

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS AMBIENTAIS

Wesley Rafael Nunes da Silva

**ANÁLISE DE ESTRATÉGIAS PARA IDENTIFICAÇÃO DE COPAS DE
ÁRVORES EM ÁREAS ARBÓREAS UTILIZANDO IMAGENS DE
AERONAVES REMOTAMENTE PILOTADAS.**

SÃO CARLOS -SP
2021

WESLEY RAFAEL NUNES DA SILVA

ANÁLISE DE ESTRATÉGIAS PARA IDENTIFICAÇÃO DE COPAS DE
ÁRVORES EM ÁREAS ARBÓREAS UTILIZANDO IMAGENS DE
AERONAVES REMOTAMENTE PILOTADAS.

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Ambientais da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciências Ambientais¹.

Orientador(a): Prof. Dr. Vandoir Bourscheidt

São Carlos-SP
2021

¹ Apoio: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES)
- Código de Financiamento 001

Nunes da Silva, Wesley Rafael

Análise de estratégias para identificação de copas de
árvores em áreas arbóreas utilizando imagens de
Aeronaves Remotamente Pilotadas / Wesley Rafael
Nunes da Silva -- 2021.
80f.

Dissertação (Mestrado) - Universidade Federal de São
Carlos, campus São Carlos, São Carlos
Orientador (a): Vandoir Bourscheidt
Banca Examinadora: Paulo Guilherme Molin, Ricardo
Dal'Agnol da Silva
Bibliografia

1. Geoprocessamento. 2. Copas de árvores. 3. RPAs. I.
Nunes da Silva, Wesley Rafael. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática
(SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Ronildo Santos Prado - CRB/8 7325



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Ciências Ambientais

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Wesley Rafael Nunes da Silva, realizada em 26/05/2021.

Comissão Julgadora:

Prof. Dr. Vandoir Bourscheidt (UFSCar)

Prof. Dr. Paulo Guilherme Molin (UFSCar)

Prof. Dr. Ricardo DalAgnol da Silva (INPE)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciências Ambientais.

À Alexandra e Elias, que fizeram o impossível para assegurar o meu presente e futuro.

À Ana Claudia, que me assistiu nos momentos de ansiedade e concentração.

À Bianca, Thiago e todos os amigos, que contribuíram substancialmente para a conclusão desse manuscrito.

AGRADECIMENTOS

Agradeço ao meu orientador, Professor Dr. Vandoir Bourscheidt por ter aceitado realizar esse trabalho. Seus conselhos foram fundamentais para me guiar nesse percurso intenso do mestrado e suas instruções permitiram que o projeto de pesquisa evoluísse.

Agradeço ao Programa de Ciências Ambientais da Universidade Federal de São Carlos, cujo corpo docente e secretarias me ofereceram todo suporte para que eu pudesse realizar a minha pesquisa.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

RESUMO

Nunes da Silva, W. R. – ANÁLISE DE ESTRATÉGIAS PARA IDENTIFICAÇÃO DE COPAS DE ÁRVORES EM ÁREAS ARBÓREAS UTILIZANDO IMAGENS DE AERONAVES REMOTAMENTE PILOTADAS. 2021 – 80p: Dissertação (Mestrado) – Universidade Federal de São Carlos, Centro de Ciências Biológicas e da Saúde, Programa de Pós-Graduação em Ciências Ambientais, São Carlos – 2021.

O monitoramento de áreas florestais é uma ferramenta essencial para a conservação ambiental, tanto pelo controle de supressão da vegetação nativa como no acompanhamento dos processos de restauração, havendo constantes aprimoramentos das estratégias e técnicas utilizadas na sua execução. Para tanto, faz-se necessário conhecer a estrutura dessas áreas visando ter um melhor entendimento de suas dinâmicas e distribuição. Nesse cenário, o uso das ferramentas de geoprocessamento, principalmente através da técnica de segmentação das copas das árvores, surge como um grande aliado para estudar a vegetação de forma detalhada, auxiliando nos estudos da estrutura da floresta e, conseqüentemente, na sua conservação. No entanto, sua aplicabilidade é muitas vezes limitada pela qualidade das imagens disponíveis para análise e, por vezes, pelas técnicas utilizadas para interpretação dos dados. As aeronaves remotamente pilotadas (RPAs) vêm se tornando uma excelente alternativa em termos de qualidade dos dados, permitindo a obtenção de imagens de alta resolução e com baixo investimento. Além disso, com a facilitação do acesso às técnicas de aprendizagem de máquinas, é possível utilizar modelos de redes neurais para processar dados de maneira inovadora e eficiente. Neste trabalho foram utilizadas imagens obtidas através de RPA para avaliar diferentes técnicas de segmentação de copas de árvores, a fim de encontrar a mais adequada para identificar e contar indivíduos arbóreos. Os seguintes métodos foram avaliados: segmentação de filtro de moda, segmentação multiresolution, segmentação watershed e segmentação com redes neurais para aprendizagem profunda. As imagens foram obtidas em áreas com diferentes fisionomias vegetais (silvicultura de *Pinus* e mata ciliar), a fim de avaliar seu impacto nos resultados. Os resultados foram comparados usando um conjunto de dados de validação (verdade terrestre) gerados a partir de amostras independentes dessas mesmas imagens, a fim de observar a precisão das técnicas. Os resultados apontaram que o filtro de moda gerou polígonos que não correspondem às copas das árvores, possivelmente devido à intensa homogeneização das imagens, o que enfatiza seu uso como ferramenta intermediária de segmentação, não como produto final. As segmentações multiresolution e watershed, por outro lado, mostraram bom potencial para detectar copas de árvores na área de *Pinus*, porém seus parâmetros devem ser ajustados de acordo com a área-alvo, o que dificulta o uso e exige a execução do algoritmo múltiplas vezes. Em última análise, a segmentação gerada pelo modelo de aprendizado profundo (U-Net) teve problemas para conectar as bordas das copas das árvores e mostrou um resultado sem polígonos individuais para cada árvore. No entanto, a ferramenta mostrou um bom potencial para detectar clareiras dentro da floresta, com a rotulagem utilizada neste projeto. Os produtos gerados apresentaram resultados melhores na área de Silvicultura em comparação à área de mata ciliar, situação que pode estar relacionada com a diferença de adensamento das copas das árvores entre essas regiões. Finalmente, sugere-se a condução de experimentos adicionais, utilizando ainda outros sensores (como LiDAR), para verificar se há uma melhora na qualidade da segmentação.

Palavras-chave: copas de árvores; geoprocessamento; sensoriamento remoto; RPAs

ABSTRACT

The monitoring of forest areas is an essential tool for environmental conservation, either by controlling the suppression of native vegetation or by monitoring the restoration processes, with constant improvements in the strategies and techniques used in their execution. Therefore, it is necessary to know the structure of these areas in order to have a better understanding of their dynamics and distribution. In this scenario, the use of geoprocessing tools, especially through the technique of segmentation of the treetops, appears as an essential tool to study vegetation in detail, assisting in the study of the structure of the forest and its conservation. However, its applicability is often limited by the quality of the images available for analysis and, sometimes, by the techniques used for data interpretation. Remotely piloted aircraft (RPAs) are becoming an excellent alternative in terms of data quality, allowing the acquisition of high-resolution images at low cost. In addition, with easier access to machine learning techniques, it is possible to use neural network models to process data in an innovative and efficient way. In this work, images obtained through RPA were used to evaluate different techniques for segmentation of tree crowns, aiming to find the most suitable to identify and count tree individuals. The following segmentation methods were evaluated: mode filter, multiresolution, watershed and segmentation with neural networks for deep learning. The images were obtained in areas with different plant physiognomies (Pinus silviculture and riparian forest), in order to assess their impact on the results. The results were compared using a set of validation data (terrestrial truth) generated from independent samples of these same images, to observe the accuracy of the techniques. The results showed that the polygons from the mode filter do not correspond to the crowns of the trees, possibly due to the intense homogenization of the images, which emphasizes its use as an intermediate segmentation tool. The multiresolution and watershed segmentations, on the other hand, showed good potential to detect treetops in the Pinus area, however, its parameters must be adjusted according to the target area, which makes it difficult to use and requires the algorithm to be executed multiple times. Ultimately, the segmentation generated by the deep learning model (U-Net) had problems connecting the edges of the treetops and showed a result without individual polygons for each tree. However, with the labeling used in this project, the tool showed good potential for detect gaps within the forest. In all cases, results were better in Silviculture compared to the area of Riparian Forest, a situation that may be related to the difference in the density of the treetops between these regions. Finally, additional experiments are necessary, using other sensors (such as LiDAR) and methods, to check if there is an improvement in the quality of the segmentation.

Keywords: canopy; geoprocessing; remote sensing; RPAs

LISTA DE FIGURAS

Figura 1 – Sobreposição de Matrizes	19
Figura 2 – Exemplo de Variação da Resolução Espacial	20
Figura 3 – Bandas Individuais de Imagem Multiespectral	21
Figura 4 – Resolução Espacial e Temporal para Aplicações Específicas	21
Figura 5 – Exemplo de Resolução Radiométrica	22
Figura 6 – Diferença no Comportamento Espectral de Alvos Distintos	24
Figura 7 – Comparação Entre Tons de Cinza por Média e rec601 luma	25
Figura 8 – Conversão para Tons de Cinza	26
Figura 9 – Representação 2D do PCA	27
Figura 10 – Exemplos de Filtros Lineares	27
Figura 11 – Exemplo de aplicação do Filtro de Mediana	28
Figura 12- Elementos Estruturantes em Tons de Cinza	30
Figura 13 – Exemplo de Rede Neural	31
Figura 14 – Exemplo de estrutura de uma Rede Neural Convolutacional.	36
Figura 15 – Arquitetura U-Net	37
Figura 16 – Figura de Localização	38
Figura 17 – Organograma de Execução da Pesquisa	39
Figura 18 - Etapas Realizadas no Agisoft Metashape	40
Figura 19 - Verdade Terrestre Vetorizada na Silvicultura	41
Figura 20 - Verdade Terrestre Vetorizada na Mata Ciliar	42
Figura 21 – Recortes do RGB e DSM das Áreas de Estudo	46
Figura 22 – Ruídos Presente nas Imagens em Tons de Cinza	46
Figura 23 – Fatiamento da Copa em Várias Camadas	47
Figura 24 – Exemplo de Operação Morfológica de Abertura	47
Figura 25 – Valores médios da diferença entre duas imagens da série	48
Figura 26 – Valores médios da diferença entre duas imagens da série após o filtro gaussiano	48
Figura 27 – Áreas Recortadas para o Treinamento da Rede Neural	50
Figura 28 – Par Imagem RGB e Mascara Vetorizada	51
Figura 29 – Variação da Perda e Acurácia do Modelo em Função das Épocas	52
Figura 30 - Variação dos Índices de Similaridade em Diferentes Situações	53
Figura 31 – Resultado da Segmentação por Filtro de Moda PC1	54
Figura 32 – Resultado da Segmentação por Filtro de Moda PC2	54
Figura 33 – Declividade	55
Figura 34 – Resultado da Segmentação Multiresolution	56
Figura 35 – Comparação entre as Scan Lines de Cada Operação morfológica	57
Figura 36 – Resultados da Segmentação Watershed	58
Figura 37 – Resultado da Segmentação Utilizando a U-Net	59

LISTA DE QUADROS

Quadro 1 – Exemplos de Operações Morfológicas	29
Quadro 2 – Comparação entre várias técnicas de segmentação	33
Quadro 3 – Resultados na Literatura Para as Técnicas de Segmentação	34
Quadro 4 – Mínimas Locais em Função do Tamanho da Kernel.....	49

LISTA DE TABELAS

Tabela 1 – Razão entre a área da amostra e o número de polígonos por técnica de segmentação	60
Tabela 2 – Índices de Dice e Jaccard para a área de <i>Pinus</i>	60
Tabela 3 – Índices de Dice e Jaccard para a área de Mata Ciliar.....	61
Tabela 4 - Comparações dos Índices de Jaccard no Pinus.....	61
Tabela 5 - Comparação dos Índices de Jaccard na Mata Ciliar.....	61

TABELA DE ABREVIATURAS E SIGNIFICADOS

RPA - Aeronaves Remotamente Pilotadas

SMM - Sistemas Móveis de Mapeamento

LiDAR – Light Detection and Ranging

I.A. - Inteligencia Artificial

ML - Machine Learning

PDI - Processamento Digital de Imagem

PCA - Principal Components Analysis

DEM - Modelo Digital de Elevação

DSM - Modelo Digital de Superfície

DTM - Modelo Digital de Terreno

CNN - Rede Neural Convolutiva

SIG - Sistema de Informação Geográfica

SUMÁRIO

1 INTRODUÇÃO	13
2 OBJETIVOS	18
2.1 OBJETIVO GERAL	18
2.2 OBJETIVOS ESPECÍFICOS	18
3 EMBASAMENTO TEÓRICO	19
3.1 IMAGENS DIGITAIS	19
3.2 PROCESSAMENTO DE IMAGEM	22
3.2.1 Análise de Componentes Principais	26
3.2.2 Filtros	27
3.3 REDES NEURAIS	30
3.4 SEGMENTAÇÃO	32
3.4.1 PCA e filtro de moda	34
3.4.2 Multiresolution	35
3.4.3 Watershed	35
3.4.4 Aprendizagem profunda	36
4 METODOLOGIA	38
4.1 ÁREA DE ESTUDO	38
4.2 CARACTERÍSTICAS DO RPA, CÂMERA	39
4.3 AQUISIÇÃO E PREPARAÇÃO DE DADOS	40
4.4 MÉTODOS DE SEGMENTAÇÃO	42
4.4.1 Segmentação de Filtro de Moda	42
4.4.2 Multiresolution Segmentation	43
4.4.3 Watershed Segmentation	44
4.4.4 Segmentação por Rede Neural – Aprendizagem Profunda	49
4.5 ÍNDICES PARA AVALIAÇÃO DOS RESULTADOS: SØRENSEN–DICE E COEFICIENTE JACCARD	53
5 RESULTADOS	54
5.1 FILTRO DE MODA	54
5.2 MULTIREOLUTION	55
5.3 WATERSHED	56
5.4 U-NET	58
5.5 CONTAGEM DOS SEGMENTOS: ÍNDICE DE SØRENSEN–DICE E ÍNDICE DE JACCARD	59
6 DISCUSSÃO	62
6.1 IMAGENS DO RPA	62
6.2 FILTRO DE MODA	63
6.3 MULTIREOLUTION	64
6.4 WATERSHED	65
6.5 U NET	68

7 CONCLUSÃO	71
REFERÊNCIAS	73

1 INTRODUÇÃO

As florestas do Brasil sofreram alterações significativas durante o processo histórico de ocupação antrópica, evidenciadas pela enorme redução de áreas de mata atlântica (ROSS, 2006; NEVES, 2006) e pela expansão agropecuária, especialmente no Cerrado (ROCHA *et al.*, 2012; SILVA *et al.*, 2014; SILVA e MIZARRA, 2011) e na Amazônia (FEARNSIDE, 2004; SOARES-FILHO *et al.*, 2005).

A crescente demanda por recursos naturais e a degradação dos ambientes, podem levar à redução ou esgotamento desses bens (HASSLER, 2005) e à alteração dos serviços ecossistêmicos - que auxiliam na manutenção da vida e, quando degradados, trazem consequências extremamente graves para a sociedade (CAMPANILI e SCHÄFFER, 2010), como a redução da quantidade de água potável, estoques de pescado, impacto na regulação do clima e etc. - reforçando a necessidade de estratégias para viabilizar a conservação e garantir a manutenção destes recursos e serviços.

Uma ferramenta muito útil para auxiliar na conservação dos recursos naturais e do ambiente é o monitoramento em macro escala, sendo crucial para acompanhar as mudanças na vegetação e auxiliar na elaboração de estratégias de conservação, pois evidencia as alterações ambientais causadas por ações antrópicas, sejam elas positivas ou negativas para o ecossistema (DILLENBURG, 2007; KRONKA *et al.*, 2005; MORAES *et al.*, 2018). Especificamente, o monitoramento ambiental em macro escala permite identificar áreas de desmatamento, acompanhar o desenvolvimento de indivíduos plantados, seja para o agronegócio ou em projetos de restauração, verificar a implementação de planos de manejo e, por fim, o monitoramento também fornece dados importantes para cientistas, subsidia a elaboração de políticas e facilita o entendimento da sociedade acerca da disponibilidade dos recursos naturais (LOVETT *et al.*, 2004; BARRETT e CUTIS, 2013).

Outra abordagem importante para a conservação dos recursos naturais está no manejo sustentável de florestas, o qual possibilita a extração controlada de madeira, equilibrando interesses econômicos quando executado de forma eficiente (Beckline *et al.*, 2017).

Nesse contexto, o sensoriamento remoto apresenta-se como importante ferramenta para realizar o monitoramento do ambiente, pois sua capacidade de cobertura e amostragem temporal facilita a coleta de dados em escala regional e global (ZHANG *et al.*, 2003; MYNEMI, 1997; KRONKA, 2005; MORAES *et al.*, 2018).

O sensoriamento remoto pode ser utilizado a partir de diversas plataformas: satélites, as aeronaves, as RPAs (Aeronaves Remotamente Pilotadas), os SMMs (Sistemas Móveis de Mapeamento) e as plataformas estáticas (TOTH e JÓZKÓW, 2016).

Entre as plataformas citadas, os satélites são amplamente utilizados para os mais variados estudos em ciências da terra, fauna e flora, recursos hídricos e uso e ocupação dos solos (CAMPBELL e WYNNE, 2011; DIAZ-DELGADO *et al.*, 2017). É necessário ressaltar a importância de dados de alta resolução espacial e com baixo tempo de revisita para as análises de dados ecológicos, entretanto sua aquisição pode representar um alto investimento, pois esses dados normalmente não estão disponíveis em fontes gratuitas de imagens de satélites (NAGENDRA e ROCCHINI, 2008; SEELAN, 2003; GRÜN, 2000).

Como alternativa destacam-se os RPAs que, assim como os satélites, podem ser utilizados para aquisição de dados de alta resolução confiáveis, porém de maneira mais econômica (CALVARIO, 2017). Seu uso tem ganhado espaço no sensoriamento remoto nos últimos anos, pois é de interesse comum à ciência e aplicações em geral (TOTH e JÓZKÓW, 2016). A facilidade que essa plataforma traz para a aquisição de dados apresenta inúmeras possibilidades, como na agricultura de precisão (JORGE e INAMASU, 2014), no monitoramento de vegetação (AASEN *et al.*, 2015), inspeção de florestas (FARIA e COSTA, 2015), resposta à desastres (DANIEL *et al.*, 2009), mapeamento de território (FUYI *et al.*, 2012), modelos tridimensionais gerados por fotogrametria (LISEIN *et al.*, 2013), entre outros.

Entre as possibilidades, vale destacar o potencial dos RPAs para o monitoramento de florestas. Tang e Shao (2015) compilou algumas dessas oportunidades de estudo, evidenciando: pesquisar florestas, encontrar e mapear clareiras entre as copas, mensurar a altura das copas das árvores em florestas, encontrar vida selvagem e servir de apoio em floresta de manejo intenso.

No entanto, para obter informações assertivas acerca das imagens adquiridas por sensoriamento remoto (Satélites ou RPAs), há sempre a necessidade de realizar procedimentos para corrigir as informações e tratá-las para que o resultado adequado seja alcançado. Especialmente para RPAs, o processamento das imagens permite verificar se os dados calibrados (pré-processados) são de qualidade e, assim, podem ser utilizados (KOC-SAN *et al.*, 2018). Em geral, técnicas são utilizadas para gerar produtos intermediários ou finais a partir das imagens do sensoriamento, como é o

caso da filtragem (KOC-SAN *et al.*, 2018) ou o processamento digital de imagem (CALVARIO, 2017; KOC-SAN *et al.*, 2018; JENSEN, 2015).

Essa última técnica vem sendo amplamente utilizada por pesquisadores para a visualização e testes de hipótese, de modo que os tipos mais comuns de processamento de imagem são: correlação radiométrica e geométrica, aprimoramento de imagem, fotogrametria (LISEIN, 2013; CSILLIK *et al.*, 2018), extração paramétrica e não paramétrica de informação e a análise de dados por inteligência artificial (I.A.) (JENSEN, 2015). Essa manipulação dos dados possibilita a criação de novas informações, viabilizando novas interpretações do conjunto de dados e facilitando seu processo de classificação.

A constante melhoria dos sensores e o aumento da qualidade e da frequência da captura de dados, somados com a crescente demanda por informação tratada e resultados rápidos, torna proeminente utilizar meios que excedam a capacidade humana para a interpretação dos dados. Uma alternativa está no uso de inteligência artificial (I.A.), que pode abrir caminho para análises complexas de dados em tempo viável, uma vez que há constante otimizações de algoritmos acompanhadas de elevado crescimento do poder computacional. Um ramo da I.A. que pode ser considerado uma ferramenta importante para processar grandes quantidades de informações de dados de sensoriamento remoto é o chamado aprendizado de máquina (*Machine Learning* - ML) (SEGARAN 2007, ZHANG *et al.* 2016).

Na literatura, há diversas ocorrências do uso de técnicas de processamento de imagem e aprendizado de máquina em dados gerados por sensoriamento remoto com as mais diversas finalidades, sendo possível citar a contagem de árvores (coqueiros) em imagens de alta resolução (APARNA *et al.*, 2018); monitoramento ambiental a partir de classes geradas a partir do processamento nas redes neurais (THOMAZELLA *et al.*, 2018); contagem de árvores na citricultura (CSILLIK *et al.*, 2018); controle de plantação de agave no México (CALVARIO, 2017); e detecção e extração de copas de árvores em áreas urbanas (BHAVANA *et al.*, 2018).

A combinação de processamento de imagem e aprendizado de máquina abre espaço para o desenvolvimento de novas técnicas de monitoramento ambiental. Questões como mudanças na estrutura da vegetação podem ser observadas em tempo reduzido, viabilizando respostas mais rápidas e precisas à problemas ambientais (ZHANG, 2014). É evidente o potencial em utilizar técnicas de processamento de imagem e aprendizado de máquinas em imagens de alta resolução

para a extração de informações. Essa característica, somada a capacidade de aquisição de dados com incomparável resolução espacial e temporal e a baixo custo dos RPAs, pode resultar em um conjunto de técnicas capazes de viabilizar pesquisas científicas, monitoramentos de áreas naturais, análises da arborização urbana e elaborar novas estratégias para a agricultura de precisão.

É possível citar exemplos como o de Zhang (2014), que discute como as mudanças na estrutura da vegetação podem ser observadas em tempo reduzido, viabilizando respostas mais rápidas e precisas à problemas ambientais; o de Jorge e Inamasu (2014), abordando a aplicabilidade do RPA na agricultura de precisão e; como o de Thomazella et al.(2018), que trata sobre o monitoramento de áreas em que a ocupação está sujeita à restrições legais; são fundamentais para compreender o potencial existente para o aprimoramento e uso de técnicas mais modernas no processamento de imagens de alta resolução usando RPAs.

Além dessas pontuações, a possibilidade de utilizar informações tridimensionais derivadas da fotogrametria das imagens obtidas por RPAs, e de instalar equipamentos como LiDAR e câmeras hiperspectrais nessas aeronaves, as tornam uma das técnicas mais baratas e promissoras para o monitoramento de pequenas áreas de restauro florestal, permitindo o monitoramento da estrutura de copas de árvores e estimação de biomassa a partir de sistemas mais baratos (ALMEIDA, 2019; ZAHAWI et al, 2015).

Amparado nessas novas perspectivas para a captura e processamento de informações, é possível pensar em novas abordagens para a aquisição de dados de áreas verdes arborizadas, sejam elas naturais ou artificiais. Uma das abordagens para o levantamento de informações em florestas está na contagem do número de árvores que aparecem em imagens de sensoriamento remoto (DI LEO, 2017). Assim, mudanças naturais ou antrópicas poderiam ser identificadas e localizadas com precisão, permitindo investigar as origens dessa alteração e, caso necessário, tomar ações rápidas (GOMES e FERREIRA, 2007).

Pensando na contagem de indivíduos arbóreos, diferentes técnicas podem ser aplicadas para a sua execução. Um dos métodos mais comuns é a segmentação, que permite a extração de objetos comuns na imagem, ou seja, identificação individual de cada copa de árvore nesse caso. Esse procedimento caracteriza-se por particionar a imagem em um conjunto de regiões de interesse (SHAPIRO e STOCKMAN, 2001).

Um meio para realizar a segmentação das copas das árvores está na utilização de algoritmos para identificar e extrair árvores únicas do dossel. Existem diversas abordagens na literatura com nível variável de complexidade (e qualidade dos resultados), tais como: homogeneização de pixels ou redução de redundância (GOMES E FERREIRA, 2007); análise de imagem em várias escalas para detecção de similaridade (KE E QUACKENBUSH, 2007, ONISHI E ISE, 2018); segmentação suportada por regras fuzzy e movimento browniano (ERIKSON, 2004); processamento e adaptação de imagens para uso com ferramentas gerais de segmentação, como segmentação de bacias hidrográficas (DI LEO 2017, JING et al. 2012, NORDIN et al. 2019); e o uso de aprendizado profundo, especificamente Redes Neurais Convolucionais (CSILLIK *et al.* 2018, LI et al. 2016, APARNA et al. 2018).

Porém, independentemente do método utilizado, é necessário assegurar qualidade no processo de segmentação e que o mesmo ocorra com o maior nível de acerto possível, pois além de impactar na contagem dos indivíduos, pode aumentar os erros encontrados em procedimentos de classificação dos elementos detectados. Além disso, cada método de extração de características (independentemente do tipo e complexidade) possui limitações e dificuldades que aparecem quando um algoritmo é aplicado em diferentes tipos de florestas (plantação ou natural). Sabe-se, por exemplo, que as propriedades ópticas da floresta para o sensoriamento remoto dependem de vários fatores externos e internos com limitações de detecção variáveis (GUYOT et al. 1989).

Dessa forma, esse estudo pretende avançar no entendimento e aplicação de técnicas de segmentação e classificação para o delineamento das copas das árvores dentro do dossel por meio de imagens geradas por RPA. Tomando como base diferentes tipos de vegetação (e dosséis), o estudo busca evidenciar o potencial e as limitações associadas com os métodos automáticos existentes em diferentes tipos de cobertura. Também será verificado a complexidade no uso dessas ferramentas, visto a variedade de profissionais que atuam na área ambiental e que não necessariamente possuem conhecimentos aprofundados em processamento de imagens. Espera-se que os resultados auxiliem na resposta quanto a viabilidade do uso dessas ferramentas e para qual direção seria possível caminhar com o aprimoramento das técnicas aqui usadas.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Diante do exposto, a pesquisa tem como objetivo avaliar a facilidade de uso e qualidade do resultado de quatro metodologias de segmentação para a detecção e contagem de copas de árvores, em imagens georreferenciadas adquiridas por RPAs para diferentes tipos de vegetação, buscando evidenciar a mais adequada para as condições das áreas estudadas.

2.2 OBJETIVOS ESPECÍFICOS

- Analisar a aplicabilidade de imagens de RPAs para a detecção e contagem de indivíduos arbóreos;
- Verificar o desempenho das técnicas de segmentação em duas áreas com vegetação distintas;
- Verificar as diferenças entre os resultados obtidos e apontar os principais problemas, limitações e qualidades de cada uma das técnicas aplicadas;
- Sugerir pesquisa e aprimoramento futuro das técnicas utilizadas.

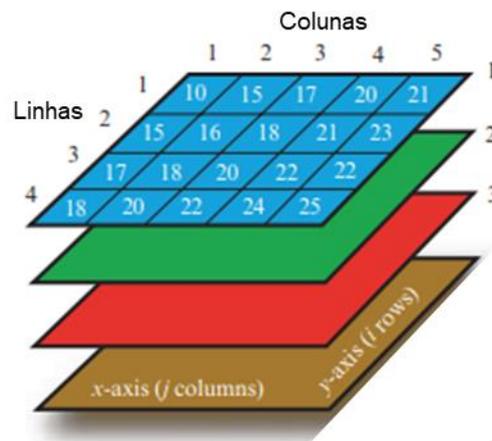
3 EMBASAMENTO TEÓRICO

Antes de aplicar as técnicas presentes nessa pesquisa, foi realizada uma revisão dos conceitos básicos que nortearam a implementação e análise das ferramentas utilizadas. Os conceitos descritos integram o geoprocessamento, tópico que faz parte da formação de profissionais na área ambiental (DEPARTAMENTO DE CIÊNCIAS AMBIENTAIS, 2018; ESCOLA DE ENGENHARIA DE SÃO CARLOS, 2017), mas também traz informações de áreas especializadas, sem relação direta com questões ambientais (DEPARTAMENTO DE COMPUTAÇÃO, 2019).

3.1 IMAGENS DIGITAIS

As imagens digitais são representadas em formato de matriz de informações individuais, denominada pixel (CAMPBELL e WYNNE, 2011). Essas informações são geradas a partir da variação de intensidade de emissão (seja por reflexão ou outra forma) de radiação eletromagnética de uma superfície captada por um determinado sensor (FITZ, 2008; CAMPBELL e WYNNE, 2011; JENSEN, 2015). Essa variação é alocada em uma matriz representa a superfície em tons de cinza e, quando sobrepostas as matrizes (Figura 1) que constituem as faixas do vermelho, verde e azul, resulta numa imagem (raster) colorida convencional (JENSEN, 2015).

Figura 1 – Sobreposição de Matrizes



Fonte: adaptado JENSEN, 2015

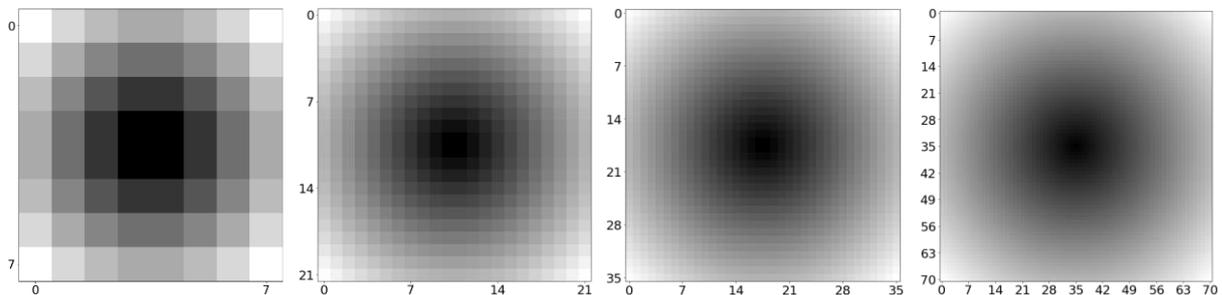
O nível de detalhes com a qual a variação de intensidade consegue ser detectada e armazenada depende da capacidade do sensor em captar, armazenar e distribuir os dados resultantes da superfície. Essa capacidade está relacionada com

as diferentes resoluções (nível de detalhe) que a imagem pode ter e divide-se em quatro tipos: espacial, radiométrica, espectral e temporal.

A resolução espacial é a medida da menor separação linear ou angular entre dois objetos que pode ser resolvida por sensoriamento remoto (JENSEN, 2015), ou seja, define o quanto um pixel representa da imagem capturada pelo sensor. Assim, dada uma sequência de imagens de um mesmo alvo, mas com resoluções espaciais diferentes, a que possuir menor tamanho de pixel, apresentará maior nível de detalhes (Figura 2).

Outra consideração é que o aumento da resolução, ou seja, maior número de pixels, resulta em um aumento na quantidade de informação que será armazenada e processada, causando impactos na performance em algoritmos. Soma-se que uma maior resolução não necessariamente atende às necessidades do usuário para uma determinada tarefa. Portanto, a escolha da resolução espacial é um fator que deve ser pensado com antecedência considerando os usos da imagem.

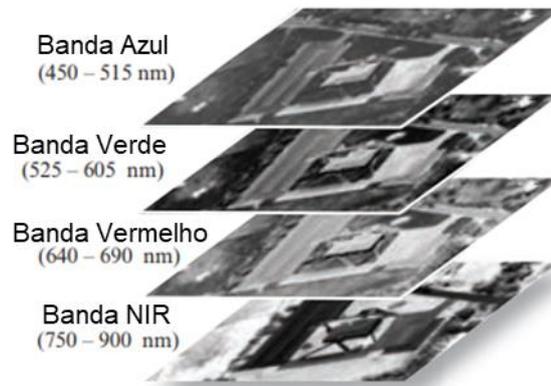
Figura 2 – Exemplo de Variação da Resolução Espacial



Fonte: elaborado pelo autor

A resolução espectral denota a capacidade do sensor de capturar faixas do espectro eletromagnético (CAMPBELL e WYNNE, 2011) e transformar em uma imagem digital. Comumente na fotografia as imagens capturadas são formadas por 3 faixas do espectro eletromagnético: o vermelho, o azul e o verde. No entanto, no sensoriamento remoto é comum capturar faixas além do espectro visível, como o infravermelho próximo (Figura 3), que pode auxiliar na composição de imagens para o mapeamento de vegetação (FITZ, 2008). Além disso, há aplicações que fazem uso de sensores hiperespectrais, como por exemplo o método usado por Nordin et al. (2019), visando capturando dezenas a centenas de faixas dentro do espectro e aplicando em um processo de segmentação de copas de árvores.

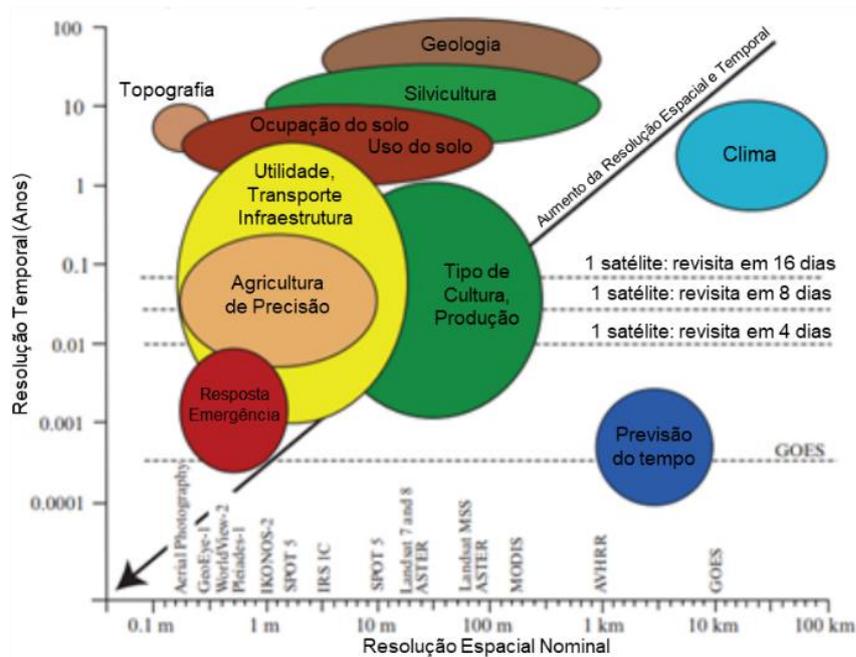
Figura 3 – Bandas Individuais de Imagem Multiespectral



Fonte: JENSEN (2015)

A resolução temporal está relacionada ao tempo existente entre a captura de duas imagens do mesmo alvo e, no caso do sensoriamento remoto, deriva do tempo de revisita de um satélite (FITZ, 2008; JENSEN, 2015). Assim, quanto menor for o tempo entre essas capturas, maior será a resolução temporal. Em algumas aplicações, uma resolução temporal alta pode ser fundamental para auxiliar na tomada de decisão, por exemplo, o monitoramento de desastres ambientais como queimadas (Figura 4).

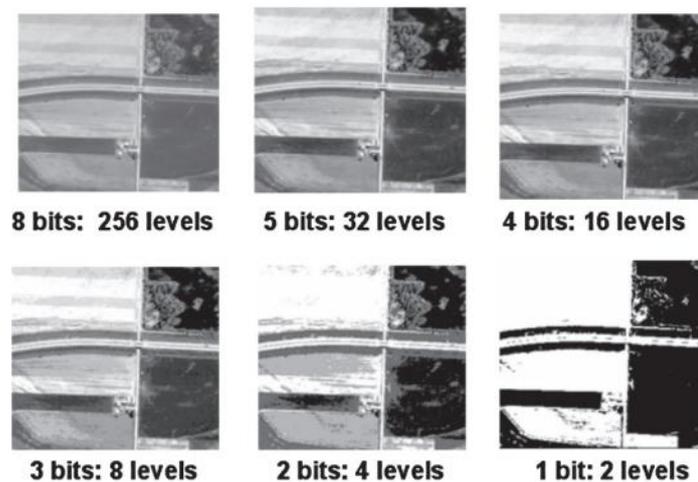
Figura 4 – Resolução Espacial e Temporal para Aplicações Específicas



Fonte: JENSEN (2015)

A resolução radiométrica do sensor é a capacidade do sensor detectar e armazenar a variação de intensidade da energia emitida pelo alvo (Figura 5). É comum que as imagens digitais tenham cada pixel com 8 bits e, conseqüentemente, valores de intensidade que vão de 0 a 255, no entanto, é possível trabalhar com valores maiores ou menores.

Figura 5 – Exemplo de Resolução Radiométrica



Fonte: CAMPBELL e WYNNE (2011)

Essas resoluções impactam diretamente no volume de informação que uma determinada imagem carrega. Por mais que pareça interessante trabalhar com os produtos mais detalhados do mercado, há uma série de limitações técnicas quanto a capacidade de gerar esse tipo de conteúdo e, no final, interpretá-lo. A resposta para um problema não necessariamente estará no material mais detalhado, pois seu preço pode ser muito alto e não atender as expectativas do trabalho além de trazer dificuldades para o desenvolvimento do projeto.

3.2 PROCESSAMENTO DE IMAGEM

Há algoritmos que permitem, para uma determinada finalidade, modificar as informações presentes nas imagens digitais a partir de operações nos valores dos pixels que a compõe. O nome dado para esse procedimento é Processamento Digital de Imagem (PDI) e aparece como uma ferramenta importante para auxiliar a visualização e modificação dos atributos das imagens (CAMPBELL e WYNNE, 2011). O PDI faz uso do poder computacional e de conceitos matemáticos, facilitando a

análise de imagens, uma vez que algumas características podem fugir à nossa percepção quando analisados diretamente (SZELISK, 2011).

Evidentemente que quando se fala em poder computacional, é necessário refletir sobre o aumento do poder de processamento disponível às pessoas nos últimos anos. Fixado um mesmo algoritmo, o tempo necessário, para processar essas informações, é reduzido substancialmente. Além disso, conceitos matemáticos já conhecidos estão se tornando cada vez mais fáceis de serem implementados, por conta de linguagens de programação mais intuitivas e da crescente demanda pela integração do conhecimento computacional com outras áreas do conhecimento.

Os principais exemplos de tipos de PDI incluem processamento de imagem (correção radiométrica e geométrica), aprimoramento de imagem, operação morfológica, fotogrametria e processamento de imagem estereoscópica, segmentação, reconhecimento, sistema especialista (e.g. árvores de decisão) e análise de imagem por rede neural e análise de dados hiperspectrais (JENSEN, 2015; SZELISK, 2011).

O PDI se beneficia amplamente dessas novas facilidades, com exemplos que vão de uma série de algoritmos sendo executados automaticamente em aplicativos de celular até ferramentas integradas a softwares de uso profissional e acadêmico. Ao direcionar essa aplicação para a área de geoprocessamento, encontra-se um conjunto de softwares que permitem a análise de imagens digitais coletadas por sensores remotos (CAMPBELL e WYNNE, 2011), como por exemplo: ERDAS ER Mapper (ERDAS), EASI/PACE, ENVI, ERDAS Imagine (parte do Hexagon Group, Suécia), GRASS GIS, IDRISI, Trimble ECOGNITION etc.

Mesmo com essa variedade de softwares, os usuários precisam de ferramentas para modificar as imagens que não estão implementadas. Visando um resultado específico, é possível implementar novas ferramentas que atendam a necessidade do usuário, caso o software de suporte e ofereça portabilidade das informações

Um exemplo de recurso que pode permitir a implementação de ferramentas está na integração da linguagem de programação Python com diversos softwares de processamento de imagem (TATEOSIAN, 2015; TOMS, 2015). Essa linguagem é considerada de fácil aprendizado, possui grande número de usuários e um conjunto de bibliotecas com diversas ferramentas que facilitam o seu uso para o processamento de imagem (TATEOSIAN, 2015). Além disso, é comum existir implementações de ferramentas de PDI em Python, tornando mais simples a

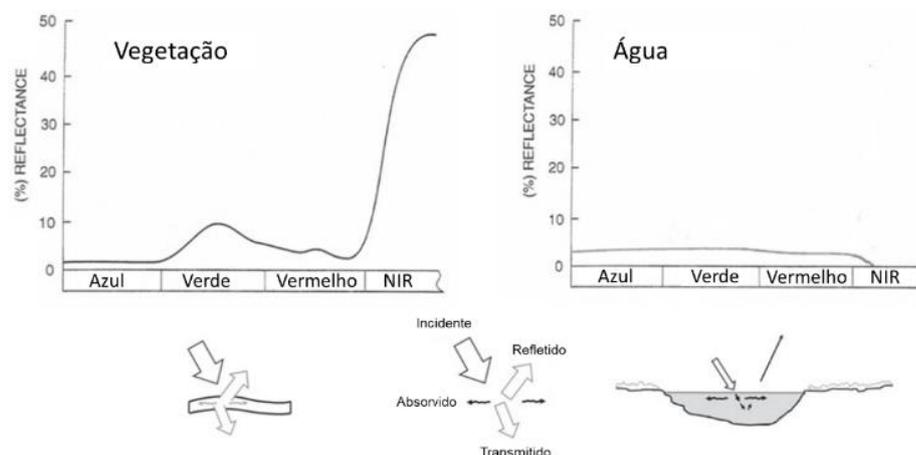
execução de algumas rotinas dentro da interface da própria linguagem. No entanto, é necessário ressaltar a adaptação do usuário pois, por mais que a simplificada que uma linguagem seja, especializar uma ferramenta pode dificultar o acesso a usuários que não possuem maior experiência com lógica de programação.

Independente da escolha do usuário quanto à programação de rotinas para o processamento de imagem ou o uso direto em um software, é necessário refletir na existência de conceitos que estão presentes no tratamento de imagens georreferenciadas. Uma série de algoritmos do PDI são de uso comum no sensoriamento remoto, pois problemas como a necessidade de reduzir ruídos ou realçar detalhes e elementos da imagem são recorrentes.

Entretanto, algumas ferramentas do PDI foram projetadas de forma que imagens com três ou mais bandas (tridimensionais) não são suportadas, sendo necessário combiná-las e simplificá-las em uma matriz bidimensional em tons de cinza ou binária. Isso ocorre, pois, imagens em tons de cinza são boas o suficiente para serem utilizadas em uma série de tarefas, não sendo necessário o uso de imagens coloridas que tornariam o processo mais difícil e custoso (FISHER *et al.*, 2003).

Vale ressaltar que, apesar de ser possível processar apenas um canal de uma imagem, é importante ponderar sobre a variação no comportamento espectral dos alvos. Uma mesma superfície possui assinaturas espectrais únicas (Figura 6) e essas diferenças podem ser úteis no processamento da imagem. Além disso, a diferença no comportamento espectral entre alvos diferentes também tem um papel importante, pois auxilia processos de identificação, e segmentação, dos mesmos.

Figura 6 – Diferença no Comportamento Espectral de Alvos Distintos



Fonte: adaptado CAMPBELL e WYNNE (2011)

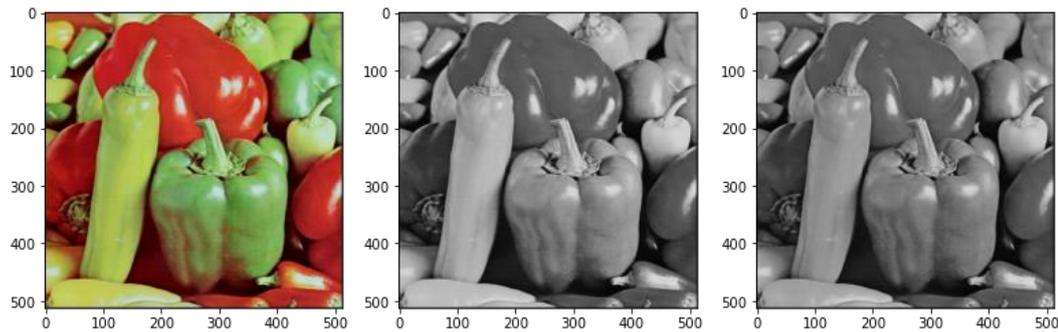
Assim é fundamental avaliar os impactos da perda de informações derivadas dos canais descartados, frente ao uso de uma abordagem que requer uma imagem em tons de cinza. Também deve ser considerado como equilibrar a presença de cada banda na nova imagem, pois o olho humano percebe as cores de formas diferentes e essa percepção pode ser considerada na análise.

Um método comumente utilizado para transformar imagens coloridas com três canais (RGB) para tons de cinza é a ponderação de cada canal visando equilibrar a luminosidade na imagem (rec601 luma):

$$Y'_{601} = 0.299R' + 0.587G' + 0.114B'$$

Diferente de uma média que não equilibra as características de luminosidade, a imagem resultante mantém suas características originais (Figura 7).

Figura 7 – Comparação Entre Tons de Cinza por Média e rec601 luma



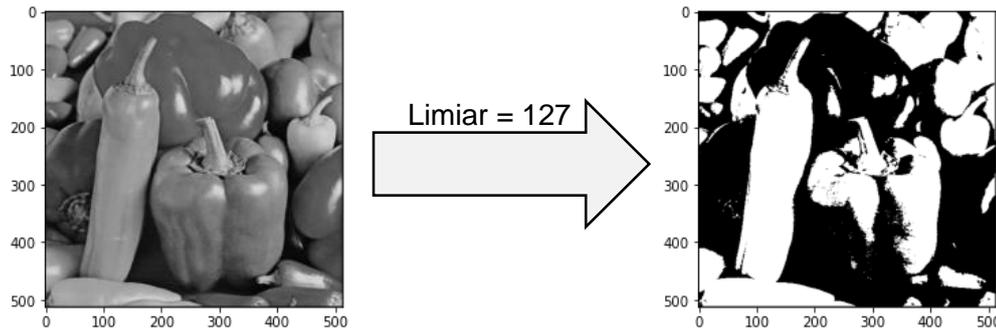
Apesar da semelhança, há uma sutil variação de iluminação. A imagem ao centro possui Média = 107 e Desvio Padrão = 46. Já a imagem da direita possui Média = 115 e Desvio Padrão = 53. Fonte: Autor

Outro alvo comum de ferramentas de PDI, como os filtros não lineares e as operações morfológicas, é a imagem binária (SZELISK, 2011). Para converter uma imagem em tons de cinza em binária é necessário utilizar uma função para traçar um limiar dividindo os valores dos pixels em dois grupos (Figura 8).

$$\theta(f, t) = \begin{cases} 1 & \text{se } f \geq t, \\ 0 & \text{caso contrário} \end{cases}$$

Esse produto resalta ou reduz atributos da imagem, possibilitando ao operador aplicar algumas ferramentas, como segmentação e operações morfológicas, para modificar esses conjuntos.

Figura 8 – Conversão para Tons de Cinza



Fonte: elaborado pelo autor

Independente da qualidade ou da complexidade da imagem, o PDI é uma abordagem para alterar as suas características e, para isso, faz-se o uso de ferramentas com um amplo contexto de aplicação, pois são utilizadas em diversas áreas do conhecimento.

3.2.1 Análise de Componentes Principais

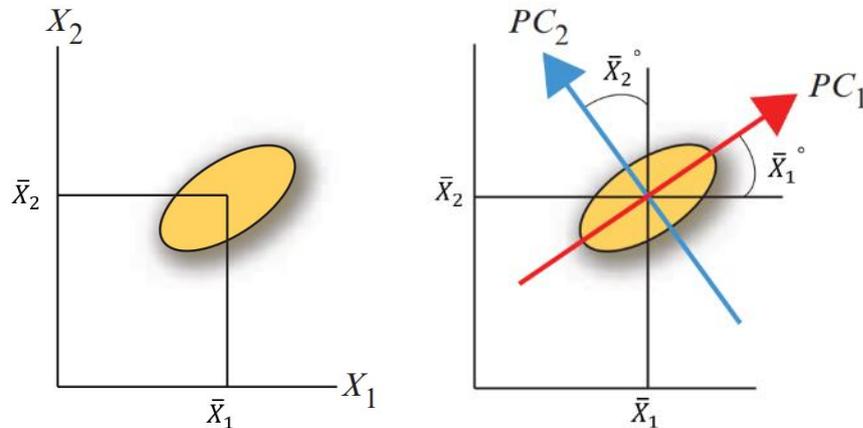
A Análise de Componentes Principais (*Principal Components Analysis* - PCA) é uma estratégia que identifica a combinação linear ótima dos canais, evidenciando as principais fontes de variação dos valores dos pixels na imagem (CAMPBELL e WYNNE, 2011). Em outras palavras, o seu uso está atrelado à redução de redundância do conjunto de dados a fim de indicar os componentes que melhor explicam a variação dos dados dos pixels.

Tem seu uso relacionado à redução de dimensionalidade (número de canais/bandas de um conjunto de dados), trazendo economia de processamento futuro caso o produto tenha dados tão bons quanto a imagem original (JENSEN, 2015). Assim, um exemplo de aplicação do PCA está no seu uso para a redução de dimensionalidades de conjunto de dados hiperespectrais (NORDIN et al., 2019; JENSEN, 2015).

O seu funcionamento pode ser melhor explicado ao considerar duas bandas, X_1 e X_2 , como eixos de um gráfico onde seus dados se encontram distribuídos e com média rotuladas como \bar{X}_1 e \bar{X}_2 (Figura 9). A partir dessa apresentação, os componentes principais são diagonais que cortam os dados, calculadas em função de eixos perpendiculares às médias. Dessa forma, o primeiro componente principal (PC₁) tende a explicar mais de 90% da variância dos dados, o segundo componente (PC₂) mais

uma parte, em torno de 2 a 10% dos dados restantes, e assim sucessivamente para os outros componentes principais, caso ocorra (JENSEN, 2015).

Figura 9 – Representação 2D do PCA



Fonte: adaptado JENSEN (2015)

3.2.2 Filtros

Os filtros, em PDI, são técnicas que modificam detalhes e atributos das imagens, podendo realçar bordas, suavizar a imagem ou remover ruídos sem alterar a sua média (SANTOS *et. al.*, 2010, SZELISK, 2011). Para isso, utiliza-se funções lineares que alteram os valores de uma imagem para gerar um resultado desejado. Um exemplo é o uso de uma função de soma ponderada em um conjunto de pixels de entrada, onde os pesos estão organizados em uma máscara (ou kernel).

O processo de usar uma kernel para filtrar vizinhos, ou seja, modificar os valores próximos ao pixel central, é denominado convolução (SZELISK, 2011). Alguns exemplos de filtros lineares são: filtro de passa-faixa (e.g. Sobel), passa-baixa, gaussiano e etc. (Figura 10).

Figura 10 – Exemplos de Filtros Lineares



Fonte: adaptado SZELISK (2011)

Há também filtros não lineares, em que suas saídas não são funções lineares das suas entradas. No caso, esses filtros minimizam/realçam ruídos e

suavizam/realçam bordas, alterando a média da imagem (SANTOS *et. al.*, 2010). Um exemplo simples é o filtro mediano (Figura 11), que substitui cada entrada pela mediana das entradas vizinhas.

Figura 11 – Exemplo de aplicação do Filtro de Mediana



Fonte: adaptado Nixon e Aguado (2012)

Além dos filtros não lineares serem aplicáveis em imagens em tons de cinza e coloridas, eles são extensivamente usados para processar imagens binárias, sendo as operações binárias mais comuns chamadas operações morfológicas (SZELISK, 2011).

Segundo Di Leo (2017), a ideia por trás da morfologia matemática (que origina as operações morfológicas) é examinar e modificar a imagem utilizando um elemento estruturante (kernel). A modificação irá variar de acordo com o tipo de operador morfológico utilizado, sendo esses: dilatação, erosão e seus derivados (Quadro 1).

A operação morfológica pode ser utilizada em conjuntos binários, pois a sua lógica é facilmente aplicável sob essas condições (DI LEO, 2017; SZELISK, 2011). No entanto, é possível trabalhar com conjuntos gradientes, constituindo assim as operações morfológicas em tons de cinza.

Para que essa transição funcione, a imagem deve ser interpretada como um gradiente de relevo, onde os valores dos pixels serão tratados como a altura relativa que aquele ponto possui (DI LEO, 2017). O nome dado para essa interpretação é Umbra, e essa abordagem permite que as operações morfológicas tratem a oscilação do pico dos valores como um sinal, dilatando ou erodindo sua superfície (Sternberg, 1986).

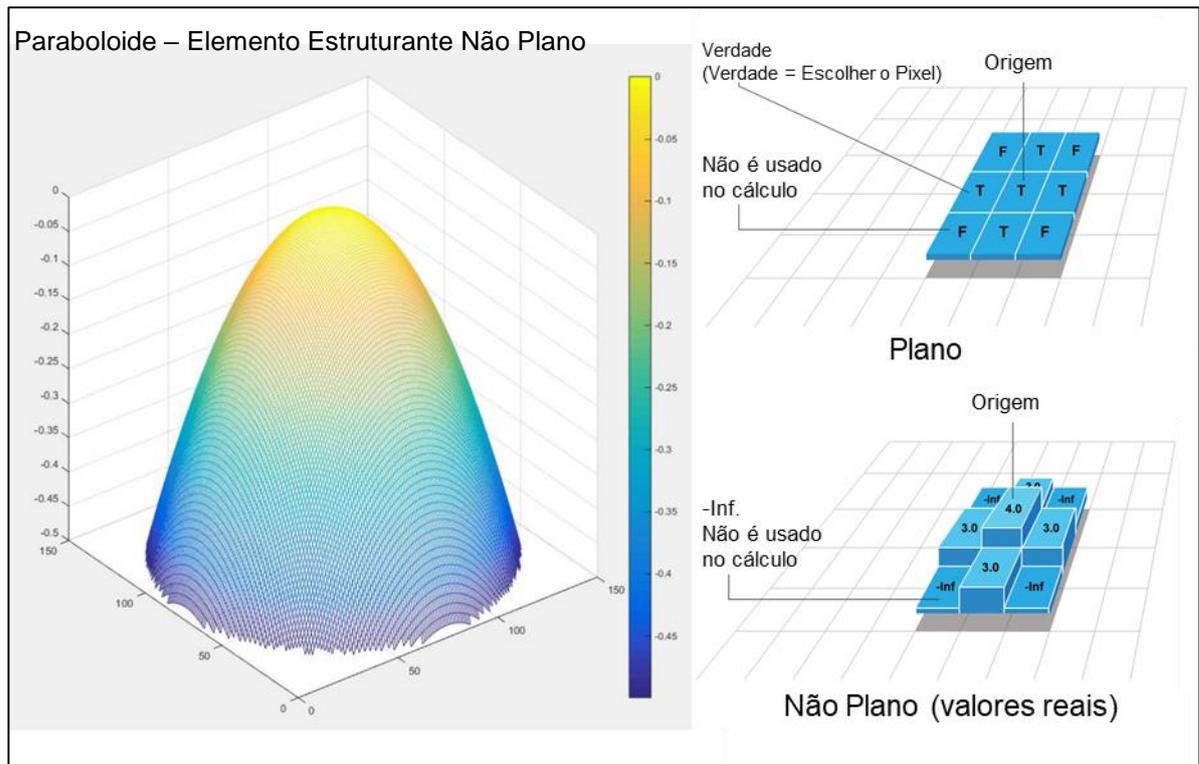
Quadro 1 – Exemplos de Operações Morfológicas

Operação morfológica	Notação	Exemplos
Imagem Original	Kernel quadrada de 5x5	
Dilatação	$I(x,y) \oplus S = \begin{cases} 1 & \text{se } S \text{ hits } I \\ 0 & \text{caso contrário} \end{cases}$	
Erosão	$I(x,y) \ominus S = \begin{cases} 1 & \text{se } S \text{ fits } I \\ 0 & \text{caso contrário} \end{cases}$	
Abertura	$I \circ S = (I \ominus S) \oplus S$	
Fechamento	$I \cdot S = (I \oplus S) \ominus S$	

Fonte: adaptado Szelisk (2011) e Di Leo (2017)

Os elementos estruturantes também se adequam ao gradiente, sendo classificados em dois tipos: planos e não planos (DI LEO, 2017). Um exemplo de elemento estruturante plano é um disco, constituindo um cilindro que modifica a Umbra em função da sua área e altura. Já os não planos possuem variação gradiente dos valores que compõe sua kernel, formando, por exemplo, um parabolóide (Figura 12).

Figura 12- Elementos Estruturantes em Tons de Cinza



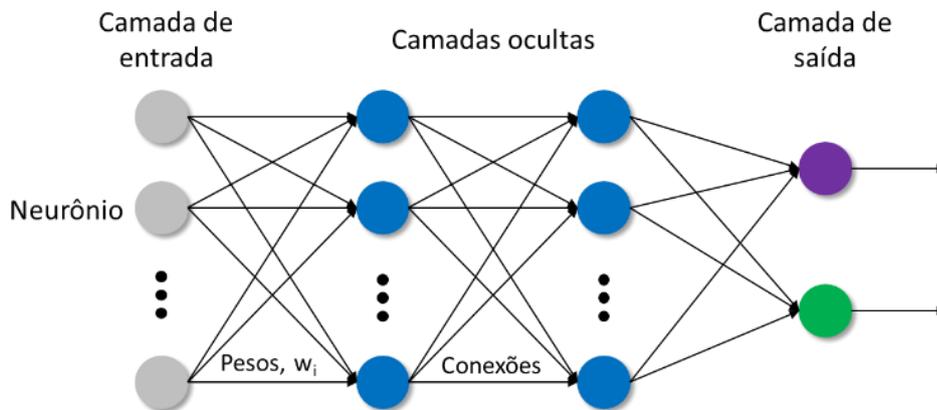
Fonte: adaptado Di Leo (2017) e MathWorks (2020)

3.3 REDES NEURAIS

As redes neurais são estruturas de neurônios que tentam simular o processo de aprendizagem humano através de conexões formadas entre os dados de entrada e saída (CAMPBELL e WYNNE, 2011), podendo resolver, por exemplo, problemas de segmentação e classificação de imagem (LI et al, 2016; LÄNGKVIST *et al.*, 2016). O aprendizado da rede neural não ocorre como uma sequência de execução de um determinado script, mas sim a partir de ajustes dos pesos que conectam os neurônios (WOLFGANG, 2017).

As redes neurais são constituídas na sua forma mais básica por três elementos (Figura 13): Uma camada de entrada que recebe a imagem; uma camada de saída formada por classes que serão analisadas e comparadas pela inteligência artificial; e uma (ou mais) camadas ocultas, que serão responsáveis por criar a associação da imagem da entrada com as classes na saída (CAMPBELL e WYNNE, 2011; JENSEN, 2015).

Figura 13 – Exemplo de Rede Neural



Fonte: elaborado pelo autor

O uso de neurônios na camada oculta permite a simulação de padrões não lineares nos dados, sendo que a quantidade de neurônios e conexões nessas camadas são questionamentos básicos para o aprimoramento da rede neural (JENSEN, 2015). Nos casos onde há um grande aumento ordenado da quantidade de neurônios e conexões na camada oculta, a terminologia aprendizado da máquina passa a ser aprendizagem profunda (Deep Learning).

É necessário treinar a rede neural para que o processo de aprendizagem do modelo aconteça, possibilitando extrair informações úteis, como classes de um uso e ocupação do solo. O treinamento é constituído pela entrada de um conjunto de dados que possuem um par já classificado, tornando necessário um preparo prévio de um banco de dados. Essa máscara, que é o dado classificado, é fundamental para que a rede neural verifique seus resultados e erros, ajustando os seus pesos até atingir um valor cada vez melhor de previsão.

Segundo Jensen (2015), uma vantagem da rede neural é a capacidade da mesma aprender a ignorar ruídos e erros presentes nos dados utilizados no treinamento. Assim, mesmo que os operadores sejam fundamentais para a escolha e amostragem dos dados, a rede conseguirá excluir vieses pessoais do processo de treinamento.

As redes neurais apresentaram boa performance na classificação de dados de sensoriamento remoto, no entanto, melhoras na precisão costumam ser pequenas, sendo necessário um grande esforço para o treinamento dos dados e um alto consumo de tempo computacional (CAMPBELL e WYNNE, 2011).

Um grande problema existente nas redes neurais é a dificuldade em encontrar uma explicação compreensiva para os resultados gerados pelo modelo, pois os ajustes do modelo estão atrelados aos pesos dos neurônios nas camadas ocultas (JENSEN, 2015). Por essa razão as redes neurais são consideradas uma caixa preta (QIU e JENSEN, 2004).

Portanto, é necessário muito cuidado com o uso dos resultados da rede neural em situações-problema críticas, pois a falta de capacidade de explicação do modelo adiciona um fator de imprevisibilidade não aceitável em alguns casos (CAMPBELL e WYNNE, 2011).

3.4 SEGMENTAÇÃO

Segmentação é o processo de identificação de traços que subdividem a imagem em regiões com valores espectrais de pixels parecidos ou dentro de uma tolerância (JENSEN, 2015). Sua aplicação em muitas tarefas é representar regiões comuns e relevantes em uma imagem, como plantações, áreas urbanas e floretas em uma imagem de satélite (SHAPIRO e STOCKMAN, 2001).

A segmentação tem dois objetivos (SHAPIRO e STOCKMAN, 2001): O primeiro é decompor a imagem em partes para análise futura. Simplificando, o ambiente pode estar bem controlado de tal forma que o processo de segmentação confiavelmente extrai apenas as partes necessárias para a análise futura.

Em alguns casos a segmentação pode ocorrer de forma trivial por conta de atributos evidentemente destacados na imagem. Mas há casos mais complexos em que pode ser necessário realizar tratamento prévio da imagem, como, por exemplo, o uso de operações morfológicas e filtro gaussiano para remoção de ruído visando facilitar a execução de técnicas de segmentação baseada em vales (DI LEO, 2017).

O segundo objetivo da segmentação é realizar uma mudança de representação na imagem. Os pixels da imagem precisam ser organizados em unidades de alto nível que serão ou mais significantes ou mais eficientes para análise futura (ou os dois).

As estratégias de segmentação têm suas abordagens categorizadas baseado nas propriedades da imagem, podendo ser delineadas a partir da detecção da sua 1 - descontinuidade ou 2 – similaridade (KAUR e KAUR, 2014; GONZALEZ e WOODS, 2002). Além disso, as técnicas de segmentação podem ser classificadas em três grupos (KAUR e KAUR, 2014; DEY et al, 2010):

- Técnica de segmentação estrutural: que se baseia na informação estrutural de uma porção requerida da imagem.
- Técnica de segmentação estocástica: que funciona a partir dos valores dos pixels ao invés de informações de uma estrutura ou região dos pixels.
- Técnicas híbridas: usam os conceitos das duas técnicas anteriores.

Kaur e Kaur (2014) compilou as vantagens e desvantagens de algumas técnicas de segmentação, apresentados no Quadro 2.

Quadro 2 – Comparação entre várias técnicas de segmentação

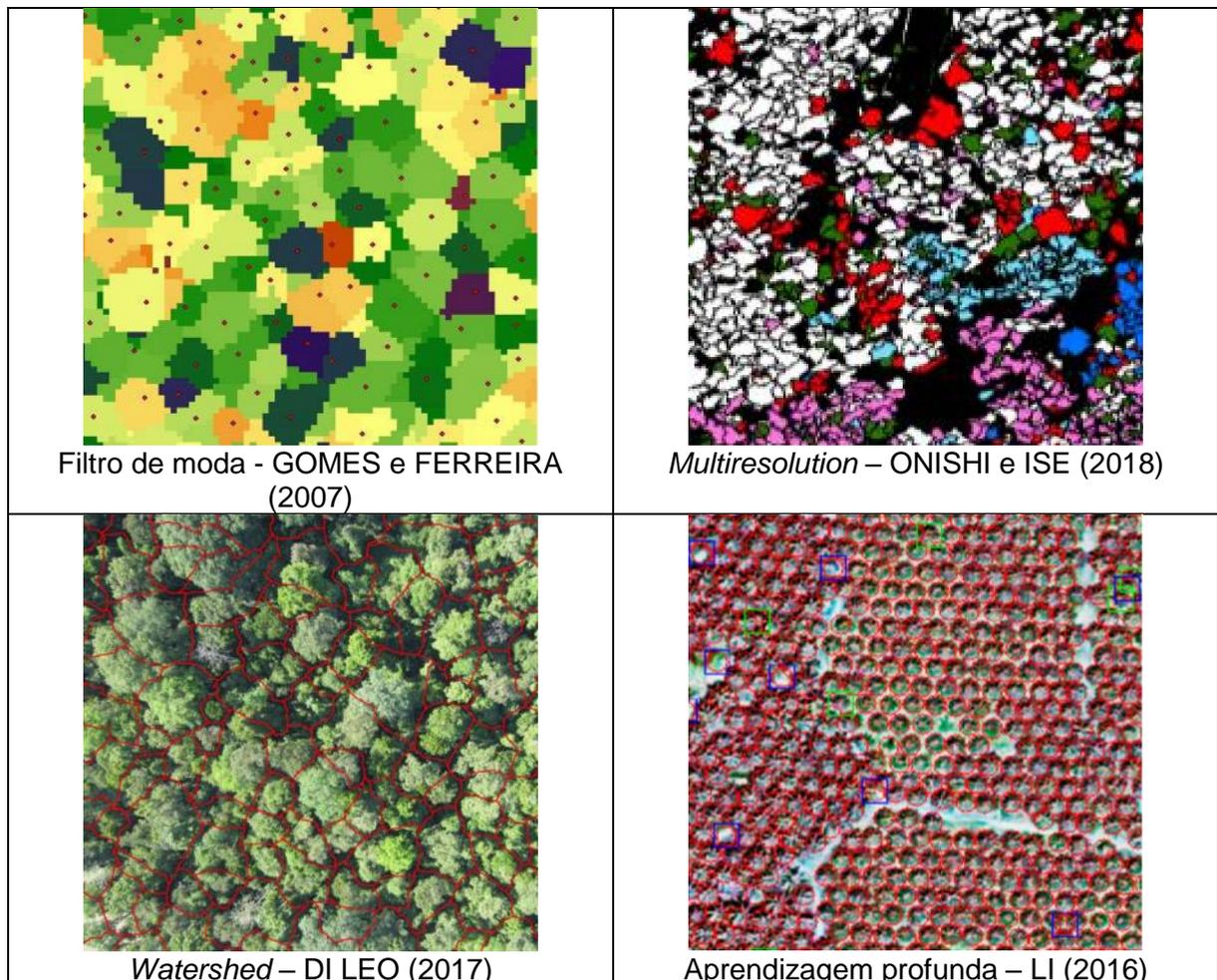
Técnica de Segmentação	Descrição	Vantagem	Desvantagem
Método Limiar (Thresholding)	Faz uso dos picos do histograma da imagem para encontrar valores para o limiar	Não precisa de informação prévia, método mais simples.	Alta dependência dos picos, detalhes espaciais não são considerados.
Método Baseado nas Bordas (Edge Based)	Detecção da descontinuidade da imagem	Bom para imagens que possuem melhor contraste entre objetos	Inadequado para bordas detectadas erroneamente ou em grande quantidade.
Método baseado na região (Region Based)	Divide a imagem em regiões homogêneas	Mais imune à ruído, útil quando é fácil definir um critério de similaridade.	Método caro em termos de tempo e memória computacional.
Método de Aglomeração (Clustering)	Divide a imagem em aglomerações (clusters) homogêneos.	A fuzzy usa associações de classes parciais. Mais útil para problemas reais.	Determinar a função de adesão não é fácil
Método Watershed	Se baseia na interpretação topológica	Os resultados são mais estáveis, os limites detectados são contínuos	Cálculo complexo de gradientes e grande quantidade de parâmetros
Método Baseado em Equação Diferencial Parcial	Com base no funcionamento de equações diferenciais	Método mais rápido, melhor para aplicativos de tempo crítico	Elevada complexidade computacional
Método Baseado em Redes Neurais	Baseado na simulação do processo de aprendizagem para a tomada de decisão	Não há necessidade de escrever programas complexos	Tempo de treinamento alto, elevado consumo de recursos computacionais, “caixa preta”

Fonte: adaptado KAUR e KAUR (2014)

Segundo SHAPIRO e STOCKMAN (2001), a chance de um único método de segmentação funcionar para todos os problemas parece ser mínima, tornando necessário que o operador que irá trabalhar a imagem seja capaz de escolher o melhor método para uma aplicação ou para resolver o problema.

Imagens como as fotografias áreas das copas das árvores possuem características específicas que podem dificultar a segmentação. Dependendo do tipo de floresta, a imagem pode ser homogênea com pouco ou nenhum detalhe que permita a fácil identificação de cada indivíduo. Na literatura há exemplos de métodos que tentam trabalhar sob esse tipo de conjunto de dados. Abaixo são descritos alguns métodos de segmentação de copas de árvores encontrados na literatura e que serão avaliados nesse trabalho. Um recorte com os principais resultados obtidos pelos autores mencionados a seguir é apresentado no Quadro 3.

Quadro 3 – Resultados na Literatura Para as Técnicas de Segmentação



Fonte: elaborado pelo autor

3.4.1 PCA e filtro de moda

No estudo realizado por Gomes e Ferreira (2007), a segmentação das copas foi gerada a partir da Análise dos Componentes Principais (PCA) de uma imagem pancromática do satélite Quickbird, seguida do uso de um filtro para homogeneizar os

valores dos pixels próximos. Os autores utilizaram para a análise a imagem pancromática, a primeira e a segunda componentes principais da PCA, e um duplo processamento usando um filtro de moda 7x7 a fim de se obter as copas das árvores (Quadro 3).

Gomes e Ferreira (2007) apontaram que o filtro de moda, aplicado em uma imagem pancromática e na primeira componente principal da PCA, pode ser uma boa ferramenta para auxiliar a individualizar as copas de árvores nas imagens do Quickbird. Nessa pesquisa, foi testada essa mesma técnica em imagens com 3 bandas geradas pelo RPAs (RGB).

3.4.2 Multiresolution

Outra estratégia analisada, descrita por ONISHI e ISE (2018), foi o algoritmo “Multiresolution Segmentation” (BAATZ e SCHÄPE, 2000). Nessa abordagem os autores fizeram o uso de fotografias tiradas pelo RPA combinadas com o modelo de elevação e de declividade para gerar um mapa com a segmentação das copas das árvores.

Esse algoritmo tem como componentes a decisão heurística para determinar a mescla dos objetos identificados na imagem a cada etapa do processamento e a definição da homogeneidade dos objetos da imagem para computar o grau de ajuste para cada par de objetos na imagem (BAATZ e SCHÄPE, 2000).

Os ajustes acontecem em diferentes escalas e resultam na segmentação por elementos comuns de acordo com as suas semelhanças. Ao combinar o Modelo Digital de Elevação, a imagem RGB e a Declividade, ONISHI e ISE (2018), obtiveram a segmentação das copas das árvores mostradas na Quadro 3.

3.4.3 Watershed

O processo de segmentação descrito por Di Leo (2017) e Jing et al. (2012), faz uso do modelo de bacia hidrográfica, sendo um algoritmo morfológico matemático que deriva de uma analogia ao fluxo de rios em bacias hidrográficas (BEUCHER, 1992). Di Leo (2017) aplicou o modelo no topo das árvores fazendo uma inversão da imagem, transformando as copas em vales, com os separadores mais escuros convertidos em picos, que posteriormente foram usados para gerar a segmentação das copas das árvores (Quadro 3). Este modelo invertido é virtualmente preenchido até que as bacias

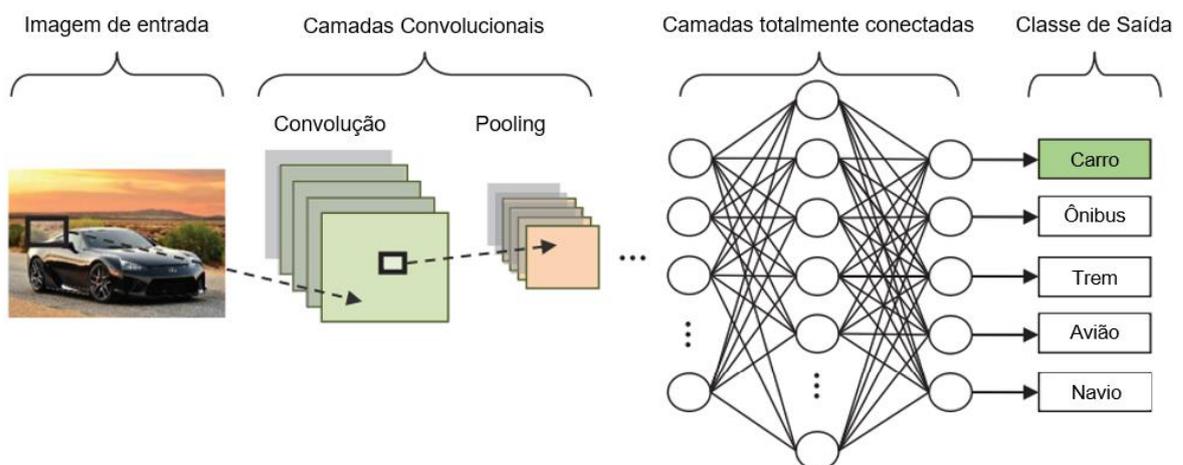
hidrográficas sejam evidentes, gerando assim as linhas de segmentação (WANG ET AL., 2004).

3.4.4 Aprendizagem profunda

Nos últimos anos, o uso de técnicas de aprendizado profundo em sensoriamento remoto tem crescido e se apresentado como uma ferramenta promissora para pesquisas em sensoriamento remoto (ZHU et al. 2017, MARMANIS et al. 2016, MARMANIS et al. 2018). Li et al. (2016) usaram um modelo de aprendizado profundo para segmentar palmeiras em uma área com alguma sobreposição de copas de árvores (Quadro 3) e obtiveram resultados interessantes.

Dentro da aprendizagem profunda é possível citar o rápido desenvolvimento de Redes Neurais Convolucionais (Convolutional Neural Networks - CNN), a qual atingiu resultados impressionantes em rotinas da visão computacional como a detecção e segmentação em imagens (ZHANG et al., 2018). Para realizar essa tarefa, segundo Rawat e Wang (2017), a CNN faz uso de um conjunto de conexões que carregam a informação seguindo uma rota de entrada à saída, passando por camadas que realizam algum nível de processamento no dado até chegar, por exemplo, em uma classificação da informação (Figura 14). Essa estrutura permite que as camadas interpretem as representações das feições no dado de entrada e execute uma compreensão de alto nível da imagem (Hinton et al., 2012; Rawat e Wang, 2017).

Figura 14 – Exemplo de estrutura de uma Rede Neural Convolucional.

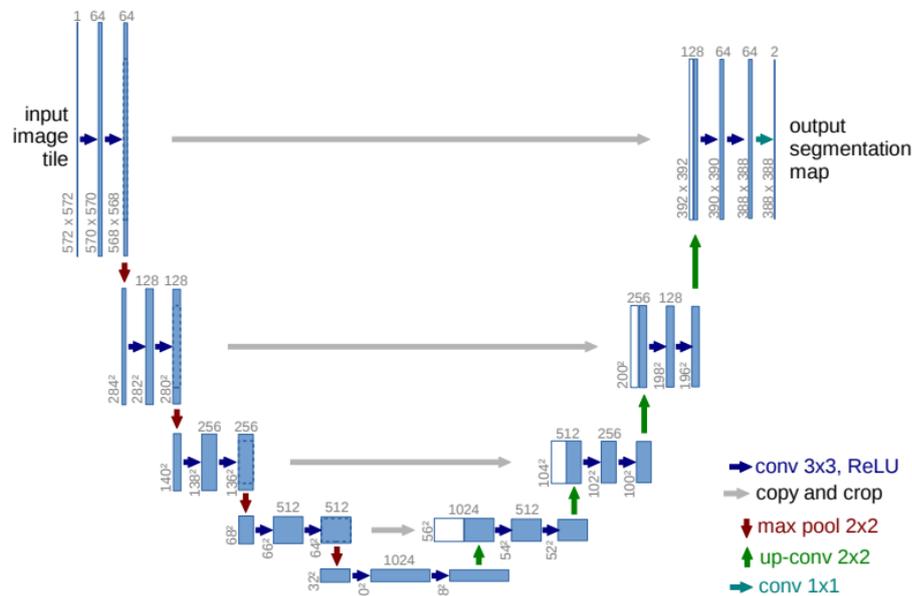


Fonte: Adaptado de Rawat e Wang (2017)

A mudança na organização das camadas nas redes neurais permite a formação de estruturas que lidam de formas diferentes com os dados. Um exemplo específico

de estrutura de CNN é a U-Net (RONNEBERGER et al., 2015). Trata-se de uma CNN que foi projetada para processar dados multidimensionais como sinais e sequências, imagens, espectrogramas de áudio, vídeos ou imagens volumétricas (LECUN et al. 2015). Segundo RONNEBERGER et al. (2015), essa estrutura de rede neural possui a capacidade de realizar mapeamentos complexos na imagem similares a precisão humana, além de possibilitar aplicação em experimentos com pequeno número de imagens disponíveis para o treinamento da rede neural. A arquitetura da U-Net pode ser vista na Figura 15.

Figura 15 – Arquitetura U-Net



Exemplo de modelo de arquitetura U-Net. Fonte: Ronneberger (2015)

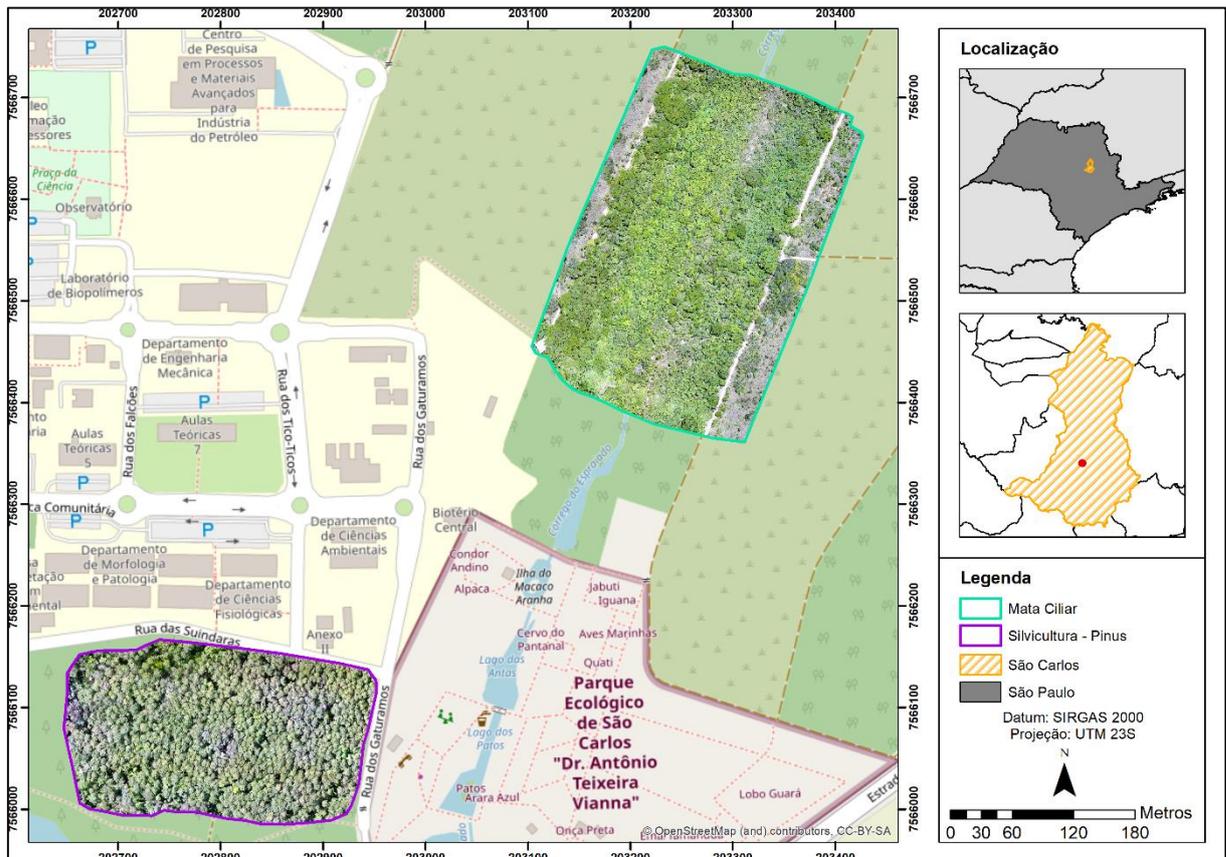
Para Falk (2018), a U-NET não tem a função de corrigir identificações de baixa qualidade estabelecidas por humanos, mas sim atuar como uma ferramenta que permite rotular grandes conjuntos de dados e, conseqüentemente, economizar o tempo despendido para o trabalho manual.

4 METODOLOGIA

4.1 ÁREA DE ESTUDO

O estudo foi realizado no município de São Carlos, interior do estado de São Paulo, contemplando duas áreas verdes, sendo: uma silvicultura de *Pinus* e mata ciliar (Figura 16).

Figura 16 – Figura de Localização



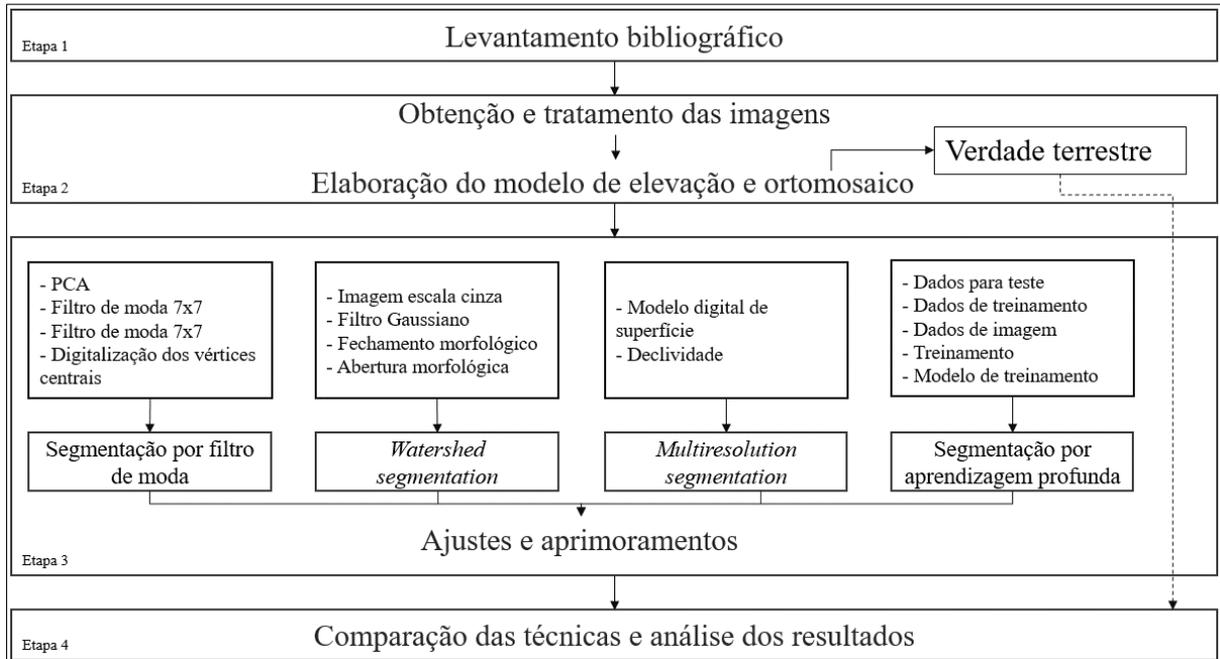
Fonte: elaborado pelo autor

A área classificada como silvicultura-pinus se caracteriza por apresentar predominância de *Pinus elliotti* var. *Engelm.* No estrato arbóreo e sub-bosque em regeneração com espécies nativas da região (LIMA, 2017). Já a área de Mata Ciliar pertence a margem de APP das nascentes do Córrego do Espriado, o qual apresenta boa parte de sua extensão no Parque Ecológico “Dr. Antônio Teixeira Vianna” e é utilizado para captação de água visando o tratamento e abastecimento urbano (PAESE, 1997).

As informações levantadas e processadas nas técnicas de segmentação tiveram seus resultados analisados e comparados com a verdade terrestre (conforme

descrito na seção 4.3), sendo que esse processo ocorreu dentro de quatro etapas, possibilitando conclusões quanto as qualidades e limitações das ferramentas estudadas (Figura 17).

Figura 17 – Organograma de Execução da Pesquisa



Fonte: elaborado pelo autor

4.2 CARACTERÍSTICAS DO RPA, CÂMERA

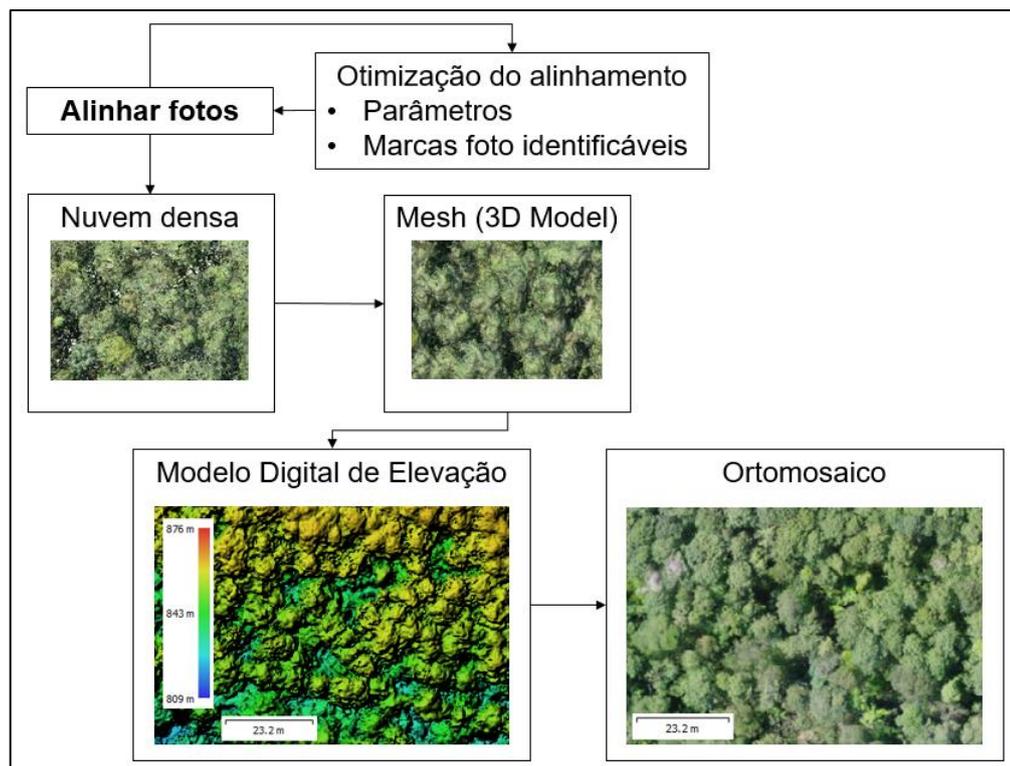
O RPA utilizado no estudo é um DJI Phantom 4 Advanced com um sensor/câmera CMOS de 1/2.3 polegadas e 20 MP acoplada a sistema de estabilização motorizada de três eixos (roll, pitch e yaw) para capturar imagens estabilizadas. Com uma capacidade de voo máxima de 30 minutos por bateria, o RPA também traz embutido um sensor de posicionamento por satélite de dupla-banda (GPS e GLONASS), permitindo gerar imagens com informações adicionais de Latitude e Longitude (além das informações de orientação já inclusas), que são gravadas diretamente no EXIF (Exchangeable Image File Format) de cada arquivo. Essas informações são essenciais para a geração do Modelo Digital de Superfície (DSM) e do ortomosaico utilizados no trabalho, conforme os procedimentos descritos a seguir.

4.3 AQUISIÇÃO E PREPARAÇÃO DE DADOS

Dois voos a 60 metros² foram realizados no dia 25 de outubro de 2019, levantando dados em duas áreas: silvicultura de *Pinus* e mata ciliar. Segundo os dados disponibilizados pela estação do INMET (Instituto Nacional de Meteorologia), situada a poucas centenas de metros das áreas, nesse dia o tempo, às 09h, apresentava temperaturas de 25,6°C, com umidade de 59% e ventos de 1m/s, com direção de 36°.

As imagens resultantes foram processadas usando Agisoft Metashape (Figura 18) para gerar um ortomosaico 2D e um DSM no formato GeoTIFF. Esses dados foram georreferenciados pelo próprio software a partir dos dados do próprio GPS da RPA e não foi realizada nenhum tipo de interpolação pelos operadores. A resolução espacial (GSD) é de 1,5 cm para o ortomosaico e 6 cm para o DSM.

Figura 18 - Etapas Realizadas no Agisoft Metashape



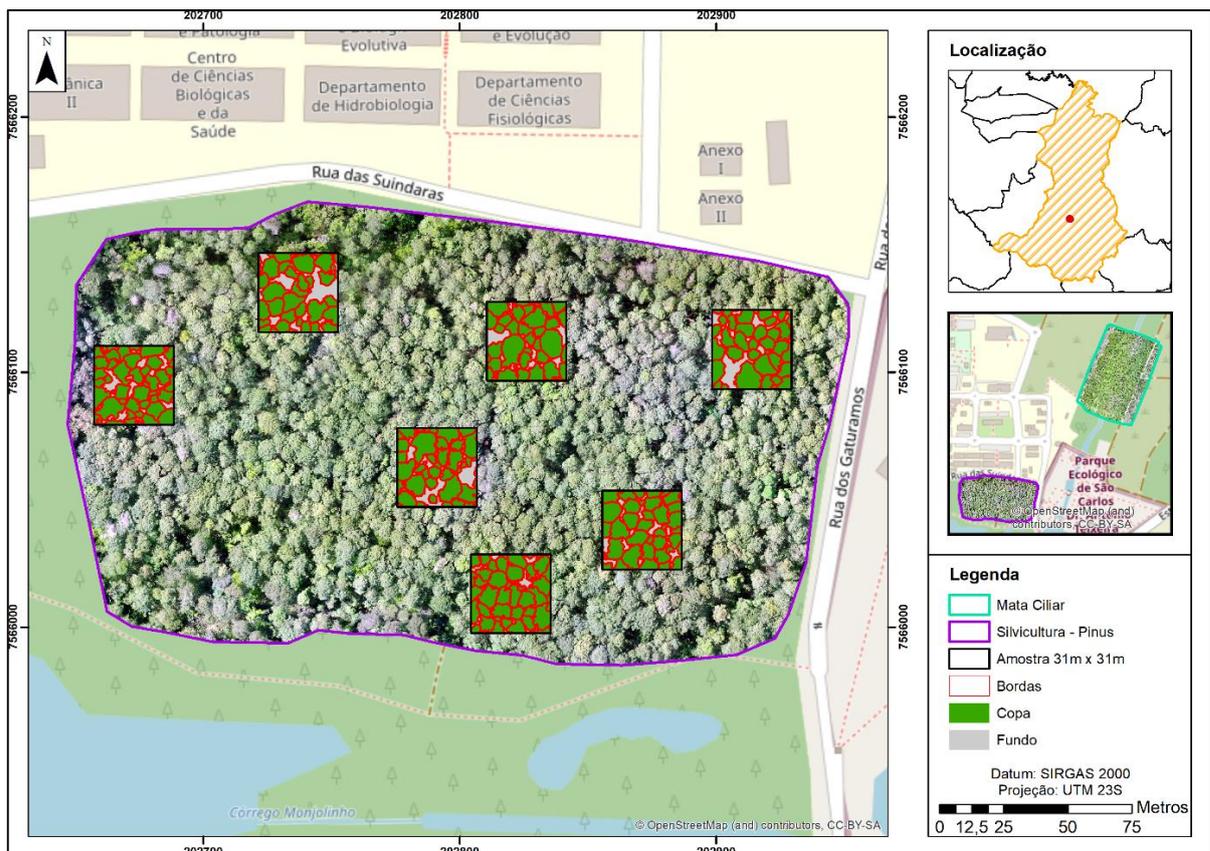
Fonte: elaborado pelo autor

² Em ambos os casos, o ponto de partida ficava a 60 metros, permitindo que o sobrevoo fosse feito a alturas um pouco maiores. Ressalta-se que essa foi a altura máxima autorizada pelo DECEA nas datas específicas.

Para preservar as proporções e a densidade de pixels para as próximas etapas de aplicação da metodologia, o ortomosaico foi reamostrado para a resolução espacial do DSM e ambos tiveram seu sistema de projeção alterado de geográfico não projetado (WGS 84) para UTM 23S (SIRGAS 2000) usando ArcGIS Pro.

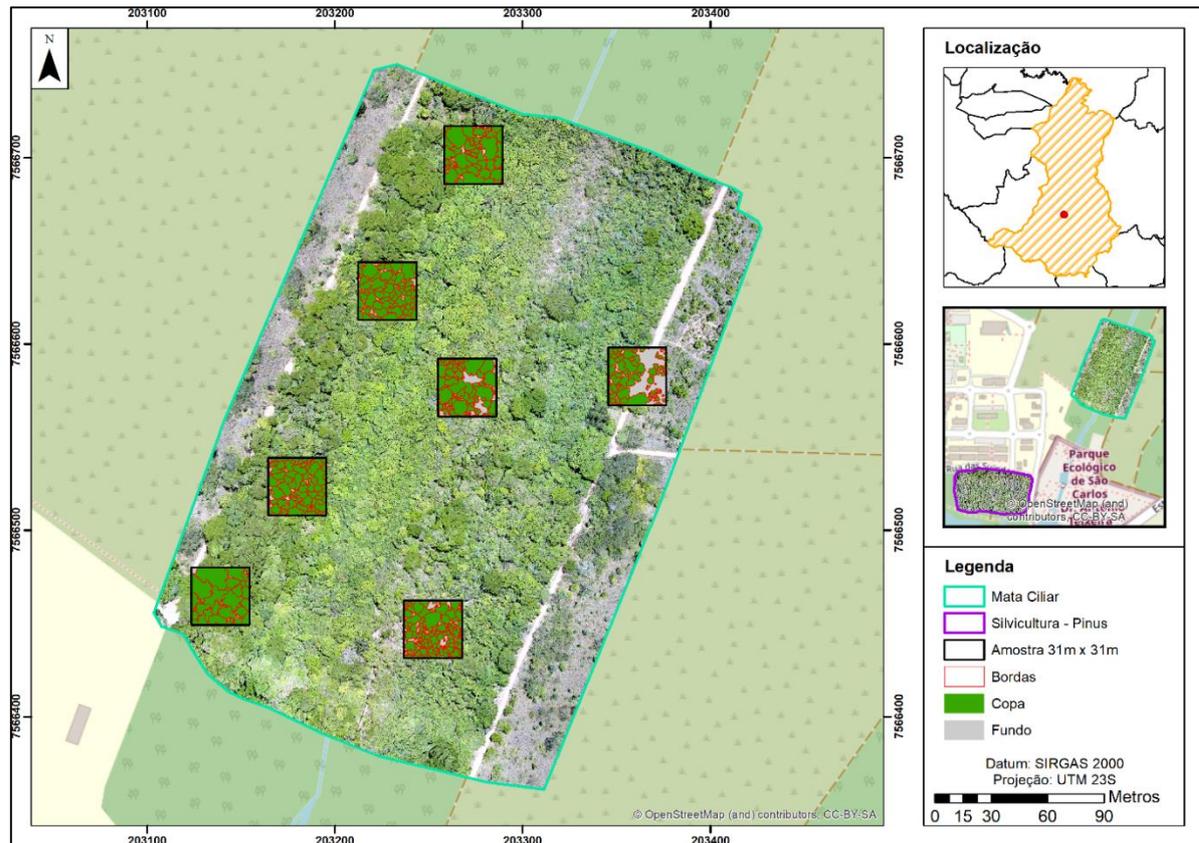
Dentro da imagem de cada área sobrevoada foram extraídos sete conjuntos de amostras com área de 961m² em diferentes regiões da imagem que incorporasse as diferentes características das áreas sobrevoadas. Essas amostras serviram de base para criação da verdade terrestre, por meio da identificação e vetorização das copas de cada árvore. Com base no ortomosaico, foram delinear delineadas as copas e clareiras, resultando em polígonos de referência para medidas de precisão e máscaras para algoritmo de aprendizado profundo, utilizando 3 categorias: copa, clareira e borda (Figuras 19 e 20).

Figura 19 - Verdade Terrestre Vetorizada na Silvicultura



Fonte: elaborado pelo autor

Figura 20 - Verdade Terrestre Vetorizada na Mata Ciliar



Fonte: elaborado pelo autor

É importante ressaltar que essa técnica para elaborar a verdade de campo não é completamente precisa, pois sua qualidade é apenas tão boa quanto a capacidade do operador de identificar cada uma das copas das árvores a partir da imagem. Uma mesma amostra, vetORIZADA por operadores igualmente habilidosos, poderá resultar em um produtos com alguma diferença.

4.4 MÉTODOS DE SEGMENTAÇÃO

Os métodos de segmentação foram baseados em técnicas encontradas na literatura e adaptados para as ferramentas disponíveis para a condução do estudo. Um recorte dos resultados de cada referência são apresentados no Quadro 3

4.4.1 Segmentação de Filtro de Moda

Gomes e Ferreira (2007) apontaram que o filtro de moda, aplicado em uma imagem pancromática e na primeira componente principal da PCA, pode ser uma boa ferramenta para auxiliar a individualizar as copas de árvores nas imagens do

Quickbird. Nessa pesquisa, foi testada essa mesma técnica em imagens com 3 bandas geradas pelo RPAs (RGBe, , para aplicá-la aos dados da área de estudo, foram executados os seguintes passos:

- Amostragem do ortomosaico e do DSM de 0,06 m para 0,6 m, visando manter a mesma resolução utilizada no estudo realizado por Gomes e Ferreira (2007) e processar utilizando os mesmos tamanhos de filtro;
- Recorte do ortomosaico da área de estudo para separar a área arbórea de outras partes da imagem;
- Execução, no ArcGIS Pro, da ferramenta PCA visando combinar o ortomosaico reamostrado e o DSM para, posteriormente, exportar o primeiro e segundo componente principal para um GeoTIFF (adaptado Gomes e Ferreira, 2007). Segundo Campbell e Wynne (2011), o PCA identifica a combinação linear ótima dos canais originais que podem explicar a variação dos pixels dentro de uma imagem. Assim, o uso da PC1 pode gerar um resultado de segmentação diferente se comparado com imagem multiespectral original;
- Dupla aplicação do filtro QGIS Majority com um modo de pesquisa circular e raio de 7 pixels, padronizando os dados semelhantes nas áreas que representam unidades de copas de árvores, segmentando a imagem. O filtro aplicado serve para a homogeneização dos pixels próximos para uma mesma cor, sendo que o filtro de moda permite preencher as lacunas entre os polígonos resultantes da conversão vetor-raster.
- Transformação do raster resultante em polígono usando a ferramenta QGIS raster to vector.

4.4.2 Multiresolution Segmentation

Essa técnica, utilizada por ONISHI e ISE (2018), foi replicada no presente estudo através dos seguintes passos:

- Criação de um modelo digital de superfície (DSM) a partir das fotografias das imagens recolhidas pelo RPA. O DSM serviu, simultaneamente, para criar o modelo de declividade e de entrada para a segmentação com os seus valores de altura;

- Geração da declividade a partir do modelo digital de superfície. Como a declividade mostra a taxa máxima de mudança de elevação entre as células, torna-se possível enfatizar as bordas das árvores (ONISHI e ISE, 2018);
- Uso do algoritmo Multiresolution Segmentation da versão trial do Software eCognition 9.x para criar o mapa com as copas das árvores, combinando o ortomosaico com o modelo de elevação e a declividade.

A ferramenta que realizou esse procedimento pode ser encontrada já implementada em softwares de terceiros, cabendo a sua aplicação no conjunto de dados de interesse, com parâmetros a ser testados a fim de viabilizar o melhor ajuste.

4.4.3 Watershed Segmentation

As etapas para aplicar esta técnica de segmentação neste estudo foram realizadas no Python (3.x) e consistiram na adaptação do modelo de Di Leo (2017) e Jing et al. (2012), descrito anteriormente, da seguinte forma:

- Conversão do ortomosaico para tons de cinza para permitir seu uso nas operações morfológicas;
- Adição do Modelo Digital de Elevação normalizado entre 0-255 por meio de uma média ponderada com a imagem em tons de cinza;
- Aplicação de um filtro gaussiano para reduzir o ruído e suavizar a imagem;
- Aplicação do filtro Gaussiano para redução do nível de ruído nas folhas. Folhas, segundo GRANT (1987), não são refletoras puramente difusas ou especulares, mas apresentam características dos dois tipos na sua estrutura e composição química, assim, a aplicação desse filtro reduz a aleatoriedade natural dos padrões de refletância das folhas;
- Fechamento morfológico com um elemento de estrutura parabólica para a redução do nível de ruído dos galhos. Esse processo ocorre através de uma operação matemática morfológica que resulta no fechamento de buracos na imagem;
- Abertura morfológica com um elemento de estrutura em formato de disco para eliminar o “leftover shallow mínima”. Diferentemente do fechamento, a abertura suaviza os contornos e elimina saliências finas;
- Inversão de valores de pixels para as copas se tornarem “vales”;

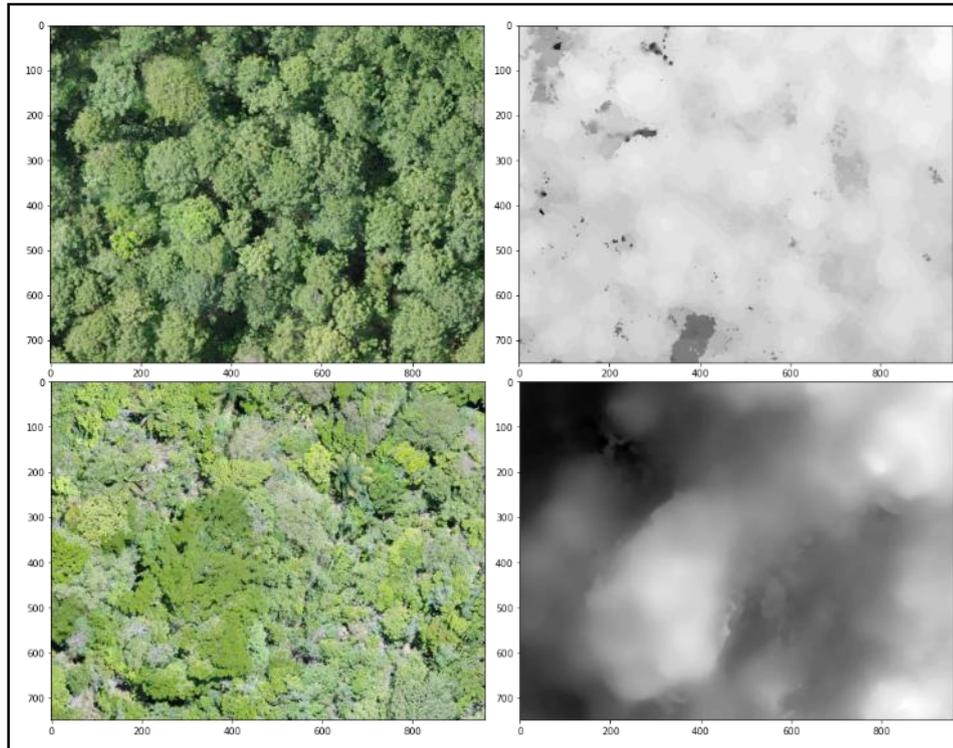
- Definição de limiares para gerar marcadores das mínimas locais utilizados na função que realiza a segmentação: esse limiar foi definido a partir de uma análise dos valores máximos e mínimos encontrados para as copas das árvores e de como as bordas são identificadas na imagem. Assim, o limiar ideal é aquele que consegue separar os valores que representam os picos das árvores do resto da sua estrutura.
- Uso de ferramenta de segmentação *watershed* e reatribuição de informações espaciais/geográficas na imagem para uso na plataforma GIS;
- Transformação do raster resultante em polígono usando a ferramenta ArcGis raster to Polygon.

A segmentação por Watershed utiliza uma série de filtros para o preparo da imagem, gerando subprodutos que mudam de acordo com os parâmetros utilizados. Assim, inicialmente a imagem e o modelo de elevação (Figura 21) foram carregados e combinados em uma única imagem em tons de cinza, utilizando os coeficientes da Rec. 601 luma para combinar o RGB e uma ponderação para combinar o resultado do RGB com o DSM. Essa ponderação foi testada empiricamente para que o DEM não fosse predominante na imagem:

$$Y'' = Y'_{601} \times 0.8 + DSM \times 0.2$$

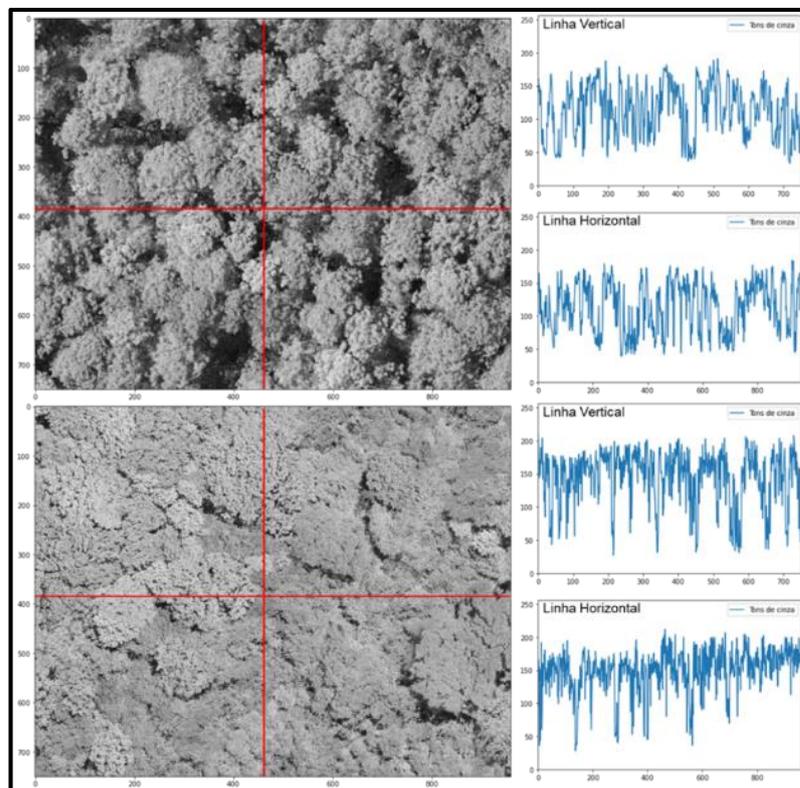
A imagem resultante apresentou alto grau de ruído para a identificação das unidades copas de árvores (Figura 22), gerado pela dispersão luminosa das folhas e galhos (Di Leo, 2017). Pensando em uma etapa inicial do processamento da imagem, foi utilizado um filtro gaussiano que é composto por dois parâmetros: tamanho do filtro e o sigma.

Figura 21 – Recortes do RGB e DSM das Áreas de Estudo



As duas imagens da parte superior são recortes do *Pinus* e as duas inferiores pertencem à Mata Ciliar. Fonte: elaborado pelo autor

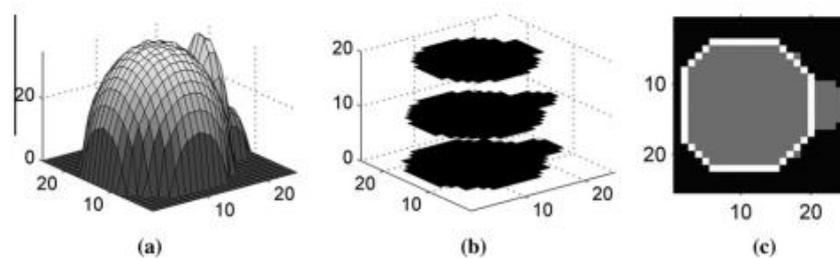
Figura 22 – Ruídos Presente nas Imagens em Tons de Cinza



A imagem e o par de gráficos na parte superior pertencem a área de *Pinus* e o conjunto da parte inferior pertence à Mata Ciliar. Fonte: elaborado pelo autor

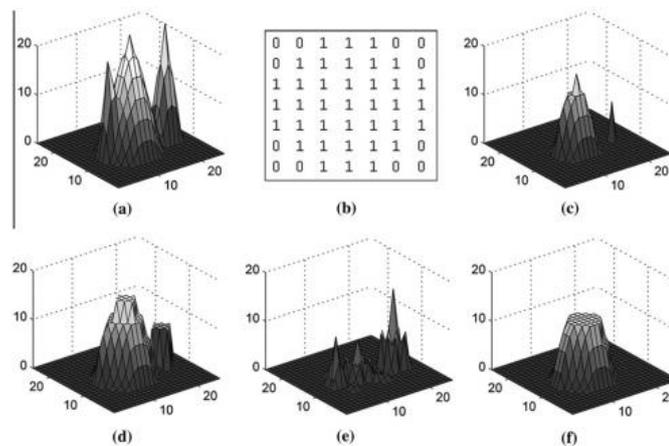
O tamanho foi definido utilizando um dos procedimentos empregados por Jing *et al.* (2012), para encontrar o tamanho médio, em pixel, das copas das árvores. Para tanto, a copa é tratada como meio elipsoide, o qual pode ser fatiado em várias camadas com formato de disco (Figura 23) a partir de uma operação morfológica de abertura utilizando um elemento estruturante no formato de disco (Figura 24) (Soille, 1999).

Figura 23 – Fatiamento da Copa em Várias Camadas



Fonte: Di Leo (2017)

Figura 24 – Exemplo de Operação Morfológica de Abertura



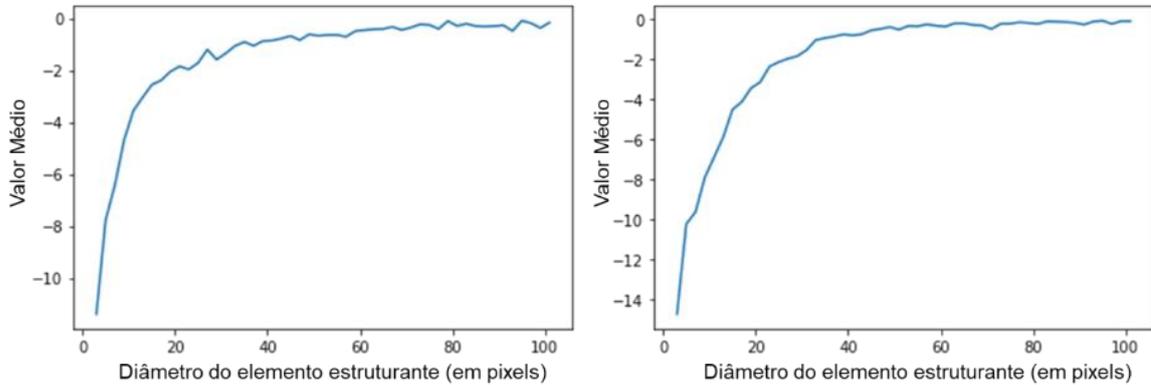
Fonte: Di Leo (2017)

O diâmetro da maior fatia pode ser encontrado por meio do cálculo da média da diferença entre imagens morfológicamente abertas, utilizando kernels de tamanhos distintos e progressivos formando uma série (Figura 25).

No entanto, é necessário ressaltar que a presença de ruído na imagem dificultou encontrar uma curva que mostrasse uma variação evidente no tamanho das árvores semelhante ao encontrado na literatura. Portanto, foram testados diferentes

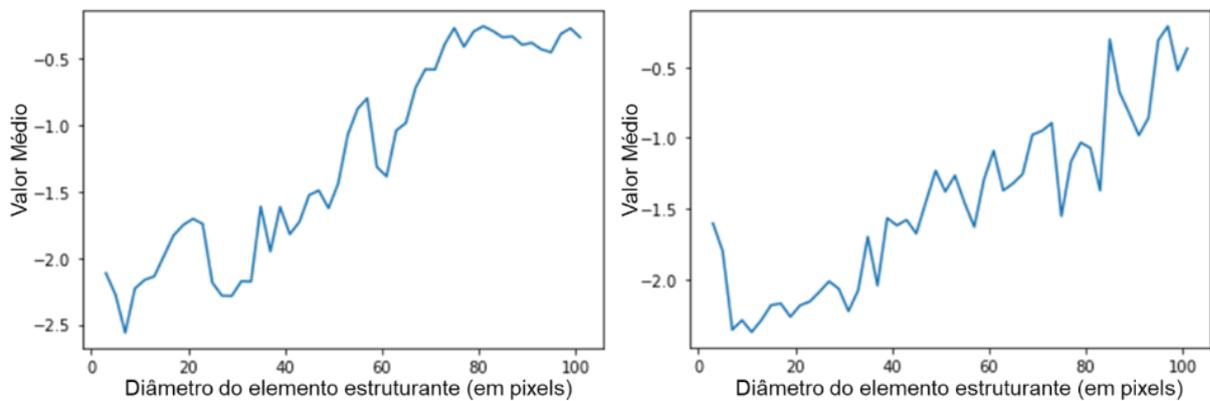
tamanhos de kernel e sigma suficientemente pequenos para um filtro gaussiano inicial, que reduziu o ruído e permitiu identificar as propriedades mais evidentes das copas das árvores (Figura 26).

Figura 25 – Valores médios da diferença entre duas imagens da série



Fonte: elaborado pelo autor

Figura 26 – Valores médios da diferença entre duas imagens da série após o filtro gaussiano



Fonte: elaborado pelo autor

Com esses dados, foi possível encontrar as mínimas locais e, dessa forma, estimar o tamanho dos principais elementos que compõe a imagem que, no caso, podem ser as copas das árvores. Jing *et al.* (2012) detectou múltiplas mínimas locais, representando diferentes tamanhos de copas de árvores, sendo que, esse efeito também foi detectado nas áreas de *Pinus* e de mata ciliar (Quadro 4).

Quadro 4 – Mínimas Locais em Função do Tamanho da Kernel

<i>Pinus</i>				Y			
Antes do Filtro		Depois do Filtro		Antes do Filtro		Depois do Filtro	
Tamanho	Valor	Tamanho	Valor	Tamanho	Valor	Tamanho	Valor
23	-1,970	7	-2,562	41	-0,809	7	-2,355
29	-1,591	29	-2,285	51	-0,521	11	-2,371
37	-1,062	33	-2,176	55	-0,359	19	-2,263
47	-0,843	37	-1,950	61	-0,373	31	-2,224
51	-0,677	41	-1,819	71	-0,494	37	-2,043
57	-0,712	49	-1,624	81	-0,239	41	-1,618
69	-0,450	61	-1,384	91	-0,272	45	-1,675
77	-0,411	71	-0,579	97	-0,236	51	-1,379
81	-0,296	77	-0,408			57	-1,628
87	-0,320	85	-0,337			63	-1,372
93	-0,490	89	-0,393			75	-1,552
99	-0,377	95	-0,449			83	-1,371
						91	-0,982
						99	-0,523

Fonte: elaborado pelo autor

A partir de comparações das medidas, foi possível detectar que, para uma imagem de 0,06 m de resolução, o tamanho das copas nas áreas estudadas era de aproximadamente 67 pixels (quatro metros).

4.4.4 Segmentação por Rede Neural – Aprendizagem Profunda

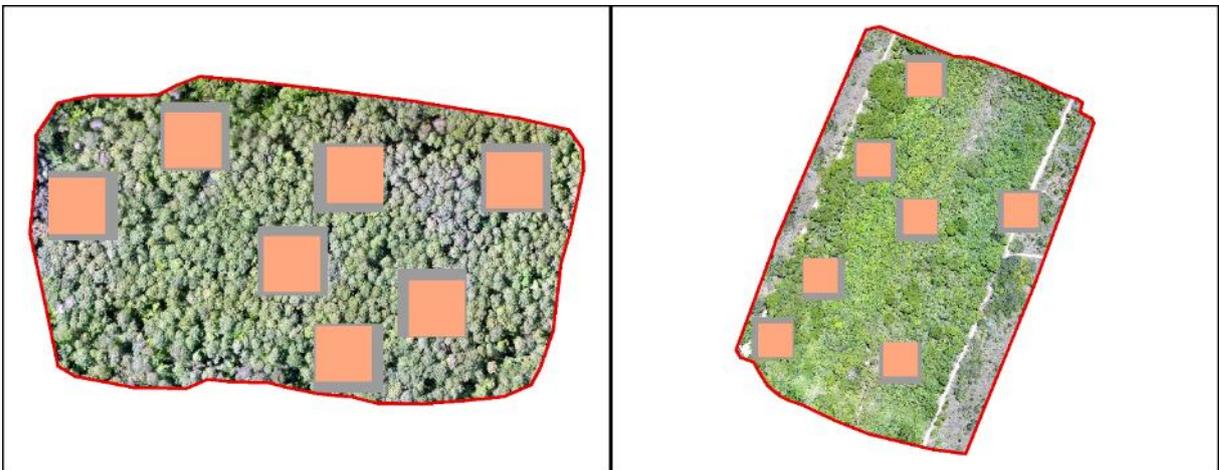
Para realizar a segmentação das copas das árvores com esse método, utilizamos uma Rede Neural Convolutiva (CNN) com uma arquitetura denominada U-NET, dentro do TensorFlow 2 com Keras API (Python 3.x).

O conjunto de dados de treinamento usados para aprendizagem profunda em visão computacional geralmente tem milhares (e.g. CIFAR-10) a milhões (e.g. ImageNet) de imagens rotuladas. Neste estudo, não foi possível criar tamanha quantidade de imagens para treinamento e, para superar essas limitações, utilizamos o aumento de dados (*Data Augmentation*) (REN et al. 2018, YU et al. 2017) com base em poucas amostras, como uma solução alternativa (RONNEBERGER et al. 2015).

A partir da biblioteca implementada por Buslaev (2020), tanto as imagens como as máscaras foram transformadas utilizando espelhamentos, rotações e sutis distorções para aumentar para 3 mil o número amostras de treinamento. É necessário ressaltar que não foram realizadas transformações nas cores da imagem, mantendo o padrão RGB para todas as amostras. As amostras foram divididas em 2 grupos, um para o treinamento e outro para validação do modelo.

Para realizar o treinamento do modelo de rede neural, sete áreas foram recortadas de cada área de estudo (Figura 27). Os retângulos utilizados para a extração são os mesmos utilizados para elaborar a verdade terrestre, possuindo lados de 31 metros e, considerando que a imagem RGB possui resolução espacial de 0,06 metros, resultando em 14 imagens com resolução próxima a 512x512 pixels. Cada imagem foi recortada e ajustada, para se transformarem em quatro imagens com resolução de 256 x 256, sendo geradas 56 imagens para o treinamento do modelo de rede neural.

Figura 27 – Áreas Recortadas para o Treinamento da Rede Neural



Os quadrados na cor laranja indicam as áreas que foram digitalizadas manualmente para criar a verdade terrestre. Já os quadrados em cinza (ligeiramente maiores), indicam as áreas que foram desconsideradas para a predição da rede neural. Fonte: elaborado pelo autor

Além disso, com o objetivo de realizar uma segmentação semântica da imagem, foram elaboradas máscaras que pudessem indicar o que cada elemento da imagem representa. Dessa forma, dentro do ArcGis foi criado um novo shapefile com duas classes: Copa e Fundo. As copas são os elementos que, no momento da inspeção visual da imagem, foram identificados e vetorizados como copas de árvores.

Já o fundo é tudo aquilo que não é copa de árvore e não faz parte do dossel (Figura 28).

Figura 28 – Par Imagem RGB e Mascara Vetorizada



Exemplo tirado a partir da amostra da área de Pinus. Fonte: elaborado pelo autor

Os contornos do dado vetorial das copas foram extraídos e convertidos em linhas, posteriormente, foi criado um buffer de 15 centímetros das linhas e, o polígono resultante foi combinado com o dado vetorial original. Como resultado, o produto gerado é uma máscara vetorial que possui três classes de informação: a copa, o que não é copa e a borda.

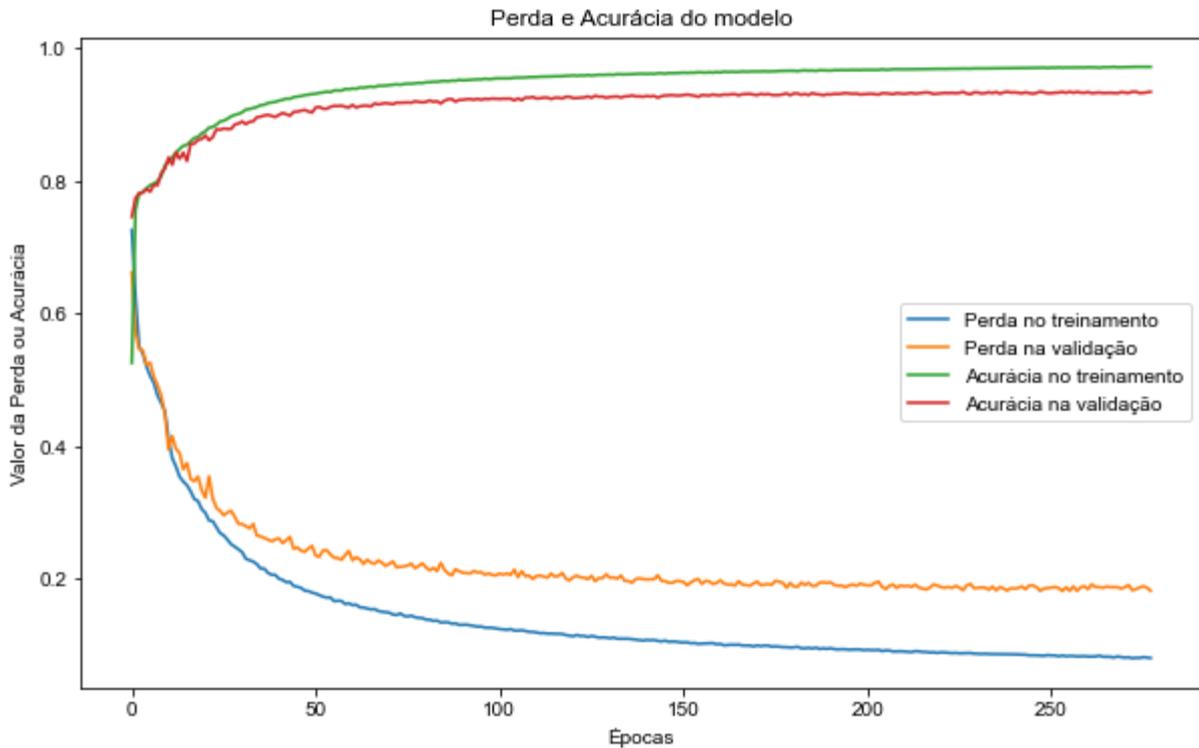
O modelo de rede neural utilizado trabalha com dados no formato raster, portanto, foi necessário converter a máscara vetorial, atribuindo os valores 0, 1 e 2 para cada classe. Esses rasters foram processados utilizando a mesma técnica para recortar a imagem RGB, resultando em pares imagem-máscara, utilizáveis no treinamento da rede neural.

As redes neurais usam grandes quantidades de dados para permitir um treinamento minimamente razoável do modelo, o que tornam as 56 imagens disponíveis insuficientes para esse procedimento. De acordo com Ronneberger et al. (2015), uma forma de LiDAR com a falta de dados para o treinamento da rede neural está no uso de *Data Augmentation*.

O treinamento foi implementado com condição de parada, sendo interrompido caso passassem 5 épocas (*epochs*) sem uma melhora no valor de perda (*loss*) dos dados de treinamento, já a condição que determina o melhor modelo, foi definida pela perda dos dados de validação. Para aferir a perda foi utilizado o índice de Jaccard (ou Coeficiente de Similaridade de Jaccard), ao comparar o resultado da saída do modelo

de rede neural com a máscara do resultado esperado. A partir disso, o treinamento do modelo durou 270 épocas (Figura 29).

Figura 29 – Variação da Perda e Acurácia do Modelo em Função das Épocas



Fonte: elaborado pelo autor

Um modelo temporário foi criado e treinado para identificar 3 classes na imagem: Copa, Borda e fundo, onde a última classe inclui tudo o que não está nas outras duas, inclusive clareiras dentro da floresta. Os procedimentos usados para gerar a segmentação foram:

- Amostragem de um total de 56 recortes amostrais, sendo 28 em cada tipo de vegetação. Essas 56 resultam da divisão de cada amostra das Figuras 18 e 19 em 4 partes de mesma área;
- Criação de conjuntos de máscaras para as imagens das 56 áreas que serão utilizadas para treinar a classificação semântica no modelo;
- Uso de *data augmentation* para aumentar substancialmente a quantidade de imagens e realizar o treinamento do modelo;
- Predição das áreas de estudo utilizando o modelo treinado.

4.5 ÍNDICES PARA AVALIAÇÃO DOS RESULTADOS: SØRENSEN–DICE E COEFICIENTE JACCARD

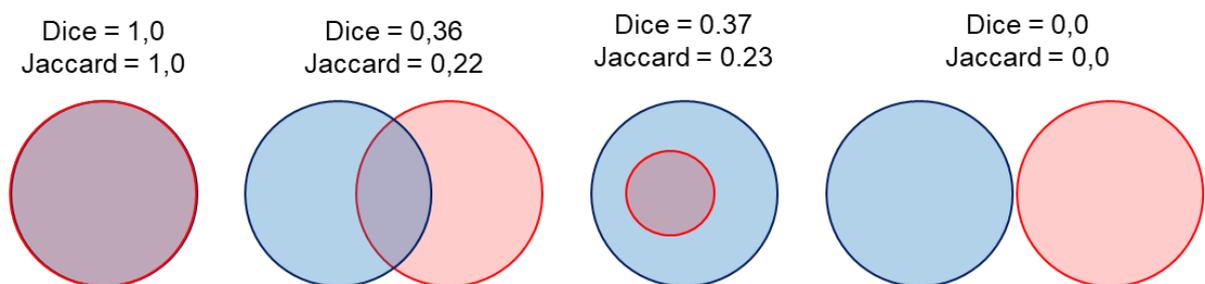
Para mensurar a qualidade dos polígonos desenhados pelos segmentadores, foram utilizados dois índices de similaridade: Sørensen–Dice e Jaccard (Figura 30). O índice de Sørensen–Dice, ou coeficiente de similaridade de Dice, é uma estatística utilizada para mensurar a similaridade entre duas amostras, a partir da seguinte equação:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

O coeficiente de similaridade de Jaccard tem aplicação parecida com o coeficiente de Dice, no entanto é mais sensível a conjuntos de dados não balanceados.

$$J(A, B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

Figura 30 - Variação dos Índices de Similaridade em Diferentes Situações



Fonte: elaborado pelo autor

Os dois índices foram aplicados para comparar a similaridade das bordas dos polígonos gerados pela ferramenta de segmentação e a verdade terrestre, pois, como as linhas tendem a ser mais estreitas e sua sobreposição mais difícil, o resultado tende a ser mais exigente quanto a sobreposição das bordas. Diferente do que ocorreria se a comparação fosse feita em função das áreas dos polígonos, onde qualquer sobreposição poderia indicar uma alta taxa de acerto das técnicas de segmentação.

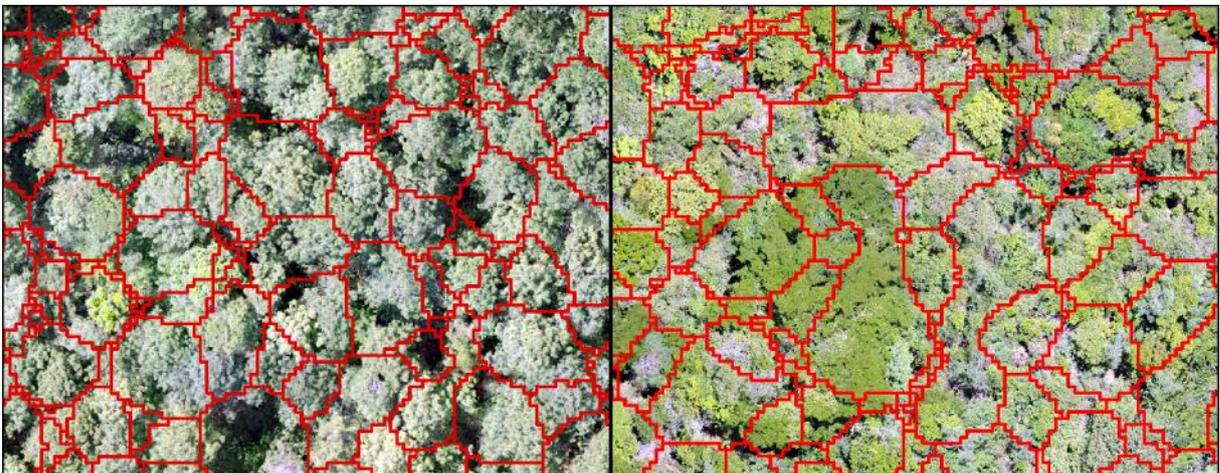
5 RESULTADOS

5.1 FILTRO DE MODA

Esses dados fizeram parte da Análise de Componente Principal (PCA), no entanto, diferente da metodologia de referência, foi adicionado na análise o Modelo Digital de Elevação. Na execução da ferramenta que gera o PCA, foram extraídos o primeiro e segundo componente da imagem.

Os dois componentes foram, de forma independente, processados por meio de uma dupla aplicação do filtro de moda com kernel 7x7 para evidenciar polígonos que supostamente representam as copas das árvores (Figura 31 e 32).

Figura 31 – Resultado da Segmentação por Filtro de Moda PC1



A imagem da esquerda é um recorte do *Pinus* e a da direita um recorte da Mata Ciliar. Fonte: elaborado pelo autor

Figura 32 – Resultado da Segmentação por Filtro de Moda PC2

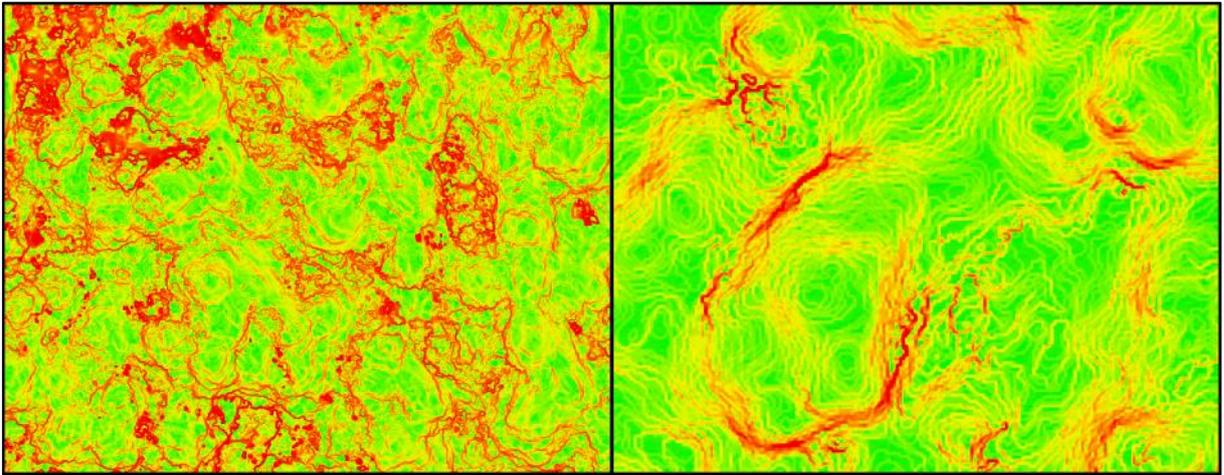


A imagem da esquerda é um recorte do *Pinus* e a da direita um recorte da Mata Ciliar. Fonte: elaborado pelo autor

5.2 MULTIREOLUTION

O processo realizado nas duas áreas de estudo foi composto pelos mesmos planos de informações utilizado por Onishi e Ise (2018), sendo esses a imagem RGB, Modelo Digital de Superfície (DSM) e Declividade (Figura 33). Os dois últimos foram normalizados para a escala de 0-255 para manter o mesmo intervalo de contraste da imagem RGB.

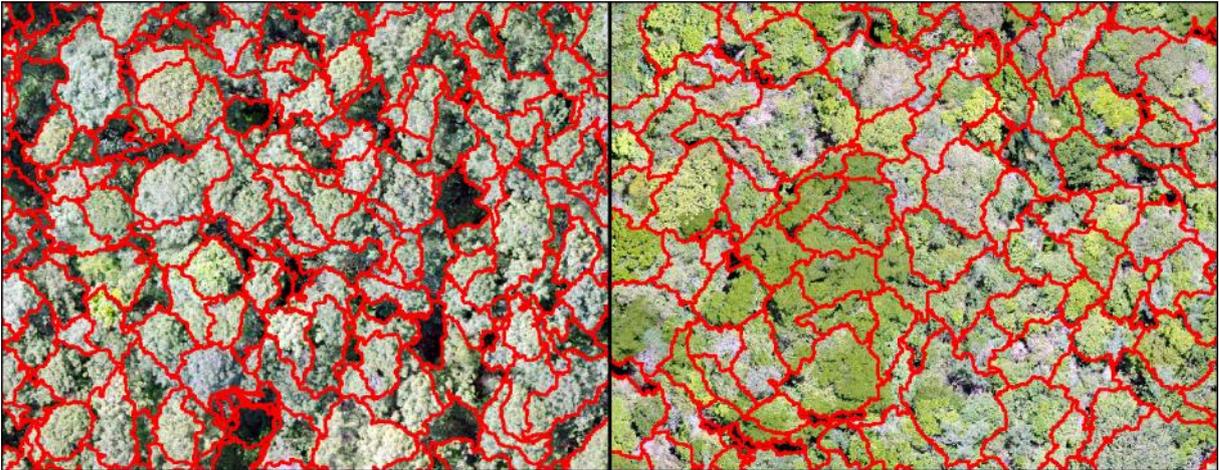
Figura 33 – Declividade



As duas imagens são declividades geradas a partir do DEM. A da esquerda pertence à área de *Pinus* e da direita à Mata Ciliar. O intervalo de valores vão de 0 a 90 graus. Fonte: elaborado pelo autor

Esses três planos de informação foram processados utilizando a ferramenta de segmentação multiresolution (Figura 34) e empiricamente foram testados os parâmetros cor e tamanho, até que fosse encontrado um resultado razoável (0,4 e 100 respectivamente), pois, a ferramenta se mostrou sensível em relação a homogeneidade das copas e, conseqüentemente, gerava muitos ou poucos segmentos com sutis mudanças nos parâmetros.

Figura 34 – Resultado da Segmentação Multiresolution



A imagem da esquerda é um recorte do *Pinus* e a da direita um recorte da Mata Ciliar. Fonte: elaborado pelo autor

5.3 WATERSHED

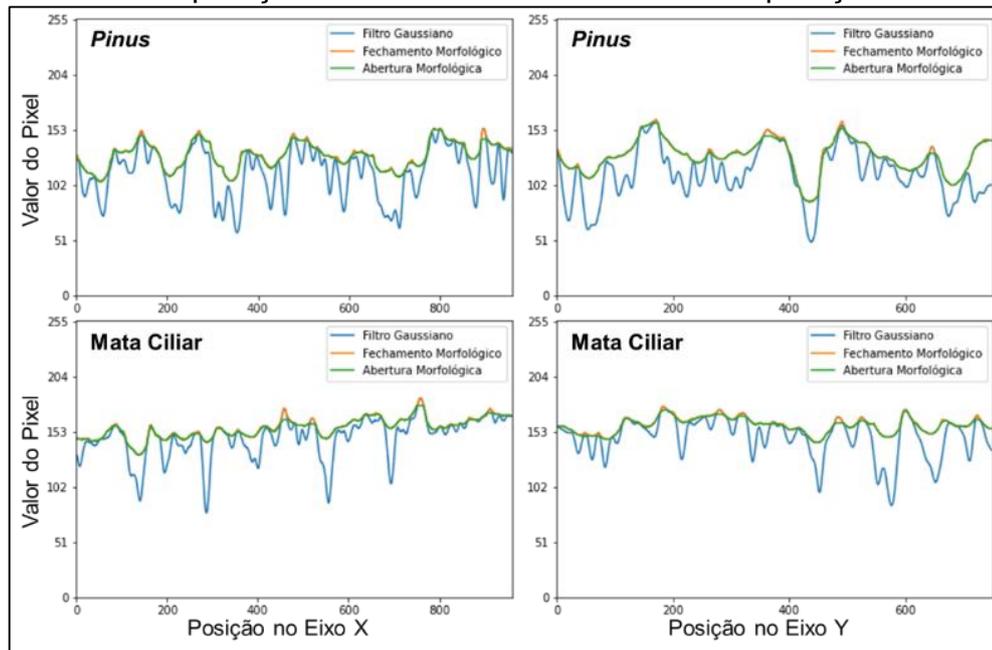
Após a imagem ser devidamente preparada, foi possível identificar um tamanho médio, em pixels, das copas das árvores, o qual foi utilizado no filtro gaussiano para a redução do ruído gerado pela dispersão da luz por conta da orientação aleatória das folhas (Di Leo, 2017). Assim, a imagem foi suavizada, dando às copas uma característica mais homogênea.

Esse resultado foi utilizado como entrada para a operação morfológica de abertura, utilizando um paraboloide como elemento estruturante. O paraboloide, segundo Di Leo (2017), deve ser estreito o suficiente para caber entre as copas das árvores e não nos espaços existentes dentro das copas que separam os galhos. Conseqüentemente, o ruído derivado da ramificação existente em uma copa de árvore tende a ser removido, transformando a copa em um elemento único.

A imagem foi morfológicamente aberta, utilizando um elemento estruturante no formato de disco plano, pois, segundo Di Leo (2017), a transformação Watershed é muito sensível a mínimas locais e a operação morfológica de fechamento não é suficiente pra otimizar a segmentação. O objetivo era tornar as áreas onde há copas de árvores planas, sendo necessário mais uma operação morfológica.

A imagem morfológicamente fechada é processada, utilizando um kernel no formato de disco plano, suavizando a imagem que será utilizada na segmentação (Figura 35).

Figura 35 – Comparação entre as Scan Lines de Cada Operação morfológica

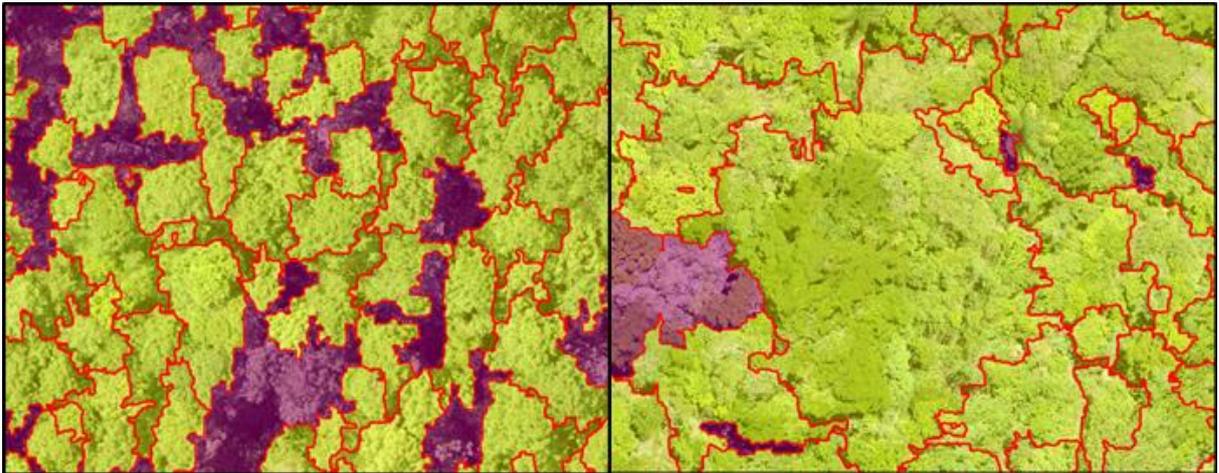


Fonte: elaborado pelo autor

A imagem morfológicamente aberta entra na cadeia de processos para realizar a segmentação por Watershed, que consiste na localização dos pontos mais baixo da bacia (imagem) e na segmentação da imagem. Para isso, a imagem é invertida e convertida para uma imagem binária utilizando um limiar que separa as copas das clareiras.

Esse produto passa por uma transformação de distância para, posteriormente, mapear e rotular cada elemento conectado na imagem transformada, que representam as copas das árvores. Por fim, as marcações e a imagem entram na ferramenta Watershed para gerar o produto segmentado (Figura 36).

Figura 36 – Resultados da Segmentação Watershed



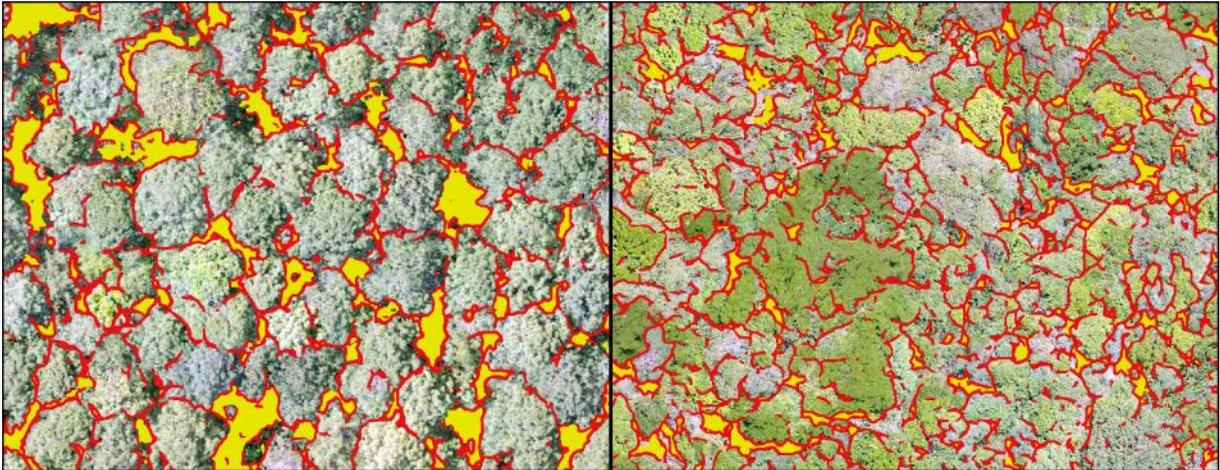
A imagem da esquerda é um recorte do *Pinus* e a da direita um recorte da Mata Ciliar. As regiões amareladas são os polígonos classificados como copas de árvores enquanto as regiões roxas são as áreas que não são copas. Fonte: elaborado pelo autor

5.4 U-NET

Após atingir um nível de acurácia e de perda considerados satisfatórios (0,9342 e 0,1816, respectivamente), o modelo treinado foi utilizado para predição do resto da área de estudo, a qual a rede neural não teve acesso enquanto treinava. Para isso, a imagem total da área de estudo foi cortada em partes menores para entrar no modelo criado.

A medida em que o modelo foi gerando os resultados para cada parte, essas foram sendo alocadas em memória para serem remontadas em uma única imagem contínua. Dessa forma, foi possível processar e organizar os resultados em duas imagens, que enquadravam as áreas de *pinus* e de mata ciliar, além de reinserir as informações geográficas originais (Figura 37).

Figura 37 – Resultado da Segmentação Utilizando a U-Net



Fonte: elaborado pelo autor

5.5 CONTAGEM DOS SEGMENTOS: ÍNDICE DE SØRENSEN-DICE E ÍNDICE DE JACCARD

Os resultados das segmentações foram recortados utilizando as 14 áreas onde as copas das árvores já haviam sido manualmente desenhadas, sendo 7 na área de *pinus* e 7 na mata ciliar. Essa decisão foi tomada para aproveitar um maior número de polígonos já vetorizados para análise. No entanto, não foi incluído na comparação a segmentação utilizando U-Net, pois, por conta da forma que a técnica foi aplicada, não foi possível criar polígonos conectados que representasse cada copa de árvore.

Para outras técnicas, os polígonos que representavam os elementos da imagem, ou já possuíam identificação própria no raster facilitando a conversão para dado vetorial, ou já eram resultados no formato vetoriais. Dessa forma, para cada uma das 14 amostras citadas, a quantidade de polígonos foi enumerada em cada um dos planos de informação das segmentações. Por fim, com o intuito de realizar uma comparação entre as metodologias, foi calculada a razão entre a área de cada amostra e o número de polígonos. Esses valores foram resumidos em função das técnicas conforme o apresentado na TABELA 1.

Tabela 1 – Razão entre a área da amostra e o número de polígonos por técnica de segmentação

	<i>Pinus</i>		Mata Ciliar	
	Total	Média	Total	Média
Verdade	95,60	13,66	51,06	7,29
Multiresolution	91,75	13,11	134,94	19,28
PCA_B1	198,62	28,37	169,93	24,28
PCA_B2	271,29	38,76	165,01	23,57
Watershed	271,93	38,85	634,51	90,64

O Total indica a soma de todos os polígonos encontrados nas 7 amostras. Assim, a média é esse total dividido pelas 7 amostras. Fonte: elaborado pelo autor

Os polígonos de cada segmentação também foram convertidos para rasters binários que indicavam as bordas e as áreas interna de cada polígono, com o objetivo de verificar a semelhança entre as segmentações. Para tanto, os polígonos foram convertidos para linhas, servindo como dados para a elaboração de um buffer de 15 centímetros (aproximadamente 3 pixels) que indicariam as bordas dos elementos segmentados, sejam ele copas (em sua maioria) ou não (e.g. clareiras).

Essas imagens binárias foram comparadas com a verdade terrestre para calcular o coeficiente Sørensen–Dice e coeficiente de Jaccard nas áreas de *Pinus* (Tabela 2) e de Mata Ciliar (Tabela 3). Assim, as informações das 14 amostras foram resumidas por técnica de segmentação utilizada, em função das duas grandes áreas. As tabelas apresentam, resumidamente, o quão semelhante os resultados da segmentação são da verdade terrestre nas duas áreas (7 amostras para cada).

Tabela 2 – Índices de Dice e Jaccard para a área de *Pinus*.

Segmentação	Média -	Desvio	Média -	Desvio
	Dice	Padrão	Jaccard	Padrão
PCA_B1	0,14516	0,01362	0,07831	0,00792
PCA_B2	0,1326	0,01971	0,07111	0,01129
Multiresolution	0,32252	0,01625	0,19236	0,01151
Watershed	0,32031	0,04008	0,19128	0,02885

Quanto mais próximo o valor for de 1.0, melhores são os resultados. Fonte: elaborado pelo autor

Tabela 3 – Índices de Dice e Jaccard para a área de Mata Ciliar.

Segmentação	Média -	Desvio	Média -	Desvio
	Dice	Padrão	Jaccard	Padrão
PCA_B1	0,16585	0,01904	0,09052	0,01126
PCA_B2	0,15773	0,01919	0,08572	0,01117
Multiresolution	0,27653	0,03315	0,16083	0,023
Watershed	0,25028	0,07023	0,14461	0,04582

Quanto mais próximo o valor for de 1.0, melhores são os resultados. Fonte: elaborado pelo autor

Os métodos Watershed e Multiresolution apresentaram os maiores valores tanto no índice de Jaccard quanto no Dice para a área de Pinus e na Mata Ciliar. Para verificar se essa diferença era significativa, foi realizado o teste estatístico de Kruskal-Wallis para identificar qual dos algoritmos eram diferentes foi realizada a comparação par a par de Nemenyi. A partir dos testes foi possível confirmar que havia diferença significativa entre os algoritmos, pois o teste de Kruskal-Wallis apontou que havia diferença significativa para a área de Pinus ($H=20.59$ e $p=0.0001$) e na Mata Ciliar ($H=16.66$ e $p=0.0008$). Por fim, a partir da Tabela 4 e 5 foi encontrado que havia diferença do Watershed e Multiresolution em comparação aos outros métodos e, não havia diferença significativa entre esses dois.

Tabela 4 - Comparações dos Índices de Jaccard no Pinus

	Multiresolution	PCA_B1	PCA_B2	Watershed
Multiresolution	1.000000	0.032935	0.006386	0.999754
PCA_B1	0.032935	1.000000	0.959073	0.042539
PCA_B2	0.006386	0.959073	1.000000	0.008732
Watershed	0.999754	0.042539	0.008732	1.000000

Tabela 5 - Comparação dos Índices de Jaccard na Mata Ciliar

	Multiresolution	PCA_B1	PCA_B2	Watershed
Multiresolution	1.000000	0.039105	0.006386	0.882629
PCA_B1	0.039105	1.000000	0.944125	0.228564
PCA_B2	0.006386	0.944125	1.000000	0.063716
Watershed	0.882629	0.228564	0.063716	1.000000

6 DISCUSSÃO

6.1 IMAGENS DO RPA

A alta qualidade da imagem torna visualmente perceptível as diferenças de características entre os dosséis das duas áreas de estudo que impactam diretamente a qualidade da segmentação. A área de pinus possui árvores com copas bem definidas e com distribuição homogênea no dossel, enquanto a área de mata ciliar demonstra uma maior heterogeneidade de tipos de copas de árvores seguido por intensa sobreposição das copas. Mesmo assim, foi possível delinear manualmente as copas das árvores. Além disso, é necessário ressaltar que a área de Pinus estudada não retrata exatamente uma área de plantio comercial, significando que o uso das imagens do RPA para a segmentação de copas com uma distribuição definida poderia resultar em um produto de maior qualidade.

Foi detectado ainda que a grande homogeneidade existente no dossel, não pensando em espécies de árvores, mas sim na textura da imagem nas áreas dificultou a criação do ortomosaico, pois, as imagens aéreas no centro das vegetações foram difíceis de alinhar diretamente no Agisoft METASHAPE, sendo necessários ajustes manuais. Esse problema pode estar relacionado com a altura (relativamente baixa) do voo e com a ausência de pontos de controle para referência na elaboração do ortomosaico.

Houve uma limitação no experimento também no que se refere ao uso de bandas do infravermelho próximo, decorrente de um problema no GPS da câmera MAPIR que, por uma falha de conectividade, não gravou as coordenadas geográficas nas imagens. Segundo KE e QUACKENBUSH (2011), a banda do infravermelho próximo poderia contribuir no processo de separar a área florestada da área não florestada, facilitando o processo de isolamento das copas para posterior segmentação. No entanto, por conta dessa dificuldade, não foi possível gerar o ortomosaico desse conjunto de dados, impedindo o uso desse plano de informação e de índices de vegetação na análise.

Apesar das dificuldades, as imagens RGB resultantes possuíam resolução alta o suficiente para viabilizar uma margem de reamostragem, para atender as demandas de cada metodologia de segmentação. A reamostragem se mostrou, na verdade, uma necessidade, pois, tanto em algumas etapas da metodologia usada no Watershed quanto no treinamento da rede neural, houveram problemas com falta de memória para imagens com 1,5 cm e 3 cm de resolução. No entanto, questiona-se a

necessidade de uma resolução elevada para o delineamento de copas, abrindo a necessidade de averiguar em que momento a resolução passa a ser um fator que limita a qualidade da segmentação.

Devido a limitações de recursos, não foi possível avaliar o impacto da variação da resolução nos resultados da segmentação, ou analisar a qualidade dos polígonos e comparar a economia de tempo de processamento. Além disso, segundo Campbell e Wynne (2011), o aumento da resolução espacial não necessariamente é sinônimo de qualidade, pois pode resultar no aumento da mistura dos pixels.

6.2 FILTRO DE MODA

A partir dos resultados encontrados nas tabelas 1, 2 e 3, é possível observar que o uso do PCA, aliado ao filtro de moda, apresentou as maiores limitações no processo de segmentação entre os métodos testados. Mesmo que os valores dos índices sejam sutilmente melhores na área de mata ciliar quando comparados com os resultados na área de *Pinus*, a contagem de polígonos é proporcionalmente inferior. Os polígonos gerados não correspondem bem às posições das copas e não apresentam um padrão de distinção de árvores.

Durante o processamento da imagem, foram feitas várias tentativas de melhorar o resultado: remoção do DSM da análise, teste com imagens de resoluções diferentes e inclusão da declividade no modelo. Porém, os testes não apresentaram uma melhora na segmentação.

Isso pode ser explicado pelo fato de o próprio PCA remover o efeito da redundância presente nas imagens combinadas, separando os componentes que melhor explicam a variação de informação. Ainda assim, essa variação não foi o suficiente para destacar as copas individuais, culminando na perda de maiores detalhes das bordas na homogeneização gerada pelo filtro de moda.

É possível que o PCA não tenha contribuído para a melhoria da qualidade da segmentação, efeito também percebido por Nordin et al. (2019). Nesse trabalho, o PCA foi aplicado em imagens hiperespectrais, como um processo intermediário de refinamento dos dados, para reduzir o número de bandas e separar as informações mais relevantes para serem utilizadas como entrada em uma segmentação Watershed, no entanto, não foram encontradas diferenças claras na qualidade do resultado com e sem o uso do PCA.

Já no processo metodológico descrito por Gomes e Ferreira (2007), os autores constataram que a técnica utilizada apresenta problemas para detectar árvores mais baixas, com forte correlação negativa com a verdade terrestre. Além disso, esses autores encontraram um bom potencial no filtro de moda para ajudar a individualizar a copa da árvore, pois reduz a redundância ao suavizar os pixels que representam a copa da árvore. Porém, o filtro também pode reduzir as bordas e, conseqüentemente, dificultar o delineamento da copa da árvore.

O filtro de moda, bem como filtros passa-baixa, são comumente usados como uma ferramenta intermediária para remoção e homogeneização de ruído para preparar uma imagem (SUN e HE, 2008; DI LEO, 2017; JING et al., 2012; NORDIN et al., 2019). Por fim, o uso do filtro de moda para a simplificação total da imagem pode acabar apenas removendo as informações das bordas das copas dentro de imagens que são naturalmente homogêneas. Dessa forma, as performances dos componentes foram baixos e próximos, indicando que as diferentes variações que cada componente sintetiza, podem não ter sido bem aproveitadas pelo filtro de moda.

Uma consideração importante é que a aplicação do PCA e do filtro de moda é simples e acessível, tornando a ferramenta de segmentação mais fácil de ser utilizada. Por mais que tenham sido utilizados dois softwares (ArcGIS e QGIS) para processar as imagens, as ferramentas são de simples execução e possuem baixo tempo de processamento. No entanto, mesmo com toda a praticidade, os resultados obtidos nessa pesquisa desencorajam o uso dessa ferramenta para problemas semelhantes.

6.3 MULTIREOLUTION

Como apresentado nos resultados, a segmentação multiresolution apresentou índices superiores em comparação ao PCA com filtro de moda (Tabela 2 e 3) e pouco superior, apesar de muito próxima, ao Watershed. Concomitantemente essa metodologia também resultou em um maior número de polígonos desenhados em ambas as áreas (TABELA 1) onde, no *Pinus*, se aproximou da verdade terrestre. Apesar disso, a área de mata ciliar se mostrou uma barreira maior para a segmentação da imagem, devido a elevada homogeneidade da área.

Ainda que a performance tenha sido superior, a partir da inspeção da imagem é possível detectar áreas onde uma única copa foi mais (super) ou menos (sub) segmentada. Esse efeito acaba sendo mais evidente nas áreas onde as copas estão

sobrepostas, dificultando detectar os limites de cada copa (Figura 24). Além disso, os parâmetros utilizados na ferramenta podem ter dificultado que copas maiores fossem segmentadas por um único polígono, resultando em um único indivíduo com vários segmentos (Figura 24).

Destaca-se que esse método de segmentação já vem implementado em softwares como Trimble eCognition e Spring e, apesar da diferença de performance entre os softwares, pode ser executado de forma prática e com pouco treinamento do usuário.

As dificuldades de execução estão relacionadas a ajustes dos parâmetros para a segmentação, que demandam tempo para serem corretamente definidos. Essa característica cria uma especificidade desses parâmetros para o conjunto de dados, tornando necessário reajustá-los de acordo com a área analisada. Além disso, a possibilidade de adicionar pesos para cada banda acaba adicionando mais parâmetros que impactam diretamente na qualidade da segmentação.

Oniishe e Ise (2018) alcançou uma segmentação satisfatória para seu objetivo, utilizando os parâmetros escolhidos na base da tentativa e erro, além de fazer revisão manual dos polígonos antes de executar a classificação das copas. Assim, mesmo com a definição de parâmetros ótimos para a área estudada, foram realizadas edições em alguns polígonos que continham mais de uma copa.

Há, no entanto, espaço para melhoria dos resultados obtidos por meio da segmentação multiresolution, a partir de mais testes acerca dos parâmetros utilizados ou da implementação de uma ferramenta que automatize o processo de ajuste e comparação dos polígonos criados com a verdade terrestre. Portanto, a performance desse método de segmentação, somada a possibilidade de aprimoramento dos resultados, a torna pertinente a futuras pesquisas.

6.4 WATERSHED

A segmentação watershed apresentou resultados nos índices de similaridade parecidos com os encontrados no multiresolution, no entanto, com uma contagem muito inferior na quantidade de polígonos desenhados. Foram criados apenas 8% dos polígonos totais na área de mata ciliar, valor extremamente baixo se comparado aos encontrados com qualquer outra metodologia, indicando a segmentação de múltiplos elementos em um único polígono.

Esse efeito pode ser explicado, novamente e em partes, pela alta homogeneidade da área de mata ciliar somado a três novos fatores: a dificuldade para traçar um limiar que conseguisse separar bem os valores dos pixels que representem os picos das copas; a influência da variação do relevo no DSM; e a existência de caminhos de areia que confundiam a segmentação do algoritmo após as imagens serem convertidas para tons de cinza.

O primeiro fator resulta da diferença na reflexão das cores nas copas das árvores, pois, nem todas são igualmente verdes. Equilibrar essas sutis diferenças sem definir um limiar para a operação binária que ignore as bordas em regiões onde as copas estão sobrepostas e com maior brilho, se provou um desafio.

No trabalho realizado por WANG et al (2004), foi utilizado um método de detecção de blob, visando detectar as bordas (por meio do Laplaciano do Gaussiano), antes de criar a imagem binária a ser utilizada no watershed, o que pode ter contribuído na segmentação ao indicar, preliminarmente, possíveis regiões que possuem copas. No entanto, os autores também encontraram problemas de inconsistências decorrentes da homogeneidade gerada pela imagem em tons de cinza, o que dificulta a detecção das bordas.

A kernel utilizada para a operação morfológica de fechamento pode não ter tido o desempenho esperado para reduzir o ruído das copas sem remover as bordas. Di Leo (2017) implementou a segmentação e os filtros utilizando o software MATLAB e, apesar da expectativa de que a implementação em Python geraria resultados parecidos, foram encontradas dificuldades que vão desde operações morfológicas idênticas gerando produtos diferentes até limitações quanto ao tipo de variável que elas suportam.

O segundo fator, mais perceptível na mata ciliar, deriva da variação do relevo, impactando na altura final das copas das árvores, pois, a região por onde o rio passa é mais baixa e, conseqüentemente, tende a abaixar a altura das copas. Para tentar reduzir esse efeito, foi considerado remover o DSM das operações, mas como o objetivo da pesquisa também era avaliar a facilidade em utilizar a ferramenta sem forçar grandes ajustes para cada área estudada, optou-se por manter o modelo ponderando um valor baixo para o mesmo.

Uma solução para esse problema estaria no uso da altura das árvores extraídas, a partir da diferença entre o modelo digital de terreno e o modelo digital de superfície, mas não foi possível testar esses dados pela dificuldade de extrair um

modelo de altura (DSM – DTM) a partir da técnica de fotogrametria, o que poderia ser superado por meio de um DEM obtido com GNSS de precisão e técnicas RTK/PPK (Real Time Kinematic/Post Processed Kinematic), ou então obtidos através de dados LiDAR. Além disso, outra possibilidade está na extração da altura dos indivíduos arbóreos utilizando uma série temporal de imagens coletadas pelo ARP, permitindo identificar a variação da estrutura de uma área vegetada.

ZAKI et al (2015), que também utilizaram a segmentação watershed para delinear as copas das árvores, detectaram que o delineamento de florestas com estrutura complexa é mais complexo do que em áreas com árvores de copas bem definidas, corroborando com os resultados aqui encontrados. Para facilitar o processo de segmentação, os autores também consideraram a possibilidade de utilizar um modelo de altura gerado via LiDAR.

Um exemplo de aplicação desse modelo foi implementado por Wan Mohd Jaafar et al (2018), obtendo uma taxa de acerto superior à 80% no resultado da segmentação gerada a partir do watershed. No entanto, além da execução de ajustes para refinar a qualidade da segmentação, houve uso de parâmetros manualmente configurados, o que caminha na contra mão de uma técnica mais genérica e prática para o delineamento das copas das árvores.

O terceiro fator está atrelado à dificuldade do algoritmo em identificar o plano de fundo (área sem árvore) da imagem, pois, além de ignorar incorretamente áreas que deveriam ser consideradas como copas, as áreas que pertenciam ao plano de fundo eram segmentadas. Assim, isolar todas as copas das árvores e evitar o uso do plano de fundo da imagem pode melhorar o resultado neste caso.

Destaca-se que houve uma relação diretamente proporcional entre o tempo de processamento da imagem e o tamanho da kernel resultante na operação morfológica. Além disso, a quantidade de memória RAM utilizada se tornou um teto para a operação, pois quando era utilizada a imagem inteira para a análise, o processo era interrompido por falta de memória RAM para kernels com tamanhos próximos e superior a 60.

Por essa razão, o processo de encontrar o tamanho das copas foi realizado em um recorte de toda área de estudo, facilitando não apenas a apresentação gráfica dos resultados, como também permitindo tamanhos maiores de kernel.

Uma alternativa que modifica os dados de entrada, e não os processos e funções utilizadas, está na reamostragem da imagem original para reduzir seu

tamanho. No entanto, é necessário verificar atentamente o impacto da reamostragem na segmentação da imagem para saber até onde é aceitável a perda de informações brutas prévias às que já ocorrem nos processos utilizados na metodologia. Outra questão extremamente pertinente é a dificuldade para executar essa segmentação, pois, ao todo, são quatro passos com diferentes parâmetros, apenas para o preparo da imagem. Ao incluir a dificuldade atrelada à implementação dessas ferramentas e do próprio watershed na linguagem Python, abre-se questionamentos quanto a viabilidade do seu uso para pessoas sem o devido treinamento na linguagem.

Em um contexto de monitoramento ambiental, onde sobram áreas e faltam recursos, implementar a segmentação watershed desde o início, dentro de uma linguagem de programação, pode não ser a abordagem ideal para atender as demandas de profissionais que precisam de resultados rápidos e não possuem domínio em computação. No entanto, essa situação não é melhor com o uso das ferramentas já implementadas no SIG. Foram testadas ferramentas de segmentação Watershed do Orfeu e SAGA, no QGIS, antes de tentar uma abordagem no Python e, por mais que exista uma interface mais acessível, ainda há uma série de parâmetros que devem ser ajustados para melhorar a qualidade da segmentação. Além disso, a documentação não é detalhada o suficiente para facilitar um entendimento pleno dessas ferramentas, diferente do que ocorre no com Python. Outra questão é a rigidez de alguns recursos no SIG, como por exemplo, as operações morfológicas em tons de cinza não permitiam a personalização dos elementos estruturantes não planos.

Ainda assim, é necessário ressaltar a importância de fazer um uso integrado do Python com o Sistema de Informação Geográfica (SIG), porque as possibilidades vão além da segmentação, indo em direção a automação de processos e implementação de ferramentas para resolver problemas específicos. Considerar a capacitação profissional ou formação de parcerias com pessoal treinado pode ser um procedimento fundamental para a gestão de áreas verdes.

6.5 U NET

Os resultados não permitiram a adequada segmentação das copas, isto é, não geraram uma delimitação clara das copas das árvores, pois, os segmentos não apresentaram conectividade. Mesmo com o alto valor de acurácia e baixo valor de perda, não é possível, e nem correto, utilizar esses valores para fazer uma comparação direta com os outros algoritmos de segmentação testados. Isso ocorre

porque os dois índices utilizados dentro do modelo foram calculados de forma diferente, utilizando nas comparações dos resultados da segmentação a matriz tridimensional que representa as categorias das máscaras utilizadas para o treinamento e validação. No entanto, por mais que não tenha sido possível alcançar o objetivo de separar cada indivíduo, foi encontrado um claro delineamento entre as áreas de dossel e clareiras.

Ao realizarmos uma inspeção visual, no SIG, da sobreposição do modelo gerado pela U-Net e a imagem RGB, é possível perceber uma alta qualidade na identificação de muitas áreas. Além disso, é perceptível uma grande diferença na quantidade e qualidade de clareiras encontradas ao comparar o resultado do watershed com o da U-Net.

O limiar, que separa as copas das clareiras/não copas no watershed, depende diretamente do valor do pixel, e não é gerado a partir de uma interpretação da região da imagem. Assim, o problema do watershed, já citado, derivado dos erros em áreas que possuíam alto valor no pixel, não ocorreu na rede neural. Inclusive, abre-se a possibilidade de experimentar combinar as duas ferramentas, ou seja, utilizar o resultado da U-Net para separar o dossel das clareiras/não copas para depois utilizar o watershed para a segmentação mais fina.

O problema da não conectividade das bordas ocorreu em áreas onde o modelo errou na classificação. No caso, os erros na classe “borda de copa” se mostraram o fator importante e prejudicial ao objetivo do estudo. No fim, é mais aceitável que, quando o modelo erre, ele erre as outras classes do que deixe de tentar desenhar as bordas na imagem.

É possível ainda que a influência da falta de amostras de treinamento e a falta de pesos para dar mais importância às bordas das copas tenham contribuído para o resultado encontrado. Além disso, a limitação pode estar atrelada ao fato método utilizado realizar uma segmentação semântica e, apesar da ideia de colocar a borda da copa como uma classe da segmentação, está pode não ser a solução ideal se comparada a uma abordagem que faz uso da segmentação por instância, como a utilizada por BRAGA et. al. (2020) para segmentar copas de árvores. Outra questão é que a mudança de textura das copas na área estudada é muito sutil, o que pode ter dificultado no processo de diferenciação das copas. Porém, a resposta real para esse problema adentra um campo da ciência que estuda o motivo pelo qual a rede neural escolhe um determinado resultado. Afinal, as informações estão escondidas dentro

da camada oculta da rede neural, nas milhares/milhões de conexões e parâmetros distribuídos dentro do modelo, reforçando o título negativo de caixa preta dado às redes neurais profundas (Qiu and JENSEN, 2004). Além da imprevisibilidade, as redes neurais sofrem de suscetibilidade a ataques que forçam uma previsão errônea, colocando ainda mais sua confiabilidade em risco (PAPERNOT et al. 2017).

Em relação a dificuldade de implementação, há duas considerações: a primeira é a facilidade que existe em utilizar um modelo definido e conhecido, pois este muito provavelmente estará disponível em algum repositório online ou sua estrutura poderá ser facilmente implementada a partir da sua descrição. Essa característica facilita para operadores que não são especialistas em redes neurais fazer uso dessas ferramentas sem ter que se aprofundar em toda a complexidade por trás da teoria.

Já a segunda consideração surge da dificuldade que o operador enfrentará em adequar o conhecimento de programação ao nível necessário, pois não é trivial o manuseio e ajuste necessário para os dados. Essa situação se agrava quando o problema possui características que fogem aos tipos de análises comumente utilizados pela comunidade, pois demandará habilidade do operador para o modelo ao problema ou tempo até uma biblioteca que atenda a necessidade seja criada e divulgada.

A especificidade do problema ao qual o modelo de rede neural foi empregado também é um fator a ser considerado. Como já citado, a segmentação de copas não é trivial visto a homogeneidade dessas áreas. Assim, o modelo poderia ter um desempenho superior caso fosse aplicado em áreas mais heterogêneas e com vegetação mais esparsa. Brandt et al (2020), realizou um estudo utilizando redes neurais para segmentação de copas na Savanna, obtendo excelentes resultados que permitiram mapear centenas de milhões de árvores em regiões hiperáridas, áridas e semiáridas, reafirmando o potencial da combinação entre aprendizagem profunda e imagens de resolução espacial muito alta.

Outra possibilidade de aplicação está no mapeamento de clareiras, o que resulta em informações úteis para estudos da estrutura e dinâmica de florestas. Um exemplo foi explorado por Dalagnol et al (2021), que mostrou ser possível utilizar as informações das clareiras para estimar as mudanças na estrutura no dossel das florestas relacionadas a árvores quebradas ou removidas.

Por fim, os dados utilizados nos SIG podem ser úteis para pesquisas em redes neurais, pois é comum haver imagens de satélites devidamente rotuladas (e.g. uso e

ocupação) e validadas por operadores humanos, condições interessantes para o treinamento da rede.

Ao considerar as qualidades e limitações de todas as técnicas utilizadas, é possível perceber que a segmentação utilizando multiresolution apresentou o melhor resultando juntamente com a menor dificuldade para implementação. No entanto, faz-se necessário considerar as potencialidades existentes no uso das CNN para a segmentação das imagens, visto que com o avanço do poder computacional e otimização dos algoritmos será possível realizar implementações de modelos de redes neurais mais robustos e versáteis.

7 CONCLUSÃO

A partir das análises realizadas foi possível perceber que os modelos de segmentação utilizados precisam de muitos ajustes para o seu funcionamento e, ainda assim, haverá imprecisões no delineamento de áreas verdes homogêneas. Por razões que vão desde imprecisão de limitações do algoritmo até alta sensibilidade de parâmetros, não foi possível criar uma segmentação com alto grau de acerto.

A segmentação utilizando o filtro de moda gerou o pior resultado dentre os estudados, com elevadas imprecisões na forma e quantidade de polígonos delineados. A segmentação multiresolution e o Watershed geraram resultados imprecisos, mas consideravelmente melhores do que a citada anteriormente, oferecendo indícios de uma detecção de bordas mais coerentes à distribuição das copas. Apesar de exaustivos teste, ainda há espaço para realizar mais ajustes nos parâmetros dessas segmentações até alcançar um valor ótimo.

A segmentação usando um modelo de rede neural de aprendizagem profunda não gerou o resultado esperado, o que não permitiu sua comparação dentro dos parâmetros estabelecidos. No entanto, apresentou características que indicam o potencial da ferramenta para aplicações em SIG, em dados mais heterogêneos e/ou com maior quantidade de imagens para treinamento e validação.

Há uma clara possibilidade de incluir mais procedimentos de pré-processamento nas imagens, ou até mesmo uma segmentação intermediária para separar o dossel do plano de fundo, antes de segmentar as copas. Além disso, a adição de outras camadas de informações, como o NIR, deve ser experimentado, pois

o comportamento espectral das copas resultaria em destaques na imagem que poderiam facilitar a identificação dos elementos da imagem.

Percebeu-se também a necessidade de conhecimento técnico mais aprofundado para o uso de alguns segmentadores, pois muitos processos se tornaram personalizáveis em um ambiente como o do Python, possibilitando executar filtros e operações com parâmetros não disponíveis diretamente nas plataformas de SIG.

Conclui-se que as imagens de alta resolução coletadas por RPAs permitem uma boa detecção, por humanos, das copas das árvores, situação claramente oposta à grande parte das imagens de satélites de resolução mais grosseira. Além disso, a praticidade para a coleta das informações é grande e depende do operador. No entanto, dependendo do objetivo do usuário, quantidade de imagens coletadas e processamentos realizados, o custo computacional pode se tornar um desafio. Condição que ficou evidente para a segmentação, que precisa ser melhor explorada e mais otimizada em imagens de alta resolução.

Portanto, visando facilitar o processo de segmentação, torna-se interessante encontrar um limiar de resolução espacial que permita realizar a segmentação das copas com alta qualidade sem tornar o processo computacionalmente custoso; experimentar outros métodos de implementação de redes neurais, como por exemplo segmentação por instância; realizar experimentos utilizando imagem de infravermelho próximo ou tecnologias como LiDAR para incorporar mais informações sobre as características das copas das árvores. Esses aprofundamentos poderiam resultar em uma segmentação mais coesa com a realidade e, conseqüentemente, aumentar a confiança na automação desse processo.

REFERÊNCIAS

AASEN, Helge; BURKART, Andreas; BOLTEN, Andreas; *et al.* Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 108, p. 245–259, 2015.

APARNA, P; MP, Harshita; SOUNDER, Harshitha; *et al.* CNN Based Technique for Automatic Tree Counting Using Very High Resolution Data. *In: 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C)*. Bangalore, India: IEEE, 2018, p. 127–129. Disponível em: <<https://ieeexplore.ieee.org/document/8437104/>>. Acesso em: 3 jan. 2019.

BAATZ, Martin; SCHÄPE, Arno. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. **Proceedings of Angewandte Geographische Informationsverarbeitung XII**, p. 12–23, 2000.

BARRETT, E. C.; CURTIS, D. E. **Introduction to environmental remote sensing**. 1. ed. London: Routledge, 1999.

BECKLINE, Mukete. Perspectives of Remote Sensing and GIS Applications in Tropical Forest Management. *American Journal of Agriculture and Forestry*, v. 5, n. 3, p. 33, 2017.

BEUCHER, S. THE WATERSHED TRANSFORMATION APPLIED TO IMAGE SEGMENTATION. **Scanning microscopy-supplement**, p. 27, 1992.

BHAVANA, B. L.; SRIDEVI, N.; HEBBAR, Dr. Ramachandrea. Tree Crown Detection and Extraction from High Resolution Satellite Images in an Urban Area. *In: 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C)*. Bangalore, India: IEEE, 2018, p. 8–11. Disponível em: <<https://ieeexplore.ieee.org/document/8436619/>>. Acesso em: 12 mar. 2019.

BRAGA, G., José R.; PERIPATO, Vinícius; DALAGNOL, Ricardo; *et al.* Tree Crown Delineation Algorithm Based on a Convolutional Neural Network. **Remote Sensing**, v. 12, n. 8, p. 1288, 2020.

BRANDT, Martin; TUCKER, Compton J.; KARIRYAA, Ankit; *et al.* An unexpectedly large count of trees in the West African Sahara and Sahel. **Nature**, v. 587, n. 7832, p. 78–82, 2020.

BUSLAEV, Alexander; IGLOVIKOV, Vladimir I; KHVEDCHENYA, Eugene; *et al.* Alumentations: fast and flexible image augmentations. **Information**, v. 11, n. 2, p. 125, 2020.

CALVARIO, Gabriela; SIERRA, Basilio; ALARCÓN, Teresa; *et al.* A Multi-Disciplinary Approach to Remote Sensing through Low-Cost UAVs. **Sensors**, v. 17, n. 6, p. 1411, 2017.

CAMPANILI, Maura; SCHÄFFER, Wigold Bertolo. **Mata Atlântica: manual de adequação ambiental**. Brasília: MMA/SBF, 2010. (Biodiversidade, 35).

CAMPBELL, James B.; WYNNE, Randolph H. **Introduction to Remote Sensing**. 5. ed. [s.l.]: Guilford Publications, 2011.

CSILLIK, Ovidiu; CHERBINI, John; JOHNSON, Robert; *et al.* Identification of Citrus Trees from Unmanned Aerial Vehicle Imagery Using Convolutional Neural Networks. **Drones**, v. 2, n. 4, p. 39, 2018.

DALAGNOL, Ricardo; WAGNER, Fabien H.; GALVÃO, Lênio S.; *et al.* Large-scale variations in the dynamics of Amazon forest canopy gaps from airborne lidar data and opportunities for tree mortality estimates. **Scientific Reports**, v. 11, n. 1, p. 1388, 2021.

DANIEL, Kai; DUSZA, Bjoern; LEWANDOWSKI, Andreas; *et al.* AirShield: A system-of-systems MUAV remote sensing architecture for disaster response. *In*: **2009 3rd Annual IEEE Systems Conference**. Vancouver, BC, Canada: IEEE, 2009, p. 196–200. Disponível em: <<http://ieeexplore.ieee.org/document/4815797/>>. Acesso em: 3 jan. 2019.

DE ALMEIDA, Danilo R. A.; STARK, Scott C.; VALBUENA, Ruben; *et al.* A new era in forest restoration monitoring. **Restoration Ecology**, v. 28, n. 1, p. 8–11, 2020.

DEPARTAMENTO DE CIÊNCIAS AMBIENTAIS. **PROJETO PEDAGÓGICO DO CURSO BACHARELADO EM GESTÃO E ANÁLISE AMBIENTAL**. São Carlos: [s.n.], 2018. Disponível em: <<http://www.prograd.ufscar.br/cursos/cursos-oferecidos-1/gestao-e-analise-ambiental/gestao-e-analise-ambiental-projeto-pedagogico-2019>>. Acesso em: 30 dez. 2020.

DEPARTAMENTO DE COMPUTAÇÃO. **BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO - PROJETO PEDAGÓGICO**. São Carlos: [s.n.], 2019. Disponível em: <<http://bcc.dc.ufscar.br/wp-content/uploads/2019/01/PPC-do-BCC-vFinal.pdf>>. Acesso em: 30 dez. 2020.

DEY, Vivek; ZHANG, Yun; ZHONG, Ming. A review on image segmentation techniques with remote sensing perspective. *In*: Vienna, Austria: IAPRS, 2010, v. XXXVIII, p. 12.

DI LEO, Eugenio. **Individual Tree Crown detection in UAV remote sensed rainforest RGB images through Mathematical Morphology**. Mestrado, University of Amsterdam, 2017.

DIAZ-DELGADO, Ricardo; LUCAS, Richard; HURFORD, Clive. **The Roles of Remote Sensing in Nature Conservation**. 1. ed. New York, NY: Springer Berlin Heidelberg, 2017.

DILLENBURG, Angélica Karina. A IMPORTÂNCIA DO MONITORAMENTO AMBIENTAL NA AVALIAÇÃO DA QUALIDADE DE UM RIO – ESTUDO DE CASO - MERCEDES, PR. **Revista Urutágua**, v. 12, p. 10, 2007.

DUFOUR, S.; BERNEZ, I.; BETBEDER, J.; *et al.* Monitoring restored riparian vegetation: how can recent developments in remote sensing sciences help? **Knowledge and Management of Aquatic Ecosystems**, n. 410, p. 10, 2013.

ERIKSON, Mats. **Segmentation and Classification of Individual Tree Crowns**. Acta Universitatis Agriculturae Sueciae, Suécia, 2004.

ESCOLA DE ENGENHARIA DE SÃO CARLOS. **PROJETO POLÍTICO-PEDAGÓGICO: CURSO DE GRADUAÇÃO EM ENGENHARIA AMBIENTAL**. São Carlos: [s.n.], 2017. Disponível em: <<http://www5.eesc.usp.br/coc-ambiental/files/PPP%20-%20Ambiental%20-%20atualizacao%20-%202017.pdf>>. Acesso em: 30 dez. 2020.

FARIA, Rodrigo Ribeiro de; COSTA, Marledo Egidio. A inserção dos veículos aéreos não tripuláveis (drones) como tecnologia de monitoramento no combate ao dano ambiental. **Revista Ordem Pública**, v. 8, n. 1, p. 81–103, 2015.

FEARNSIDE, Philip M. Desmatamento na Amazônia: dinâmica, impactos e controle. **Acta Amazonica**, v. 36, n. 3, p. 395–400, 2006.

FISHER, Robert; PERKINS, Simon; WALKER, Ashley; *et al.* Image processing learning resources. **Hypermedia Image Precessing Reference**, 2003. Disponível em: <[https://homepages.inf.ed.ac.uk/rbf/HIPR2/gryimage.htm#:~:text=A%20grayscale%20\(or%20graylevel\)%20image,be%20provided%20for%20each%20pixel.](https://homepages.inf.ed.ac.uk/rbf/HIPR2/gryimage.htm#:~:text=A%20grayscale%20(or%20graylevel)%20image,be%20provided%20for%20each%20pixel.)>. Acesso em: 5 jan. 2021.

FITZ, Paulo Roberto. **Geoprocessamento sem complicação**. 1. ed. São Paulo, Brasil: Oficina de Textos, 2008.

FUYI, Tan; BOON CHUN, Beh; MAT JAFRI, Mohd Zubir; *et al.* Land cover/use mapping using multi-band imageries captured by Cropcam Unmanned Aerial Vehicle Autopilot (UAV) over Penang Island, Malaysia. *In*: CARAPEZZA, Edward M.; WHITE, Henry J. (Orgs.). Edinburgh, United Kingdom: [s.n.], 2012, p. 85400S. Disponível em: <<http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.974911>>. Acesso em: 15 fev. 2019.

GOMES, Priscila Brochado; FERREIRA, Marcos Cesar. Estimation of tree crown number in a quickbird image using an image processing method. *In*: **2007 IEEE International Geoscience and Remote Sensing Symposium**. Barcelona, Spain: IEEE, 2007, p. 4354–4355. Disponível em: <<http://ieeexplore.ieee.org/document/4423816/>>. Acesso em: 12 mar. 2019.

GONZALEZ, Rafael C.; WOODS, Richard E. **Digital Image Processing**. 2. ed. [s.l.]: Prentice Hall, 2002.

GRANT, Lois. Diffuse and specular characteristics of leaf reflectance. **Remote Sensing of Environment**, v. 22, n. 2, p. 309–322, 1987.

GRÜN, Armin. Potential and limitations of highresolution satellite imagery. **ETH Zurich**, p. 13, 2000.

GUYOT, Gérard; GUYON, Dominique; RIOM, Jacques. Factors affecting the spectral response of forest canopies: A review. **Geocarto International**, v. 4, n. 3, p. 3–18, 1989.

HASSLER, Márcio Luís. A importância das Unidades de Conservação no Brasil. **Sociedade & Natureza**, v. 17, n. 33, 2005.

HINTON, Geoffrey E.; SRIVASTAVA, Nitish; KRIZHEVSKY, Alex; et al. Improving neural networks by preventing co-adaptation of feature detectors. arXiv:1207.0580 [cs], 2012. Disponível em: <<http://arxiv.org/abs/1207.0580>>. Acesso em: 14 jul. 2021.

IBGE, INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Mapa de Biomas e de Vegetação do Brasil. Disponível em: <<http://www.ibge.gov.br/>>.

JENSEN, John R. **Introductory Digital Image Processing: A Remote Sensing Perspective**. 4. ed. [s.l.]: Prentice Hall, 2015.

JING, Linhai; HU, Baoxin; NOLAND, Thomas; et al. An individual tree crown delineation method based on multi-scale segmentation of imagery. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 70, p. 88–98, 2012.

JORGE, Lúcio André de Castro; INAMASU, Ricardo Yassushi Inamasu. Uso de veículos aéreos não tripulados (VANT) em Agricultura de Precisão. In: **Agricultura de precisão: Resultados de um Novo Olhar**. [s.l.]: Embrapa, 2014, p. 109–134.

KAUR, Dilpreet; KAUR, Yadwinder. Various Image Segmentation Techniques: A Review. p. 6, 2014.

KE, Yinghai; QUACKENBUSH, Lindi J. FOREST SPECIES CLASSIFICATION AND TREE CROWN DELINEATION USING QUICKBIRD IMAGERY. In: Tampa, Florida: [s.n.], 2007, p. 7–11.

KE, Yinghai; QUACKENBUSH, Lindi J. A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. **International Journal of Remote Sensing**, v. 32, n. 17, p. 4725–4747, 2011.

KOC-SAN, Dilek; SELIM, Serdar; ASLAN, Nagihan; et al. Automatic citrus tree extraction from UAV images and digital surface models using circular Hough transform. **Computers and Electronics in Agriculture**, v. 150, p. 289–301, 2018.

KRONKA, Francisco José Nascimento; NALON, Marco Aurélio; MATSUKUMA, Ciro Koiti; et al. Monitoramento da vegetação natural e do reflorestamento no Estado de São Paulo. **XII Simpósio Brasileiro de Sensoriamento Remoto - SBSR**, p. 1569–1576, 2005.

LÄNGKVIST, Martin; KISELEV, Andrey; ALIREZAIE, Marjan; et al. Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks. **Remote Sensing**, v. 8, n. 4, p. 329, 2016.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015.

LI, Weijia; FU, Haohuan; YU, Le; et al. Deep Learning Based Oil Palm Tree Detection and Counting for High-Resolution Remote Sensing Images. **Remote Sensing**, v. 9, n. 1, p. 22, 2016.

LIMA, Raul Sampaio De. **Fatores de Variação das Estimativas de Precipitação Interna em uma Floresta Secundária**. Mestrado, Universidade Federal de São Carlos, São Carlos/SP, 2017.

LISEIN, Jonathan; PIERROT-DESEILLIGNY, Marc; BONNET, Stéphanie; *et al.* A Photogrammetric Workflow for the Creation of a Forest Canopy Height Model from Small Unmanned Aerial System Imagery. **Forests**, v. 4, n. 4, p. 922–944, 2013.

LOVETT, Gary M.; BURNS, Douglas A.; DRISCOLL, Charles T.; *et al.* Who needs environmental monitoring? **Frontiers in Ecology and the Environment**, v. 5, n. 5, p. 253–260, 2007.

MARMANIS, D.; SCHINDLER, K.; WEGNER, J.D.; *et al.* Classification with an edge: Improving semantic image segmentation with boundary detection. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 135, p. 158–172, 2018.

MARMANIS, D.; WEGNER, J. D.; GALLIANI, S.; *et al.* SEMANTIC SEGMENTATION OF AERIAL IMAGES WITH AN ENSEMBLE OF CNNs. **ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. III–3, p. 473–480, 2016.

MATHWORKS. Structuring Elements. Disponível em: <https://se.mathworks.com/help/images/structuring-elements.html>. Acesso em: 20 out. 2020.

MORAES, Douglas Rafael; FERREIRA NETO, Luiz; COSTA, Merilene do Socorro; *et al.* Monitoramento de áreas embargadas por desmatamento ilegal. **Revista Brasileira de Cartografia**, v. 70, n. 5, p. 1595–1617, 2018.

MYNENI, Ranga B.; KEELING, C. D.; TUCKER, C. J.; *et al.* Increased plant growth in the northern high latitudes from 1981 to 1991. **Nature**, v. 386, n. 6626, p. 698–702, 1997.

NAGENDRA, Harini; ROCCHINI, Duccio. High resolution satellite imagery for tropical biodiversity studies: the devil is in the detail. **Biodiversity and Conservation**, v. 17, n. 14, p. 3431–3442, 2008.

NEVES, Ana Carolina Marzullo. **DETERMINANTES DO DESMATAMENTO NA MATA ATLÂNTICA: Uma Análise Econômica**. Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro, 2006.

NIXON, MS; AGUADO, AS. Basic image processing operations. **Feature Extraction & Image Processing for Computer Vision**. Elsevier, p. 83–136, 2012.

NORDIN, Sitinor Atikah; ABD LATIF, Zulkiflee; OMAR, Hamdan. Individual tree crown segmentation in tropical peat swamp forest using airborne hyperspectral data. **Geocarto International**, v. 34, n. 11, p. 1218–1236, 2019.

ONISHI, Masanori; ISE, Takeshi. Automatic classification of trees using a UAV onboard camera and deep learning. **arXiv preprint arXiv:1804.10390**, p. 9, 2018.

PAESE, Adriana. **Caracterização e Análise Ambiental do Campus da Universidade Federal de São Carlos (UFSCar)**, São Carlos, SP. Mestrado, Universidade Federal de São Carlos, São Carlos/SP, 1997.

PAPERNOT, Nicolas; MCDANIEL, Patrick; GOODFELLOW, Ian; *et al.* Practical Black-Box Attacks against Machine Learning. *In: Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*. Abu Dhabi United Arab Emirates: ACM, 2017, p. 506–519. Disponível em: <<https://dl.acm.org/doi/10.1145/3052973.3053009>>. Acesso em: 5 set. 2020.

QIU, F; JENSEN, J R. Opening the black box of neural networks for remote sensing image classification. p. 22, 2004.

RAWAT, Waseem; WANG, Zenghui. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*, v. 29, n. 9, p. 2352–2449, 2017.

REN, Yun; ZHU, Changren; XIAO, Shunping. Small Object Detection in Optical Remote Sensing Images via Modified Faster R-CNN. **Applied Sciences**, v. 8, n. 5, p. 813, 2018.

ROCHA, Genival Fernandes; FERREIRA, Laerte Guimarães; FERREIRA, Nilson Clementino; *et al.* DETECÇÃO DE DESMATAMENTOS NO BIOMA CERRADO ENTRE 2002 E 2009: PADRÕES, TENDÊNCIAS E IMPACTOS. **Revista Brasileira de Cartografia**, v. 63, n. 3, p. 9, 2012.

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-Net: Convolutional Networks for Biomedical Image Segmentation. **arXiv:1505.04597 [cs]**, 2015. Disponível em: <<http://arxiv.org/abs/1505.04597>>. Acesso em: 4 set. 2020.

ROSS, Jurandyr. **Ecogeografia do Brasil**. 1. ed. [s.l.]: Oficina de Textos, 2006.

SANTOS, Alexandre Rosa dos; PELUZIO, Telma Machado de oliveira; SAITO, Nathália Suemi. **Spring 5.1. 2: Passo a passo: Aplicações Práticas**. Alegre - Espírito Santo: [s.n.], 2010.

SEELAN, Santhosh K; LAGUETTE, Soizik; CASADY, Grant M; *et al.* Remote sensing applications for precision agriculture: A learning community approach. **Remote Sensing of Environment**, v. 88, n. 1–2, p. 157–169, 2003.

SEGARAN, Toby. **Programming Collective Intelligence**. [s.l.]: O'Reilly Media, Inc., 2007.

SHAPIRO, L; STOCKMAN, G. Computer vision prentice hall. **Inc., New Jersey**, 2001.

SILVA, Adriana Aparecida; MIZIARA, Fausto. AVANÇO DO SETOR SUCROALCOOLEIRO E EXPANSÃO DA FRONTEIRA AGRÍCOLA EM GOIÁS. **Pesquisa Agropecuária Tropical**, v. 41, n. 3, 2011. Disponível em: <<http://www.revistas.ufg.br/index.php/pat/article/view/11054>>. Acesso em: 4 set. 2020.

- SILVA, J.B.L.; FERREIRA, W.L.; ALMEIDA, K.N.S.; *et al.* Evolução Temporal do Desmatamento e Expansão Agrícola entre 1984 a 2010 na Sub-Bacia do Rio Uruçuí-Preto, Piauí. **Revista Engenharia na Agricultura - REVENG**, v. 22, n. 3, p. 254–261, 2014.
- SOARES-FILHO, Britaldo Silveira; NEPSTAD, Daniel Curtis; CURRAN, Lisa; *et al.* Cenários de desmatamento para a Amazônia. **Estudos Avançados**, v. 19, n. 54, p. 137–152, 2005.
- SOILLE, Pierre. Morphological phase unwrapping. **Optics and lasers in Engineering**, v. 32, n. 4, p. 339–352, 1999.
- SUN, Ying; HE, Guo-jin. Segmentation of High-Resolution Remote Sensing Image Based on Marker-Based Watershed Algorithm. *In: 2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery*. Jinan Shandong, China: IEEE, 2008, p. 271–276. Disponível em: <<http://ieeexplore.ieee.org/document/4666397/>>. Acesso em: 4 set. 2020.
- SZELISKI, Richard. Image processing. *In: Computer Vision*. [s.l.]: Springer, 2011, p. 87–180.
- TANG, Lina; SHAO, Guofan. Drone remote sensing for forestry research and practices. **Journal of Forestry Research**, v. 26, n. 4, p. 791–797, 2015.
- TATEOSIAN, Laura. **Python For ArcGIS**. Raleigh, North Carolina: Springer, 2015.
- THOMAZELLA, R.; CASTANHO, J. E.; DOTTO, F.R.L.; *et al.* Environmental Monitoring Using Drone Images and Convolutional Neural Networks. *In: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*. Valencia: IEEE, 2018, p. 8941–8944. Disponível em: <<https://ieeexplore.ieee.org/document/8518581/>>. Acesso em: 3 jan. 2019.
- TOMS, Silas. **ArcPy and ArcGIS–Geospatial Analysis with Python**. Birmingham: Packt Publishing Ltd, 2015.
- TOTH, Charles; JÓZKÓW, Grzegorz. Remote sensing platforms and sensors: A survey. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 115, p. 22–36, 2016.
- WAN MOHD JAAFAR, Wan; WOODHOUSE, Iain; SILVA, Carlos; *et al.* Improving Individual Tree Crown Delineation and Attributes Estimation of Tropical Forests Using Airborne LiDAR Data. **Forests**, v. 9, n. 12, p. 759, 2018.
- WANG, Le; GONG, Peng; BIGING, Gregory S. Individual Tree-Crown Delineation and Treetop Detection in High-Spatial-Resolution Aerial Imagery. **Photogrammetric Engineering & Remote Sensing**, v. 70, n. 3, p. 351–357, 2004.
- WOLFGANG, Ertel. **Introduction to artificial intelligence**. 2. ed. Germany: Springer, 2017.
- YAKUBOVSKIY, Pavel. Segmentation models. **GitHub repository**, 2019. Disponível em: <https://github.com/qubvel/segmentation_models>.

YU, Xingrui; WU, Xiaomin; LUO, Chunbo; *et al.* Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework. **GIScience & Remote Sensing**, v. 54, n. 5, p. 741–758, 2017.

ZAHAWI, Rakan A.; DANDOIS, Jonathan P.; HOLL, Karen D.; *et al.* Using lightweight unmanned aerial vehicles to monitor tropical forest recovery. **Biological Conservation**, v. 186, p. 287–295, 2015.

ZAKI, Nurul Ain Mohd; LATIF, Zulkiflee Abd; ZAINAL, Mohd Zainee; *et al.* Individual tree crown (ITC) delineation using watershed transformation algorithm for tropical lowland dipterocarp. In: 2015 **International Conference on Space Science and Communication (IconSpace)**. Langkawi, Malaysia: IEEE, 2015, p. 237–242. Disponível em: <<http://ieeexplore.ieee.org/document/7283795/>>. Acesso em: 5 maio 2021.

ZHANG, Jianming; JIN, Xiaokang; SUN, Juan; *et al.* Spatial and semantic convolutional features for robust visual object tracking. *Multimedia Tools and Applications*, v. 79, n. 21–22, p. 15095–15115, 2020.

ZHANG, Liangpei; ZHANG, Lefei; DU, Bo. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. **IEEE Geoscience and Remote Sensing Magazine**, v. 4, n. 2, p. 22–40, 2016.

ZHANG, Xiaoyang; FRIEDL, Mark A.; SCHAAF, Crystal B.; *et al.* Monitoring vegetation phenology using MODIS. **Remote Sensing of Environment**, v. 84, n. 3, p. 471–475, 2003.

ZHANG, ZhaoXuan. **Native Vegetation Classification Using Remote Sensing Techniques: A Case Study of Dairy Flat Regrowth Bush by Using the AUT Unmanned Aerial Vehicle**. Auckland University of Technology, 2014.

ZHU, Xiao Xiang; TUIA, Devis; MOU, Lichao; *et al.* Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. **IEEE Geoscience and Remote Sensing Magazine**, v. 5, n. 4, p. 8–36, 2017.