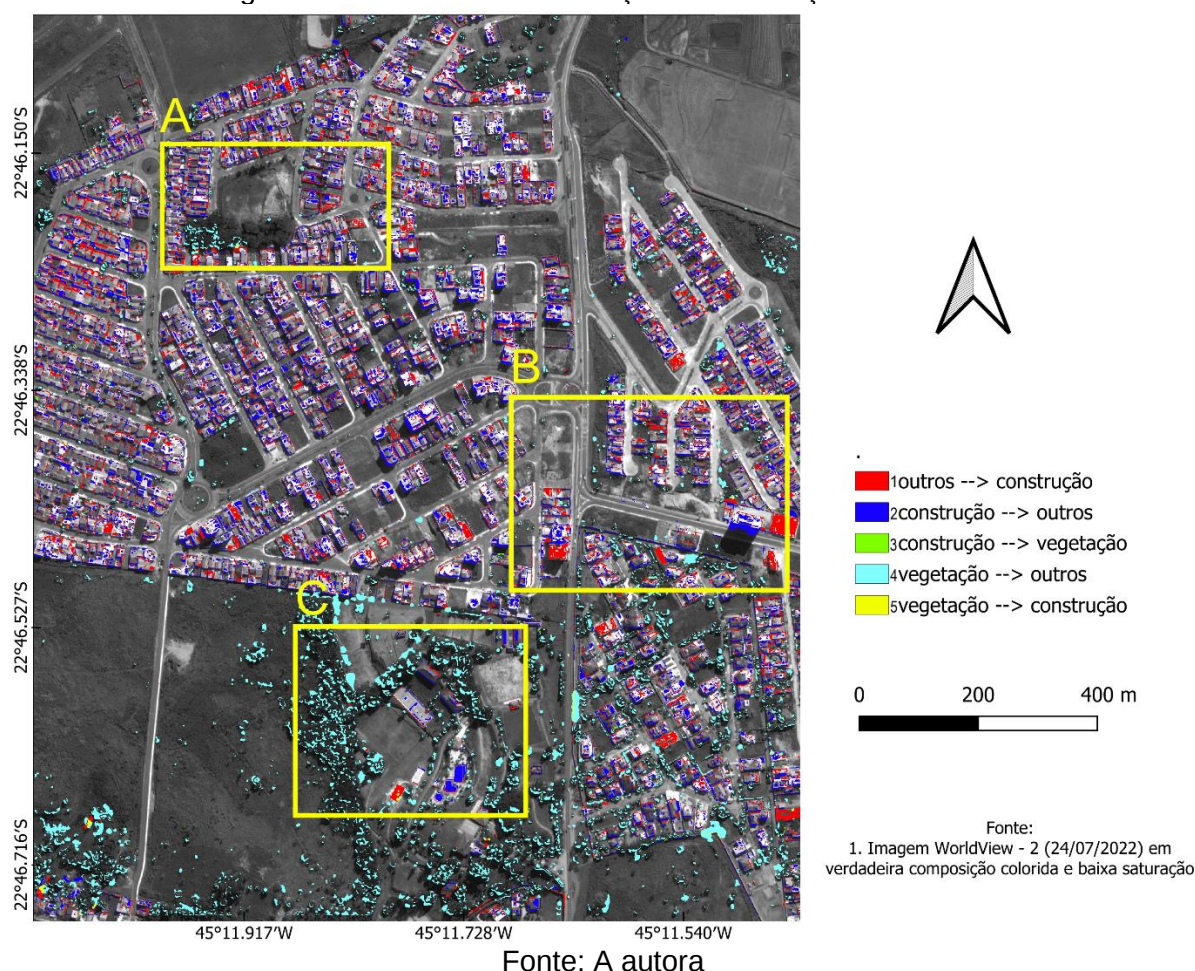
A análise visual também revela uma maior confusão entre as classes *construções* e *outros*, assim como entre *vegetação arbórea* e *outros*, como indicado pelas matrizes de confusão normalizadas da *área 1*. Essas confusões podem ser atribuídas à heterogeneidade espectral entre a classe *construções* e a classe *outros*, levando a uma semelhança em alguns pixels, como exemplificado na Figura 29. Além disso, a dificuldade em delinear com precisão a copa das árvores, devido à presença de muitos pixels de sombra e vegetação rasteira, contribui para a confusão na classificação da *vegetação arbórea*, como evidencia a Figura 30.

Após a classificação das duas áreas, nas duas datas, o procedimento seguinte é a detecção de mudanças por meio de um conjunto de regras, como está definido na Figura 21 do Capítulo de Material e Métodos.

Assim como o processo de classificação, a detecção de mudanças é feita para

a *área 1*, utilizada no treinamento da rede e para *área 2*, não utilizada no treinamento da U-Net. A Figura 41 mostra o resultado da detecção de mudanças em toda a *área 1*, tendo como base a imagem de 2022 em composição colorida verdadeira, mas com baixa saturação para realçar a camada da detecção de mudanças. Nessa figura, estão demarcadas, em amarelo, 3 regiões (A, B e C) que estão ampliadas nas Figuras 42, 43 e 44, para melhor análise de algumas das mudanças identificadas.

Figura 41 - Resultado da detecção de mudanças na *área 1*.

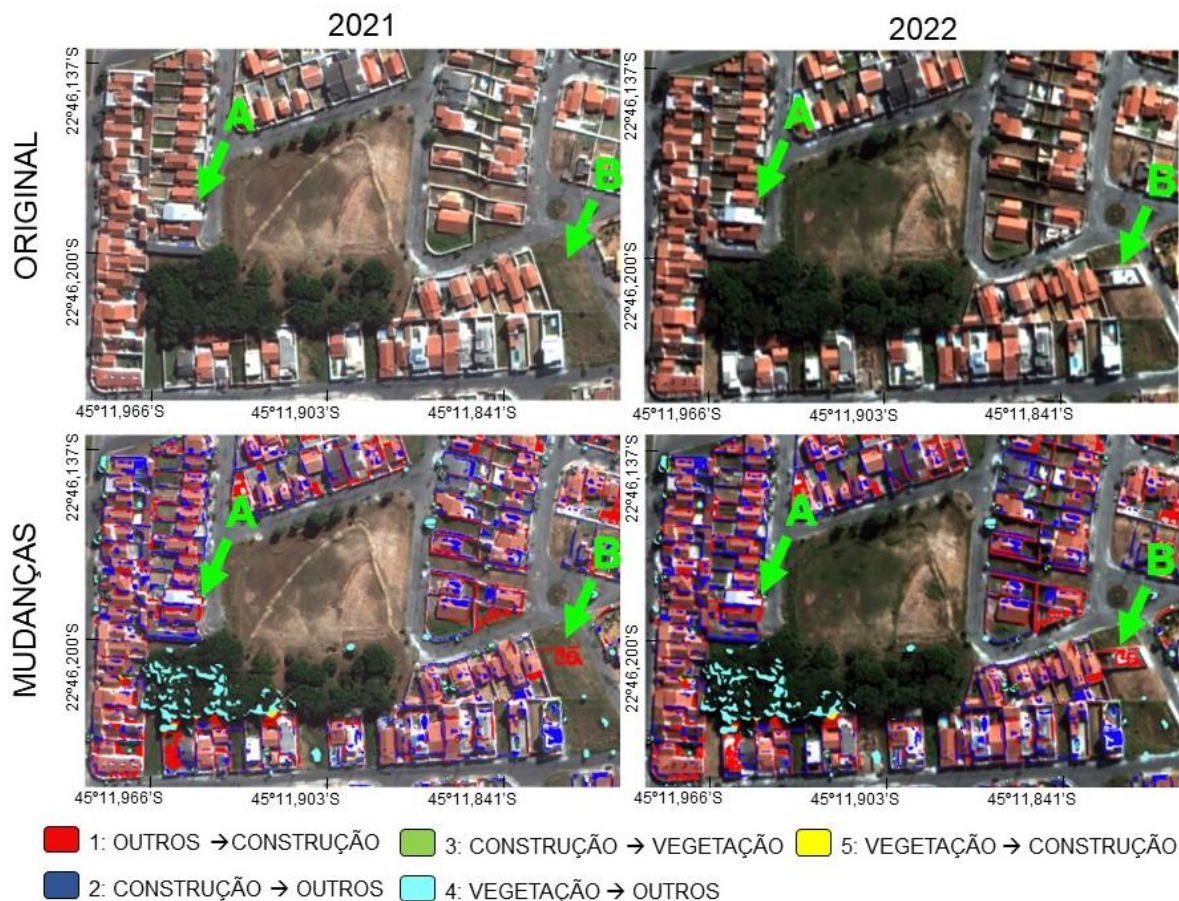


Fonte: A autora

A região A, ampliação de parte da *área 1*, é apresentada na Figura 42, com o resultado da detecção de mudanças, tendo como base as imagens de 2021 e 2022 em composição colorida verdadeira. Nessa figura pode-se observar que há a detecção de novas *construções* em vermelho, como mostra a seta B. No entanto, também é visível, que outras áreas não pertinentes a nenhuma das classes consideradas, também são detectadas como mudanças. Na seta A esse tipo de ocorrência é visível, onde nas imagens originais as *construções* não tiveram nenhuma modificação, mas mudanças em vermelho (1: outros para construção) e azul (2: construção para outros)

são detectadas, como mostra as duas imagens abaixo na Figura 42.

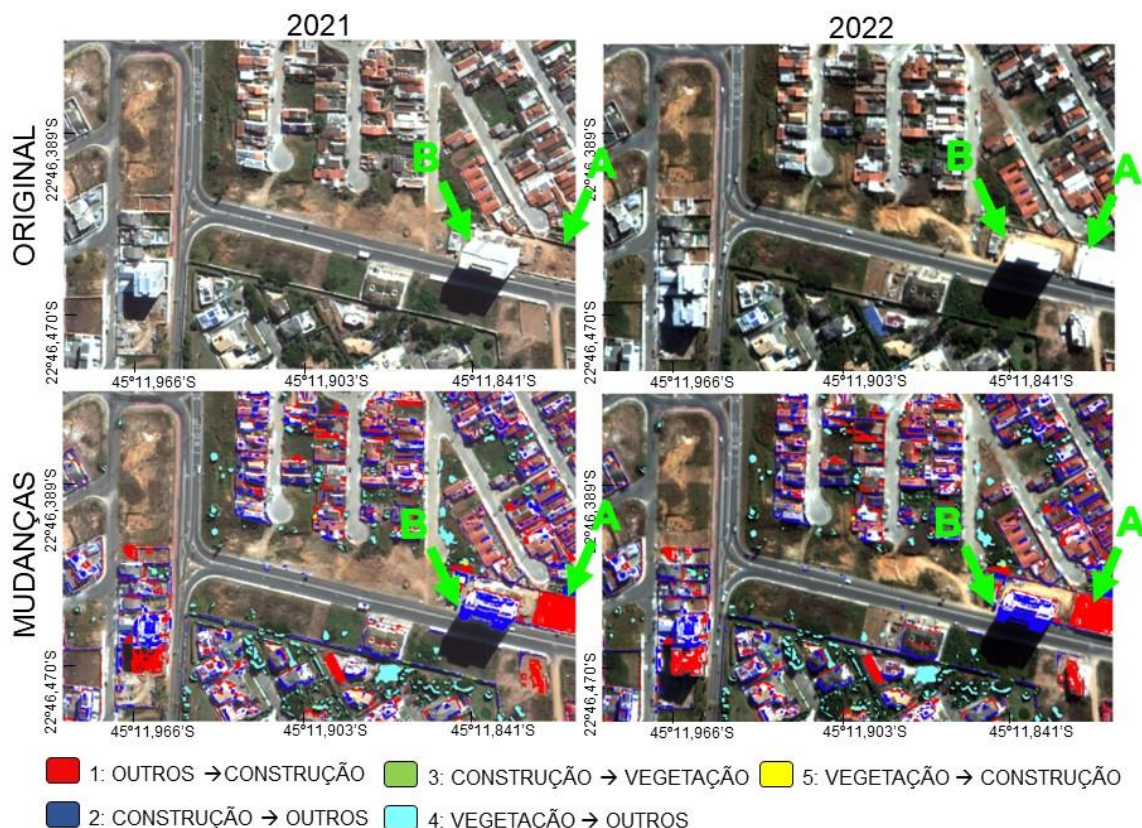
Figura 42 - Região de ampliação A: resultado da detecção de mudanças da área 1 nas duas datas.



Fonte: A autora.

A Figura 43 destaca o principal problema que surge como resultado da aplicação desta metodologia. Embora uma nova construção seja identificada em vermelho (seta A), observa-se que o edifício ao lado, que permanece inalterado ao longo do período estudado, é erroneamente detectado como mudança do tipo 2 (azul), como mostrado pela seta B. Esta ocorrência sugere que, de 2021 a 2022, a classe *construções* parece ser alterada para classe *outros*, como se houvesse ocorrido uma supressão da construção, embora este não seja o caso. Essa detecção incorreta pode ser atribuída às diferenças nos ângulos de aquisição das imagens, e esse problema se estende por toda a imagem, como mostra a Figura 43, onde áreas não relevantes também são identificadas como mudanças.

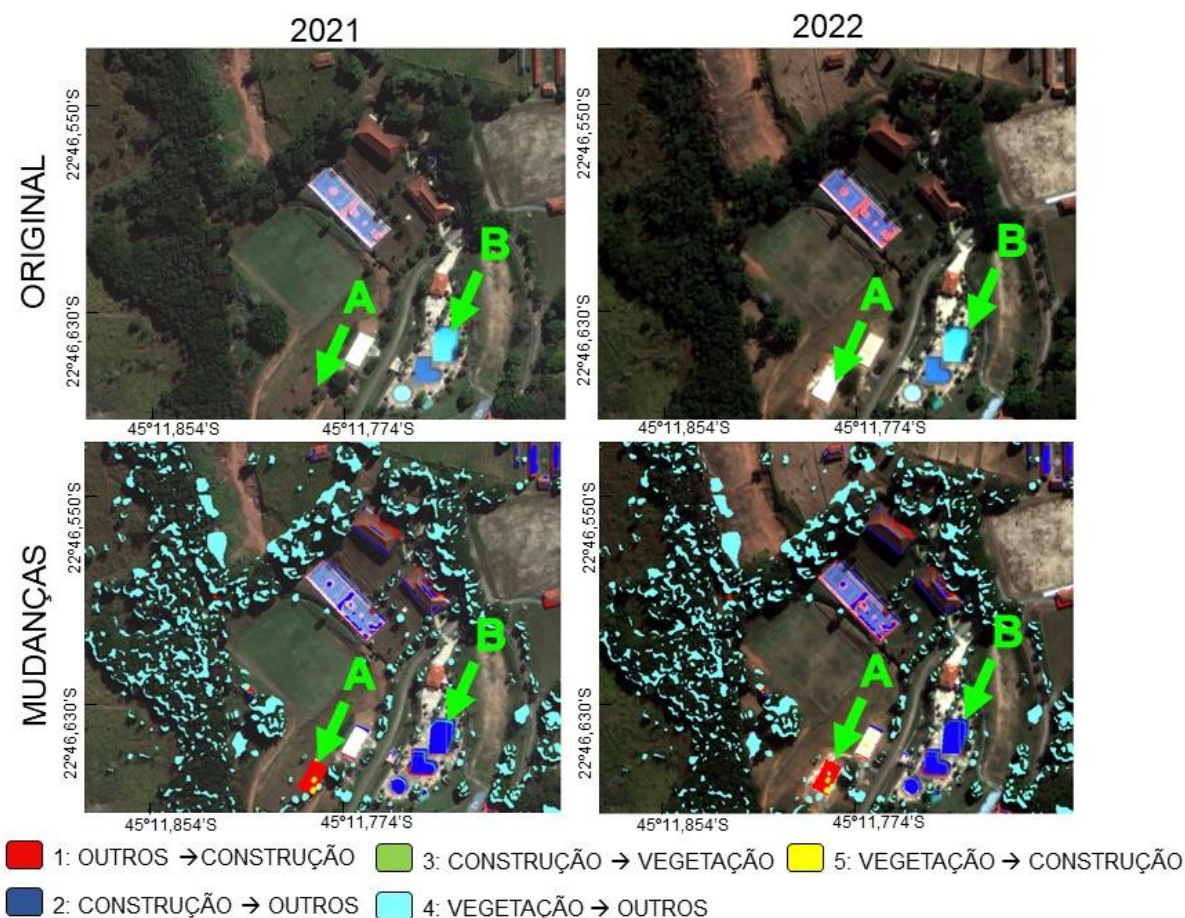
Figura 43 - Região de aproximação B: resultado da detecção de mudanças da área 1 nas duas datas.



Fonte: A autora.

Na extensão total da imagem mostrada na Figura 41, não são observadas mudanças do tipo 5 (amarelo), que compreendem locais anteriormente classificados como *vegetação arbórea* e que são modificados para *construções*. No entanto, ao analisar a ampliação C, na Figura 44, é possível destacar esse tipo específico de mudança, conforme indicado pela seta A. Além disso, é notável que o erro na classificação do objeto piscina resulta em uma detecção de mudança incorreta. Inicialmente, esse objeto é classificado de forma correta como *construções* em 2021, mas em 2022 é erroneamente categorizado como *outros*, o que ocasiona uma identificação de mudança equivocada (indicado pela seta B).

Figura 44 - Região de aproximação C: resultado da detecção de mudanças da área 1 nas duas datas.



Fonte: A autora.

Como mencionado anteriormente, assim como a classificação, a detecção de mudanças também é realizada para área 2. A Figura 45 apresenta uma visão geral de toda a detecção de mudanças nessa área, com a imagem base de 2022 em composição colorida verdadeira e baixa saturação para melhor visualização da camada de mudanças. Para uma análise mais detalhada das modificações identificadas, três regiões (A, B e C) são selecionadas e ampliadas, como ilustrado nas Figuras 46, 47 e 48.

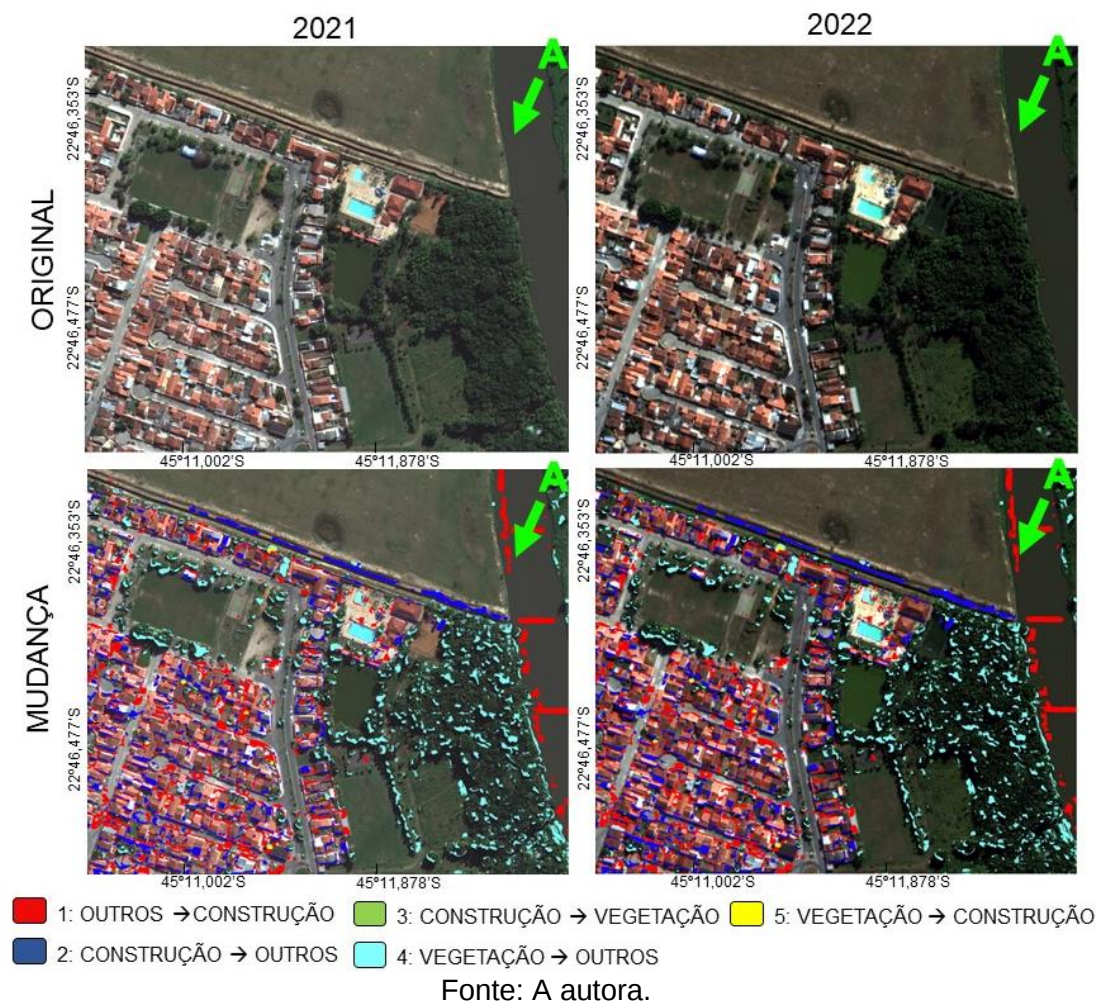
Figura 45 - Resultado da detecção de mudanças na área 2.



Fonte: A autora.

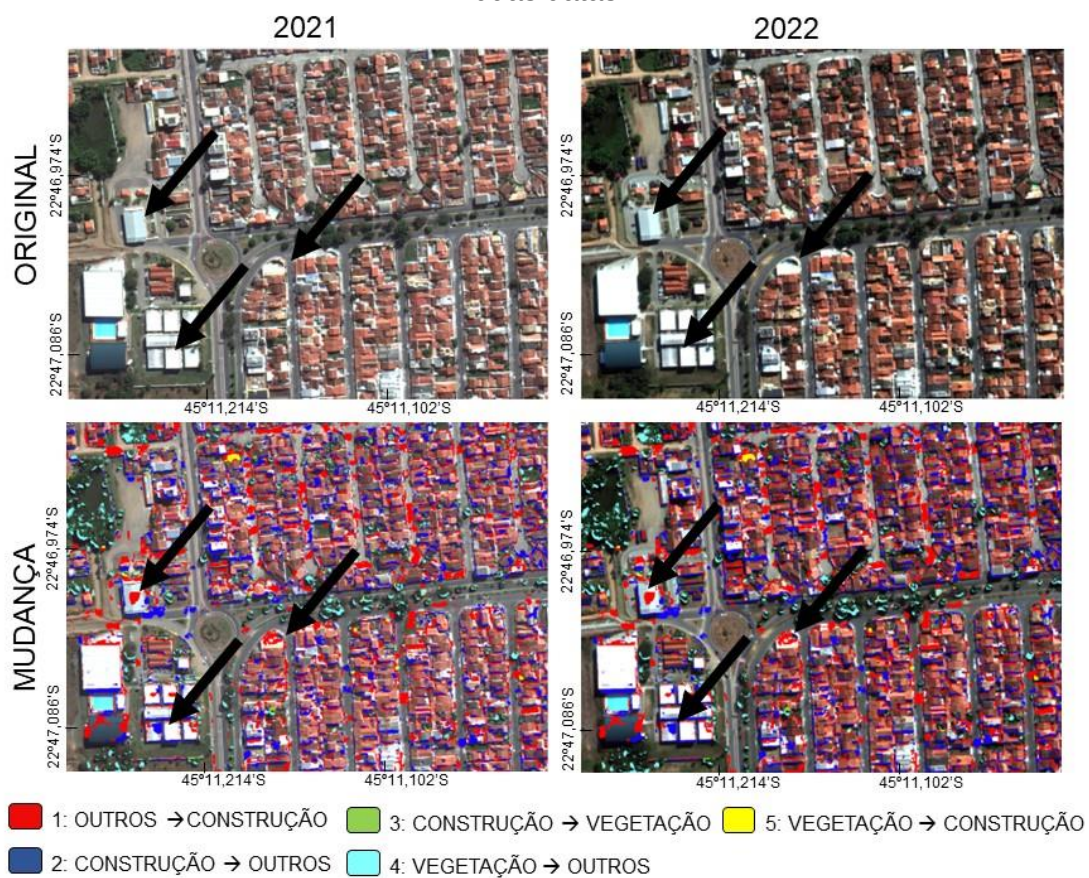
Na Figura 46, na região de aproximação A, observam-se diversas detecções decorrentes da discrepância angular entre as duas imagens e também a identificação errônea de trechos do corpo d'água como mudanças do tipo 1 (vermelho), conforme indicado pela seta A. Ou seja, o equívoco na classificação, evidenciado na Figura 38, tem um impacto direto no resultado da detecção de mudanças.

Figura 46 - Região de aproximação A: resultado da detecção de mudanças da área 2 nas duas datas.



Na Figura 47, na área de aproximação B, podem ser observados alguns pontos indicados pelas setas pretas, nos quais a classificação incorreta influencia diretamente o resultado da detecção de mudanças.

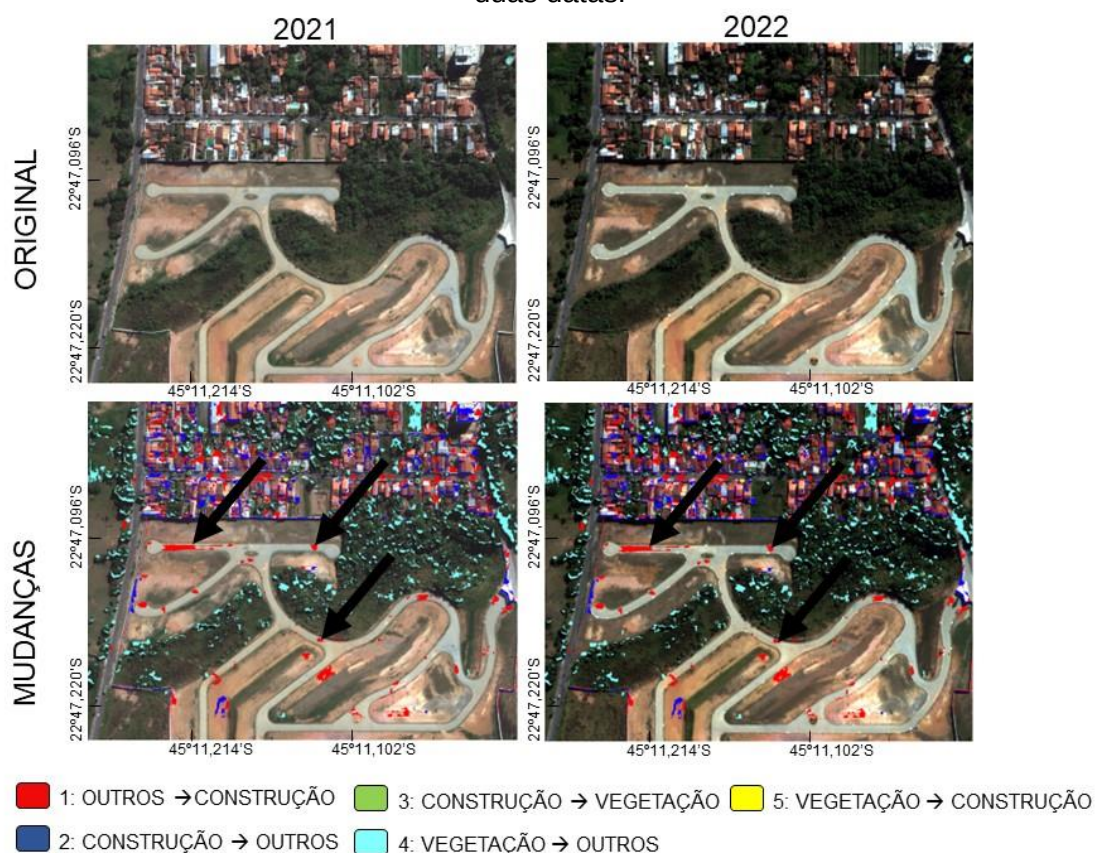
Figura 47 - Região de aproximação B: resultado da detecção de mudanças da área 2 nas duas datas.



Fonte: A autora.

Na Figura 48, na região de aproximação C da área 2, é possível observar a identificação de locais influenciados pela classificação errônea, conforme indicado pelas setas em preto.

Figura 48 - Região de aproximação C: resultado da detecção de mudanças da área 2 nas duas datas.



Fonte: A autora.

Para melhor entendimento das mudanças no uso e cobertura do solo na região de estudo, são elaboradas duas matrizes de transição, uma para a área 1 e outra para a área 2. Na Figura 49, é exibida a matriz de transição da área 1. Destaca-se que a classe com a maior proporção de área mantida é a classe *outros*, com 92,94% de persistência, equivalente a aproximadamente 70,26% do total da imagem. Em seguida, a classe *construções*, apresenta uma persistência de aproximadamente 78,28%, cerca de 12,91% da área 1; enquanto a classe *vegetação arbórea* tem uma persistência de 59,17%, correspondente a 4,68% da extensão da área.

Figura 49 - Matriz de transição da área 1.

Área de Estudo 1		24 de julho de 2022			
		Outros m ² %	Construção m ² %	Vegetação m ² %	Total m ² %
28 de julho de 2021	Outros m ² %	1055888 _{92,94}	47928 _{19,75 4,22}	32323 _{31,30 2,84}	1136140 ₁₀₀
	Construção m ² %	53159 _{4,60 21,46}	193941 _{78,28}	647 _{0,63 0,26}	247748 ₁₀₀
	Vegetação m ² %	47727 _{4,12 40,17}	782 _{0,32 0,66}	70305 _{59,17}	118816 ₁₀₀
	Total m ² %	1156776 ₁₀₀	242652 ₁₀₀	103276 ₁₀₀	1502704

% se referem aos valores de perda (linha)
% se refere aos valores de ganho (coluna)
% persistência da classe

Fonte: A autora.

Ao examinar a matriz representada na Figura 49, observa-se que as classes que apresentam as transições mais significativas são aquelas da *vegetação arbórea* para a categoria *outros*. Verifica-se uma diminuição de 40,17% na área vegetativa, que passa para a classe *outros*, e um aumento de 31,30% na *vegetação*, proveniente da categoria *outros*. Adicionalmente, a matriz indica uma redução de 21,46% na área classificada como *construções*, que migra para a categoria *outros*, enquanto há um acréscimo na área construída de 19,75%, originária da categoria *outros*.

Além disso, nota-se uma baixa incidência de transição entre as classes de *vegetação arbórea* e *construções*, corroborando com as observações da análise visual, onde esse tipo de transição representa menos de 1% da área total examinada. Ou seja, ao observar a Figura 41 percebe-se que a maioria das detecções são de *outros* para *construções* (vermelho), *construções* para *outros* (azul) e *vegetação arbórea* para *outros* (azul claro). Além disso, boa parte das mudanças detectadas são resultantes da diferença de ângulo de aquisição entre as datas, ademais não são identificadas grandes áreas modificadas.

Com relação à área 2, a classe *outros*, *construções* e *vegetação arbórea* tiveram uma persistência de 93,42%, 69,95% e 48,78%, respectivamente, sendo que cada classe corresponde a uma proporção de 72,15%, 10,66% e 3,66% da área 2.

Com relação a transição entre as classes, o mesmo comportamento da área 1 (Figura 49) é verificado na área 2 (Figura 50), isso é, as classes que mais transicionaram foram *vegetação arbórea* com *outros* e *construções* com *outros*. Nota-se que a perda da *vegetação arbórea* para a classe *outros* é de cerca de 50,14%, enquanto o ganho da *vegetação arbórea* sobre a classe *outros* é de aproximadamente

de 33,15%. Enquanto a classe *construções* tem uma perda de 29,66% sobre a classe *outros* e um ganho de 23,13% sobre a mesma classe. Também se constata que as classes *vegetação arbórea* e *construções* pouco variaram entre elas, cerca de 1%, não havendo de forma considerável então a remoção de *vegetação arbórea* para novas *construções* e vice-versa.

Figura 50 - Matriz de transição da área 2.

Área de Estudo 2		24 de julho de 2022			
		Outros m ² %	Construção m ² %	Vegetação m ² %	Total m ² %
28 de julho de 2021	Outros m ² %	3773381 _{93,42}	169162 _{23,13} _{4,19}	96572 _{33,15} _{2,39}	4039116 ₁₀₀
	Construção m ² %	236513 _{5,62} _{29,66}	557880 _{69,95}	3112 _{1,07} _{0,39}	797506 ₁₀₀
	Vegetação m ² %	197051 _{4,68} _{50,14}	4262 _{0,58} _{1,08}	191655 _{48,78}	392969 ₁₀₀
	Total m ² %	4206947 ₁₀₀	731305 ₁₀₀	291339 ₁₀₀	5229591

% se referem aos valores de perda (linha)
% se refere aos valores de ganho (coluna)
% persistência da classe

Fonte: A autora.

Portanto, a análise da matriz de transição nas duas áreas indica que há várias mudanças entre a classe *outros* e a *vegetação arbórea* e *construções*. Contudo, a observação visual constata que a maioria das mudanças observadas ocorrem devido a diferença de ângulo de aquisição entre as duas imagens. Dessa forma, não se pode inferir que essas transições resultaram de desmatamento, intensa urbanização ou demolições, por exemplo. Ademais, os pontos interpretados e disponibilizados pela GeoPixel corroboram a inexistência desses eventos.

A análise de qualidade da detecção de mudanças admite como verdade o arquivo de pontos interpretados e disponibilizados pela GeoPixel, que identificam mudanças nas áreas de *construções* e *vegetação arbórea*. Além disso, polígonos de mudanças resultantes da própria metodologia são obtidos por meio da vetorização automática das imagens e a visualização desses polígonos correspondem as Figuras 41 e 45. A vetorização produz polígonos com as seguintes classes: 1 para indicar apenas o aumento da classe *construções*, 2 para representar apenas a diminuição da classe *construções*, 3 para denotar o aumento da *vegetação arbórea* e diminuição da *construções*, 4 para mostrar a diminuição da classe de *vegetação arbórea* e 5 a diminuição da classe *vegetação arbórea* e o aumento da classe *construções*.

Em seguida, por meio da interpretação visual os indicadores VP, FP e FN das classes *construções* e *vegetação arbórea* são contabilizados. O resultado dessas três métricas para as duas classes, nas duas áreas consideradas são apresentados na Tabela 7.

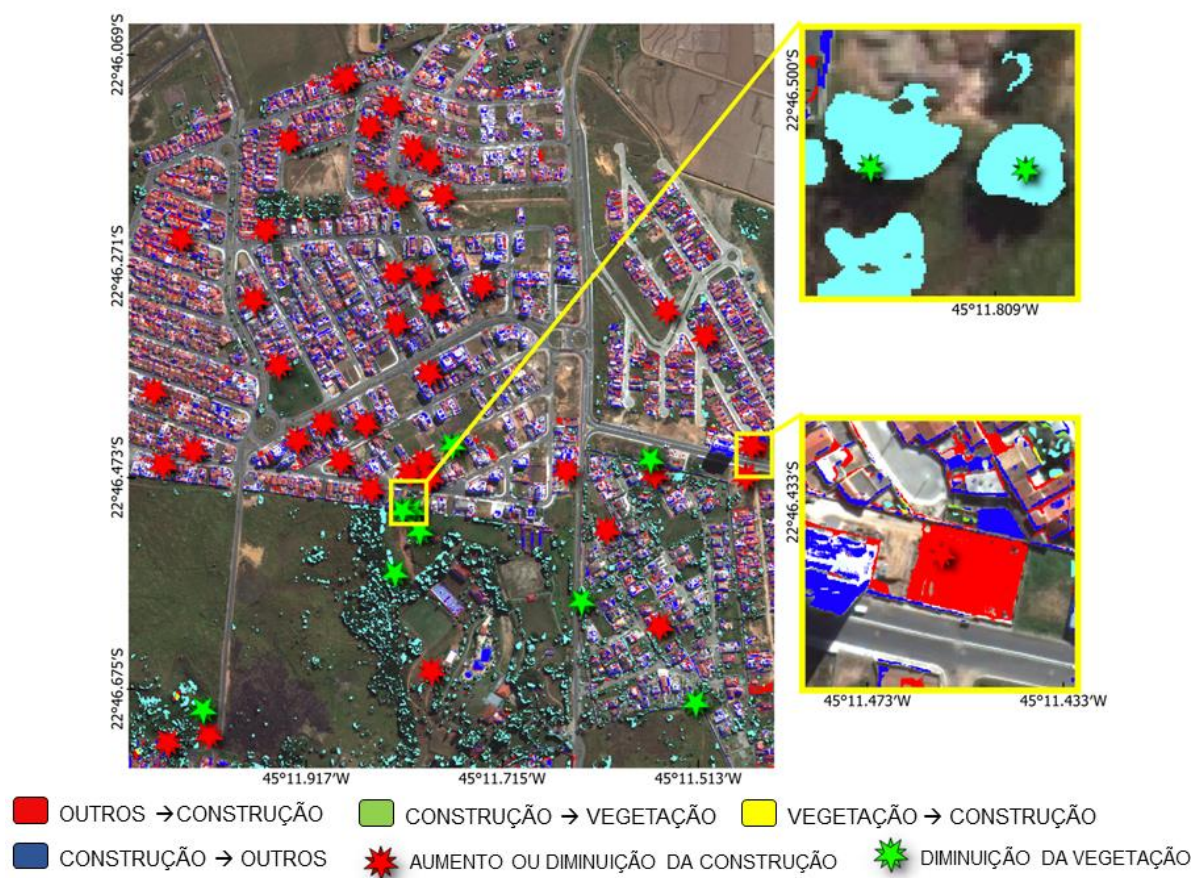
Tabela 7 - VP, FP e FN da detecção de mudanças.

		VP	FP	FN
Construções	Area 1	38	42748	4
	Area 2	19	51311	0
Vegetação Arbórea	Area 1	9	6653	0
	Area 2	8	19045	0

Fonte: A autora.

A Tabela 7 mostra que há 38 polígonos na *área 1* nos quais ocorrem variações na classe *construções*, tanto aumento quanto diminuição, ou seja, que possuem pontos de mudança dentro destes polígonos. Da mesma forma, são identificados 19 polígonos na *área 2* com pontos de mudança na classe *construções* dentro deles. Na Figura 51, é possível visualizar a distribuição desses pontos ao longo da *área 1*, onde os pontos de referência, disponibilizados pela Geopixel, são divididos em diminuição da *vegetação arbórea* representado em verde e aumento ou diminuição da *construções* representado em vermelho, além disso é possível ver também os polígonos de mudanças.

Figura 51 - Mudanças detectadas pela metodologia na área 1 e os pontos de mudanças de referência.



Fonte: A autora.

A Figura 52 mostra a distribuição dos pontos de referência ao longo do resultado da área 2.

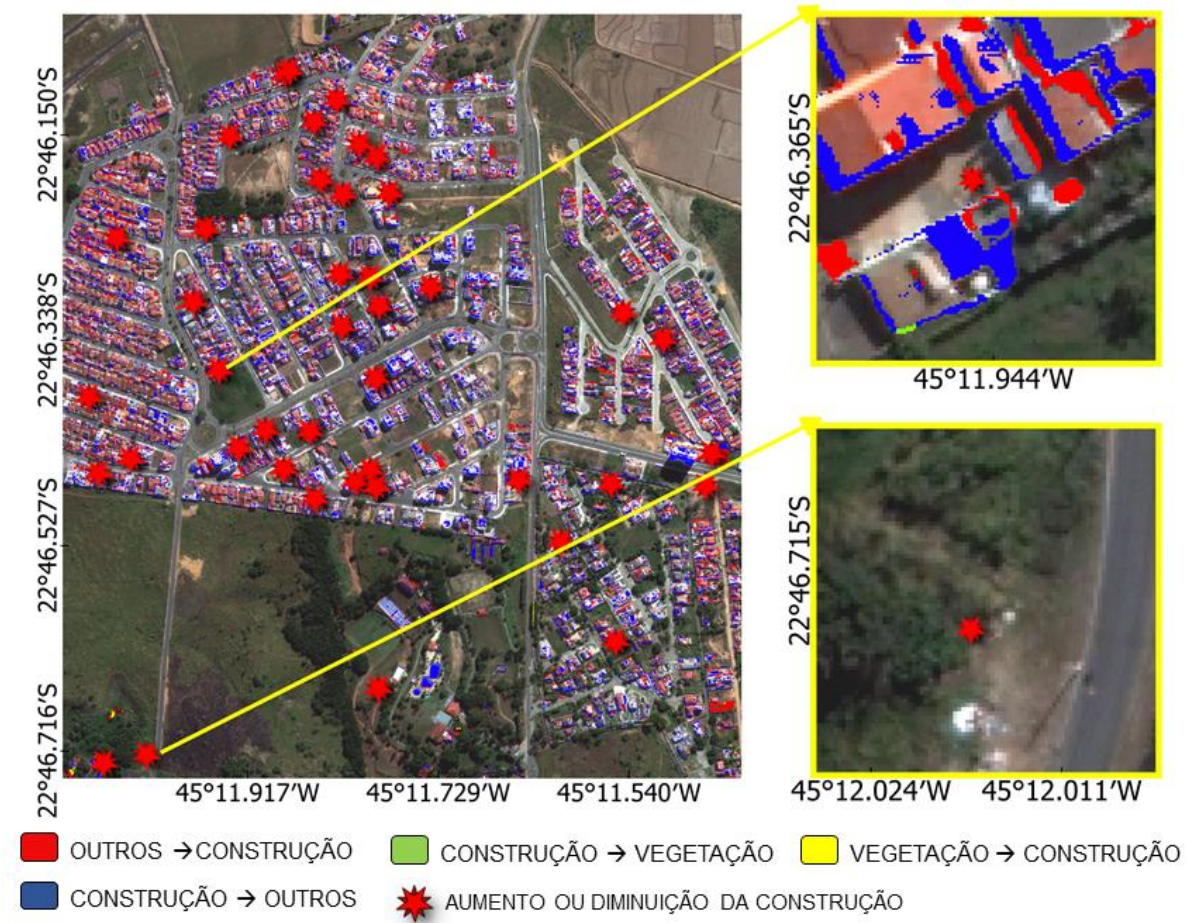
Figura 52 - Mudanças detectadas pela metodologia na área 2 e os pontos de mudanças de referência.



Fonte: A autora.

Também são contabilizados 9 polígonos na *área 1* e 8 polígonos na *área 2*, nos quais ocorre a supressão da *vegetação* e que possuem pontos de mudança dessa classe em seu interior. Além disso, observa-se que somente na classe *construções* da *área 1*, há a presença de 4 pontos de mudança que não estão associados a nenhum polígono de mudança, como exemplifica a Figura 53. Adicionalmente, cabe ressaltar que diversos polígonos são detectados como mudanças sem que estas modificações fossem identificadas pela GeoPixel.

Figura 53 -Mudanças detectadas pela metodologia na *área 1* e a não correspondência com os pontos de mudanças de referência.



Fonte: A autora.

A partir dos valores de VP, FP e FN, as métricas precisão, *recall* e *f1-score* são calculadas para as duas classes, nas duas áreas (Tabela 8).

Tabela 8 - Análise de qualidade da detecção de mudanças com as métricas, precisão, recall e *f1-score*.

		Precisão (%)	Recall (%)	F1-Score (%)
Construções	Area 1	0,09	90,48	0,18
	Area 2	0,04	100	0,08
Vegetação Arbórea	Area 1	0,14	100	0,28
	Area 2	0,04	100	0,08

Fonte: A autora.

A alta taxa de *recall*, variando entre 90,48% e 100%, indica que o modelo é capaz de identificar a maioria das mudanças reais presentes nos dados. No entanto, a baixa precisão sugere que apenas uma pequena proporção das mudanças

detectadas corresponde efetivamente a mudanças reais. O desequilíbrio entre *recall* e precisão indica que o modelo tende a gerar muitos falsos positivos (FP), possivelmente devido à diferença no ângulo de aquisição das imagens. Além disso, parte desses FP também pode ser atribuída à classificação realizada pela U-Net, que não atingiu 100% das métricas avaliadas, ou até mesmo ao registro das imagens, que pode ter incluído deformações nas imagens.

O *f1-score*, que é a média harmônica entre precisão e *recall*, fornece uma medida geral do equilíbrio entre essas duas métricas. Entretanto, valores abaixo de 1% indicam um desempenho baixo na harmonização entre precisão e *recall*. Esses resultados destacam a necessidade de aprimoramentos na metodologia apresentada.

Para melhorar os resultados, várias ações são recomendadas. A utilização de dados altimétricos diminuiria o efeito causado pela diferença no ângulo de aquisição. A ampliação do conjunto de dados melhoraria a representatividade das classes, o que pode ser alcançado aumentando o número de bandas espectrais utilizadas e expandindo a área de estudo. Além disso, sugere-se aperfeiçoar a caracterização das classes, incluindo outras categorias como sombra, vegetação rasteira e asfalto para mitigar as confusões apresentadas na classificação. Essas sugestões, que podem ser aplicadas a trabalhos futuros, têm o potencial de aprimorar a metodologia apresentada.

Importante observar que a metodologia apresentada exige atenção contínua, uma vez que a U-Net pode ter a necessidade de ser retreinada, ou seja, treinada novamente sempre que houver uma queda no desempenho do modelo, alterações significativas nos dados de entrada ou o surgimento de novas variáveis que possam impactar o resultado da classificação.

5. CONSIDERAÇÕES FINAIS

A metodologia apresentada nessa dissertação aborda a detecção de mudanças em ambientes urbanos através da análise de imagens orbitais de alta resolução espacial. A identificação de mudanças, dada a complexidade dos elementos presentes nas cenas de áreas urbanas, muitas vezes, é conduzida por interpretação visual. Nesse contexto, este estudo propõe a aplicação da Rede Neural Convolucional U-Net para classificar imagens de alta resolução espacial, do satélite *WorldView – 2*, seguida por um conjunto de regras para identificar mudanças em áreas urbanas, em duas datas distintas, para duas áreas selecionadas no município de Guaratingetá – SP.

Através dos procedimentos metodológicos desenvolvidos, é possível avaliar o comportamento da rede U-Net na detecção de alterações em áreas urbanas, especialmente nas classes de *construções* e *vegetação arbórea*, indicativas de possíveis modificações como supressão ou aumento de construções, desmatamento, moradias irregulares e descarte irregular de entulho, de modo a nortear melhores políticas públicas.

A arquitetura U-Net demonstra eficácia na classificação das classes propostas, de modo que o resultado da primeira data revela que a *área 1* alcança 89,64% de acurácia e 87,25% para o índice Kappa (Cohen, 1960). Para a segunda data, esses valores são de 81,15% e 86,02%, respectivamente, indicando uma classificação excelente de acordo com o proposto por Landis e Koch (1977). Além disso, a análise visual da *área 2* mostra que a parametrização da U-Net é capaz de classificar áreas não utilizadas no treinamento com eficiência.

A detecção de mudanças, conduzida com base nas imagens classificadas e em um conjunto de regras, apresenta bons resultados para a métrica *recall*, acima de 90%, nas duas áreas, evidenciando a capacidade do modelo em identificar mudanças. Contudo, a precisão é notoriamente baixa, inferior a 1%, sugerindo uma elevada taxa de falsos positivos nas duas áreas, devido principalmente pela diferença no ângulo de aquisição das duas datas.

Dessa maneira, o estudo evidenciou a capacidade da metodologia em detectar mudanças na área urbana. No entanto, o resultado é afetado pela identificação de alterações não relevantes, ou seja, mudanças que de fato não ocorreram e são registradas principalmente devido às diferenças nos ângulos de

aquisição entre as duas datas, resultando em muitos falsos positivos. Recomenda-se, portanto, a utilização de dados altimétricos para mitigar o efeito ocasionado pela diferença no ângulo de aquisição, visando atender às necessidades dos órgãos municipais no monitoramento de ocorrências com mais eficiência. Além disso, em futuros trabalhos, também se recomenda a ampliação do conjunto de dados e a melhoria na caracterização das classes incluindo outros alvos.

REFERÊNCIAS

ABELPRE - ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE LIMPEZA PÚBLICA E RESÍDUOS ESPECIAIS (São Paulo). **Resíduos sólidos: Manual de Boas Práticas no Planejamento**. [S. l.: s. n.], 2017. Disponível em: <https://abrelpe.org.br/residuos-solidos-manual-de-boas-praticas-no-planejamento/>. Acesso em: 21 set. 2022.

AHMED, Wafaa Shihab; KARIM, Abdul amir A. The Impact of Filter Size and Number of Filters on Classification Accuracy in CNN. **International Conference on Computer Science and Software Engineering (CSASE)**, Duhok, Iraque, p. 88-93, 18 abr. 2020. DOI 10.1109/CSASE48920.2020.9142089. Disponível em: <https://ieeexplore.ieee.org/document/9142089/authors>. Acesso em: 15 out. 2022.

ALMEIDA, José Antonio Pacheco de; OLIVEIRA, Paulo José de. Sensoriamento Remoto I. 1. ed. São Cristóvão - SE: CESAD, 2010. 113 p. Disponível em: <https://cesad.ufs.br/ORBI/public/uploadCatalogo/>. Acesso em: 6 fev. 2024.

AMORIM, Amilton; PELEGRINA, Marcos Aurélio; JULIÃO, Rui Pedro. **Cadastro e Gestão Territorial**: uma visão luso-brasileira para a implementação de sistemas de informação cadastral nos municípios. São Paulo: UNESP Digital, 2018. 131 p. DOI 10.7476/9788595462823. Disponível em: <https://books.scielo.org/id/8b9n4/pdf/amorim-9788595462823.pdf>. Acesso em: 17 jul. 2022.

ANTUNES, Alzir Felipe Buffara; HOLLATZ, Roberta Cristina Vedor. Cadastro Técnico Multifinalitário de baixo custo utilizando VANT (veículo aéreo não tripulado). **XVII SBSR - Simpósio Brasileiro de Sensoriamento Remoto**, João Pessoa, 17. ed. p. 5858 - 5864, 2015.

ASZEMI, Nurshazlyn Mohd; SELVAM, Dhanapal Durai Dominic Panneer. Hyperparameter Optimization in Convolutional Neural Network using Genetic Algorithms. **International Journal of Advanced Computer Science and Applications**, [S. l.], v. 10, n. 6, p. 269 - 278, 2019. DOI 10.14569/IJACSA.2019.0100638. Disponível em: https://www.researchgate.net/publication/334151021_Hyperparameter_Optimization_in_Convolutional_Neural_Network_using_Genetic_Algorithms. Acesso em: 4 out. 2022.

BAI, Ting; WANG, Le; YIN, DaMeng; SUN, KaiMin; CHEN, YePei; LI, WenZhuo; LI, DeRen. Deep learning for change detection in remote sensing: a review. **Geo-spatial Information Science**, [S. l.], p. 262-288, 2021. DOI 10.1080/10095020.2022.2085633. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/10095020.2022.2085633>. Acesso em: 21 jul. 2022.

BATTY, Michael; CHESHIRE, James. Cities as Flows, Cities of Flows. **Environment and Planning B Planning and Design**, [S. l.], v. 38, n. 2, p. 195 – 196, 1 abr. 2011. DOI 10.1068/b3802ed. Disponível em: <https://journals.sagepub.com/doi/10.1068/b3802ed>. Acesso em: 12 set. 2022.

BERMAN, Maxim; TRIKI, Amal Rannen; BLASCHKO, Matthew. The Lovász-Softmax Loss: A Tractable Surrogate for the Optimization of the Intersection-Over-Union Measure in Neural Networks. **IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, [S. l.], p. 4413-4421, 18 jun. 2018. DOI 10.48550/arXiv.1705.08790. Disponível em: https://openaccess.thecvf.com/content_cvpr_2018/papers/Berman_The_LovaSz-Softmax_Loss_CVPR_2018_paper.pdf. Acesso em: 13 fev. 2024.

BEZERRA, Maria Carina Maia; CHESHIRE, Pedro David Rodrigues Lima; CAVALCANTE, Augusto César Chaves; SOUSA JUNIOR, Almir Mariano de. A utilização de VANT em levantamentos cadastrais para fins de atualização de cadastro imobiliário: Estudo de caso no município de Pau dos Ferros/RN, **Investigação Científica nas Ciências Sociais Aplicadas** Natal, 2019. XVIII ENAPUR - Encontros Nacionais da Associação Nacional de Pós-Graduação e Pesquisa em Planejamento Urbano e Regional.

BISHOP, Yvonne; FIENBERG, Stephen; HOLLAND, Paul. **Discrete Multivariate Analysis: Theory and Practice**. Reino Unido: MIT, 1975. 575 p. Disponível em: <https://esl.hohoweiya.xyz/references/Discrete-Multivariate-Analysis.pdf>. Acesso em: 6 fev. 2024.

BLACK, William; DWYER, Andrew. Local versus global measures of accuracy: an important distinction for diagnostic imaging. **Medical Decision Making**, [S. l.], p. 266-273, 6 fev. 1990. DOI 10.1177/0272989X9001000404. Disponível em: <https://journals.sagepub.com/doi/10.1177/0272989X9001000404>. Acesso em: 15 fev. 2024.

BLASCHKE, Thomas; HAY, Geoffrey J.; WENG, Qihao; RESCH, Bernd. Collective Sensing: Integrating Geospatial Technologies to Understand Urban Systems—An Overview. **Remote Sensing**, [S. l.], v. 3, n. 8, p. 1743-1776, 19 ago. 2011. DOI 10.3390/rs3081743. Disponível em: <https://www.mdpi.com/2072-4292/3/8/1743>. Acesso em: 30 set. 2022.

BOLTZMANN, Ludwig. On some problems of the mechanical theory of heat. **Philosophical Magazine and Journal of Science**. 236-237, 1878 DOI: 10.1080/14786447808639502. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/14786447808639502>. Acesso em: 14/02/2024

BUCKLAND, Michael; GEY, Fredric. The relationship between Recall and Precision. **Journal of the American Society for Information Science**, [S. l.], p. 12-19, 14 jan. 1994. DOI 10.1002/(SICI)1097-4571(199401)45:1%3C12::AID-ASI2%3E3.0.CO;2-L. Disponível em: <https://www.semanticscholar.org/paper/The-Relationship-between-Recall-and-Precision-Buckland-Gey/edb199dbab5011487011e97484d2ded4fff8dd90>. Acesso em: 6 jun. 2023.

CAMPOS, Pedro Bueno Rocha; QUEIROZ FILHO, Alfredo Pereira de. Matriz de Transição na Detecção das Mudanças do Uso e Ocupação do Solo: Estudo de caso do Centro Educacional Unificado da Paz - Zona Norte de São Paulo. **O Espaço Geográfico em Análise**, Curitiba, v. 42, p. 225 - 238, dez. 2017. DOI 10.5380/raega. Disponível em: <https://revistas.ufpr.br/raega/article/view/47694/34156>. Acesso em: 31 maio 2023.

CÂNDIDO, Anny Keli Aparecida Alves; SILVA, Normandes Matos da; PARANHOS FILHO, Antonio Conceição. Imagens de Alta Resolução Espacial de Veículos Aéreos Não Tripulados (VANT) no Planejamento do Uso e Ocupação do Solo. **Anuário do Instituto de Geociências - UFRJ**, Rio de Janeiro, p. 147 - 156, 23 mar. 2015. DOI 10.11137/2015_1_147_156. Disponível em: www.anuario.igeo.ufrj.br. Acesso em: 3 maio 2023.

CANTONI, Stefano; VASSENA, Giorgio Paolo Maria. Fast Mapping and Geospatial Data Acquisition by Innovative Indoor Mapping System. **Computer Science**, [S. l.], p. 213-217, 31 jan. 2019. DOI 10.5194/isprs-archives-XLII-2-W9-213-2019. Disponível em: <https://isprs-archives.copernicus.org/articles/XLII-2-W9/213/2019/>. Acesso em: 23 ago. 2023.

CHAI, TianFeng; DRAXLER, Roland. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. **Geoscientific Model Development**, [S. l.], p. 1247-1250, 30 jun. 2014. DOI 10.5194/gmd-7-1247-2014. Disponível em: <https://gmd.copernicus.org/articles/7/1247/2014/>. Acesso em: 9 fev. 2024.

CHEN, Bin; XU, Bing; GONG, Peng. Mapping essential urban land use categories (EULUC) using geospatial big data: Progress, challenges, and opportunities. **Big Earth Data**, [S. l.], v. 3, p. 410 - 441, 27 abr. 2021. DOI 10.1080/20964471.2021.1939243. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/20964471.2021.1939243>. Acesso em: 22 maio 2023.

CHEN, Ziyi; WANG, Cheng; LI, Jonathan; XIE, Ninci; HAN, Yan. Reconstruction Bias U-Net for Road Extraction from Optical Remote Sensing Images. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, [S. l.], v. 14, p. 2284-2294, 22 jan. 2021. DOI 10.1109/JSTARS.2021.3053603. Disponível em: <https://ieeexplore.ieee.org/document/9333652>. Acesso em: 6 jul. 2022.

CHICCO, Davide; JURMAN, Giuseppe. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. **BMC Genomics**, [S. l.], v. 21, n. 6, p. 1-13, 2 jan. 2020. DOI <https://doi.org/10.1186/s12864-019-6413-7>. Disponível em: <https://bmcbgenomics.biomedcentral.com/articles/10.1186/s12864-019-6413-7>. Acesso em: 1 ago. 2023.

COHEN, Jacob. A Coefficient of Agreement for Nominal Scales. **Educational and Psychological Measurement**, [S. l.], p. 37 - 47, 1 fev. 1960. DOI 10.1177/001316446002000104. Disponível em: <https://journals.sagepub.com/doi/10.1177/001316446002000104>. Acesso em: 5 out. 2022.

COSTA, Alisson Hayasi da; SANTOS, Renato Augusto Corrêa dos; CERRI, Ricardo. Performance Analysis of Deep Neural Networks in piRNAs Classification. **XV Encontro Nacional de Inteligência Artificial e Computacional**, [S. l.], p. 12345, 22 out. 2018. DOI <https://doi.org/10.5753/eniac.2018.4449>. Disponível em: <https://sol.sbc.org.br/index.php/eniac/article/view/4449>. Acesso em: 19 fev. 2022.

DAVE, Chintan P.; JOSHI, Rahul; SRIVASTAVA, S. S. A Survey on Geometric Correction of Satellite Imagery. **International Journal of Computer Applications**, [S. l.], p. 24 - 27, 12 abr. 2015. DOI 10.5120/20389-2655. Disponível em: <https://www.semanticscholar.org/paper/A-Survey-on-Geometric-Correction-of-Satellite-Dave-Joshi/15bbb7730f6ca41b569911bcf54a347b171b9e62>. Acesso em: 10 out. 2022.

DENG, Jin S.; WANG, Ke; HONG, Yang; QI, Jia G. Spatio-temporal dynamics and evolution of land use change and landscape pattern in response to rapid urbanization. **Landscape and Urban Planning**, [S. l.], p. 187-198, 30 set. 2009. DOI 10.1016/j.landurbplan.2009.05.001. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0169204609000814>. Acesso em: 28 nov. 2022.

DENG, Li; YU, Dong. **Deep Learning: Methods and Applications**. 3-4: Now, 2014. 190 p. v. 7. ISBN 978-1-60198-815-7. DOI 10.1561/20000000039. Disponível em: <https://nowpublishers.com/article/Details/SIG-039>. Acesso em: 5 set. 2023.

DINIZ, Cesar Guerreiro. Três décadas de mudanças na planície costeira brasileira: o status dos manguezais, da aquicultura e salicultura a partir de séries temporais Landsat e técnicas de aprendizado de máquina. Orientador: Pedro Walfir Martins e Souza Filho. 2021. 92 f. Tese (Doutorado em Geologia e Geoquímica) - Instituto de Geociências, Universidade Federal do Pará, Belém, 2021. Disponível em: <http://repositorio.ufpa.br/jspui/handle/2011/13504>. Acesso em: 20/02/2024

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. In: **Satélites de Monitoramento: WorldView - 2**. [S. l.], 27 nov. 2018. Disponível em: <https://www.embrapa.br/satelites-de-monitoramento/missoes/worldview>. Acesso em: 12 abr. 2022.

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. In: **Base de dados climáticos**. [S. l.], 26 jan. 2012. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/939638/base-de-dados-climaticos>. Acesso em: 6 fev. 2024.

FACELI, Katti et al. **Inteligência artificial: uma abordagem de aprendizado de máquina**. 2011. Rio de Janeiro: LTC. Acesso em: 28 nov. 2023.

FERREIRA, Calebe; AMARAL, Ranielly. Redes Neurais Multicamadas Para Classificação e Regressão. [S. l.], 17 set. 2019. Disponível em: <https://rpubs.com/CalebeF/534504>. Acesso em: 7 jun. 2022.

FIROUZABADI, P. Zeaiean; GHANAVATI, E. Digital Approaches for Change Detection in Urban Environment. **Geographical Research**, [S. l.], v. 22, n. 1, p. 1423 - 1437, 21 mar. 2007. DOI 10.1007/s43615-021-00035-y. Disponível em: <https://www.sid.ir/filesserver/je/94620078401>. Acesso em: 12 out. 2022.

FRANÇOSO, Maria Teresa; LIMA, Luis Antônio; TREVISAN, Fernanda Lodi. Simplification of the cadastre updating for small municipalities. **International Conference Virtual City and Territory. "9º Congresso Città e Territorio Virtuale, Roma, 2, 3 e 4 ottobre 2013"**, Roma, p. 591-592, 3 out. 2013. DOI 10.5821/ctv.7933. Disponível em: <https://upcommons.upc.edu/handle/2099/16221>. Acesso em: 7 nov. 2022.

FU, TengYu; MA, Lei; LI, Manchun; JOHNSON, Brian A. Using convolutional neural network to identify irregular segmentation objects from very high-resolution remote sensing imagery. **Journal of Applied Remote Sensing**, [S. l.], v. 12, n. 2, p. 1-21, 1 maio 2018. DOI 10.1117/1.JRS.12.025010. Disponível em: <https://colab.ws/articles/10.1117/1.JRS.12.025010>. Acesso em: 17 maio 2022.

FYFE, Colin; ORTIZ-BOYER, Domingo; GARCIA-PEDRAJAS, Nicolas. Evolutionary algorithms and cross entropy. **International Journal of Knowledge-based and Intelligent Engineering Systems**, [S. l.], p. 215 - 221, 1 out. 2012. DOI 10.3233/KES-2012-00244. Disponível em: <https://dl.acm.org/doi/abs/10.3233/KES-2012-00244>. Acesso em: 10 maio 2023.

GEOPIXEL: Inteligência Municipal. [S. l.], [s. d.]. Disponível em: <https://geopixel.com.br/#>. Acesso em: 13 fev. 2024.

GHANBARI, Hamid; MAHDIANPARI, Masoud; HOMAYOUNI, Saeid; MOHAMMADIMANESH, Fariba. A meta-analysis of convolutional neural networks for remote sensing applications. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, [S. l.], v. 14, p. 3602 - 3613, 7 jan. 2021. DOI 10.1109/JSTARS.2021.3065569. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9376238>. Acesso em: 23 set. 2022.

GHOLAMALINEZHAD, Hossein; KHOSRAVI, Hossein. Pooling Methods in Deep Neural Networks, a Review. **Computer Vision and Pattern Recognition**, [S. l.], p. 1-17, 2 set. 2020. DOI 10.48550/arXiv.2009.07485. Disponível em: <https://arxiv.org/abs/2009.07485>. Acesso em: 22 out. 2022.

GONG, JianYa; LIU, Chun; HUANG, Xin. Advances in urban information extraction from high- resolution remote sensing imagery. **Science China Earth Sciences**, [S. l.], v. 63, n. 4, p. 463 - 475, 25 abr. 2020. DOI 10.1007/s11430-019-9547-x. Disponível em: <https://link.springer.com/article/10.1007/s11430-019-9547-x>. Acesso em: 17 ago. 2022.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. Deep Learning. MIT Press. 2016.

GUIRADO, Emilio; TABIK, Siham; ALCARAZ-SEGURA, Domingo; CABELLO, Javier; HERRERA, Francisco. Deep-Learning Convolutional Neural Networks for scattered shrub detection with Google Earth Imagery. **Computer Vision and Pattern Recognition**, [S. l.], p. 1 - 23, 3 jun. 2017. DOI 10.48550/arXiv.1706.00917. Disponível em: <https://arxiv.org/abs/1706.00917>. Acesso em: 5 set. 2022.

GUPTA, Neha. International Conference on recent trends in applied Sciences with Engineering Applications, 2013, India. **Artificial Neural Network** [S. l.: s. n.], 2013. v. 3. DOI 10.30780/specialissue-icaccg2020/007. Disponível em: <https://iiste.org/Journals/index.php/NCS/article/view/6063/6019>. Acesso em: 5 set. 2022.

HE, KaiMing; ZHANG, XiangYu; REN, ShaoQing; SUN, Jian. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. **Computer Vision and Pattern Recognition**, [S. l.], p. 1 - 11, 6 fev. 2015. DOI 10.48550/arXiv.1502.01852. Disponível em: <https://arxiv.org/abs/1502.01852>. Acesso em: 5 set. 2022.

HONG, Gang; ZHANG, Yun. The image registration technique for High Resolution Remote Sensing image in Hilly area. **International Society for Photogrammetry and Remote Sensing**, [S. l.], p. 1 - 6, 20 jan. 2005. Disponível em: <https://www.semanticscholar.org/paper/THE-IMAGE-REGISTRATION-TECHNIQUE-FOR-HIGH-REMOTE-IN-Hong-Zhang/3af1af6cf774b264526276745d097324c2e9a613>. Acesso em: 6 jun. 2022.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Aglomerados Subnormais**.2020. Disponível em: < <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/tipologias-do-territorio/15788-aglomerados-subnormais.html?=&t=acesso-ao-produto> >. Acesso em: 18 de abril de 2022.

IBGE - Instituto Brasileiro de Geografia e Estatística. In: **Censo Demográfico**. [S. l.], 14 mar. 2022. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/trabalho/22827-censo-demografico-2022.html>. Acesso em: 17 fev. 2024.

INPE. Sensoriamento Remoto. 2022. Disponível em: < <http://www.dsr.inpe.br/DSR/areas-de-atuacao/sensores-plataformas/sensores-opticos-alta-resolucao-espacial> >. Acesso em: 18 de abril de 2022.

ISOLA, Phillip; ZHU, Jun-Yan; ZHOU, TingHui; EFROS, Alexei A. Image-to-Image Translation with Conditional Adversarial Networks. **Conference on Computer Vision and Pattern Recognition (CVPR)**, [S. l.], p. 5967 - 5976, 21 jul. 2017. DOI 10.1109/CVPR.2017.632. Disponível em: <https://ieeexplore.ieee.org/document/8100115>. Acesso em: 3 out. 2022.

JACCARD, Paul. **Recherches embryologiques sur l'" Ephedra helvetica"**. 1894. Dissertação - Lausanne. Disponível em: <https://www.biodiversitylibrary.org/bibliography/169588>. Acesso em: 14 fev. 2024.

JANGA, Bhargavi; ASAMANI, Gokul Prathin; SUN, ZiHeng; CRISTEA, Nicoleta. A Review of Practical AI for Remote Sensing in Earth Sciences. **Remote Sensing**, [S. l.], v. 15, n. 16, p. 4112-4146, 21 ago. 2023. DOI 10.3390/rs15164112. Disponível em: <https://www.mdpi.com/2072-4292/15/16/4112>. Acesso em: 15 fev. 2024.

JHA, Debesh; HICKS, Steven; EMANUELSEN, Krister; JOHANSEN, Håvard; JOHANSEN, Dag; LANGE, Thomas de; RIEGLER, Michael; HALVORSEN, Pål. Medico Multimedia Task at MediaEval 2020: Automatic Polyp Segmentation. ArXiv preprint, [S. l.], 30 dez. 2020. DOI 10.48550/arXiv.2012.15244. Disponível em: <https://arxiv.org/abs/2012.15244>. Acesso em: 21 fev. 2024.

JINADASA, Wathsala; KAHAWALAGE, Amila C.; HALSTENSEN, Maths; SKEIE, Nils-Olav; JENS, Klaus-Joachim. Deep Learning Approach for Raman Spectroscopy. **Recent Developments in Atomic Force Microscopy and Raman Spectroscopy for Materials Characterization**, [S. l.], p. 1 - 27, 7 jan. 2022. DOI 10.5772/intechopen.99770. Disponível em: <https://www.intechopen.com/chapters/78632>. Acesso em: 15 ago. 2022.

JOSEPHINE, V L Helen; NIRMALA, A P; ALLURI, Vijaya Lakshmi. Impact of Hidden Dense Layers in Convolutional Neural Network to enhance Performance of Classification Model. **IOP Conference Series: Materials Science and Engineering**, India, p. 1 - 8, 1 abr. 2021. DOI 10.1088/1757-899X/1131/1/012007. Disponível em: <https://iopscience.iop.org/article/10.1088/1757-899X/1131/1/012007>. Acesso em: 25 ago. 2022.

KHAN, Salman H.; HAYAT, Munawar; PORIKLI, Fatih. Regularization of deep neural networks with spectral dropout. **Neural Networks**, [S. l.], v. 110, p. 82- 90, 4 fev. 2019. DOI 10.1016/j.neunet.2018.09.009. Disponível em: www.sciencedirect.com/science/article/pii/S0893608018302715. Acesso em: 10 jul. 2023

KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey. ImageNet Classification with Deep Convolutional Neural Networks. **Communications of the ACM**, [S. l.], v. 60, n. 6, p. 84 - 90, 24 maio 2017. DOI 10.1145/3065386. Disponível em: <https://dl.acm.org/doi/10.1145/3065386>. Acesso em: 9 fev. 2022.

KUCHARCZYK, Maja; HAY, Geoffrey J.; GHAFARIAN, Salar; HUGENHOLTZ, Chris H. Geographic Object-Based Image Analysis: A Primer and Future Directions. **Remote Sensing**, [S. l.], v. 12, n. 12, p. 1 - 33, 23 jun. 2020. DOI 10.3390/rs12122012. Disponível em: <https://www.mdpi.com/2072-4292/12/12/2012>. Acesso em: 19 set. 2023.

KUX, Hermann J. H.; PINHO, Carolina M. D. de; SOUZA, Iris de Marcelhas e. High-Resolution Satellite images for Urban Planning, studies in progress at INPE (National Institute for Space Research - Brazil). **ISPRS Technical Commission II Symposium**, Vienna, p. 121 - 124, 14 jul. 2006. Disponível em: <https://www.isprs.org/proceedings/XXXVI/part2/pdf/kux.pdf>. Acesso em: 12 set. 2022.

LAMBIN, Eric F.; LINDERMAN, Marc. Time series of remote sensing data for land change science. **IEEE Transactions on Geoscience and Remote Sensing**, [S. l.], v. 44, n. 7, p. 1926 - 1928, 8 jun. 2006. DOI 10.1109/TGRS.2006.872932. Disponível em: <https://ieeexplore.ieee.org/document/1645291>. Acesso em: 15 nov. 2022.

LANDIS, J. Richard; KOCH, Gary G. The Measurement of Observer Agreement for Categorical Data. **International Biometric Society**, [S. l.], v. 33, n. 1, p. 159 - 174, mar. 1977. DOI 10.2307/2529310. Disponível em: <https://www.jstor.org/stable/2529310>. Acesso em: 12 maio 2022.

LECUN, Yann; BOTTOU, Léon; BENGIO, Yoshua; HAFNER, Patrick. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, [S. l.], v. 86, n. 11, p. 2278 - 2324, 3 nov. 1998. DOI 10.1109/5.726791. Disponível em: <https://ieeexplore.ieee.org/document/726791>. Acesso em: 7 nov. 2022.

LIDER ENGENHARIA E GESTÃO DE CIDADES. **Plano Diretor Municipal**: Município de Guaratinguetá - SP. [S. l.], 1 fev. 2019. Disponível em: <https://planodiretor.guaratingueta.sp.gov.br/wp-content/uploads/2019/11/Produto-4-Diagn%C3%B3stico-Vers%C3%A3o-05.pdf>. Acesso em: 14 fev. 2024.

LIMA, Gustavo Araújo; BRAVO, Daniel Trevisan; ARAÚJO, Sidnei Alves de. Utilização de redes neurais convolucionais para a detecção de objetos em imagens aéreas adquiridas por drones / Use of convolutional neural networks to detect objects in aerial images acquired by drones. **Brazilian Journal of Development**, [S. l.], v. 6, n. 7, p. 50702 - 50713, 24 jul. 2020. DOI 10.34117/bjdv6n7-646. Disponível em: <https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/13828/11569>. Acesso em: 14 jul. 2022.

LIN, Tsung-Yi; GOYAL, Priya; GIRSHICK, Ross; HE, Kaiming; DOLLÁR, Piotr. Focal Loss for Dense Object Detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, [S. l.], v. 42, n. 2, p. 318-327, 1 fev. 2020. DOI 10.1109/TPAMI.2018.2858826. Disponível em: <https://ieeexplore.ieee.org/document/8417976>. Acesso em: 10 fev. 2024.

LIPTON, Zachary Chase; ELKAN, Charles; NARAYANASWAMY, Balakrishnan. Thresholding Classifiers to Maximize F1 Score. **Machine Learning and Knowledge Discovery in Databases**, [S. l.], v. 8725, 8 fev. 2014. DOI 10.1007/978-3-662-44851-9_15. Disponível em: https://link.springer.com/chapter/10.1007/978-3-662-44851-9_15. Acesso em: 14 fev. 2024.

LONG, Jonathan; SHELHAMER, Evan; DARRELL, Trevor. Fully Convolutional Networks for Semantic Segmentation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, [S. l.], v. 39, n. 4, p. 640 - 651, 2015. DOI 10.1109/CVPR.2015.7298965. Disponível em: <https://ieeexplore.ieee.org/document/7298965>. Acesso em: 12 set. 2023.

LOYOLA, Laura; KNOWLES, Jason; MARX, Andrew; MCALINDEN, Ryan; FLEMING, Steven. Research and Teaching Applications of Remote Sensing Integrated with GIS: Examples from the Field. **Journal of Geographic Information System**, [S. l.], v. 11, n. 6, p. 670-682, 3 dez. 2019. DOI 10.4236/jgis.2019.116041. Disponível em: <https://www.scirp.org/journal/paperinformation?paperid=97076>. Acesso em: 17 fev. 2024.

MAO, Huizi; KONG, TaeYoung; DALLY, William J. CaTDet: Cascaded Tracked Detector for Efficient Object Detection from Video. **Computer Vision and Pattern Recognition**, [S. l.], v. 10, n. 21, p. 1 - 24, fev. 2019. DOI 10.3390/app10217834. Disponível em: <https://www.mdpi.com/2076-3417/10/21/7834>. Acesso em: 16 out. 2023.

MATEOS, Carlos Javier Broncano; RUIZ, Carlos Pinilla; CRESPO, Rubén Gonzalez; SANZ, Andrés Castillo. Relative Radiometric Normalization of Multitemporal images. **International Journal of Artificial Intelligence and Interactive Multimedia**, [S. l.], v. 1, n. 3, p. 53 - 59, dez. 2010. DOI 10.9781/ijimai.2010.139. Disponível em: https://www.ijimai.org/journal/sites/default/files/IJIMAI20101_3_9. Acesso em: 24 ago. 2023.

MAXWELL, Aaron E.; WARNER, Timothy A.; FANG, Fang. Implementation of machine-learning classification in remote sensing: an applied review. **International Journal of Remote Sensing**, [S. l.], v. 39, n. 9, p. 2784–2817, 28 out. 2017. DOI 10.1080/01431161.2018.1433343. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/01431161.2018.1433343>. Acesso em: 24 nov. 2022.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **Bulletin of mathematical biophysics**, [S. l.], v. 5, p. 115- 133. 1943. DOI 10.2307/2268029. Disponível em: <https://www.cambridge.org/core/journals/journal-of-symbolic-logic/article/abs/warren-s-mcculloch-and-walter-pitts-a-logical-calculus-of-the-ideas-immanent-in-nervous-activity-bulletin-of-mathematical-biophysics-vol-5-1943-pp-115133/7DFDC43EC1E5BD05E9DA85E1C41A01BD>. Acesso em: 24 mar. 2022.

MEDIUM. Softmax vs Logsoftmax. 2021. Disponível em: <https://medium.com/@AbhiramiVS/softmax-vs-logsoftmax-eb94254445a2> >. Acesso em: 10 de janeiro de 2023.

MILER, Mario; ODOBAŠIĆ, Dražen; MEDAK, Damir. An Efficient Web-GIS Solution based on Open-Source Technologies: A Case-Study of Urban Planning and Management of the City of Zagreb, Croatia. **Facing the Challenges – Building the Capacity**, Australia, p. 1 - 16, 15 abr. 2010. DOI 10.13140/2.1.5107.8088. Disponível em: <https://oa.mg/work/2156091041>. Acesso em: 6 nov. 2023.

MOHAMMED, Nagi Zomrawi; ELHAJ, Eiman Eisa Ahmed. The Effect of Polynomial Order on Georeferencing Remote Sensing Images. **International Journal of Engineering and Innovative Technology (IJEIT)**, [S. l.], v. 2, n. 8, p. 5 - 8, 8 fev. 2013. Disponível em: https://www.ijeit.com/vol%202/Issue%208/IJEIT1412201302_02. Acesso em: 18 out. 2023.

MOTA, Suetônio. Urbanização e meio ambiente. Rio de Janeiro: ABES, 3 ed. 2003

NAIR, Vinod; HINTON, Geoffrey E. Rectified Linear Units Improve Restricted Boltzmann Machines. 2010, [S.l.: s.n.], 2010. p.807–814.

NIMER, Edmon. Climatologia do Brasil. IBGE. Departamento de Recursos Naturais e Estudos Ambientais, 1989. 422 p.

O'CONNELL, Jerome; CONNOLLY, John; VERMOTE, Eric F.; HOLDEN, Nicholas M. Radiometric normalization for change detection in peatlands: a modified temporal invariant cluster approach. **International Journal of Remote Sensing**, [S. l.], v. 34, n. 8, p. 2905 - 2924, jan. 2013. DOI doi.org/10.1080/01431161.2012.752886. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/01431161.2012.752886>. Acesso em: 25 jul. 2023.

OHATA, Arlete Tieko; QUINTANILHA, José Alberto. O uso de algoritmos de clustering na mensuração da expansão urbana e detecção de alterações na Região Metropolitana de São Paulo (1991-2002). **XII Simpósio Brasileiro de Sensoriamento Remoto**, Goiânia, Brasil, p. 647 - 655, 21 abr. 2005. Disponível em: <http://marte.sid.inpe.br/attachment.cgi/ltid.inpe.br/sbsr/2004/11.21.15.43/doc/647>. Acesso em: 11 nov. 2022.

PAN, ZhuoKun; XU, JiaShu; GUO, YuBin; HU, YueMing; WANG, GuangXing. Deep Learning Segmentation and Classification for Urban Village Using a Worldview Satellite Image Based on U-Net. **Remote Sensing**, [S. l.], v. 12, n. 10, p. 1 - 17, 15 maio 2020. DOI 10.3390/rs12101574. Disponível em: <https://www.mdpi.com/2072-4292/12/10/1574>. Acesso em: 6 abr. 2022.

PANIZZA, Andrea de Castro; FONSECA, Fernanda Padovesi. Técnicas de Interpretação Visual de Imagens. **GeoUSP: Espaço e Tempo**, [S. l.], v. 15, n. 3, p. 30 - 43, 30 abr. 2011. DOI 10.11606/issn.2179-0892.geousp.2011.74230. Disponível em: <https://www.revistas.usp.br/geousp/article/view/74230>. Acesso em: 23 maio 2022.

PASHAEI, Mohammad; KAMANGIR, Hamid; STAREK, Michael J.; TISSOT, Philippe. Review and Evaluation of Deep Learning Architectures for Efficient Land Cover Mapping with UAS Hyper-Spatial Imagery: A Case Study Over a Wetland. **Remote Sensing**, [S. l.], p. 1 - 29, 16 mar. 2020. DOI 10.3390/rs12060959. Disponível em: <https://www.mdpi.com/2072-4292/12/6/959>. Acesso em: 18 jul. 2022.

PASSO, Denilson Pereira; BIAS, Edilson de Souza; BRITES, Ricardo Seixas; COSTA, Gilson Alexandre Ostwald Pedro da. Uso do Sistema Interimage para a Identificação de Alvos Urbanos em Imagens do Satélite Worldview II. **Revista Brasileira de Cartografia**, [S. l.], v. 65, n. 6, set. 2013. DOI 10.14393/rbcv65n6-43891. Disponível em: <https://seer.ufu.br/index.php/revistabrasileiracartografia/article/view/43891>. Acesso em: 3 ago. 2022.

POUR, Amin Beiranvand; PARK, Yongcheol; PARK, Tae-Yoon; HONG, Jong Kuk; HASHIM, Mazlan; WOO, Jusun; AYOABI, Iman. Evaluation of ICA and CEM algorithms with Landsat-8/ASTER data for geological mapping in inaccessible regions. **Geocarto International**, [S. l.], p. 785-816, 9 fev. 2018. DOI 10.1080/10106049.2018.1434684. Disponível em: <https://www.tandfonline.com/action/showCitFormats?doi=10.1080%2F10106049.2018.1434684>. Acesso em: 14 fev. 2024.

PROVOST, Foster John; FAWCETT, Tom Elliott; KOHAVI, Ron. The Case against Accuracy Estimation for Comparing Induction Algorithms. **International Conference on Machine Learning**, [S. l.], p. 445 - 453, 24 jul. 1998. Disponível em: https://www.researchgate.net/publication/2373067_The_Case_Against_Accuracy_Estimation_for_Comparing_Induction_Algorithms. Acesso em: 21 jun. 2022.

REN, YongFeng; YU, YongTao; GUAN, HaiYan. DA-CapsUNet: A Dual-Attention Capsule U-Net for Road Extraction from Remote Sensing Imagery. **Remote Sensing**, [S. l.], v. 12, n. 18, 4 set. 2020. DOI 10.3390/rs12182866. Disponível em: <https://www.mdpi.com/2072-4292/12/18/2866>. Acesso em: 5 jul. 2022.

RENNÓ, Camilo Daleles. **Avaliação de medidas texturais na discriminação de classes de uso utilizando imagens SIR-C/X-SAR do Perímetro Irrigado de Bebedouro, Petrolina, PE**. 2004. 109 f. Dissertação (Mestrado) - INPE, [S. l.], 2004. Disponível em: <http://mtc-m12.sid.inpe.br/col/sid.inpe.br/jeferson/2004/03.05.11.22/doc/publicacao.pdf>. Acesso em: 13 abr. 2022.

REZATOFIGHI, Hamid; TSOI, Nathan; GWAK, JunYoung; SADEGHIAN, Amir; REID, Ian; SAVARESE, Silvio. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression. **2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**, EUA, p. 658 - 666, 2019. DOI 10.1109/CVPR.2019.00075. Disponível em: <https://ieeexplore.ieee.org/document/8953982>. Acesso em: 16 fev. 2022.

RIBEIRO, Bárbara Maria Giacom; FONSECA, Leila Maria Garcia; KUX, Hermann Johann Heinrich. Mapeamento da Cobertura do Solo Urbano utilizando imagens WorldView-II e o sistema Interimage. **Revista Brasileira de Cartografia**, [S. l.], v. 3, n. Edição Especial 40 anos, p. 51 - 63, 2011. Disponível em: <https://seer.ufu.br/index.php/revistabrasileiracartografia/article/view/43767>. Acesso em: 5 jul. 2022.

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-Net: Convolutional Networks for Biomedical Image Segmentation. **Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015**, [S. l.], p. 234 - 241, 18 nov. 2015. DOI 10.1007/978-3-319-24574-4_28. Disponível em: https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28. Acesso em: 14 maio 2021.

SHAFIQUE, Ayesha; CAO, Guo; KHAN, Zia; ASAD, Muhammad; ASLAM, Muhammad. Deep Learning-Based Change Detection in Remote Sensing Images: A Review. **Remote Sensing**, [S. l.], v. 14, n. 4, p. 871 - 911, 11 fev. 2022. DOI 10.3390/rs14040871. Disponível em: <https://www.mdpi.com/2072-4292/14/4/871>. Acesso em: 17 out. 2022.

SHAO, Yang; LUNETTA, Ross. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. **ISPRS Journal of Photogrammetry and Remote Sensing**, [S. l.], v. 70, p. 78 - 87, jun. 2012. DOI 10.1016/j.isprsjprs.2012.04.001. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0924271612000664?via%3Di> hub. Acesso em: 8 maio 2023.

SHI, WenZhong; ZHANG, Min; ZHANG, Rui; CHEN, ShanXiong; ZHAN, Zhan. Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges. **Remote Sensing**. vol. 12, 2020. DOI: 10.3390/rs12101688. Acesso em: 13 out. 2023.

SI SALAH, Hayet; GOLDIN, Sally E.; REZGUI, Abdelmounaam; EL ISLAM, Bachari Nour; AIT-AOUDIA, Samy. What is a remote sensing change detection technique?: Towards a conceptual framework. **International Journal of Remote Sensing**, [S. l.], v. 41, n. 5, 11 out. 2019. DOI 10.1080/01431161.2019.1674463. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/01431161.2019.1674463>. Acesso em: 23 jun. 2023.

SILVA, Daniel Moura da. Classificação de lavouras por aprendizagem profunda com dados de sensores remotos. 2018. 108 f., il. Trabalho de Conclusão de Curso (Bacharelado em Engenharia de Software) —Universidade de Brasília, Brasília, 2018.

SINGH, Ashbindu. Review Article Digital change detection techniques using remotely-sensed data. **International Journal of Remote Sensing**, [S. l.], v. 10, n. 6, p. 989 - 1003, 8 jul. 2010. DOI 10.1080/01431168908903939. Disponível em: <https://www.tandfonline.com/doi/pdf/10.1080/01431168908903939>. Acesso em: 3 jul. 2023.

SUDRE, Carole; LI, WenQi; VERCAUTEREN, Tom; OURSELIN, Sebastien; CARDOSO, Jorge. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. **Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support**, [S. l.], v. 10553, p. 240-248, 9 set. 2017. DOI 10.1007/978-3-319-67558-9_28. Disponível em: https://link.springer.com/chapter/10.1007/978-3-319-67558-9_28#citeas. Acesso em: 13 fev. 2024.

TATO, Ange; NKAMBOU, Roger. Improving Adam Optimizer. **6th International Conference on Learning Representations**, [S. l.], 2018. Disponível em: <https://openreview.net/pdf?id=HJfpZq1DM>. Acesso em: 16 maio 2023.

TENSORFLOW. In: **Tensores e Operações**. [S. l.], 24 jan. 2022. Disponível em: https://www.tensorflow.org/js/guide/tensors_operations?hl=pt-br. Acesso em: 6 fev. 2024.

TONDEWAD, Priyanka S.; DALE, Manisha P. Remote Sensing Image Registration Methodology: Review and Discussion. **Procedia Computer Science**, [S. l.], v. 171, p. 2390 - 2399, 2020. DOI 10.1016/j.procs.2020.04.259. Disponível em: [sciencedirect.com/science/article/pii/S1877050920312515](https://www.sciencedirect.com/science/article/pii/S1877050920312515). Acesso em: 7 jan. 2023.

UTSUMI, Alex Garcez; MARTINS DE ALMEIDA OLIVEIRA, Nubya; GUIMARÃES SOUSA, Nadia Análise de imagens Sentinel-2 no mapeamento de formações florestais no município de Uberaba-MG. **Scientia Plena**, [S. l.], v. 19, n. 12, 2023. DOI: 10.14808/sci.plena.2023.129901. Disponível em: <https://www.scientiaplena.org.br/sp/article/view/7522>. Acesso em: 27 fev. 2024.

WU, ChuYi; ZHANG, Feng; XIA, JunShi; XU, YiChen; LI, GuoQing; XIE, JiBo; DU, ZhenHong; LIU, RenYi. Building Damage Detection Using U-Net with Attention Mechanism from Pre- and Post-Disaster Remote Sensing Datasets. **Remote Sensing**, [S. l.], v. 13, n. 5, p. 905, fev. 2021. DOI 10.3390/rs13050905. Disponível em: <https://www.mdpi.com/2072-4292/13/5/905>. Acesso em: 25 maio 2022.

XU, Kelvin; BA, Jimmy; KIROS, Ryan; CHO, KyungHyun; COURVILLE, Aaron; SALAKHUTDINOV, Ruslan; ZEMEL, Richard; BENGIO, Yoshua. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. **International conference on machine learning**, [S. l.], 10 fev. 2015. DOI 10.48550/arXiv.1502.03044. Disponível em: <https://arxiv.org/abs/1502.03044v3>. Acesso em: 12 jul. 2023.

YUAN, ZiHao; LI, Ji; LI, Zhe; DING, CaiWen; REN, Ao; YUAN, Bo; QIU, QinRu; DRAPER, Jeffrey; WANG, YanZhi. Softmax Regression Design for Stochastic Computing Based Deep Convolutional Neural Networks. **GLSVLSI '17: Proceedings of the on Great Lakes Symposium on VLSI**, [S. l.], p. 467 - 470, 10 maio 2017. DOI 10.1145/3060403.3060467. Disponível em: <https://dl.acm.org/doi/10.1145/3060403.3060467>. Acesso em: 16 maio 2023.

ZHANG, XiaoDong; CHEN, GuanZhou; WANG, WenBo; WANG, Qing; DAI, Fan. Object-Based Land-Cover Supervised Classification for Very-High-Resolution UAV Images Using Stacked Denoising Autoencoders. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, [S. l.], v. 10, n. 7, p. 3373 - 3385, jul. 2017. DOI 10.1109/JSTARS.2017.2672736. Disponível em: <https://ieeexplore.ieee.org/document/7879207>. Acesso em: 19 set. 2023.

ZHANG, XinChe. The AlexNet, LeNet-5 and VGG NET applied to CIFAR-10. **2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)**, China, p. 414-419, setembro 2021. DOI: 10.1109/ICBASE53849.2021.00083.

ZHAO, Hang; GALLO, Orazio; FROSIO, Iuri; KAUTZ, Jan. Loss Functions for Image Restoration with Neural Networks. **IEEE Transactions on Computational Imaging**, [S. l.], v. 3, n. 1, p. 47 - 57, mar. 2018. DOI 10.1109/TCI.2016.2644865. Disponível em: <https://ieeexplore.ieee.org/document/7797130>. Acesso em: 14 fev. 2023.

ZHOU, WeiQi; TROY, Austin. An Object-Oriented Approach for Analyzing and Characterizing Urban Landscape at the Parcel Level. **International Journal of Remote Sensing**, [S. l.], v. 29, n. 11, p. 3119 - 3135, 19 maio 2008. DOI 10.1080/01431160701469065. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01431160701469065>. Acesso em: 10 out. 2022.

ZHU, XiaoXiang; TUIA, Devis; MOU, LiChao; XIA, GuiSong; ZHANG, LiangPei; XU, Feng; FRAUNDORFER, Friedrich. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. **IEEE Geoscience and Remote Sensing**, [S. l.], v. 5, n. 4, p. 8 - 36, 5 dez. 2017. DOI 10.1109/MGRS.2017.2762307. Disponível em: <https://ieeexplore.ieee.org/document/8113128/authors>. Acesso em: 18 jan. 2022.

ZHUANG, Xin; ENGEL, Bernard A.; XIONG, XiaoPing; JOHANNSEN, Chris J. Analysis of Classification Results of Remotely Sensed Data and Evaluation of Classification Algorithm. **Photogrammetric Engineering and Remote Sensing**, EUA, v. 61, n. 4, p. 427 - 433, maio 1995. Disponível em: https://www.asprs.org/wp-content/uploads/pers/1995journal/apr/1995_apr_427-433. Acesso em: 4 abr. 2023.

ZHUANG, Xin; ENGEL, Bernard A.; XIONG, XiaoPing; JOHANNSEN, Chris J. Analysis of Classification Results of Remotely Sensed Data and Evaluation of Classification Algorithm. **Photogrammetric Engineering and Remote Sensing**, EUA, v. 61, n. 4, p. 427 - 433, maio 1995. Disponível em: https://www.asprs.org/wp-content/uploads/pers/1995journal/apr/1995_apr_427-433. Acesso em: 4 abr. 2023.

APÊNDICE A

2021		2022				
E (m)	N (m)	E (m)	N (m)	ΔE (m)	ΔN (m)	Resíduo (m)
479840,3936	7481224,023	479836,5489	7481226,130	0,125226439	-0,213118691	0,247186645
479307,6439	7480558,923	479303,9454	7480561,298	-0,052723116	0,193362984	0,200421981
479410,5505	7480459,973	479406,7131	7480462,123	0,297988047	-0,267985825	0,400765865
479860,4441	7480878,307	479856,8249	7480880,644	-0,282406682	0,321715790	0,428082450
480794,5658	7480396,974	480790,7391	7480399,068	0,297611612	-0,023389450	0,298529292
481319,5462	7480051,700	481315,7788	7480053,512	0,269435267	-0,548355292	0,610973723
481239,3442	7481532,880	481235,7292	7481535,237	-0,110478665	-0,116078868	0,160249303
481173,0401	7481547,405	481169,3756	7481549,815	-0,006810332	-0,007010255	0,009773653
481358,7304	7480847,857	481354,9602	7480850,072	0,021548850	0,154026787	0,155526860
481374,3895	7480485,621	481370,4219	7480487,644	0,464037799	-0,101693973	0,475050254
481565,3157	7480366,066	481561,8222	7480368,312	-0,486449202	0,375007251	0,614217603
480325,7556	7481126,634	480321,9942	7481128,861	0,019208619	0,078204299	0,080528774
480844,8006	7480111,562	480841,3499	7480113,848	-0,303233217	0,303939270	0,429336073
480142,1886	7480475,318	480138,6850	7480477,514	-0,347332183	0,055638097	0,351760206
479450,4826	7482008,277	479446,7642	7482010,979	-0,012746711	0,463161383	0,463336752
479426,0649	7482064,906	479422,4208	7482067,331	-0,138393054	-0,144605703	0,200158554
479489,5943	7481789,097	479485,8139	7481791,260	0,022822457	-0,421591433	0,422208718
479991,8461	7481892,462	479988,0826	7481895,139	0,187162645	0,430995402	0,469879657
480115,9131	7481619,524	480112,1717	7481621,620	0,046258135	-0,489427802	0,491608979
481146,0889	7481106,101	481142,3527	7481108,287	-0,010726709	-0,042793970	0,044117867

Fonte: A autora.

APÊNDICE B

```
## LEITURA DAS IMAGENS E DAS MASCARAS
img_dir = '/caminho_no_drive/pasta_das_imagens'
mask_dir = '/caminho_no_drive/pasta_das_mascaras'

img_lst = os.listdir(img_dir)
mask_lst = os.listdir(mask_dir)

img = [filename for filename in img_lst if filename.startswith('imagem_')]
mask = [filename for filename in mask_lst if filename.startswith('mascara_')]

img.sort()
mask.sort()

X = []
for i in img:
    img_m = rasterio.open(os.path.join(img_dir, i))
    arrx = img_m.read()
    X.append(arrx)
X = np.array(X)

y = []
for i in mask:
    mask_m = rasterio.open(os.path.join(mask_dir, i))
    arry = mask_m.read()
    y.append(arry)
y = np.array(y)

print("Unicas classes na camada de dados: ", np.unique(y))

X = np.moveaxis(np.array(X), 1, 3)
y = np.moveaxis(np.array(y), 1, 3)

## SEPARAR EM TESTE E TREINAMENTO
n_classes = len(np.unique(y))
labels_cat = to_categorical(y, num_classes=n_classes)

X_train, X_test, y_train, y_test = train_test_split(X, labels_cat, test_size = 0.20,
random_state = 0)

IMG_HEIGHT = X_train.shape[1]
IMG_WIDTH = X_train.shape[2]
IMG_CHANNELS = X_train.shape[3]
```

```

        ## DEFINIÇÃO DO COEFICIENTE DE JACCARD (IOU)
def jacard_coef (y_true, y_pred):
    y_true_f = K.flatten(y_true)
    y_pred_f = K.flatten(y_pred)
    intersection = K.sum(y_true_f*y_pred_f)
    return (intersection + 1.0)/(K.sum(y_true_f) + K.sum(y_pred_f) - intersection + 1.0)


        ## DEFINIÇÃO DA U-NET
def unet(input_shape, num_classes):
    inputs = Input(input_shape)

    conv1 = Conv2D(64, 3, activation='relu', padding='same')(inputs)
    conv1 = Conv2D(64, 3, activation='relu', padding='same')(conv1)
    conv1 = Dropout(0.2)(conv1)
    pool1 = MaxPooling2D(pool_size=(2, 2))(conv1)

    conv2 = Conv2D(128, 3, activation='relu', padding='same')(pool1)
    conv2 = Conv2D(128, 3, activation='relu', padding='same')(conv2)
    conv2 = Dropout(0.2)(conv2)
    pool2 = MaxPooling2D(pool_size=(2, 2))(conv2)

    conv3 = Conv2D(256, 3, activation='relu', padding='same')(pool2)
    conv3 = Conv2D(256, 3, activation='relu', padding='same')(conv3)

    upconv4 = Conv2DTranspose(128,(2,2), strides=(2, 2),padding='same')(conv3)
    concat4 = concatenate([conv2, upconv4], axis=-1)
    conv4 = Conv2D(128, 3, activation='relu', padding='same')(concat4)
    conv4 = Conv2D(128, 3, activation='relu', padding='same')(conv4)
    conv4 = Dropout(0.2)(conv4)

    upconv5 = Conv2DTranspose(64, (2,2), strides=(2, 2),padding='same')(conv4)
    concat5 = concatenate([conv1, upconv5], axis=-1)
    conv5 = Conv2D(64, 3, activation='relu', padding='same')(concat5)
    conv5 = Conv2D(64, 3, activation='relu', padding='same')(conv5)
    conv5 = Dropout(0.2)(conv5)

    # Camada de saída
    output = Conv2D(num_classes, 1, activation='sigmoid')(conv5)

    model = Model(inputs=inputs, outputs=output)
    return model

input_shape = (200, 200, 4)
num_classes = 3

model = unet(input_shape, num_classes)

model.compile(optimizer='adam', loss='categorical_crossentropy',
metrics=['accuracy', jacard_coef])

```

```
model.summary()
```

```
## TREINAMENTO
```

```
history = model.fit(X_train, y_train,  
                    batch_size = 40,  
                    verbose=1,  
                    epochs=200,  
                    validation_data=(X_test, y_test),  
                    shuffle=False)
```

```
## PLOTAR GRAFICOS DE ANALISE DO TREINAMENTO
```

```
loss = history.history['loss']  
val_loss = history.history['val_loss']  
epochs = range(1, len(loss) + 1)  
plt.plot(epochs, loss, 'y', label='Perda do Treinamento')  
plt.plot(epochs, val_loss, 'r', label='Perda da Validação')  
plt.title('Perda do Treinamento e da Validação')  
plt.xlabel('Épocas')  
plt.ylabel('Perda')  
plt.legend()  
plt.show()
```

```
acc = history.history['jacard_coef']  
val_acc = history.history['val_jacard_coef']
```

```
plt.plot(epochs, acc, 'y', label='IoU do Treinamento')  
plt.plot(epochs, val_acc, 'r', label='IoU da Validação')  
plt.title('IoU do Treinamento e da Validação')  
plt.xlabel('Épocas')  
plt.ylabel('IoU')  
plt.legend()  
plt.show()
```

```
## SALVAR OU CARREGAR O MODELO
```

```
# SALVAR:
```

```
model.save('/caminho_no_drive/results.zip')
```

```
# CARREGAR:
```

```
model = load_model('/caminho_no_drive/results.zip',  
                  custom_objects={'jacard_coef':jacard_coef})
```

```
# IOU:
```

```
y_pred=model.predict(X_test)  
IOU_keras = MeanIoU(num_classes=3)  
IOU_keras.update_state(y_test, y_pred)
```

```

print("Mean IoU =", IOU_keras.result().numpy())

## ACURACIA GERAL
accuracy_metric = Accuracy()
accuracy_metric.update_state(y_test, y_pred)
accuracy_value = accuracy_metric.result().numpy()
accuracy_percentage = accuracy_value * 100
print("Acurácia Global: {:.2f}%".format(accuracy_percentage))

## KAPPA
test = y_test.reshape(-1, y_test.shape[-1])
pred = y_pred.reshape(-1, y_pred.shape[-1])

kappa = cohen_kappa_score(test.argmax(axis=1), pred.argmax(axis=1))
print("Índice Kappa:", kappa)

## MATRIZ DE CONFUSÃO NORMALIZADA
confusion = confusion_matrix(test.argmax(axis=1), pred.argmax(axis=1))

confusion_normalized = confusion.astype('float') / confusion.sum(axis=1)[:,
np.newaxis]

fig, ax = plt.subplots(figsize=(10, 10))
plt.rcParams['font.size'] = 20

sns.heatmap(confusion_normalized, ax=ax, annot=True, cmap='Blues')
ax.set_xlabel('Predito')
ax.set_ylabel('Real')
ax.set_title('Matriz de Confusão Normalizada')
plt.show()

## SALVAR PREDIÇÕES
img_dir = '/caminho_driva/pasta_imagens'
img_lst = os.listdir(img_dir)
img = [filename for filename in img_lst if filename.startswith('img_')]
img.sort()

j = 0
for i in img:
    img_m = rasterio.open('/caminho_drive/pasta_imagens/'+i)
    meta = img_m.meta
    meta.update({"count": 1})
    imgs = rasterio.open('/caminho_drive/predicao_'+str(j)+'.tif', 'w', **meta)
    test_img = img_m.read()
    test_img_input = np.expand_dims(test_img, 0)
    A = np.moveaxis(test_img_input, 1, 3)[: :, :]

```

```

prediction = (model.predict(A))
predicted_img=np.argmax(prediction, axis=3)[:,:,:]
imgs.write(predicted_img)
j=j+1

im1 = rasterio.open('/caminho/predicao2021.tif')
ar1 = im1.read()
im2 = rasterio.open('/caminho/predicao2022.tif')
ar2 = im1.read()

im_1flat = (ar1.reshape(-3, ar1.shape[-3]))[:,:]
im_1flat = im_1flat.astype(int)
im_2flat = (ar2.reshape(-3, ar2.shape[-3]))[:,:]
im_1flat = im_1flat.astype(int)

mud = []

for i,j in zip(im_1flat, im_2flat):
    if i==0 and j==1:
        m=1
    elif i==1 and j==0:
        m=2
    elif i==1 and j==2:
        m=3
    elif i==2 and j==0:
        m=4
    elif i==2 and j==1:
        m=5
    else:
        m=0
    mud.append(m)

arrs = np.array(mud)

# Assumindo que ar1 é um array 3D
shape_ar1 = ar1.shape

# Crie uma matriz 3D preenchida com zeros
arrs_3d = np.zeros(shape_ar1, dtype=np.uint8)

# Redimensione arrs para ter o mesmo número de elementos que ar1
arrs_reshaped = np.array(mud).reshape(shape_ar1)

# Preencha cada banda de arrs_3d com arrs_reshaped
for band in range(shape_ar1[0]):
    arrs_3d[band, :, :] = arrs_reshaped

# Agora arrs_3d tem a mesma forma que ar1
print(len(arrs_3d))

```

```

meta = im1.meta
imgs = rasterio.open('/caminho/mudanca.tif','w',**meta)
imgs.write(arrs_3d)

                                ## MATRIZ DE TRANSIÇÃO
# Abra as imagens de data 1 e data 2
data1 = rasterio.open('/caminho/predicao2021.tif')
data2 = rasterio.open('/caminho/predicao2022.tif')

# Leia os dados das imagens
data1_array = data1.read(1).astype(int) # Assumindo que você está interessado na
primeira banda
data2_array = data2.read(1).astype(int)

# Mapeie classes para valores inteiros usando um dicionário
class_mapping = {0: 0, 1: 1, 2: 2}

# Crie uma matriz de transição 3x3 (classes 0, 1, 2)
transition_matrix = np.zeros((3, 3), dtype=np.int32)

# Itere sobre cada pixel nas duas imagens e conte as transições
for x in range(data1.width):
    for y in range(data1.height):
        class1 = class_mapping[data1_array[y, x]]
        class2 = class_mapping[data2_array[y, x]]
        transition_matrix[class1, class2] += 1

print("Matriz de Transição:")
print(transition_matrix)

```