



Serviço Público Federal  
Ministério da Educação  
**Fundação Universidade Federal de Mato Grosso do Sul**  
Faculdade de Engenharias, Arquitetura e Urbanismo e Geografia - FAENG



**AURIMAR DA COSTA LIMA FILHO**

**MAPEAMENTO DE ÁREAS PERMEÁVEIS  
NA BACIA HIDROGRÁFICA DO CÓRREGO BANDEIRA  
USANDO APRENDIZADO PROFUNDO**

**P G R N**

**Campo Grande, MS.**

**2022**

**UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL  
FACULDADE DE ENGENHARIAS E ARQUITETURA E URBANISMO E  
GEOGRAFIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM RECURSOS NATURAIS**

**AURIMAR DA COSTA LIMA FILHO**

**MAPEAMENTO DE ÁREAS PERMEÁVEIS  
NA BACIA HIDROGRÁFICA DO CÓRREGO BANDEIRA  
USANDO APRENDIZADO PROFUNDO**

Dissertação submetida ao Programa de Pós-graduação em Tecnologias Ambientais da Universidade Federal do Mato Grosso do Sul como requisito parcial a obtenção do título de Mestre em Recursos Naturais. Geoprocessamento Aplicado.

**Orientador: Prof. Dr. Jamil Alexandre Ayach Anache  
Coorientador: Prof. Dr. José Marcato Júnior**

Aprovada em:

**Banca Examinadora:**

Prof. Dr. Jamil Alexandre Ayach Anache  
Orientador PGRN- UFMS

Prof. Dra. Ana Paula Marques Ramos  
UNOESTE

Prof. Dra. Camila Leonardo Mioto  
UFR

Prof. Dr. Roberto Macedo Gamarra  
UFMS

Prof. Dr. Paulo Tarso Sanches de Oliveira  
UFMS

**Campo Grande, MS.  
2022**

## AGRADECIMENTOS

Agradeço primeiramente a Deus por ter me dado paciência, dedicação, sabedoria e discernimento para realizar este estudo.

Agradeço a minha mãe Elizabeth da Glória França Lima e minha irmã Laíza França da Costa Lima que são pessoas mais importantes na minha vida, por terem me dado apoio, força nos momentos mais difíceis. São a razão por estar fazendo este trabalho.

Aos meus amigos e namorada que me ajudaram e que sempre estavam pertos quando precisava.

Ao orientador Prof. Dr. Jamil Alexandre Ayach Anache e coorientador Prof. Dr. José Marcato Júnior, sempre atenciosos, pacientes e presentes quando precisava.

A Prof<sup>ª</sup>. Dr<sup>ª</sup>. Ana Paula Marques Ramos, Prof<sup>ª</sup>. Dr<sup>ª</sup>. Camila Leonardo Miotto, Prof. Dr. Roberto Macedo Gamarra e Prof. Dr. Paulo Tarso Sanches de Oliveira por avaliarem este estudo e participação da banca.

Agradeço ao Programa de Pós-Graduação em Recursos Naturais (PPGRN), da Faculdade de Engenharias, Arquitetura e Urbanismo (FAENG) da Fundação Universidade Federal de Mato Grosso do Sul (UFMS).

Agradeço a CAPES pelo acesso ao Portal de Periódicos.

O presente trabalho foi realizado com apoio da Fundação Universidade Federal de Mato Grosso do Sul - UFMS/MEC - Brasil.

## SUMÁRIO

LISTA DE FIGURAS.....	VI
LISTA DE TABELAS.....	VIII
RESUMO.....	IX
ABSTRACT .....	X
<b>1 INTRODUÇÃO .....</b>	<b>11</b>
<b>2 OBJETIVO .....</b>	<b>14</b>
2.1 OBJETIVO GERAL.....	14
2.2 OBJETIVO ESPECÍFICOS .....	14
<b>3 REVISÃO BIBLIOGRÁFICA.....</b>	<b>15</b>
3.1 CLASSIFICAÇÃO SUPERVISIONADA POR <i>PIXELS</i> .....	15
3.2 DEEP LEARNING – APRENDIZADO PROFUNDO.....	16
<b>3.2.1 Rede Neural Convolutacional (CNN) .....</b>	<b>20</b>
3.2.1.1 Rede totalmente convolutacional (FCN) .....	25
3.2.1.2 U-Net .....	26
<b>3.2.2 ResNet.....</b>	<b>28</b>
3.3 CLASSIFICAÇÃO DE USO E COBERTURA DO SOLO .....	29
<b>3.3.1 Áreas permeáveis.....</b>	<b>32</b>
<b>4 METODOLOGIA .....</b>	<b>35</b>
4.1 ÁREA DE ESTUDO .....	35
4.2 COLETA E PRODUÇÃO DE DADOS.....	38
<b>4.2.1 Preparação e manipulação das imagens.....</b>	<b>38</b>
<b>4.2.2 Rotulação dos dados de treinamento.....</b>	<b>41</b>
<b>4.2.3 Exportação de dados e treinamento de modelo de aprendizado profundo.....</b>	<b>41</b>
<b>4.2.4 Classificação dos <i>pixels</i> usando aprendizado profundo.....</b>	<b>43</b>
<b>4.2.5 Avaliação de desempenho dos modelos .....</b>	<b>43</b>
<b>5 RESULTADOS E DISCUSSÃO .....</b>	<b>46</b>
5.1 TREINAMENTO DE MODELO DE APRENDIZADO PROFUNDO .....	46
5.2 MAPEAMENTO DAS ÁREAS PERMEÁVEIS NA BACIA HIDROGRÁFICA DO CÓRREGO BANDEIRA.....	51

5.2.1	Erros de classificação .....	57
6	CONCLUSÃO .....	64
7	REFERÊNCIAS BIBLIOGRÁFICAS .....	65
	APÊNDICES .....	74

## LISTA DE FIGURAS

<b>Figura 1.</b> Estrutura CNN (MENEZES, 2017). .....	23
<b>Figura 2.</b> Arquitetura U-Net (Adaptado de WAGNER et al., 2019). .....	27
<b>Figura 3.</b> Localização Bacia Hidrográfica do Córrego Bandeira. ....	36
<b>Figura 4.</b> Representação gráfica das formações do Cerrado (COUTINHO, 1978). .....	37
<b>Figura 5.</b> Áreas selecionadas para rotulação e treinamento. Imagens aéreas cedidas pela PMCG. ....	40
<b>Figura 6.</b> Áreas selecionadas para rotulação e treinamento. Imagens orbitais WorldView-2, cedidas pelo IMASUL. ....	40
<b>Figura 7.</b> Curvas de perda para treinamento e validação (ESRI, 2021). ....	45
<b>Figura 8.</b> Rotulação das imagens escolhidas na área de estudo. Imagens aéreas cedidas pela PMCG (10 cm de resolução espacial). .....	46
<b>Figura 9.</b> Rotulação das imagens escolhidas na área de estudo. Imagens orbitais WorldView-2, cedidas pelo IMASUL (50 cm de resolução espacial). ....	47
<b>Figura 10.</b> Resultados modelo de avaliação, empregando a imagem aérea (ESRI <i>ArcGIS Pro</i> , 2021). .....	48
<b>Figura 11.</b> Resultado modelo de avaliação, empregando imagem orbital WorldView-2 (ESRI <i>ArcGIS Pro</i> , 2021). .....	49
<b>Figura 12.</b> Área de estudo empregando a imagem aérea, resolução espacial 10 cm. ....	52
<b>Figura 13.</b> Área de estudo empregando imagem orbital WorldView-2, resolução espacial 50 cm. ....	53
<b>Figura 14.</b> Classificação por <i>pixels</i> na área de estudo, empregando a imagem aérea, resolução espacial 10 cm. ....	54
<b>Figura 15.</b> Classificação por <i>pixels</i> na área de estudo, empregando imagem orbital WorldView-2, cedida pelo IMASUL. ....	55
<b>Figura 16.</b> Erro de classificação na imagem aérea (resolução espacial 10 cm). Região do Campus da UFMS. ....	59
<b>Figura 17.</b> Erro de classificação na imagem aérea (resolução espacial 10 cm), próximo a nascente do Córrego Bandeira. ....	60
<b>Figura 18.</b> Erro de classificação na imagem orbital (resolução espacial 50 cm). Região do Campus da UFMS. ....	61

<b>Figura 19.</b> Erro de classificação na imagem orbital (resolução espacial 50 cm). Região do Rádio Clube Campo. ....	<b>62</b>
<b>Figura 20.</b> Erro de classificação na imagem orbital (resolução espacial 50 cm). Região da Lagoa Itatiaia. ....	<b>63</b>

## LISTA DE TABELAS

<b>Tabela 1.</b> Modelo de tabela de atributos dos arquivos <i>shapefiles</i> .....	<b>41</b>
<b>Tabela 2.</b> Resultado de área da classificação por <i>pixels</i> da imagem aérea (resolução espacial 10 cm). .....	<b>51</b>
<b>Tabela 3.</b> Resultado de área da classificação por <i>pixels</i> da imagem orbital (resolução espacial 50 cm). .....	<b>51</b>

## RESUMO

LIMA FILHO, A. C. (2022). Mapeamento de Áreas Permeáveis na Bacia Hidrográfica do Córrego Bandeira Usando Aprendizado Profundo. 2022. 75. Dissertação - Programa de Pós Graduação em Recursos Naturais. Universidade Federal de Mato Grosso do Sul, Brasil.

O *Deep Learning* é uma subárea de aprendizagem de máquina, aplicado para solucionar tarefas que envolvem processamento de imagens, visão computacional, processamento de sinais e processamento de linguagem natural. Ele consiste em uma coleção de algoritmos como o subconjunto de aprendizado da máquina (*machine learning*), especializado em aprendizagem hierárquica de conceitos de grandes dados (*big data*). O objetivo deste trabalho é identificar as áreas permeáveis da Bacia Hidrográfica do Córrego Bandeira, utilizando imagens aéreas e orbitais e métodos *Deep Learning*. Nos últimos anos, muitos estudos estão usando esta ferramenta para processar imagens de sensoriamento remoto, assim como mostrou ser eficiente em processar imagens tanto ópticas (hiperespectrais e multiespectrais) quanto de radar, apresentando resultados acurados no mapeamento de diferentes tipos de cobertura do solo. Identificar e mapear as áreas permeáveis proveniente e ou inseridas em áreas urbanas, empregando o sensoriamento remoto, é relevante para estudos hidrológicos e microclima, bem como para proteção, monitoramento e manejo das mesmas. Utilizando a rede U-Net e *backbone* ResNet-34 na imagem aérea com resolução espacial de 10 cm, a arquitetura alcançou a classificação com 55 épocas (*epochs*), acurácia de 0,84 e F1-Score de 0,88, tal como a imagem orbital, com resolução espacial de 50 cm, utilizando os mesmos modelos de arquitetura, atingiu a classificação com 27 épocas (*epochs*), acurácia de 0,796 e F1-Score de 0,86. Comparando as classificações, observa-se que a classificação empregando imagem de resolução espacial 10 cm, houve segmentação, rotulação e identificação significativa. Enquanto a classificação utilizando imagem de resolução espacial 50 cm, houve apenas segmentação, rotulação e identificação de áreas de grandes escalas. Nota-se que, as resoluções espaciais interferiram nos resultados de treinamento e classificação. Por mais que ambas são conceituadas como imagens de alta resolução, a imagem orbital, com resolução de 50 cm, demonstrou não ser adequada para este tipo de estudo, na qual conquistou uma classificação insatisfatória.

**Palavras-chave:** deep learning, sensoriamento remoto, cobertura do solo, áreas permeáveis, classificação de imagens de satélite.

## ABSTRACT

Deep Learning is a sub-area of machine learning, applied to solve tasks involving image processing, computer vision, signal processing and natural language processing. It consists of a collection of algorithms as the machine learning subset, specializing in hierarchical learning of big data concepts. The objective of this work is to identify the permeable areas of the Córrego Bandeira Watershed, using aerial and orbital images and Deep Learning methods. In recent years, many studies are using this tool to process remote sensing images, as it has shown to be efficient in processing both optical (hyperspectral and multispectral) and radar images, presenting accurate results in mapping different types of land cover. Identifying and mapping permeable areas from and/or inserted in urban areas, using remote sensing, is relevant for hydrological and microclimate studies, as well as for their protection, monitoring and management. Using the U-Net network and the ResNet-34 backbone in the aerial image with a spatial resolution of 10 cm, the architecture achieved the classification with 55 epochs (epochs), an accuracy of 0.84 and an F1-Score of 0.88, such as the orbital image, with a spatial resolution of 50 cm, using the same architectural models, reached the classification with 27 epochs (epochs), accuracy of 0.796 and F1-Score of 0.86. Comparing the classifications, it is observed that the classification using image of spatial resolution 10 cm, there was segmentation, labeling and significant identification. While the classification using 50 cm spatial resolution image, there was only segmentation, labeling and identification of large-scale areas. Note that the spatial resolutions interfered in the training and classification results. As much as both are conceptualized as high resolution images, the orbital image, with a resolution of 50 cm, proved to be unsuitable for this type of study, in which it achieved an unsatisfactory rating.

**Keywords:** deep learning, remote sensing, land cover, permeable areas, satellite image classification.

## 1 INTRODUÇÃO

Um grupo de pesquisadores teve a percepção de criar uma rede com a intenção de substituir ou evoluir sistemas e equipamentos que são manufaturados, bem como facilitar os projetos que são elaborados à mão. Esta rede possui uma tecnologia de aprendizado que potencializa muitos aspectos da sociedade moderna, desde pesquisas realizadas na web até filtragem de conteúdo em redes sociais e sites de comércio eletrônico. Por ser uma tecnologia de última geração, está cada vez mais presente em produtos como câmeras e smartphones (LECUN, BENGIO, HINTON, 2015).

Esse tipo de sistema de aprendizado é aplicado para identificar objetos em imagens, transcrever a fala em texto, combinar notícias, postagens ou produtos com o interesse do usuário gerando resultados relevantes de pesquisa. Essas aplicações fazem uso de uma classe de técnicas denominada de *Deep Learning* (LECUN, BENGIO, HINTON, 2015).

O *Deep Learning* é uma ferramenta para solucionar tarefas que envolvem reconhecimento de fala e facial, recomendações de notícias, processamento de imagens e entre outras funções. Nos últimos anos, muitos estudos estão usando-o para processar imagens de sensoriamento remoto, apresentando resultados acurados no mapeamento de diferentes tipos de cobertura do solo (KUSSUL et al., 2017).

Em muitos estudos que efetuam trabalhos semelhantes à classificação de cobertura do solo, são utilizadas imagens de satélites com baixa resolução espacial, como RapidEye, Landsat e CBERS, tendo como objetivo realizar análises de solos rurais e urbanos. Nessas obras também foi empregado o *Normalized Difference Vegetation Index* (NDVI), que em sua tradução significa Índice de Vegetação por Diferenças Normalizadas. Embora a ferramenta NDVI seja útil e de fácil uso para quantificar a vegetação, não é possível detectar pequenos detalhes devido à resolução das imagens

(MORENO-ARMENDÁRIZ et al., 2019). Outras técnicas tradicionais de sensoriamento remoto também foram e são empregadas para mapeamento de cobertura do solo.

Os métodos tradicionais empregados em sensoriamento remoto vêm sendo substituídos pelo *Deep Learning* (DL). Esse método tem a função de identificar e classificar elementos em imagens de sensoriamento remoto, por meio de programação e algoritmos desenvolvidos.

No contexto urbano, imagens com maior resolução espacial possibilitam identificar detalhes não vistos em imagens com menor resolução, como as mencionadas anteriormente. Para suprir essa demanda, atualmente se têm imagens orbitais com resolução submétrica (30 cm), além da possibilidade de obtenção de imagens com ainda maior resolução, a se destacar as obtidas a partir de plataformas aéreas tripuladas e não tripuladas (VANT – Veículos Aéreos Não-Tripulados).

A classificação de cobertura do solo com imagens de sensoriamento remoto desempenha um papel importante em muitas aplicações, como gestão de recursos, agricultura de precisão, planejamento urbano e proteção ambiental (TONG et al., 2020). Fora isso, também são encontrados em estudos na área de hidrologia urbana.

Classificar e mapear a vegetação urbana é uma tarefa ampla e técnica, pois a vegetação é o fator-chave e desempenha um papel substancial para a sobrevivência de todas as entidades vivas. Como a população está crescendo a uma taxa alarmante, a competição pelo sustento aumenta. Portanto, monitorar a mudança da cobertura do solo nos ajudaria a expandir a cobertura vegetal para atender às necessidades diárias (NIJHAWAN et al., 2018). Assim sendo, este projeto tem a finalidade de aplicar o método de *Deep Learning* para que se possa identificar e classificar as referidas áreas permeáveis por meio da utilização de imagens aéreas e orbitais. Com isso, demonstrando uma técnica que servirá para planejamento, monitoramento e proteção ambiental. Entretanto, existem

muitas variações de técnicas em *Deep Learning* e de dados de sensoriamento remoto disponíveis. Dessa forma, é relevante buscar a melhor solução computacional para obtenção das classificações de cobertura do solo urbano utilizando imagens de alta resolução.

## 2 OBJETIVO

### 2.1 OBJETIVO GERAL

O objetivo deste trabalho é identificar as áreas permeáveis da Bacia Hidrográfica do Córrego Bandeira, utilizando imagens aéreas e orbitais e métodos *Deep Learning*.

### 2.2 OBJETIVO ESPECÍFICOS

- Obter e processar imagens aéreas e orbitais de alta resolução espacial;
- Identificar amostras de treinamento dentro da bacia hidrográfica em estudo;
- Caracterizar o desempenho do modelo de *Deep Learning* aplicado em imagens de diferentes resoluções espaciais para mapear áreas permeáveis;
- Identificar as áreas permeáveis de toda área de estudo (Bacia Hidrográfica do Córrego Bandeira) utilizando técnica de *Deep Learning* treinada e validada para as condições locais.

### 3 REVISÃO BIBLIOGRÁFICA

#### 3.1 CLASSIFICAÇÃO SUPERVISIONADA POR *PIXELS*

A classificação supervisionada é frequentemente usada em métodos de classificação baseados em *pixels* (ZHANG et al., 2020). Ela redivide a imagem, estatisticamente em classes baseadas nas áreas de treinamento (training samples) fornecidas ao sistema. Algoritmos são utilizados para classificar o *pixel* (os seus *digital numbers*) de acordo com sua máxima semelhança a cada uma das diferentes classes criadas durante o processo. Durante a classificação, os *pixels* com informações duvidosas podem ser classificados posteriormente. Para que o sistema possa realizar a classificação ele deve ser “treinado”. Este “treinamento” consiste no fornecimento ao sistema amostras dos padrões que devem ser reconhecidos. Estas amostras podem ser obtidas em campo ou por meio de outras fontes, como imagens aéreas (PARANHOS FILHO et al., 2016).

No entanto, o método de classificação baseado em *pixels* tem duas grandes limitações: o primeiro é o problema dos *pixels* mistos em que os recursos de várias classes são apresentados em um único *pixel*. Geralmente ocorre na classificação de imagens de baixa e média resolução. Os problemas de *pixels* mistos podem ser resolvidos por meio do aumento da resolução espacial. A segunda limitação fundamental é que o contexto espacial dos *pixels* circundantes não é usado na classificação. Especialmente quando se lida com uma imagem de alta resolução, as diferenças nas características de textura dentro de uma área de objeto podem resultar nas diferenças entre *pixels* nessa área, reduzindo assim a precisão da classificação baseada em *pixels*. Isso é denominado de “efeito sal e pimenta” (*salt and pepper effect*) (ZHANG et al., 2020).

Apesar de que o olho humano seja um dispositivo na distinção de cores e padrões em uma imagem, o cérebro humano não tem a capacidade de processar o mesmo volume de informações na celeridade em que os sistemas digitais de processamento de

imagens possuem. Em contrapartida, não podemos afirmar que os sistemas de processamento digital de imagens possam ser sempre mais eficazes que o operador humano. Isso se deve por conta da sua experiência, além de sua visão. Um exemplo disto é a análise de imagens de áreas urbanas, pois os sistemas automáticos tradicionais têm dificuldade em reconhecer a textura urbana (PARANHOS FILHO et al., 2016). A análise das imagens é de grande importância em muitas aplicações práticas, como classificação de cobertura e uso da terra ou alteração e detecção de objetos (ZHU et al., 2017).

### 3.2 DEEP LEARNING – APRENDIZADO PROFUNDO

O *Deep Learning* é um poderoso instrumento para solucionar tarefas que envolvem processamento de imagens, visão computacional, processamento de sinais e processamento de linguagem natural (KUSSL et al., 2017) e consiste em uma coleção de algoritmos como o subconjunto de aprendizado da máquina (*machine learning*), especializado em aprendizagem hierárquica de conceitos de grandes dados (*big data*) (CAO; DRAGIĆEVIĆ; LI, 2019). Com base nos avanços recentes, o *deep learning* está provando ser um conjunto de ferramentas muito bem-sucedidas, às vezes capazes de superar até mesmo os humanos na resolução de tarefas altamente computacionais (ZHU et al., 2017). Aprendizado profundo é um aprendizado baseado em algoritmos nas redes neurais. Uma rede neural compreende neurônios ou unidade com certa ativação  $\alpha$  e parâmetros  $\theta = \{W, \gamma\}$ . Modelos de aprendizado profundo (redes) são compostos de muitas camadas que transformam dados de entrada (imagens) em saída (categorias) enquanto aprendem recursos de alto nível progressivamente (MA et al., 2019).

As unidades que não estão na camada de entrada ou saída são convencionalmente chamadas de unidades ocultas. As camadas ocultas podem ser vistas como distorção da entrada de forma não linear para que as categorias se tornem linearmente separáveis pela última camada (LECUN; BENGIO; HINTON, 2015). Uma

rede neural contendo múltiplas camadas ocultas são normalmente consideradas como uma rede neural “profunda”, daí o termo “aprendizado profundo” (MA et al., 2019).

Os métodos de aprendizagem profunda são métodos de aprendizagem de representação com múltiplos níveis de representação, obtidos compondo módulos simples, mas não lineares, que cada um transforma a representação em um nível (começando pela entrada bruta) em uma representação em um nível mais alto, levemente um nível mais abstrato. Com a composição de tais transformações suficientes, funções muito complexas podem ser aprendidas. O aprendizado de representação é um conjunto de métodos que permite que uma máquina seja alimentada com dados brutos e descubra automaticamente as representações necessárias para detecção ou classificação (LECUN; BENGIO; HINTON, 2015).

As técnicas convencionais de aprendizagem de máquina foram limitadas em sua capacidade de processar dados naturais em sua forma bruta. Durante décadas, a construção de um sistema de reconhecimento de padrões ou aprendizado de máquina exigiu engenharia cuidadosa e considerável experiência de domínio para projetar um extrator de recursos que transformasse os dados brutos (como os valores de *pixel* de uma imagem) em uma representação interna adequada ou vetor de recurso a partir do qual o subsistema de aprendizagem, muitas vezes um classificador, poderia detectar ou classificar padrões de entrada (LECUN; BENGIO; HINTON, 2015).

A opção convencional é projetar à mão extratores de bons recursos, o que requer uma quantidade considerável de habilidade de engenharia e experiência de domínio. Mas tudo isso pode ser evitado se bons recursos puderem ser aprendidos automaticamente usando um procedimento de aprendizagem de propósito geral. Esta é a principal vantagem do aprendizado profundo (LECUN; BENGIO; HINTON, 2015).

Para tarefas de classificação, camadas mais elevadas de representação amplificam aspectos da entrada que são importantes para a discriminação e suprimem variações irrelevantes. Uma imagem, por exemplo, vem na forma de uma matriz de valores de *pixel* e os recursos aprendidos da primeira camada de representação normalmente representam a presença ou ausência de bordas em orientações particulares e locais específicos na imagem. A segunda camada normalmente detecta motivos identificando arranjos específicos de bordas, independentemente de pequenas variações nas posições das bordas. A terceira camada pode montar motivos em combinações maiores que correspondem a partes de objetos familiares e camadas subsequentes detectariam objetos como combinações dessas partes. O aspecto chave do aprendizado profundo é que essas camadas de recursos não são projetadas por engenheiros humanos: elas são aprendidas com dados usando um procedimento de aprendizagem de propósito geral (LECUN; BENGIO; HINTON, 2015).

Nos últimos anos, muitos estudos estão usando esta ferramenta para processar imagens de sensoriamento remoto, assim como mostrou ser eficiente em processar imagens tanto óticas (hiperespectrais e multiespectrais) quanto de radar, apresentando resultados acurados no mapeamento de diferentes tipos de cobertura do solo (KUSSL et al., 2017).

De acordo com Yuan et al. (2020), os *softwares* de *Deep Learning* em imagens detectadas remotamente são distintos daquelas imagens naturais. As imagens detectadas remotamente possuem diversos padrões mais complicados, assim como informações espectrais de espaço-temporal. A rede neural (NN) foi agregada no sensoriamento remoto ambiental e aplicada em muitos aspectos, tais como mapeamento de cobertura de solo, recuperação de parâmetros ambientais, fusão de dados e redução de escala, construção e previsão de informações.

Como mencionado anteriormente, o *Deep Learning* foi recentemente incorporado no mapeamento da cobertura do solo e atingiu ótimos resultados, em consequência das suas superioridades na obtenção de recursos em múltiplas escalas e múltiplos níveis. Comparado aos métodos tradicionais, o método de classificação baseado em *Deep Learning* tem vantagens significativas em termos de precisão de classificação, especialmente em áreas urbanas complexas (YUAN et al., 2020).

Embora a rede neural tenha boa universalidade, uma única estrutura de rede não consegue resolver todos os problemas de sensoriamento remoto, com isso muitas estruturas foram desenvolvidas (YUAN et al., 2020). As arquiteturas de aprendizagem profunda são capazes de extrair características de alto nível, hierárquicas e abstratas, que geralmente são mais robustas ao processamento não linear (ZHU et al., 2017). Dentre as arquiteturas de *Deep Learning*, a Rede Neural Convolutacional (CNN) é mais utilizada para classificação de imagens de sensoriamento remoto (ALEM, KUMAR, 2020; BEN HAMIDA et al., 2018; BHOSLE, MUSANDE, 2019; CAO, DRAGIĆEVIĆ, LI, 2019; CAO et al., 2020; CASTELLUCCIO et al., 2015; CHEN et al., 2018a; CHEN et al., 2018b; CHENG et al., 2018; CHENG et al., 2020; CHENG, HAN, LU, 2017; DU et al., 2019; HUANG, ZHAO, SONG, 2018; KAMPPFMEYER, SALBERG, JENSSEN, 2017; KATTENBORN, EICHEL, FASSNACHT, 2019; KUSSUL et al., 2017; LI et al., 2018; LIN et al., 2020; LIU et al., 2019; LONG, SHELFHAMER, DARRELL, 2015; MA et al., 2019; MORENO-ARMENDÁRIZ et al., 2019; NIJHAWAN et al., 2017; ORTEGA ADARME et al., 2020; OSCO et al., 2021; SCOTT et al., 2017; TONG et al., 2020; TORRES et al., 2020; WAGNER et al., 2019; WEN et al., 2017; WENG et al., 2017; YU et al., 2017; YUAN et al., 2020; ZHANG et al., 2018; ZHANG et al., 2019; ZHANG et al., 2020; ZHAO, DU, EMERY, 2017; ZHU et al., 2017;)

### 3.2.1 Rede Neural Convolutacional (CNN)

Como uma das principais redes neurais profundas, a Rede Neural Convolutacional foi projetada para processar e analisar dados sensoriais ou imagens em larga escala levando em consideração à sua característica estacionária em escalas locais e globais (ZHANG et al., 2019).

As redes neurais convolucionais tornaram-se a abordagem mais comum para diferentes tarefas de análise de imagem, como classificação automática, detecção de objetos e segmentação semântica. Recentemente, as CNNs têm sido amplamente aplicadas para problemas de sensoriamento remoto que alcançam o estado da arte em muitas aplicações (TORRES et al., 2020).

A segmentação semântica é a tarefa de atribuir um rótulo de classe a cada *pixel* na imagem. Assim, a segmentação semântica tem o potencial de capturar a forma e o tamanho do objeto com mais precisão do que a detecção de objetos, o que pode ser essencial em muitas aplicações (TORRES et al., 2020).

Dentro da Rede Neural Convolutacional, camadas convolucionais e camadas de agrupamento são conectadas alternativamente para generalizar os recursos para apresentações profundas e abstratas. Normalmente, as camadas convolucionais são compostas de pesos e vieses que são aprendidos por intermédio de um conjunto de fragmentos de imagens por meio da imagem. Esses pesos são compartilhados por diferentes mapas de características, nos quais múltiplos recursos são aprendidos com uma quantidade reduzida de parâmetros e uma função de ativação, por exemplo, unidades lineares retificadas (ReLU), é seguida para fortalecer a não linearidade das operações convolucionais (ZHANG et al., 2019)

Embora as CNNs tenham sido introduzidas há muitos anos, somente nos últimos anos a implementação e o treinamento de grandes CNNs se tornou possível

devido tanto ao progresso tecnológico quanto aos resultados da pesquisa. Os principais fatores de capacitação têm sido o rápido crescimento do poder de computação acessível, especialmente das unidades de processamento gráfico (GPUs) e a difusão de grandes conjuntos de dados de imagens rotuladas para treinamento. Por outro lado, a pesquisa contribuiu com muitas novas soluções para um aprendizado mais rápido e confiável, como a unidade linear retificada (ReLU), um neurônio com uma não-linearidade simplificada que permite um treinamento muito mais rápido e técnicas para reduzir o excesso de ajuste, como o abandono ou o aumento de dados (CASTELLUCCIO *et al.*, 2015).

Uma imagem é organizada como uma matriz de valores de *pixels*. Dada uma imagem de entrada aleatória, a Rede Neural Convolutiva reconhece a classe a que pertence e os valores de probabilidade de que a entrada pertence a cada classe (Figura 6). Existem quatro tipos básicos de camadas que compõem uma Rede Neural Convolutiva e seus papéis são os seguintes (CAO; DRAGIĆEVIĆ; LI, 2019):

- *Camadas convolucionais*: são as mais importantes. Elas computam a convolução da imagem de entrada com os pesos da rede. Os neurônios, na primeira camada oculta, visualizam apenas uma pequena janela de imagem e aprendem características de baixo nível. Aqueles em camadas mais profundas visualizam (indiretamente) porções maiores da imagem e são capazes de aprender características mais expressivas por meio da combinação de características de baixo nível. Cada camada é caracterizada por alguns hiperparâmetros: o número de filtros a serem aprendidos, seu suporte espacial, a passada entre as diferentes janelas e um opcional de zero-preenchimento que controla o tamanho da saída da camada (CASTELLUCCIO *et al.*, 2015). As saídas são chamadas de mapas de recursos (CAO; DRAGIĆEVIĆ; LI, 2019). Resumindo, as primeiras camadas

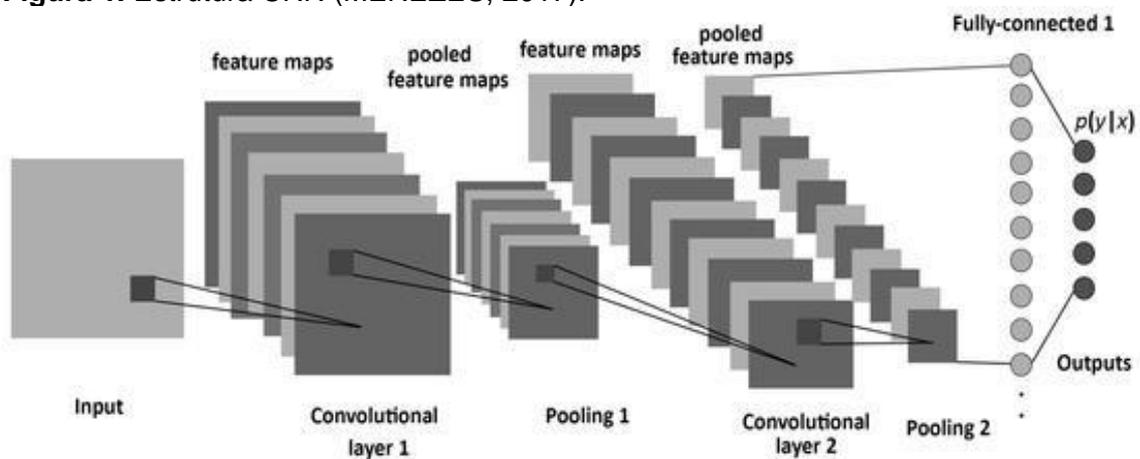
geralmente capturam características de baixo nível (como bordas, linhas e cantos) enquanto as camadas mais profundas são capazes de aprender características mais expressivas (como estruturas, objetos e formas) combinando as de baixo nível (CHENG; HAN; LU, 2017).

- *Camadas de agrupamento*: é uma camada de amostragem baixa, com o objetivo de diminuir o tamanho dos mapas de recursos e reduzir o custo computacional (CAO; DRAGIĆEVIĆ; LI, 2019). Reduzir o tamanho da camada de entrada por meio de algumas operações locais não lineares, por exemplo  $max()$ , de modo a reduzir o número de parâmetros a serem aprendidos e fornecer alguma invariância de tradução. Os hiperparâmetros mais relevantes são o suporte da janela de agrupamento e o passo entre diferentes janelas (CASTELLUCCIO *et al.*, 2015). Normalmente, após cada camada convolucional, existem camadas de agrupamento que são criadas computando alguma operação não linear local de um recurso específico sobre uma região da imagem. Esse processo garante que o mesmo resultado possa ser obtido, mesmo quando os recursos de imagem possuem pequenas traduções ou rotações, o que é muito importante para classificação de detecção de cena (CHENG; HAN; LU, 2017).
- *Camadas de normalização* (CASTELLUCCIO *et al.*, 2015) ou *camadas de unidades lineares retificadas* (ReLU) (CAO; DRAGIĆEVIĆ; LI, 2019): inspiradas por esquemas de inibição presentes nos neurônios reais do cérebro, visam a melhorar a generalização. Elas são tipicamente usadas com neurônios sigmóides (não ReLU) (CASTELLUCCIO *et al.*, 2015). Segundo Cao, Dragićević e Li (2019) as camadas executam uma função não linear (como sigmoide) na entrada. Introduzir a não linearidade ao sistema pode melhorar a eficiência computacional sem perder muita precisão. Já Cheng, Han e Lu (2017), mencionam

que essas camadas visam melhorar a generalização inspirada em esquemas de inibição apresentados nos neurônios reais do cérebro.

- *Camadas totalmente conectadas:* são normalmente usadas como as últimas camadas da rede. Ao remover as restrições, eles podem resumir melhor as informações transmitidas pelas camadas de nível inferior, tendo em vista a decisão final. Apesar da conectividade total, sua complexidade ainda é acessível graças às camadas de tamanho anterior (CASTELLUCCIO *et al.*, 2015).

**Figura 1.** Estrutura CNN (MENEZES, 2017).



Exemplificando, as camadas convolucionais funcionam como extratores de características hierárquicas, as camadas de agrupamento conduzem a queda espacial dos mapas de recursos, enquanto a camada totalmente conectada serve como classificadora para gerar as probabilidades de classificação preditiva dos dados de entrada (TONG *et al.*, 2020).

Devido a poderosa capacidade de aprender representações de características de alto nível (mais abstrata e semanticamente significativas), redes neurais convolucionais profundas estão sendo exploradas em sistemas de detecção de objetos em contraste com os métodos mais tradicionais seguidos por um classificador baseado em recursos artesanais (ZHU *et al.*, 2017).

O treinamento da CNN pode ser realizado usando algoritmos de aprendizagem supervisionados. Modelos de CNN bem treinados devem aprender com conjuntos de dados suficientemente rotulados para que os parâmetros possam ser muito ajustados. Cada processo de treinamento da CNN consiste em passe de encaminhamento, função de perda, retropropagação e atualização de parâmetros e pode ser apresentado nas seguintes seis etapas (CAO; DRAGIĆEVIĆ; LI, 2019):

1. Inicialização do modelo: todos os parâmetros foram gerados aleatoriamente;
2. Propagação para frente: a entrada passou pelas camadas de rede e calculou a saída estimada do modelo;
3. Cálculo da função da perda: a função de perda foi utilizada para avaliar a capacidade de previsão do método. A função de perda mais comum usou o erro quadrado médio (E) entre a saída esperada do modelo  $y_{esperado}$  e a saída estimada  $y_{estimado}$  e pode ser formulado da seguinte forma:

$$E = \frac{1}{n} \sum_{i=1}^n (y_{estimado} - y_{esperado})^2 \quad (1)$$

O valor de  $E$  diminui durante o processo de treinamento até atingir uma constante que foi usada como limiar para a próxima etapa. O objetivo do processo de treinamento foi minimizar a função de perda alimentando a rede de muitas entradas e a saída desejada;

4. Retropropagação: é um processo em que retrocede o derivado de erros;
5. Atualização dos parâmetros: em geral a atualização de pesos (parâmetros) pelas regras delta foi definida da seguinte forma:

$$\text{Novos pesos} = \text{Antigos pesos} - \text{Taxa derivada} \times \text{Taxa de aprendizagem} \quad (2)$$

Se o desvio for positivo, o parâmetro deve ser diminuído; se o desvio for negativo, o parâmetro deve ser aumentado; se o desvio for igual a 0, então o parâmetro será ótimo. A taxa de aprendizagem é constante e deve ser definida com base na

experiência; se muito pequeno, levará muito tempo para obter pesos ideais; se muito grande, os pesos se desviarão do ideal;

6. Itera o processo até a convergência. Com base na etapa anterior, os pesos são atualizados muito lentamente, solicitando muitas iterações para obter os pesos desejados e minimizar a função de perda. Na realidade, a CNN não processa imagens uma vez para aumentar a produtividade. Um lote de amostras é propagado por meio da rede ao mesmo tempo.

#### 3.2.1.1 Rede totalmente convolucional (FCN)

A rede totalmente convolucional é o trabalho mais importante no *deep learning* para a segmentação semântica, ou seja, a tarefa de atribuir um rótulo semântico a cada *pixel* da imagem. Para realizar essa tarefa, a saída da CNN deve ser do mesmo tamanho de *pixel* da entrada (ao contrário da única classe por imagem dos modelos acima mencionados). A rede introduz muitas ideias significativas, como (ZHU et al., 2017):

- aprendizagem ponta a ponta do algoritmo de *upsampling* por meio de uma estrutura de codificador/decodificador que primeiro reduz o tamanho da ativação e depois aumenta novamente;
- usando uma arquitetura totalmente convolucional, que permite que a rede tire imagens de tamanho arbitrário como entrada, pois não há uma camada totalmente conectada no final que exija um tamanho específico das ativações;
- introduzindo conexões de salto como uma forma de fundir informações de diferentes profundidades na rede para a inferência multiescala.

Essa rede extrai recursos e gera um mapa inicial de classificação grosseira usando um conjunto de camadas convolucionais que, devido à sua configuração interna, gera um resultado espacialmente reduzido (quando comparado à entrada original). Para restaurar a resolução original e gerar o mapa temático, esta abordagem emprega camadas

de deconvolução que aprendem como aumentar a amostragem do mapa de classificação inicial e produzir a previsão final densa (OSCO et al., 2021).

Como podemos analisar, as redes neurais convolucionais estão se tornando uma técnica mais comum para diferentes tarefas de análise de imagem, como também a classificação automática e segmentação semântica. A segmentação semântica tem a função de atribuir um rótulo de classe a cada *pixel* na imagem, assim como tem o potencial de capturar a forma e o tamanho do objeto com mais precisão, podendo ser essencial em muitas aplicações (TORRES et al., 2020).

#### 3.2.1.2 U-Net

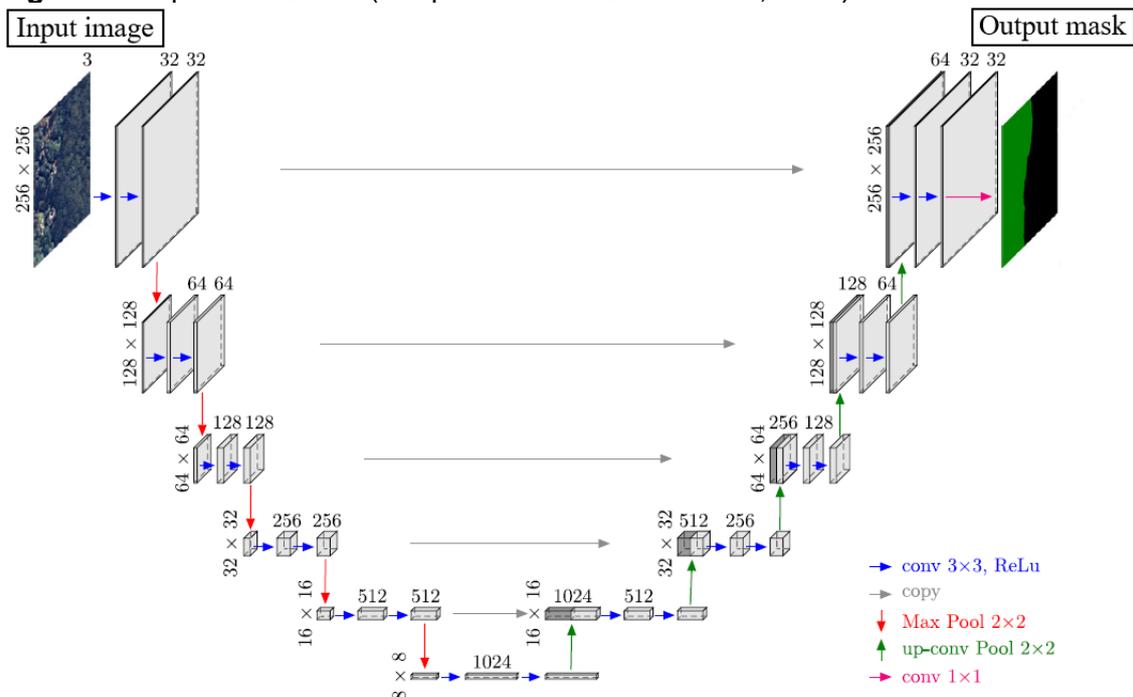
O modelo U-Net foi publicado inicialmente por (RONNEBERGER; FISCHER; BROX, 2015) para segmentação de imagens biomédicas. Posteriormente, alguns outros estudos avaliaram o potencial deste modelo, como: Classificação de cobertura e uso do solo urbano (ZHANG et al., 2018); Segmentação semântica de um espaço verde urbano (LIU et al., 2019); Segmentação semântica de uma área de cultivo de pequenos agricultores (DU et al., 2019a); Mapeamento dos tipos de floresta e perturbações na Mata Atlântica com imagens de alta resolução (WAGNER et al., 2019); Segmentação semântica de uma única espécie de árvore em ambiente urbano (TORRES et al., 2020); Classificação de cobertura do solo e detecção de objetos (ZHANG et al., 2020); Segmentação semântica de pomar de citros (OSCO et al., 2021).

U-Net é um modelo melhorado da Rede Totalmente Convolutacional, que consiste em caminho de contração simétrica e caminho expansivo. Combina recursos de baixo nível com informações espaciais detalhadas com recursos de alto nível com informações semânticas para melhorar a precisão da segmentação, e tem alcançado resultados promissores em tarefas de segmentação de uma classe (ZHANG et al., 2018). É uma técnica que auxilia o aperfeiçoamento da precisão espacial do resultado da

segmentação. A imagem de entrada é processada, pela primeira vez, por um caminho codificador – que constitui-se em camadas convolucionais e de agrupamento – que reduz a resolução espacial. Sucessivamente, é seguido por um caminho decodificador que recupera a resolução original da imagem espacial seguidas por camadas convolucionais (“*up-convolution*”). Ademais, a rede usa as chamadas conexões de salto, anexando a saída das camadas correspondentes no caminho do codificador para as entradas do caminho do decodificador (TORRES et al., 2020). De acordo com Zhang et al., 2020, a arquitetura é simples e eficiente, consiste em um caminho de contração para capturar o contexto e um caminho de expansão simétrico, permitindo uma localização precisa.

A arquitetura da U-Net apresenta um padrão em forma de U (Figura 11), onde um degrau é uma série de duas convoluções seguidas por uma camada de redução da resolução e um degrau consiste em uma série de duas convoluções seguidas por um aumento da amostragem. As conexões são feitas entre o caminho de redução da amostra e aumento da amostra em cada escala (SUDRE et al., 2017).

**Figura 2.** Arquitetura U-Net (Adaptado de WAGNER et al., 2019).



### 3.2.2 ResNet

Além das principais camadas da Rede Neural Convolutacional, as Redes Residuais (ResNet) adotam conexões residuais para melhorar o desempenho do modelo. A estrutura da conexão residual pode reduzir muito a dificuldade de otimização, além de possibilitar o treinamento de redes muito mais profundas (TONG et al., 2020).

Os autores (ZHU et al., 2017) mencionam que o propósito do ResNet é adicionar conexões de atalho que contornem duas ou mais camadas convolucionais empilhadas, executando o mapeamento de identidade. Assim, as conexões são adicionadas junto com a saída de convoluções empilhadas. Contudo, os autores (SCOTT et al., 2017) relatam que as arquiteturas ResNet empregam conexões residuais que contornam duas ou mais camadas de convolução por vez, permitindo que o erro se propague melhor para trás através da rede.

Conforme (CAO et al., 2020), os ResNets impulsionam conexões de atalho entre algumas camadas empilhadas para obter aprendizado residual, o que permite que eles se aprofundem e aprendam de forma mais eficaz. As imagens de entrada são alimentadas em primeiro lugar em uma camada convolutacional 7x7 (seguido pela normalização do lote e ativação ReLU), então uma camada de agrupamento máximo 3x3 é operada na saída. A partir deste momento, os mapas de recursos de saída são alimentados em quatro grupos de blocos residuais (ResBlocks). Os mapas de recursos de saída são ainda mais reduzidos a vetores de recursos pelo agrupamento da média global. Finalmente, uma camada totalmente conectada é acrescentada para produzir a saída. Essa saída será explorada como o recurso final da imagem extraída.

Os modelos ResNet possuem cinco versões, separadamente com 18, 34, 50, 101 e 152 camadas. Os modelos ResNet 18 e 34 são consideradas arquiteturas rasas, enquanto os modelos 50, 101 e 152 são arquiteturas muito profundas (TONG et al., 2020).

### 3.3 CLASSIFICAÇÃO DE USO E COBERTURA DO SOLO

A classificação de cobertura do solo no domínio sensoriamento remoto desempenha um papel importante que incluem campos de aplicação de planejamento urbano, detecção de mudanças e monitoramento ambiental nas tecnologias de sensoriamento remoto (ALEM; KUMAR, 2020).

Na história moderna, o crescimento mundial da população trouxe desafios relevantes para a sociedade e para o meio ambiente, como a crescente demanda por moradias, alimentos, recursos naturais e serviços básicos. Devido a isso, a necessidade de um planejamento e gestão eficiente do uso do solo está aumentando, não apenas para eliminar os efeitos negativos das decisões históricas de uso do solo, mas também para tornar as comunidades futuras mais saudáveis e sustentáveis. Portanto, há a necessidade de métodos computacionais mais avançados para analisar dados geoespaciais da superfície terrestre para quantificar e entender melhor a dinâmica complexa dos processos de mudança do uso e cobertura do solo (CAO; DRAGIĆEVIĆ; LI, 2019).

Porém, há uma confusão de nomenclatura na área de obtenção de cartas de cobertura e uso do solo. “Cobertura do solo” (*Land Cover*) não possui o mesmo significado de “Uso do solo” (*Land Use*). Por exemplo, a classificação de uma imagem pode identificar um campo e com informações adicionais pode-se chegar à conclusão de que este campo é utilizado como pastagem. Logo, a imagem mostra uma cobertura do solo de campo, enquanto o uso do solo é pastagem ou alguma outra prática de manejo. Dessa forma, pode ser conveniente tratar as duas definições de modo distinto (PARANHOS FILHO et al., 2016).

De acordo com (CAO; DRAGIĆEVIĆ; LI, 2019), o uso do solo representa o uso humano do meio ambiente natural para fins econômicos, urbanos, recreativos,

conservatórios e governamentais. A cobertura do solo representa uma região da superfície terrestre coberta por características físicas como vegetação, água, floresta e entre outros.

Há ainda áreas de uso misto, sendo que a cada ano ou estação do ano podem apresentar um tipo de cobertura diferente. Desse modo, é necessário obter informações complementares para identificar o uso do solo com outras imagens do mesmo local, em outras épocas, ou até mesmo estudo de campo (PARANHOS FILHO et al., 2016).

Para obter informações significativas dos processos de uso e cobertura do solo é necessário adquirir um conjunto de dados de sensoriamento remoto. Esses conjuntos de dados fornecem coberturas de escalas regionais a globais. A interpretação desses dados é uma das principais formas de entender o status e as mudanças nos ambientes naturais e construídos. Nos últimos anos, os sensores e técnicas de sensoriamento remoto tornaram-se cada vez mais sofisticados. Eles podem fornecer um grande volume de conjunto de dados com alta qualidade e resolução espacial fina (CAO; DRAGIĆEVIĆ; LI, 2019).

Conforme Huang, Zhao e Song (2018), métodos de mapeamento de cobertura e uso do solo urbano dependem da interpretação de fotos aéreas e levantamentos de campo, sendo ambos trabalhosos e demorados. Com o desenvolvimento de tecnologias de sensoriamento remoto, um grande número de imagens de alta resolução espacial cobrindo uma área urbana podem ser obtidos por sensores instalados nas aeronaves ou satélites. Embora seja útil extrair informações da cobertura e uso do solo de tais imagens, muitas vezes contém vários tipos de cobertura do solo com características espaciais, espectrais e geométricas distintas, tornando o mapeamento automático um desafio.

Além disso, os detalhes ricos obtidos nas imagens de alta resolução, como forma geométrica e o conteúdo estrutural dos objetos, trazem mais desafios à classificação

de cobertura do solo, bem como as diversas condições de imagem que geralmente levam a distorções fotográficas, variações de escala e mudanças de iluminação, o que muitas vezes reduz seriamente a separabilidade entre diferentes classes de cobertura do solo (TONG et al., 2020).

A classificação da cobertura do solo com imagens de satélite ou aéreas visa associar cada *pixel* em uma imagem com uma categoria de cobertura de solo pré-definida. Para isso, as abordagens de classificação utilizando informações espectrais têm sido intensamente estudadas (TONG et al., 2020).

Uma infinidade de algoritmos de classificação da cobertura do solo foi desenvolvida, com sólidos fundamentos teóricos, baseados nas propriedades espectrais e espaciais dos *pixels*. Entretanto, a tarefa torna-se cada vez mais difícil à medida que o nível de abstração aumenta, passando de *pixels* a objetos, e depois cenas (CASTELLUCCIO et al., 2015).

Rotular uma imagem de acordo com um conjunto de categorias semânticas é o objetivo da classificação da cena. Este é um problema muito desafiador, pois as coberturas do solo que caracterizam uma determinada classe podem apresentar uma grande variabilidade e os objetos podem aparecer em diferentes escalas e orientações (CASTELLUCCIO et al., 2015).

Recentemente, à medida que o volume de dados de imagem de sensoriamento remoto acessível aumenta tremendamente, a análise de imagem de cena de alto nível, como a classificação de imagens de cena, tem atraído especial interesse de pesquisa. Existem dois tipos de tarefas de classificação de imagem na literatura de sensoriamento remoto: (a) classificação em nível de *pixel*, ou seja, classifica *pixels* em uma imagem de sensoriamento remoto e (b) classificação em nível de imagem, ou seja,

classificar imagens individuais em um conjunto de dados de imagem de sensoriamento remoto (YU *et. al.*, 2017).

O primeiro tipo de tarefa de classificação de imagem visa atribuir rótulos categóricos a *pixels* dentro de uma imagem e o processamento pode ser feito em termos de operações baseadas em *subpixel*, *pixels* ou subconjunto de *pixels* (por exemplo, região de objetos) (YU *et. al.*, 2017).

O segundo tipo de classificação de imagem visa classificar cada imagem de cena em uma categoria cênica, por exemplo, rio, floresta, agricultura, solo exposto, e entre outros, e este trabalho foca nesse tipo de classificação. Uma maneira amplamente aceita de desenvolver tais algoritmos de classificação é treinar um determinado modelo de classificação dado dados de treinamento suficientes, ou seja, imagens de sensoriamento remoto com rótulos de classe de cena conhecidos. O modelo treinado é então usado para estimar os rótulos da classe de cena de imagens de sensoriamento remoto desconhecidas (YU *et al.*, 2017).

### **3.3.1 Áreas permeáveis**

Com o crescimento de áreas urbanas, por meio de parcelamento do solo, o cenário paisagístico vem sofrendo alterações, prejuízos e comprometendo a qualidade ambiental urbana (DE BORTOLO; ANTUNES RODRIGUES; BORGES, 2018). Embora a urbanização ofereça muitas vantagens sociais e tecnológicas, o custo desse privilégio é a destruição de áreas rurais e ambientes naturais. Devido a expansão urbana, os ambientes naturais estão diminuindo e o caso mais influente disso é o desaparecimento das florestas. As florestas abrigam um número incrível de espécies de animais e insetos, e desempenham um papel maciço na regulação da atmosfera terrestre. O sequestro de carbono das florestas é um elemento vital na atenuação das concentrações globais de gases de efeito estufa (LIN *et al.*, 2020).

Os ambientes naturais, ou áreas verdes, em meios urbanos é fundamental para o nosso bem estar. Alguns estudos revelam que paisagens naturais influenciam no humor, níveis de estresse e auxilia no combate a algumas doenças, em razão destes ambientes serem mais atraentes para praticar exercícios e atividades ao ar livre (MORENO-ARMENDÁRIZ et al., 2019).

Além disso, a área permeável é uma área coberta por vegetação que permite e facilita a passagem da água da superfície para o interior do solo. Essa passagem podemos denominar de infiltração, sendo essencial para a manutenção do ciclo hidrológico, promovendo a recarga dos aquíferos e o abastecimento dos cursos hídricos (TUCCI, 2001).

Não obstante, a cobertura vegetal possui um relevante papel na drenagem urbana, pois intercepta as gotas da água pluvial e interfere no processo precipitação-vazão, retardando o escoamento superficial e consequentemente reduzindo as vazões máximas. Além disso, reduz a erosão do solo e evita o carreamento de partículas (TUCCI, 2001).

Ademais, em áreas que possui cobertura vegetal, irradia uma quantidade menor de calor para o ambiente, ou seja, locais com presença de vegetação tende a apresentar valores mais baixos de temperatura do ar, comparado com locais que possui áreas impermeáveis, bem como a ausência de indivíduos arbóreos (SOUZA, 2017).

As árvores são componentes dominante da paisagem natural urbana, desempenham um papel importante na melhoria do ambiente ecológico urbano por meio de filtragem do ar, regulação do microclima, redução de ruído e amenização da qualidade da água (WEN et al., 2017).

Como a vegetação urbana é essencial para os municípios, é pertinente que cada cidade possua um planejamento e manejo da paisagem natural urbana. Para isso, é

significante na tomada de decisão o detalhamento da distribuição espacial, assim como as informações detalhadas (espécies, tipos de habitat e entre outros) (WEN et al., 2017).

Usualmente, a classificação detalhada da vegetação urbana, assim como as áreas permeáveis, pode ser realizada por meio de levantamentos de solo, imagens aéreas ou interpretação de sensoriamento remoto. Entretanto, levantamentos de solo convencionais podem ser demorados e custosos devido à complexidade da cena urbana, dinâmica da paisagem e restrições de acessibilidade para áreas privadas, enquanto a observação direta da cobertura do solo por meio de imagem aérea ou dados de satélite pode permitir a eficaz classificação de árvores. Imagens aéreas têm sido a principal fonte de dados para a classificação detalhada de árvores urbanas em diversas pesquisas (WEN et al., 2017).

Diante do que foi exposto, entende-se que área permeável é apenas aquela área revestida por vegetação (grama, arbustos ou árvores). Porém, áreas com solo exposto também são consideradas como áreas permeáveis. Por mais que não possui a cobertura vegetal, ela também não possui cobertura impermeável (piso cerâmico, telha cerâmica, asfalto). Portanto, a água superficial tem a capacidade de infiltrar em solo exposto.

## 4 METODOLOGIA

### 4.1 ÁREA DE ESTUDO

O estudo é abordado na Bacia Hidrográfica do Córrego Bandeira (figura 3) situada à sudeste do perímetro urbano do Município de Campo Grande, abrangendo os bairros Carlota e Dr. Albuquerque, e parte dos bairros Parati, Piratininga, América, Pioneiros, Universitário, Jardim Paulista, Rita Vieira, TV Morena, Vilas Boas, São Lourenço e Tiradentes (PLANURB, 2020).

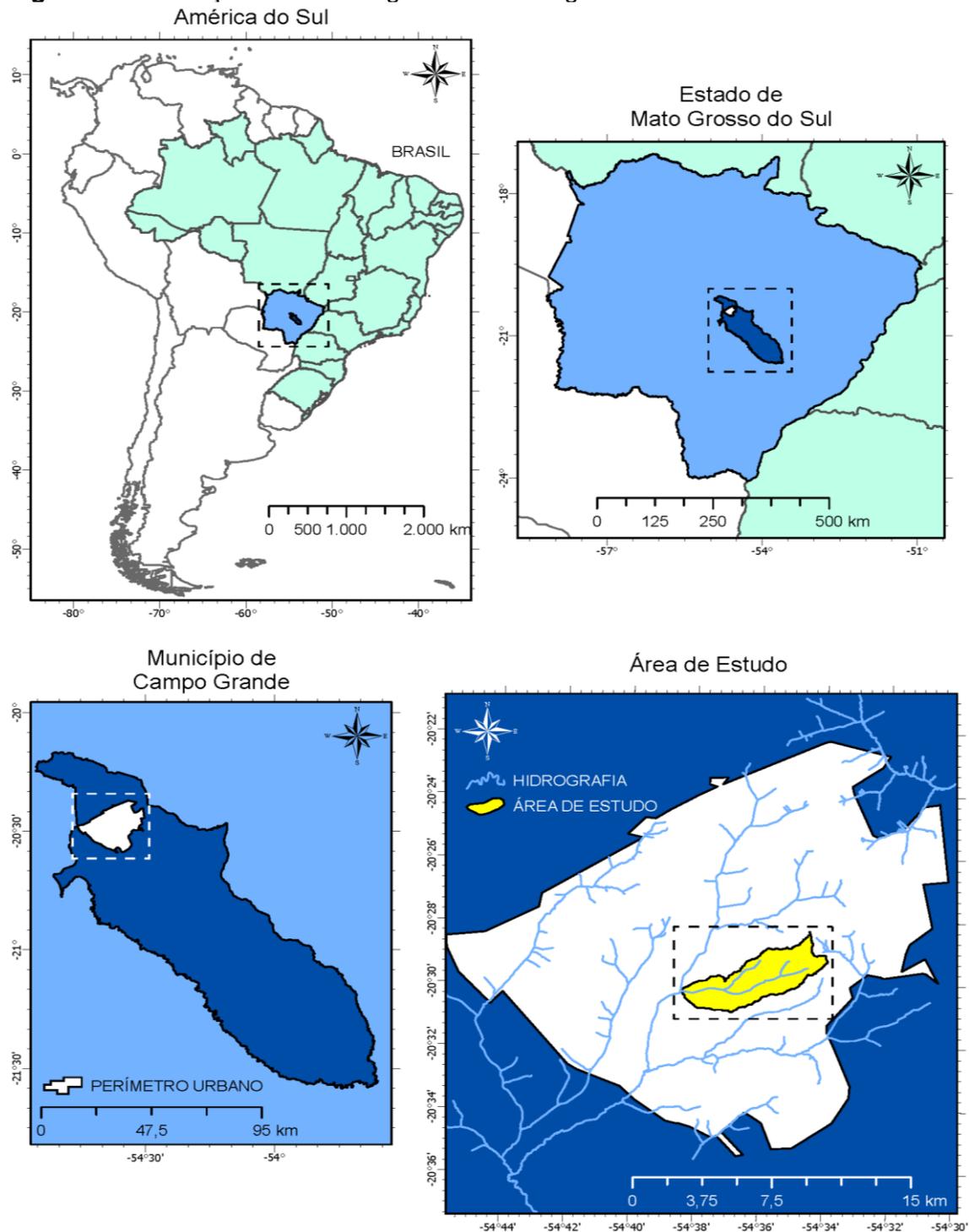
A Bacia Hidrográfica do Córrego Bandeira possui área de aproximadamente 13,30 km<sup>2</sup>. Está inserida na Bacia Hidrográfica do Rio Anhanduí e consequentemente na Bacia Hidrográfica do Rio Paraná. Inclusive, compõe os Córregos Bandeira, Portinho Pache e Cabaça, o Lago do Rádio Clube Campo e o Lago do Amor, visto que esses corpos hídricos vêm sofrendo processo acelerado de assoreamento e degradação (PLANURB, 2020).

Além dos corpos hídricos, a Bacia contempla o Parque Doutor Anísio de Barros, formado por uma lagoa natural, nomeada de Lagoa Itatiaia (CAMPO GRANDE, 2003), o Parque Linear do Córrego Cabaça, nomeado como Parque Presidente Juscelino Kubitschek (CAMPO GRANDE, 2010) e a Reserva Particular do Patrimônio Natural (RPPN) da Universidade Federal de Mato Grosso do Sul (UFMS), reconhecida como Unidade de Conservação (UC) pela Deliberação CECA/MS n. 002, de 12 de fevereiro de 2003, e a sua constituição teve como propósito “a preservação das condições naturais primitivas, semiprimitivas, recuperadas ou em recuperação, para manutenção de seu aspecto paisagístico ou preservação do ciclo biológico da flora e fauna nativa” (PLANURB, 2019).

A área em questão faz parte do segundo maior bioma da América do Sul, o Cerrado. É composto por um conjunto de formas de vegetação que é exposto conforme

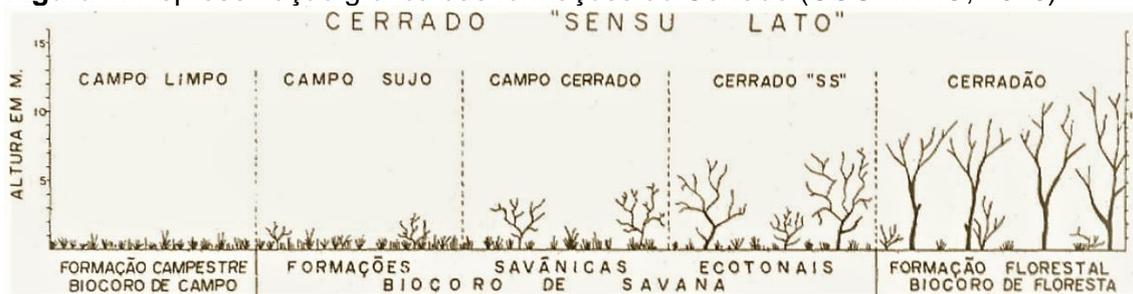
um gradiente de biomassa, diretamente relacionado à fertilidade do solo. Suas principais fisionomias são: Campo Limpo, Campo Sujo, Cerrado, Cerradão, além da presença da Floresta Aluvial (mata ciliar) e áreas de Tensão Ecológica, caracterizadas pela proximidade Cerrado/Floresta Estacional Semidecidual e pelas áreas formadas por ações antrópicas utilizadas para agropecuária (PLANURB, 2019).

**Figura 3.** Localização Bacia Hidrográfica do Córrego Bandeira.



Conforme Coutinho (1978) o Cerrado possui um complexo de formações, sendo: formação campestre, constituída de campo limpo; formações savânicas, representada por campo sujo, campo cerrado e cerrado “s.s.”; e formação florestal, caracterizada basicamente pelo cerradão (figura 4).

**Figura 4.** Representação gráfica das formações do Cerrado (COUTINHO, 1978).



O Bioma Cerrado é considerado uma das maiores biodiversidade de espécies, sendo reconhecida também como a savana mais rica do mundo, porém apresenta uma extrema abundância de espécies endêmicas e sofre uma excepcional perda de habitat. Entretanto, é rica em avifauna, compreendendo cerca de 837 espécies, 199 espécies de mamíferos, 1200 espécies de peixes, 180 espécies de répteis, 150 espécies de anfíbios e é refúgio de 13% das borboletas, 35% das abelhas e 23% dos cupins dos trópicos (MMA, 2020).

Em relação à geologia, geomorfologia e pedologia, a área de estudo, conforme a Carta Geotécnica (1991), abrange todas as unidades homogêneas, porém, em maior parte se encontra nas unidades I e III. A unidade I é caracterizada por basalto e arenitos intertrapeanos da Formação Serra Geral e são encontrados latossolos vermelho-escuro e roxo, possuindo baixa a média permeabilidade e baixa a média susceptibilidade a erosão. Já na unidade III, é caracterizada por arenitos intertrapeanos da Formação Serra Geral e possui a presença de latossolos amarelo e vermelho-escuro, com a presença de nível de água muito próximo a superfície (até 3m), relevo plano, solos orgânicos e alta erodibilidade (PMCG, 1991).

Conforme dados apresentados pelo Perfil Socioeconômico de Campo Grande (PLANURB, 2019) e a classificação de Koppen, o clima da região situa-se em uma faixa de transição entre o subtipo Cfa, mesotérmico úmido sem estiagem ou pequena estiagem, e o subtipo Aw, tropical úmido com estação chuvosa no verão e seca no inverno. A precipitação média acumulada nos anos de 2009 a 2018 foram 1688,65mm.

De acordo com o Plano Diretor de Desenvolvimento Urbano Ambiental de Campo Grande (PDDUA), instituída pela Lei Complementar n. 341, de 4 de dezembro de 2018, a Bacia está inserida totalmente na Macrozona 2 (MZ2) que é definida como adensamento prioritário, com densidade demográfica líquida prevista de até 240 habitantes por hectare e densidade demográfica de até 55 habitantes por hectare, e dividida entre as Zonas Ambientais 3 e 4 (ZA3 e ZA4).

## 4.2 COLETA E PRODUÇÃO DE DADOS

O levantamento de imagens aéreas do ano de 2013 e imagens orbitais Worldview-2 do ano de 2016, com resolução espacial de 10 e 50 cm, tal como disponibilizadas pela Prefeitura Municipal de Campo Grande (PMCG) e pelo Instituto de Meio Ambiente de Mato Grosso do Sul (IMASUL), respectivamente.

O procedimento para a concepção dos dados equivale a cinco etapas, sendo: (i) Preparação e manipulação das imagens; (ii) Segmentação e classificação manual das imagens; (iii) Exportação de dados de treinamento para aprendizagem profunda; (iv) Treinamento de modelo de aprendizado profundo; e (v) Classificação dos *pixels* usando aprendizado profundo.

### 4.2.1 Preparação e manipulação das imagens

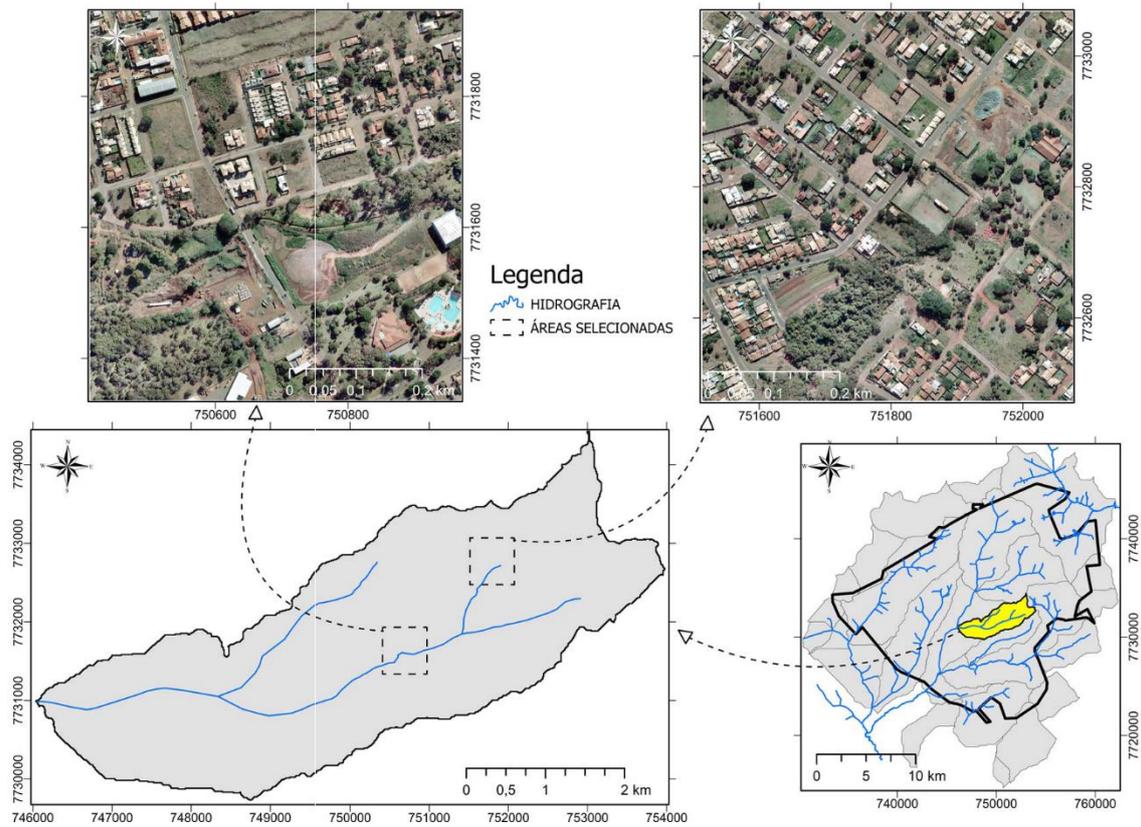
As imagens aéreas, com resolução espacial de 10 cm, capturadas em 2013 por veículos aéreos tripuláveis, foram disponibilizadas em diversos quadrantes, totalizando 1.324 imagens no tamanho 5.619 x 5.946 *pixels* e em formato *Enhanced*

*Compression Wavelet* (.ecw). Porém, foram selecionadas apenas duas imagens, das quais estão inseridas na Bacia Hidrografia do Córrego Bandeira (Figura 5). Entretanto, o conjunto de ferramentas de *Deep Learning* requer que as imagens estejam no formato *Tagged Image File Format* (.tiff). Logo, foi necessária a conversão das imagens.

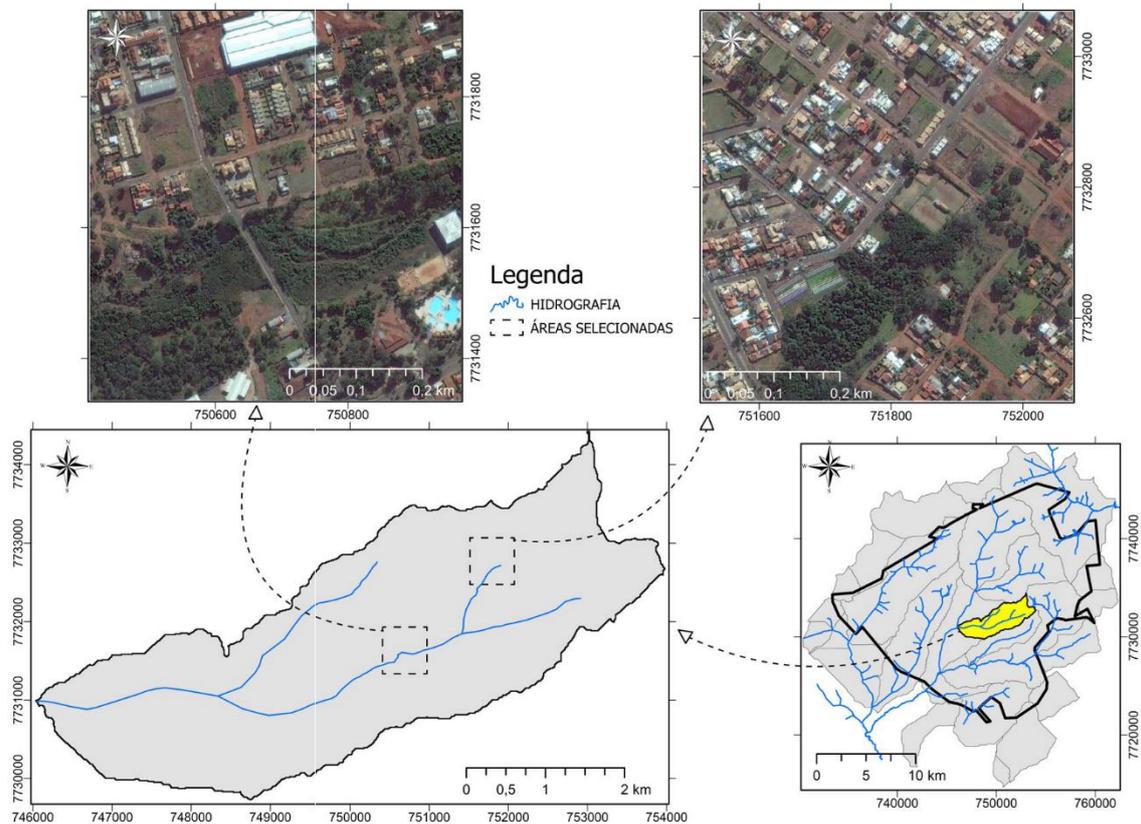
As imagens orbitais, com resolução espacial de 50 cm, capturadas no ano de 2016 pelo satélite WorldView-2, foram disponibilizadas em cinco imagens, sendo duas no tamanho de 316.441 x 225.089 *pixels*; uma no tamanho 106.046 x 112.206 *pixels* e duas no tamanho 99.148 x 330.866 *pixels*, e em formato “.ecw”. No entanto, foi selecionada apenas uma imagem, no tamanho 316.441 x 225.089 *pixels*. Além disto, a imagem por contemplar uma extensa área, foi necessário extrair dois fragmentos, dispondo a mesma semelhança de dimensão que as imagens aéreas. Contudo, apresentando o tamanho de 1.125 x 1.190 *pixels* e em formato “.tiff”, conforme a figura 6. Os procedimentos das imagens acima foram realizados pelo *software QGIS* versão 3.16 (QGIS Development Team, 2009).

Ainda assim, essas imagens orbitais foram disponibilizadas já abrangendo a fusão de imagens, conhecido também como *pansharpening*, ou seja, são imagens que contém a fusão das bandas espectrais com a banda pancromática. Logo, as imagens WorldView-2 contém 8 bandas espectrais que possuem resolução espacial de 2 m e uma banda pancromática que possui resolução espacial de 50 cm (LIN et al., 2015).

**Figura 5.** Áreas selecionadas para rotulação e treinamento. Imagens aéreas cedidas pela PMCG.



**Figura 6.** Áreas selecionadas para rotulação e treinamento. Imagens orbitais WorldView-2, cedidas pelo IMASUL.



#### 4.2.2 Rotulação dos dados de treinamento

A segmentação e classificação manual em imagens aérea ou orbital solicita que o intérprete seja paciente, atencioso e meticuloso, em virtude da dificuldade de contornar, com precisão, a cobertura e uso do solo, devido a heterogeneidade e detalhes que a maior resolução espacial permite observar. Dessa forma, as áreas consideradas permeáveis foram classificadas em três categorias: Vegetação Arbórea; Vegetação Rasteira; e Solo Exposto. Foram gerados um arquivo, em formato *shapefile* (.shp), para cada imagem, portanto dois arquivos. No arquivo relacionado à imagem aérea, foram gerados no total 3.104 polígonos, sendo: 1.064 na categoria Vegetação Arbórea; 1.194 na categoria Vegetação Rasteira; 846 na categoria Solo Exposto. No arquivo pertinente à imagem orbital, foram gerados no total 438 polígonos, sendo: 199 na categoria Vegetação Arbórea, 149 na categoria Vegetação Rasteira e 90 na categoria Solo Exposto. Abaixo, na tabela 1, segue o modelo de tabela de atributos dos arquivos desenvolvidos. Para esse procedimento foram utilizadas as imagens aéreas e orbitais e manuseadas no *software QGIS* versão 3.16.

**Tabela 1.** Modelo de tabela de atributos dos arquivos *shapefiles*.

OBJECTID	SHAPE	CLASSNAME	CLASSVALUE
1	Polígono	VEGETAÇÃO ARBÓREA	1
2	Polígono	VEGETAÇÃO RASTEIRA	2
3	Polígono	SOLO EXPOSTO	3

#### 4.2.3 Exportação de dados e treinamento de modelo de aprendizado profundo

Para este e ademais métodos foram empregados o *software ArcGIS Pro* versão 2.8 (ESRI, 2021). Este programa foi selecionado por já possuir o conjunto de ferramentas *Deep Learning*. Dessa forma, não foi necessário manusear arquivos e trabalhar com algoritmos em outros programas. Entretanto, antes de iniciar as atividades, é necessário fazer *download* e instalação do *plug-in Deep Learning*.

No programa, no campo de geoprocessamento, no conjunto de ferramentas *Deep Learning*, encontra-se a ferramenta Exportar Dados de Treinamento para Aprendizagem Profunda. Nesta ferramenta selecionamos a imagem, em formato “.tiff”, em raster de entrada (*input raster*); o arquivo concebido em *shapefile* (.shp) na classe de recurso de entrada (*input features class*); tipo de formato da imagem (.tiff) (*image format*); tamanho dos fragmentos (*tile size*), sendo o padrão pelo programa 256 x 256 *pixels*, bem como empregado por alguns pesquisadores (ALEM, KUMAR, 2020; CAO, DRAGIĆEVIĆ, LI, 2019; CHENG et al., 2018; CHENG et al., 2020; CHENG, HAN, LU, 2017; KAMPFMEYER, SALBERG, JENSSEN, 2017; WAGNER et al., 2019; WENG et al., 2017; YU et al., 2017); e o formato de metadados (*metadata format*) utilizamos o fragmento classificado (*classified tiles*), em virtude deste formato ser usado principalmente para classificação de *pixels*. Sendo assim, executamos a exportação de dados.

Na sequência é empregada a ferramenta de treinamento. Nesta ferramenta foi utilizado o tipo de modelo (*model type*) U-Net, e no modelo de suporte principal (*backbone model*) ResNet-34. Além disso, o modelo foi treinado para no máximo 100 épocas (*epochs*) (ORTEGA ADARME et al., 2020; ZHANG et al., 2018), tamanho do lote (*batch size*) 8 (ZHANG et al., 2020); taxa de aprendizagem (*learning rate*) 0,0001 (KATTENBORN, EICHEL, FASSNACHT, 2019; TORRES et al., 2020; WAGNER et al., 2019) e 20% de validação (DOS SANTOS et al., 2019; WAGNER et al., 2019; ZHANG et al., 2020). Uma vez que épocas (*epochs*) representa o número de iterações necessárias para analisar todo o conjunto de dados do trabalho com um número amostras por lote; o tamanho do lote (*batch size*) influencia a eficácia da convergência de perda e, por sua vez, a fase de atualização do sistema, o que o torna um parâmetro muito crítico e basicamente dependente dos dados (BEN HAMIDA et al., 2018); e validação significa o

percentual de amostras de treinamento que serão usadas para validar o modelo. O tamanho do lote (*batch size*) foi definido levando em consideração a demanda e a disponibilidade de memória da GPU, assim como o modelo de suporte principal (*backbone model*).

#### 4.2.4 Classificação dos *pixels* usando aprendizado profundo

Após o treinamento, na ferramenta de classificação dos *pixels*, seleciona a imagem de entrada (*input raster*), em formato “.tiff”, define a pasta e o nome do classificador de saída (*output classifier raster*) e por último indica a definição de modelo (*model definition*), que no caso é o arquivo do modelo de treinamento, em formato “.dlpk” ou “.emd” (este último formato é aceito apenas no *software ArcGIS Pro*) e posteriormente executar a ferramenta.

#### 4.2.5 Avaliação de desempenho dos modelos

Para avaliar o desempenho da classificação da cobertura do solo em cada tipo de modelo *Deep Learning* empregado e nas diferentes imagens testadas, será utilizada a métrica de Acurácia e F1-Score. Segundo Moreno-Armendáriz et al. (2019), as métricas utilizadas para avaliar o desempenho desse tipo de rede são calculadas usando as seguintes equações:

$$Accuracy = \frac{\sum_{i=1}^l \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}}{l} \quad (3)$$

$$Precision = \frac{\sum_{i=1}^l tp_i}{\sum_{i=1}^l tp_i + fp_i} \quad (4)$$

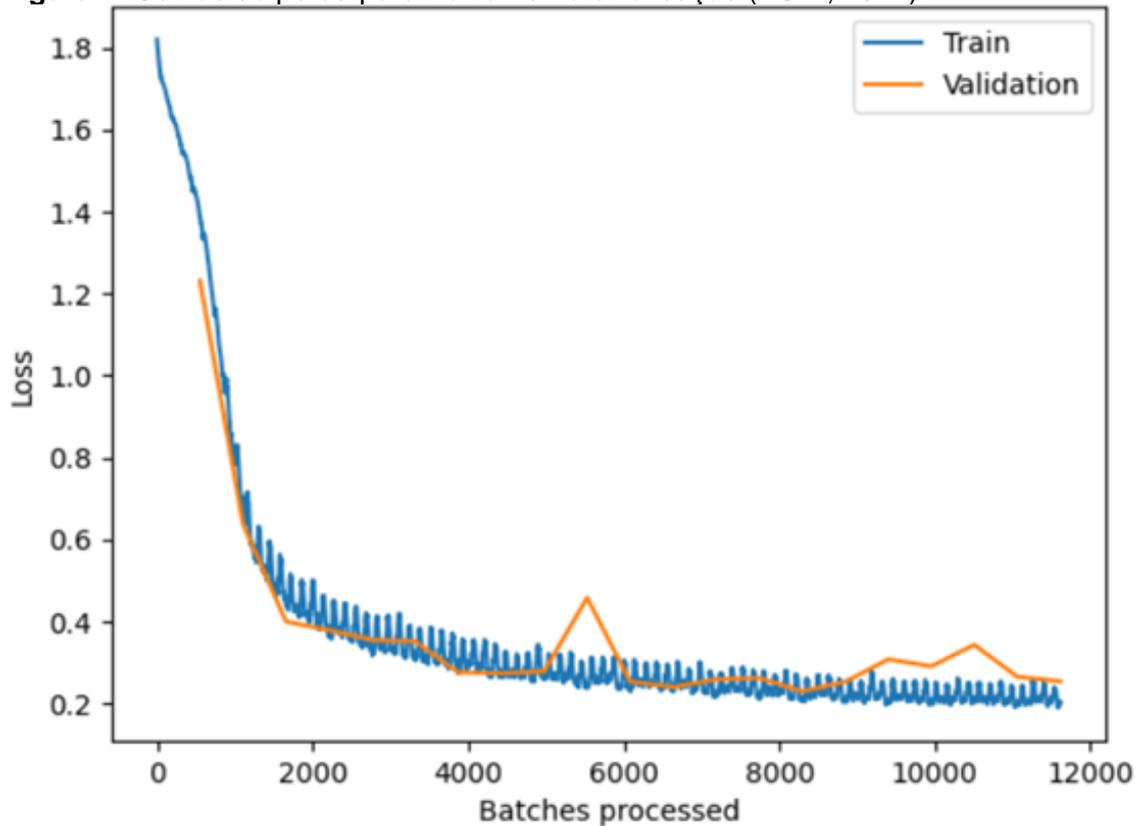
$$Recall = \frac{\sum_{i=1}^l tp_i}{\sum_{i=1}^l tp_i + fn_i} \quad (5)$$

$$F_{1\_score} = 2 \times \frac{(Precision \times Recall)}{Precision + Recall} \quad (6)$$

Em que:  $tp$  significa verdadeiros positivos,  $tn$  verdadeiros negativos,  $fp$  falsos positivos e  $fn$  falso negativos (MORENO-ARMENDÁRIZ et al., 2019). Dessa forma, *precision* é o número de resultados positivos corretos divididos pelo número de todos os resultados positivos retornados pelo classificador, *recall* é o número de resultados positivos corretos divididos pelo número de todas as amostras relevantes, e *F1* representa a média harmônica da *precision* e *recall* (DU et al., 2019).

Além disso, a figura 7 apresenta um exemplo do gráfico que inclui curvas de perda para treinamento e validação. Perda menor é melhor, então é esperado ver curvas que diminuam com o tempo e se nivelam. A perda de treinamento mostra o quão bem o modelo foi aprendido, e a perda de validação mostra o quão generalizado foi o aprendizado. Neste tipo de gráfico, pode verificar se há ajuste insuficiente e excessivo do modelo revisando as curvas de perda, bem como identificar modelos de ajuste excessivo com uma perda de validação que diminui até um ponto, mas depois começa a subir ou nivelar em relação à perda de treinamento. Um bom ajuste é geralmente identificado pela perda de treinamento e validação que diminuem até um ponto de estabilidade com uma lacuna mínima entre os dois valores finais de perda (ESRI, 2021).

**Figura 7.** Curvas de perda para treinamento e validação (ESRI, 2021).



Ainda há o coeficiente de pontuação *Dice*, que é uma medida de sobreposição amplamente usada para avaliar o desempenho da segmentação (SUDRE et al., 2017).

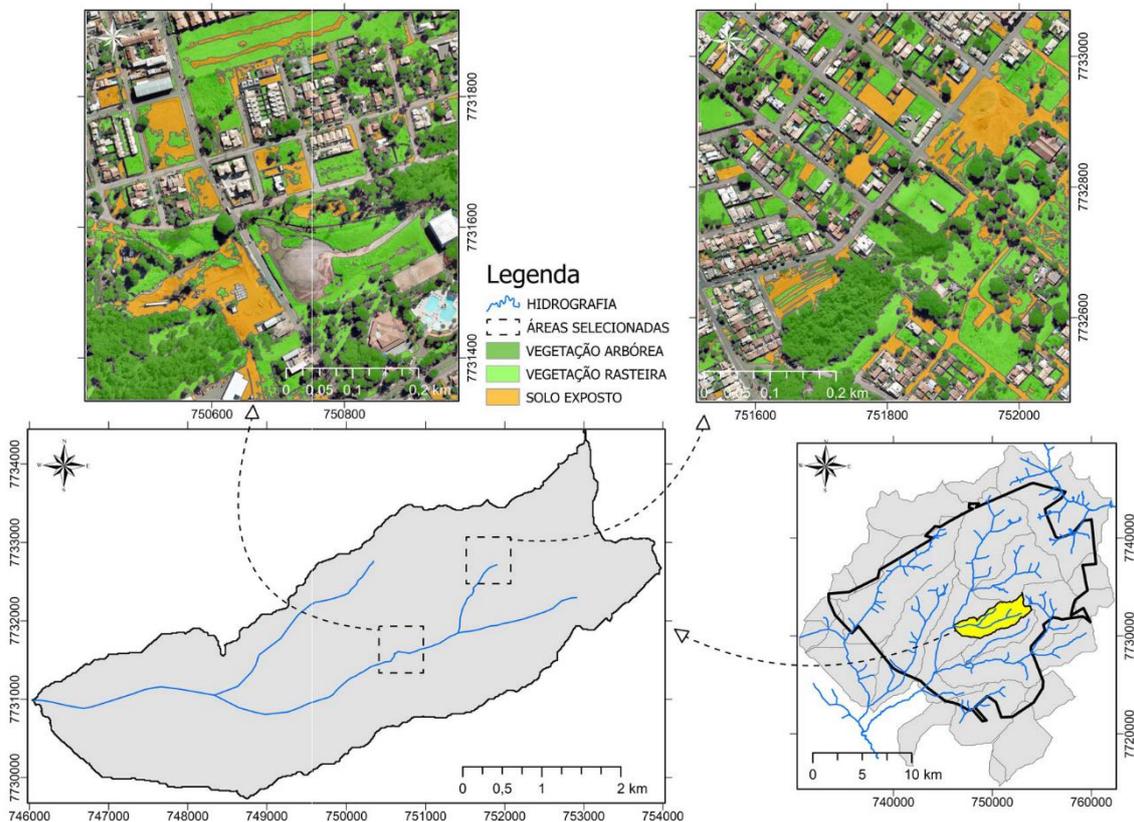
Para efetuar todas essas fases foi utilizado um computador *notebook* com uma CPU Intel(R) Core(TM) i7-8750H CPU@2.20GHz, memória de 16 GB e placa gráfica NVIDIA GeForce GTX 1060 e 14 GB de memória gráfica.

## 5 RESULTADOS E DISCUSSÃO

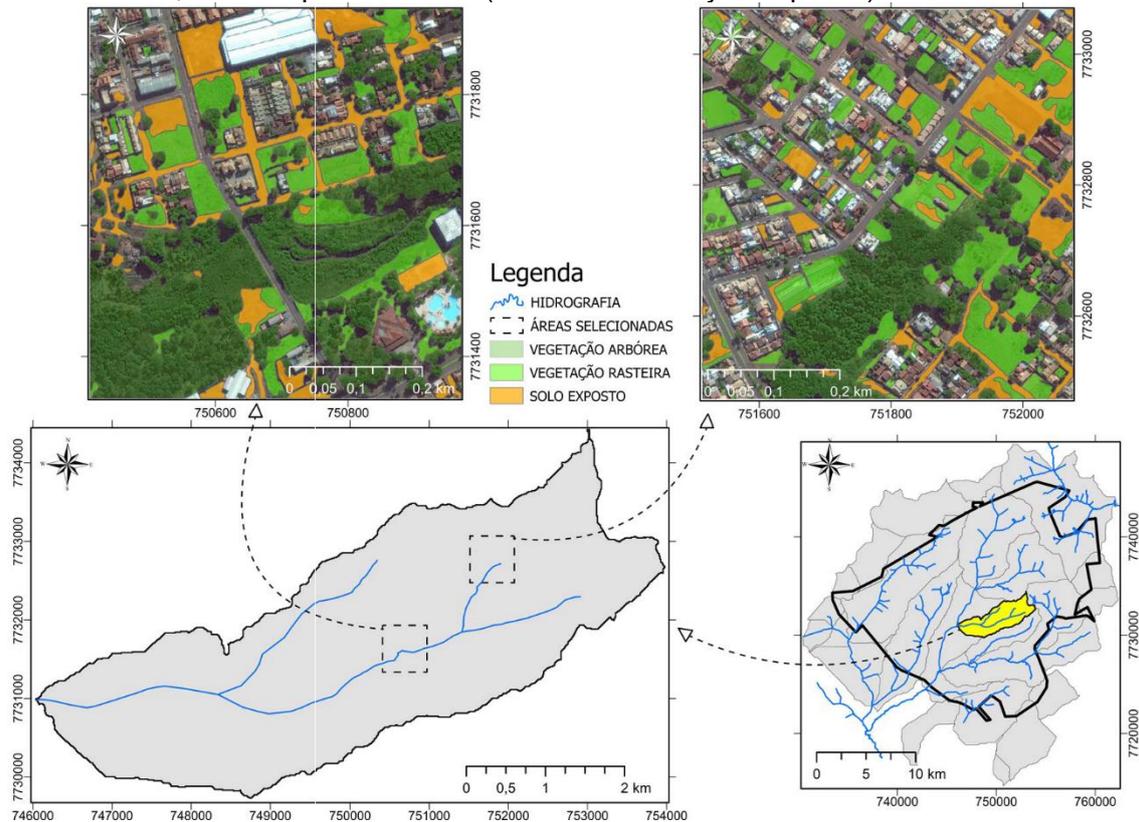
### 5.1 TREINAMENTO DE MODELO DE APRENDIZADO PROFUNDO

Conforme mencionado no capítulo 4.2.2, foram elaborados arquivos *shapefile* com as rotulações das imagens selecionadas na área de estudo (figuras 8 e 9), classificando as áreas permeáveis nas categorias: Vegetação Arbórea; Vegetação Rasteira e Solo Exposto. Que, posteriormente, foram aplicados nas ferramentas de exportação de dados e treinamento do modelo *Deep Learning*.

**Figura 8.** Rotulação das imagens escolhidas na área de estudo. Imagens aéreas cedidas pela PMCG (10 cm de resolução espacial).



**Figura 9.** Rotulação das imagens escolhidas na área de estudo. Imagens orbitais WorldView-2, cedidas pelo IMASUL (50 cm de resolução espacial).



Há muitos trabalhos que utilizaram segmentação semântica automática para separar e classificar as categorias desejadas (BEN HAMIDA et al., 2018; OSCO et al., 2021; TONG et al., 2020; ZHANG et al., 2020), tal como alguns dispuseram o trabalho de realizar delineamento manualmente como foi efetuado (TORRES et al., 2020; WAGNER et al., 2019). Neste trabalho não houve a segmentação semântica devido ao propósito de ser rotulado e classificado apenas as áreas permeáveis. Isto é, se processar a imagem operando a segmentação semântica, ela separará a imagem por completa, não apenas o desejado. Portanto, a rotulação e classificação foi realizada manualmente.

Na sequência, podemos visualizar os resultados dos modelos de avaliação e curvas de perda para treinamento e validação (figuras 10 e 11).

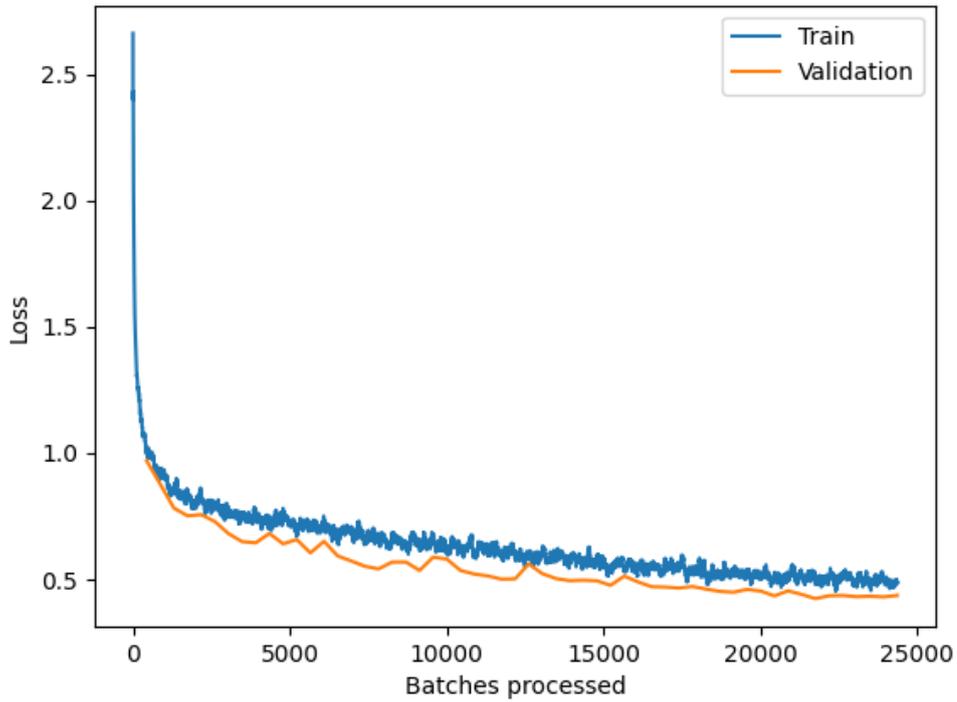
**Figura 10.** Resultados modelo de avaliação, empregando a imagem aérea (ESRI ArcGIS Pro, 2021).

**UnetClassifier**

**Backbone:** resnet34

**Learning Rate:** 1.0000e-04

**Training and Validation loss**



**Analysis of the model**

**Per class metrics:**

	NoData	VEGETACAO_RASTEIRA	VEGETACAO_ARBOREA	SOLO_EXPOSTO
<b>precision</b>	0.908035	0.765096	0.818990	0.830409
<b>recall</b>	0.860870	0.879870	0.813881	0.729066
<b>f1</b>	0.883823	0.818479	0.816428	0.776445

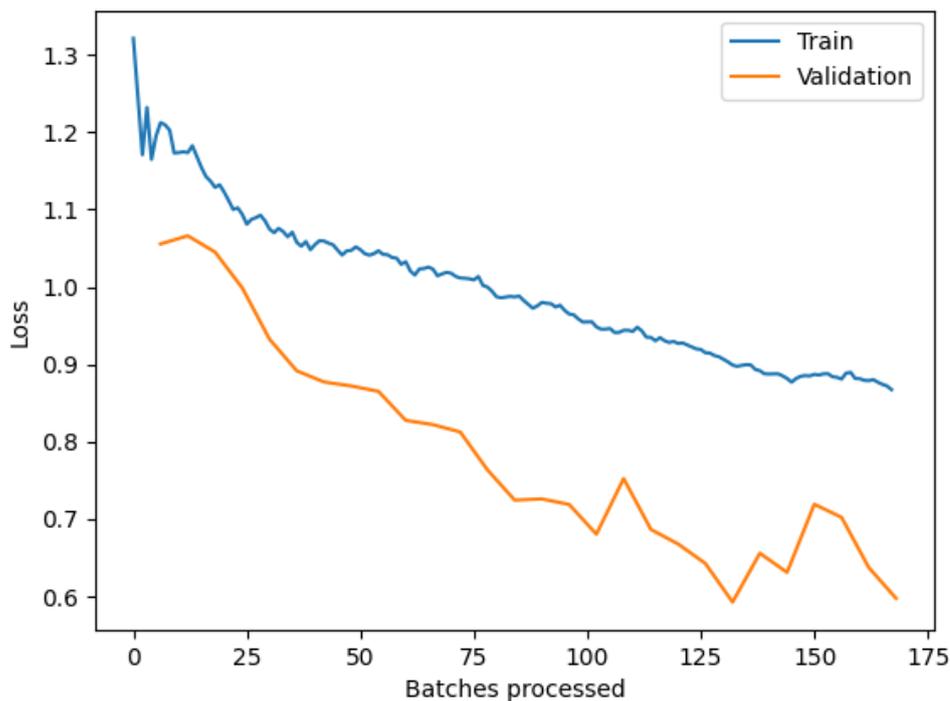
**Figura 11.** Resultado modelo de avaliação, empregando imagem orbital WorldView-2 (ESRI ArcGIS Pro, 2021).

**UnetClassifier**

**Backbone:** resnet34

**Learning Rate:** 1.0000e-04

**Training and Validation loss**



**Analysis of the model**

**Per class metrics:**

	NoData	VEGETACAO_ARBOREA	VEGETACAO_RASTEIRA	SOLO_EXPOSTO
<b>precision</b>	0.814724	0.811192	0.335745	0.090113
<b>recall</b>	0.914015	0.753756	0.132903	0.006519
<b>f1</b>	0.861518	0.781420	0.190427	0.012159

Considerando os resultados, podemos analisar que a classificação utilizando o modelo U-Net e ResNet-34 na imagem aérea com 10 cm de resolução espacial, obteve 55 épocas (*epochs*), Acurácia de 0,84 e F1-Score de 0,88. Além disso, houve o tempo de treinamento de 4 horas, 14 minutos e 32 segundos, e a sua curva de perda para treinamento e validação apresenta uma diminuição até um ponto de estabilidade com uma lacuna mínima entre os dois valores finais de perda. Entretanto, a

classificação utilizando os mesmos modelos de arquitetura na imagem orbital, com 50 cm de resolução espacial, atingiu 27 épocas (*epochs*), Acurácia de 0,796 e F1-Score de 0,86, e tempo de treinamento de 2 minutos e 24 segundos. Ademais, a figura 11 que apresenta o quadro de perda de treinamento e validação demonstra que as perdas não houve um ponto de estabilidade e muito menos uma lacuna mínima entre os dois valores finais de perda. Os resultados mais detalhados sobre o treinamento podem ser consultados nos Apêndices A (imagem de 10 cm de resolução espacial) e B (imagem de 50 cm de resolução espacial).

Alguns trabalhos corroboram resultados de Acurácia e F1-Score acima de 0.90 ou 90% (OSCO et al., 2021; TORRES et al., 2020; WAGNER et al., 2019) são satisfatórios para mapeamentos de uso de cobertura do solo. Comparando esses valores com os resultados deste trabalho, podemos dizer que não alcançamos valores sublimes. Porém, não são considerados como insatisfatórios. Há estudos que variam esse valor entre 0.70 ou 70% a 0.98 ou 98% (DU et al., 2019; LI et al., 2017; MORENO-ARMENDÁRIZ et al., 2019; ZHANG et al., 2018). Além disto, há estudos que indicam treinamento abaixo de 30 épocas (MORENO-ARMENDÁRIZ et al., 2019; TONG et al., 2020; WAGNER et al., 2019; ZHANG et al., 2020), bem como acima de 50 épocas (ORTEGA ADARME et al., 2020; TORRES et al., 2020; ZHANG et al., 2018). Dessa maneira, os valores evidenciam que estão dentro dos parâmetros aceitos pela comunidade científica, comparados a outros trabalhos, o que diferencia são os quadros de perda de treinamento e validação.

## 5.2 MAPEAMENTO DAS ÁREAS PERMEÁVEIS NA BACIA HIDROGRÁFICA DO CÓRREGO BANDEIRA

Nas imagens seguintes (figuras 12 e 13) são mapas que apresentam a área de estudo utilizando as imagens aérea e orbital, respectivamente. Imagens das quais foram utilizadas para a ferramenta de classificação. As próximas imagens/mapas (figuras 14 e 15) são resultados das classificações por *pixels* usando o aprendizado profundo.

Ademais, após a classificação realizada por *pixels*, os resultados foram convertidos de *.tiff* para *shapefile* e, dessa maneira, capazes de demonstrar a quantidade de área de cada categoria (tabelas 2 e 3). É possível perceber que houveram grandes diferenças entre as distribuições de área entre os rótulos mapeados nas duas imagens analisadas.

**Tabela 2.** Resultado de área da classificação por *pixels* da imagem aérea (resolução espacial 10 cm).

CLASSNAME	ÁREA (ha)	%	% ÁREA BACIA
VEGETAÇÃO ARBÓREA	235,66	32,82%	17,72%
VEGETAÇÃO RASTEIRA	375,60	52,30%	28,24%
SOLO EXPOSTO	106,86	14,88%	8,03%
TOTAL	718,12	100%	53,99%

**Tabela 3.** Resultado de área da classificação por *pixels* da imagem orbital (resolução espacial 50 cm).

CLASSNAME	ÁREA (ha)	%	% ÁREA BACIA
VEGETAÇÃO ARBÓREA	114,78	34,74%	8,63%
VEGETAÇÃO RASTEIRA	209,43	63,39%	15,75%
SOLO EXPOSTO	6,19	1,87%	0,47%
TOTAL	330,40	100%	24,85%

Figura 12. Área de estudo empregando a imagem aérea, resolução espacial 10 cm.

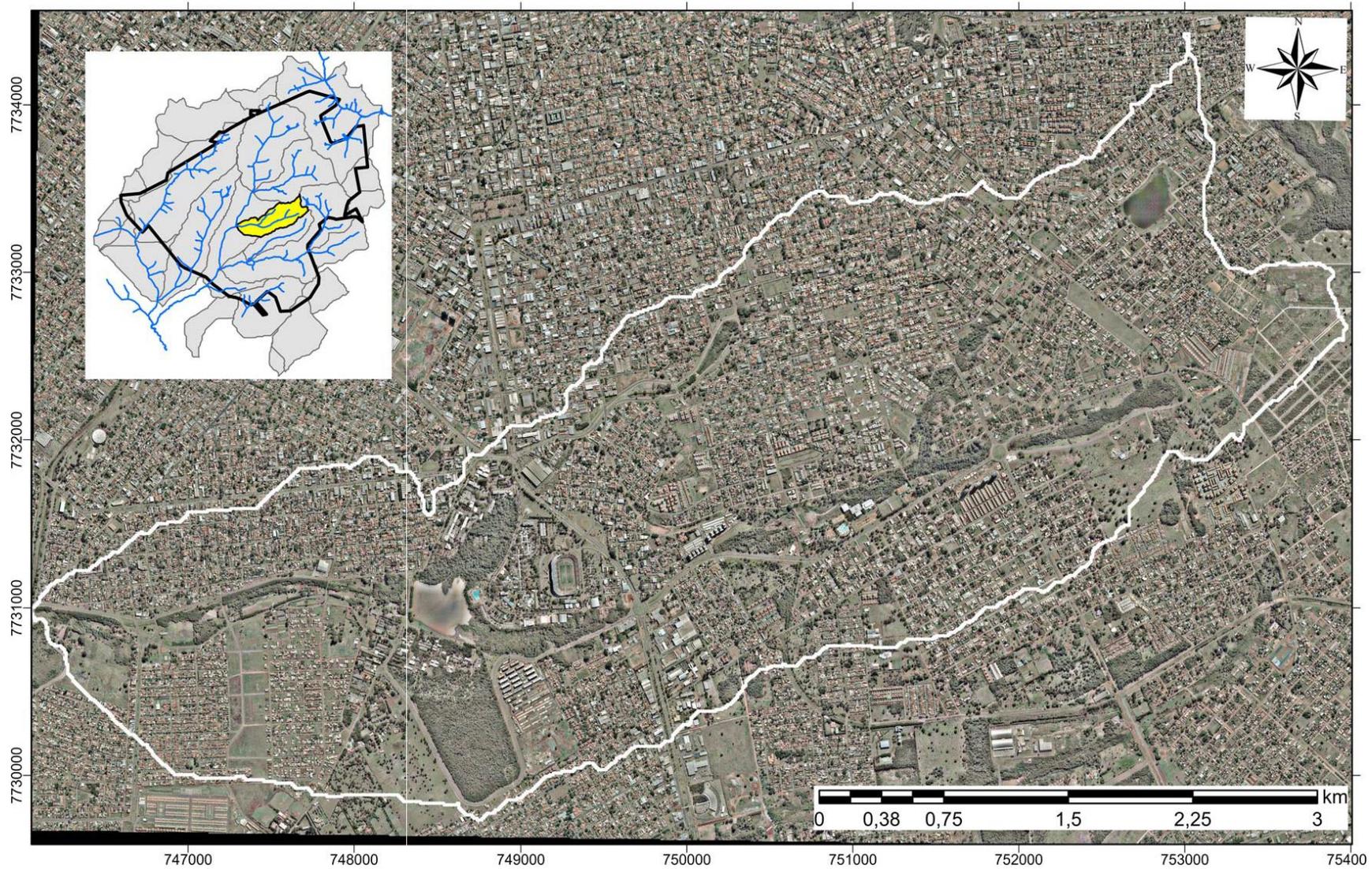


Figura 13. Área de estudo empregando imagem orbital WorldView-2, resolução espacial 50 cm.

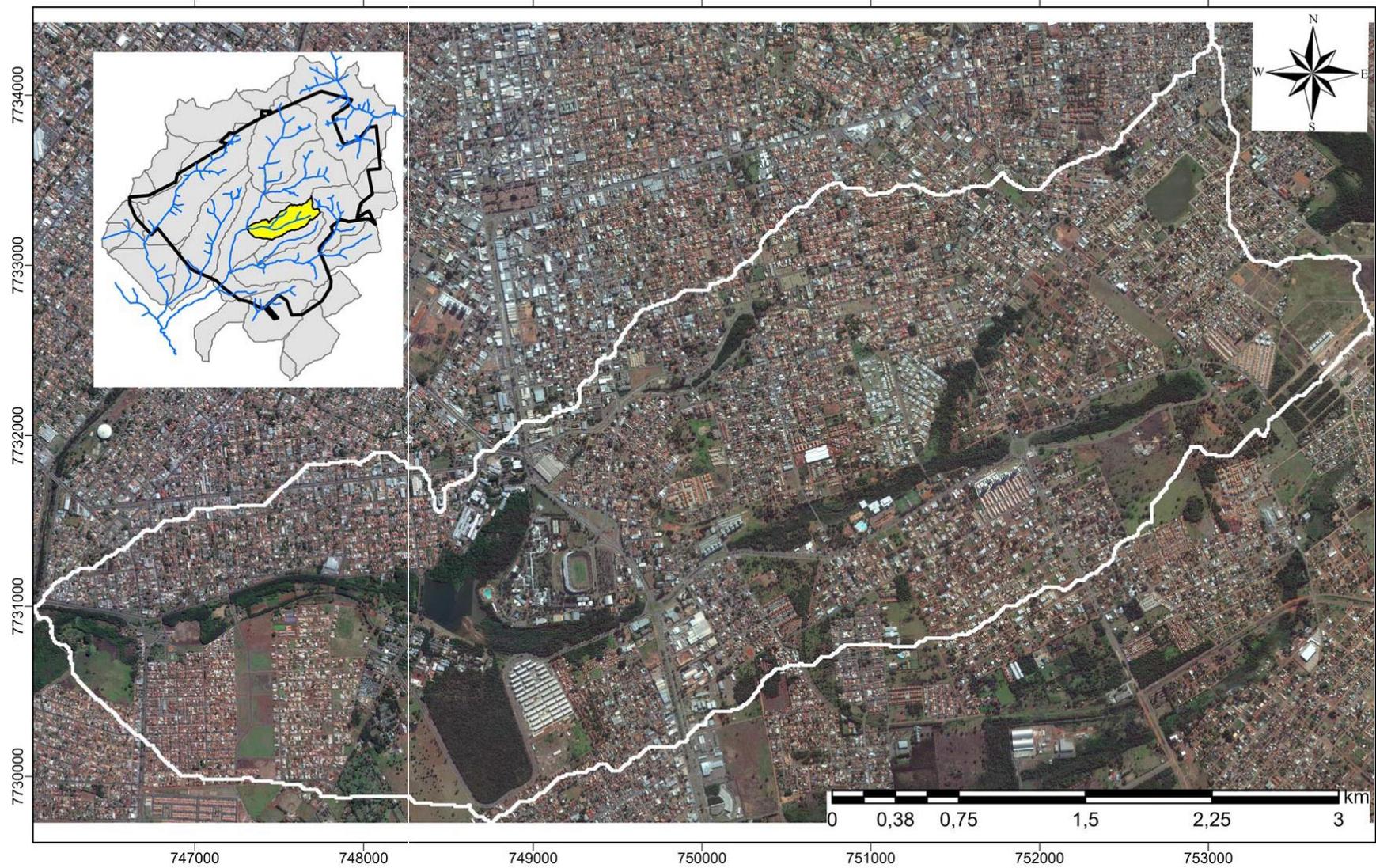


Figura 14. Classificação por *pixels* na área de estudo, empregando a imagem aérea, resolução espacial 10 cm.

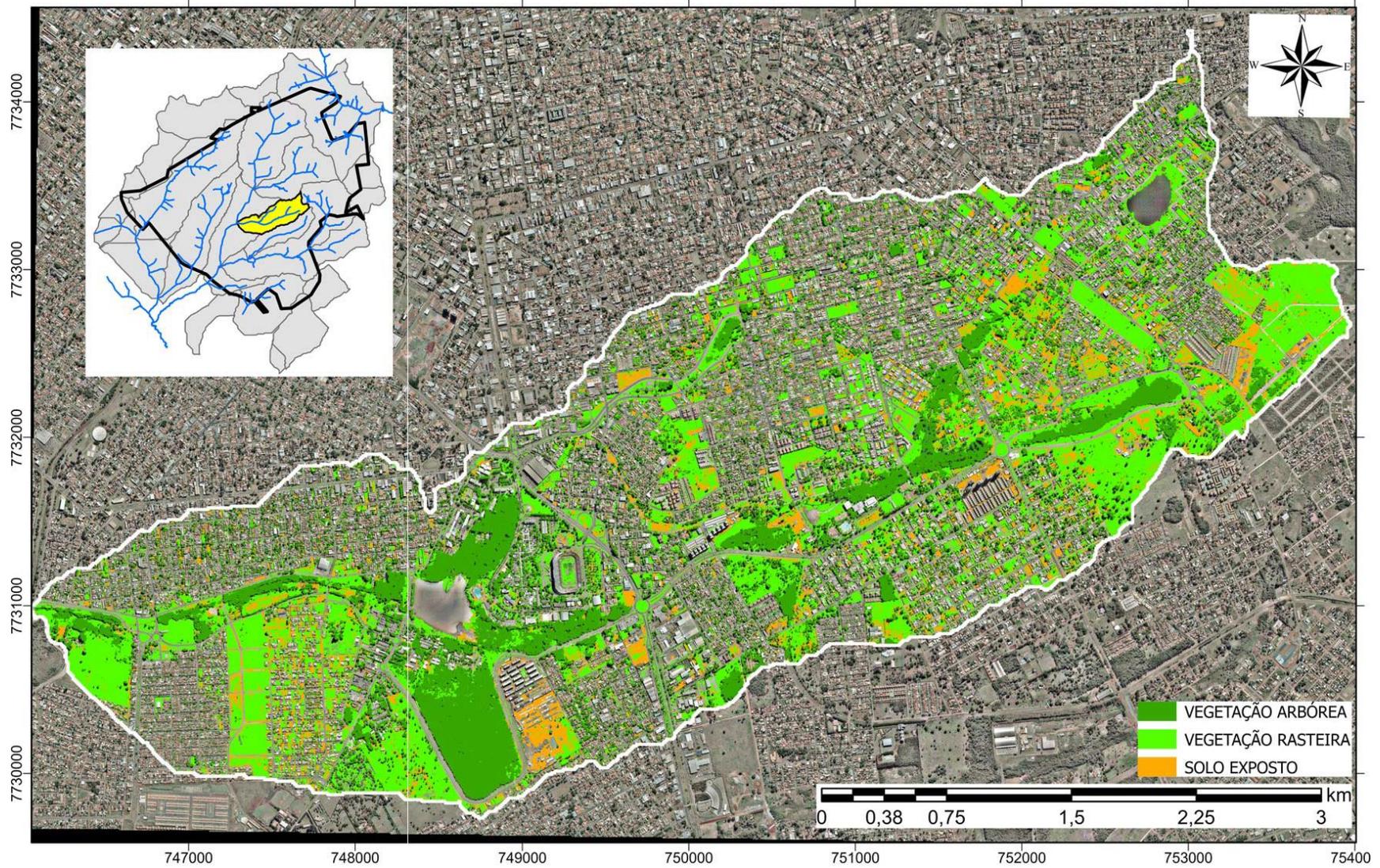
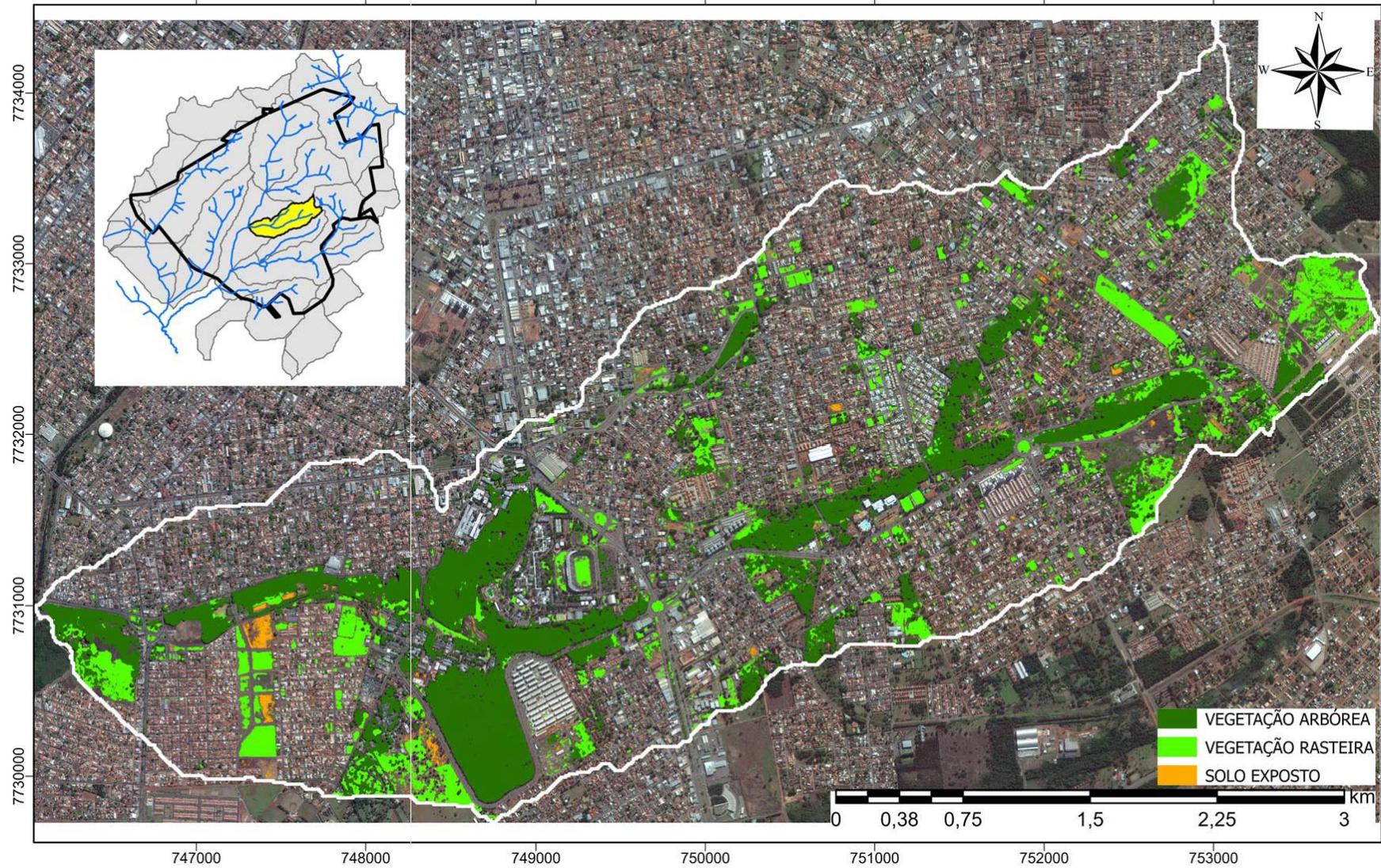


Figura 15. Classificação por *pixels* na área de estudo, empregando imagem orbital WorldView-2, cedida pelo IMASUL.



Comparando as imagens 14 e 15, bem como as tabelas 2 e 3, podemos observar que a classificação executada empregando imagem de resolução espacial 10 cm, houve uma segmentação, rotulação e identificação mais detalhada e aparentemente, mais abrangente. Enquanto a classificação utilizando imagem de resolução espacial 50 cm, houve apenas segmentação, rotulação e identificação de menor detalhamento. Como a imagem orbital já foi disponibilizada com a fusão de imagens, não sabemos que técnica foi utilizada para realizar essa fusão. Isso pode ter afetado na classificação.

Em alguns estudos, os pesquisadores aplicaram esse mesmo modelo de arquitetura. Em TORRES et al., 2020, eles usaram o aprendizado profundo com intuito de identificar uma única espécie de árvore em ambiente urbano, utilizando imagens de VANT de alta resolução espacial. No entanto, não usaram somente a arquitetura U-Net. Eles fizeram uma comparação entre diferentes modelos de Rede Totalmente Convolutiva, em busca de demonstrar o melhor modelo para a segmentação semântica da copa da árvore. Neste caso, o modelo U-Net obteve segundo melhor desempenho, perdendo para a arquitetura FC-DenseNet. Portanto, o uso do modelo U-Net combinado com imagens de resolução espacial maiores que as utilizadas neste trabalho podem corroborar para produtos de classificação melhores.

Em DU et al., 2019, utilizaram o método aprendizado profundo para mapear sistemas de agricultura familiar de pequenos agricultores, em Baodi, uma cidade da China, por meio de imagens orbital WorldView-2, com resolução espacial de 1 m. Eles também fizeram comparação entre modelos de rede neural convolutiva, bem como os modelos de aprendizado de máquina. Novamente, a arquitetura U-Net apresenta o segundo melhor desempenho. No entanto, mostra ser um modelo melhor que os de aprendizado de máquina.

Já WAGNER et al., 2019, realizaram um estudo na região da Mata Atlântica, localizado no Estado de São Paulo, usando imagens orbital WorldView-3, com resolução espacial de 30 cm. Eles delimitaram, manualmente, imagens identificando florestas naturais, florestas plantadas com eucalipto e copa de uma única espécie de árvore. Em seguida, utilizaram o método U-Net para mapear os mesmos. Eles obtiveram bons resultados, a arquitetura U-Net segmentou com precisão a floresta natural e a plantação de eucalipto. Porém, a espécie de árvore não atingiu essa precisão, pois identificaram que haveria necessidade de mais amostras de treinamento.

Deste modo, vemos que a arquitetura U-Net é empregada em diversos tipos de estudos, mas não é considerada como um dos melhores modelos existentes em *Deep Learning*, dependendo das classes a serem identificadas. Neste caso é aconselhável um melhoramento, um avanço nesta arquitetura.

### **5.2.1 Erros de classificação**

Neste item serão apresentados alguns erros de classificação que foram encontrados de forma não sistemática, mas que exemplificam as dificuldades dos classificadores no trabalho de identificação das áreas permeáveis nas imagens analisadas. Nas imagens 16 e 17 foram as classificações usando a imagem aérea, resolução espacial de 10 cm, e nas imagens 18, 19 e 20, foram classificações efetuadas usando a imagem orbital, resolução espacial de 50 cm. Podemos analisar pequenas falhas nas imagens 16 e 17 e falhas relativamente maiores nas imagens 18, 19 e 20.

No geral, os padrões previstos incluem alguns falsos negativos e falsos positivos, ou seja, falta de segmentação dentro dos polígonos de referência ou segmentação fora dos polígonos de referência (KATTENBORN; EICHEL; FASSNACHT, 2019). Em outras palavras podemos dizer que o falso negativo é aquela região que foi totalmente ou parte classificada como uma categoria, visto que não é. Por

exemplo, as regiões de lagoas que foram totalmente ou parte classificada como vegetação.

Já o falso positivo é a área na qual era para ter ocorrido a classificação, porém não ocorreu.

Ficou totalmente ou em partes sem classificação.

**Figura 16.** Erro de classificação na imagem aérea (resolução espacial 10 cm). Região do Campus da UFMS.

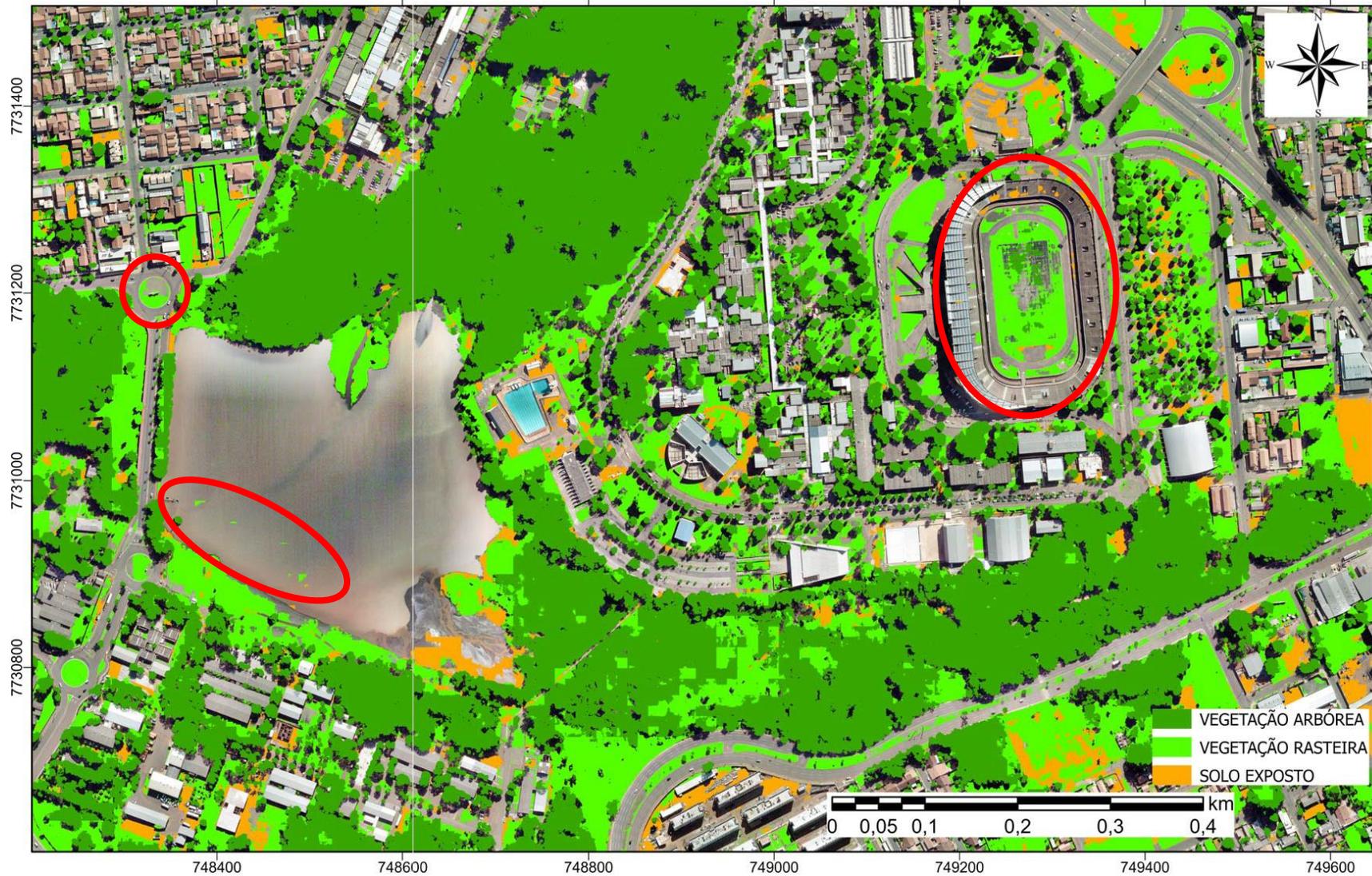


Figura 17. Erro de classificação na imagem aérea (resolução espacial 10 cm), próximo a nascente do Córrego Bandeira.



**Figura 18.** Erro de classificação na imagem orbital (resolução espacial 50 cm). Região do Campus da UFMS.

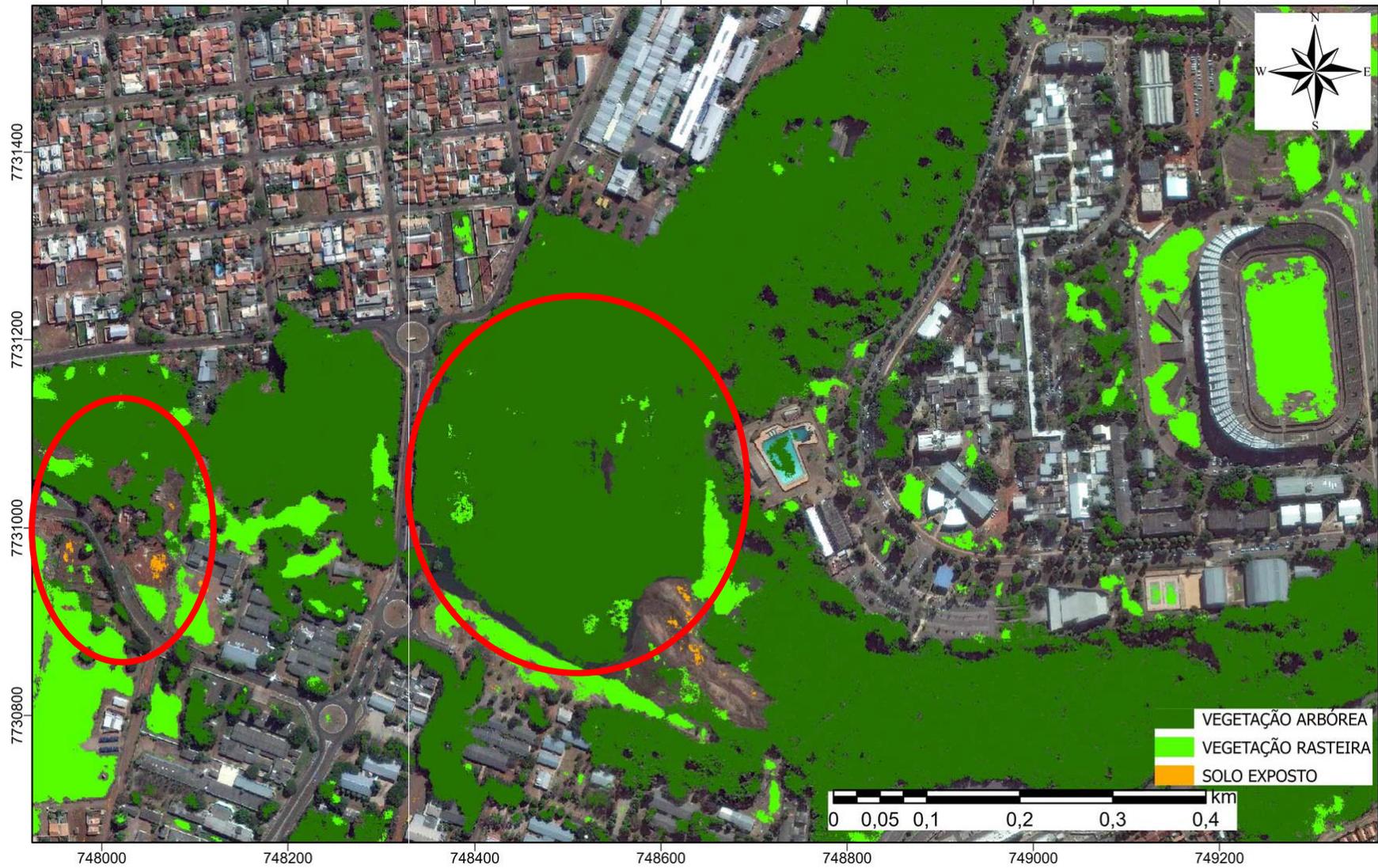
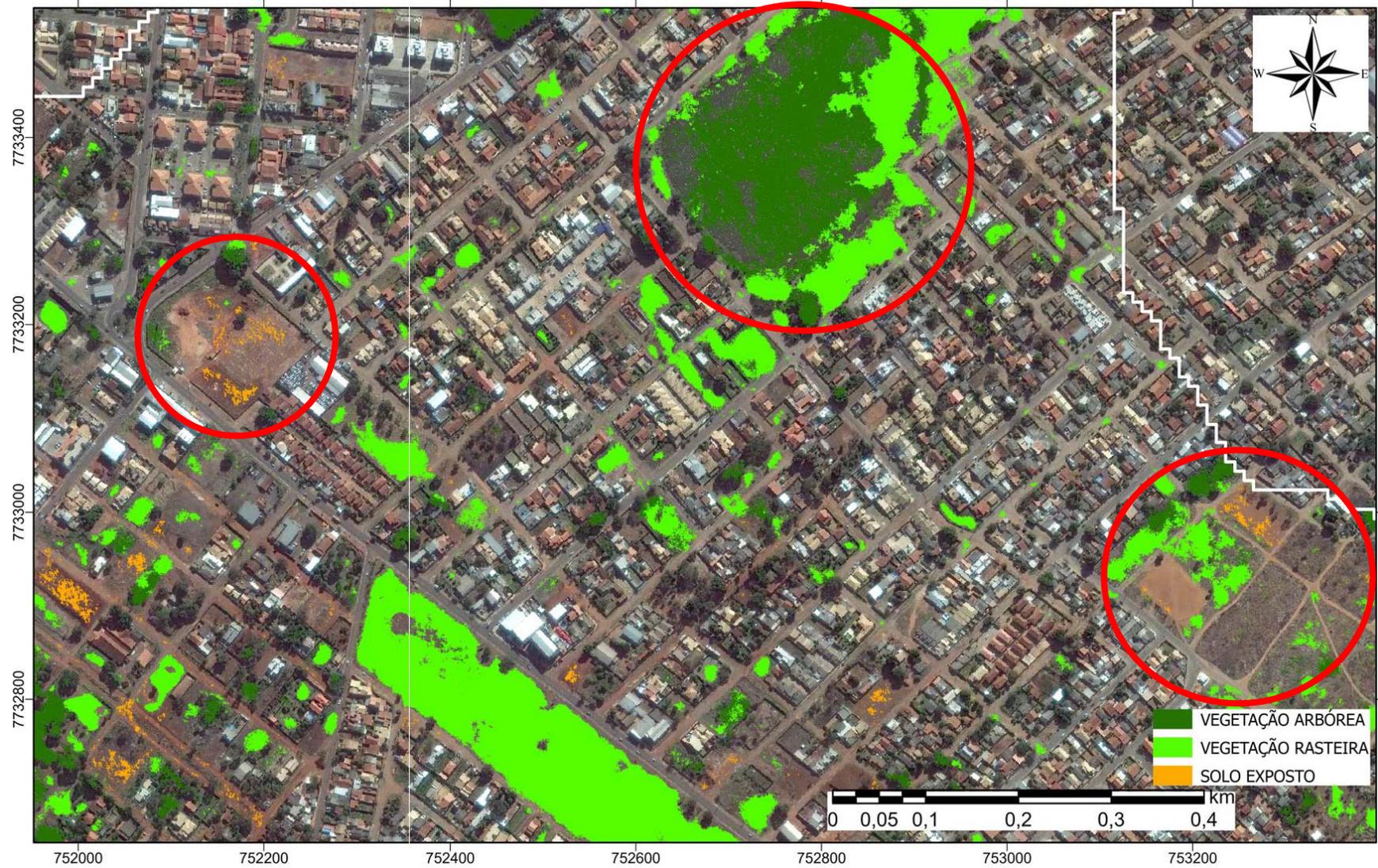


Figura 19. Erro de classificação na imagem orbital (resolução espacial 50 cm). Região do Rádio Clube Campo.



Figura 20. Erro de classificação na imagem orbital (resolução espacial 50 cm). Região da Lagoa Itatiaia.



## 6 CONCLUSÃO

Com base nos avanços tecnológicos, o *deep learning* está provando ser um conjunto de ferramentas muito bem-sucedidas, às vezes capazes de superar até mesmo os humanos na resolução de tarefas altamente computacionais. Manusear esse método tem suas vantagens em termos de classificação, especialmente em áreas complexas. Identificar e mapear as áreas permeáveis proveniente e ou inseridas em áreas urbanas, empregando o sensoriamento remoto, é relevante para estudos hidrológicos e microclima, bem como para proteção, monitoramento e manejo das mesmas.

Considerando os resultados de treinamento e os trabalhos tratados aqui, podemos atestar que a arquitetura aplicada não é vista como o melhor modelo de *Deep Learning* já utilizada. Não obtive uma precisão 100%, mas próximo disso. No entanto, ela foi capaz de segmentar e classificar uma extensa área de estudo em algumas horas. Algo que se fosse realizado manualmente poderia levar dias e dias a ser executado.

Comparando as classificações entre as imagens área e orbital, podemos analisar que as resoluções espaciais também interferiram tanto nos resultados de treinamento quanto na classificação. Por mais que ambas são conceituadas como imagens de alta resolução, a imagem orbital, com resolução de 50 cm, demonstrou ser rudimentar para esse tipo de estudo, na qual conquistou uma fraca classificação. Assim sendo, é recomendável priorizar imagens de melhores resoluções.

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA MUNICIPAL DE MEIO AMBIENTE E PLANEJAMENTO URBANO

(PLANURB). **Perfil Socioeconômico de Campo Grande**. PLANURB. 26. Ed.

Ver. Campo Grande, 2019. Disponível em:

<<http://www.campogrande.ms.gov.br/planurb/downloads/perfil->

[socioeconomico-de-campo-grande-edicao-2019/](http://www.campogrande.ms.gov.br/planurb/downloads/perfil-socioeconomico-de-campo-grande-edicao-2019/)> Acesso em 30.abr.2020.

AGÊNCIA MUNICIPAL DE MEIO AMBIENTE E PLANEJAMENTO URBANO

(PLANURB). **Educação Ambiental para a Bacia Hidrográfica do Córrego**

**Bandeira**. Disponível em:

<<http://www.campogrande.ms.gov.br/planurb/downloads/folheto-educativo->

[bandeira/](http://www.campogrande.ms.gov.br/planurb/downloads/folheto-educativo-bandeira/)> Acesso em: 30.abr.2020.

AGÊNCIA MUNICIPAL DE MEIO AMBIENTE E PLANEJAMENTO URBANO

(PLANURB). **Taxa de Relevância Ambiental**. Disponível em:

<<http://www.campogrande.ms.gov.br/planurb/downloads/taxa-de-relevancia->

[ambiental-artigo-lima-filho-a-c-et-al/](http://www.campogrande.ms.gov.br/planurb/downloads/taxa-de-relevancia-ambiental-artigo-lima-filho-a-c-et-al/)> Acesso em: 27.abr.2020.

AGÊNCIA MUNICIPAL DE MEIO AMBIENTE E PLANEJAMENTO URBANO

(PLANURB). **Carta Geotécnica – Revisão e Atualização – Produto 4 –**

**Resumo Executivo**. Disponível em:

<<http://www.campogrande.ms.gov.br/planurb/carta-geotecnica/>> Acesso em:

17.mar.2021.

ALEM, A.; KUMAR, S. Deep Learning Methods for Land Cover and Land Use

Classification in Remote Sensing: A Review. **Institute of Electrical and**

**Electronics Engineers (IEEE)**, 2020.

- BEN HAMIDA, A. et al. 3-D deep learning approach for remote sensing image classification. **IEEE Transactions on Geoscience and Remote Sensing**, v. 56, n. 8, p. 4420–4434, 2018.
- BHOSLE, K.; MUSANDE, V. Evaluation of Deep Learning CNN Model for Land Use Land Cover Classification and Crop Identification Using Hyperspectral Remote Sensing Images. **Journal of the Indian Society of Remote Sensing**, v. 47, n. 11, p. 1949–1958, 2019.
- CAMPO GRANDE. **Lei Complementar n. 341, de 4 de dezembro de 2018**. Institui o Plano Diretor de Desenvolvimento Urbano Ambiental de Campo Grande (PDDUA) e dá outras providências. Diário Oficial de Campo Grande, Campo Grande, n. 5.539, p. 1-89, 3 abr. 2019. Suplemento.
- CAMPO GRANDE. **Lei n. 4.108, de 04 de dezembro de 2003**. Denomina de “Doutor Anísio de Barros” o parque formado pela Lagoa do Itatiaia, localizado no Jardim Itatiaia, nesta capital. Diário Oficial de Campo Grande, Campo Grande, v. 6, n.1460, p.1, 5 dez. 2003.
- CAMPO GRANDE. **Lei n. 4.894, de 09 de setembro de 2010**. Denomina de Presidente Juscelino Kubitschek, o parque linear do córrego Cabaça, nesta capital. Diário Oficial de Campo Grande, Campo Grande, v. 13, n. 3113, p.3, 13 set. 2010.
- CAO, C.; DRAGIĆEVIĆ, S.; LI, S. Land-use change detection with convolutional neural network methods. **Environments - MDPI**, v. 6, n. 2, 2019.
- CAO, R. et al. Deep learning-based remote and social sensing data fusion for urban region function recognition. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 163, p. 82–97, 2020.

- CASTELLUCCIO, M. et al. **Land Use Classification in Remote Sensing Images by Convolutional Neural Networks**. 2015. Disponível em: <<http://arxiv.org/abs/1508.00092>> Acesso em: 4.8.2020.
- CHEN, G. et al. Symmetrical Dense-Shortcut Deep Fully Convolutional Networks for Semantic Segmentation of Very-High-Resolution Remote Sensing Images. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 11, n. 5, p. 1633–1644, 2018a.
- CHEN, L. C. et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 40, n. 4, p. 834–848, 2018b.
- CHENG, G. et al. Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 13, p. 3735–3756, 2020.
- CHENG, G. et al. When Deep Learning Meets Metric Learning: Remote Sensing Image Scene Classification via Learning Discriminative CNNs. **IEEE Transactions on Geoscience and Remote Sensing**, v. 56, n. 5, p. 2811–2821, 2018.
- CHENG, G.; HAN, J.; LU, X. Remote Sensing Image Scene Classification: Benchmark and State of the Art. **Proceedings of the IEEE Institute of Electrical and Electronics Engineers Inc.**, 2017.
- COUTINHO, L. M. O conceito de cerrado. **Revista Brasileira de Botânica**, v.1, n. 1, p. 17-23, 1978.
- CRÓSTA, A. P. Processamento Digital de imagens de sensoriamento remoto. Campinas, SP: **IG/UNICAMP**, 1992. 170p.

- DE BORTOLO, C. A.; ANTUNES RODRIGUES, H. L.; BORGES, M. G. Identification of Urban Green Areas from Satellite Images Worldview-II: The Case of Squares in the City of Montes Claros - MG. **GEO UERJ**, n. 32, 2018.
- DU, Z. et al. Smallholder crop area mapped with a semantic segmentation deep learning method. **Remote Sensing**, v. 11, n. 7, 2019.
- ESRI. ArcGIS Pro 2.7.3. **Copyright © 1995-2020 Esri**. All rights reserved. Published in the United States of America. 2021.
- ESRI. **Assess point cloud training results**. Disponível em: <<https://pro.arcgis.com/en/pro-app/latest/help/data/las-dataset/assessing-point-cloud-training-results.htm>> Acesso em 28.jun.2021.
- HUANG, B.; ZHAO, B.; SONG, Y. Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. **Remote Sensing of Environment**, v. 214, p. 73–86, 2018.
- KAMPFFMEYER, M.; SALBERG, A.-B.; JENSSEN, R. Urban land cover classification with missing data using deep convolutional neural networks. **2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)**, p. 5161–5164, 2017.
- KATTENBORN, T.; EICHEL, J.; FASSNACHT, F. E. Convolutional Neural Networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery. **Scientific Reports 2019 9:1**, v. 9, n. 1, p. 1–9, 27 nov. 2019.
- KUSSUL, N. et al. Deep learning classification of land cover and crop types using remote sensing data. **IEEE Geoscience and Remote Sensing Letters**, v. 14, n. 5, p. 778-782, 2017.

- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep Learning. **Nature Publishing Group**, v. 521, p. 436-444, 2015.
- LI, Y. et al. Deep learning for remote sensing image classification: A survey. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, v. 8, n. 6, 2018.
- LIN, C. H. et al. Classification of the tree for aerial image using a deep convolution neural network and visual feature clustering. **Journal of Supercomputing**, v. 76, n. 4, p. 2503–2517, 2020.
- LIU, W. et al. An Automatic Extraction Architecture of Urban Green Space Based on DeepLabv3plus Semantic Segmentation Model. **2019 IEEE 4th International Conference on Image, Vision and Computing, ICIVC**, 2019.
- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. **IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, 2015.
- MA, L. et al. Deep learning in remote sensing applications: A meta-analysis and review. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 152, p. 166-177, 2019.
- MENEZES, G. V. **Tipificação de Carcaça Bovina usando Redes Neurais Convolucionais**. Disponível em: <<https://docplayer.com.br/76478212-Tipificacao-de-carcaca-bovina-usando-redes-neurais-convolucionais.html>>  
Acesso em: 5.jul.2021.
- MINISTÉRIO DO MEIO AMBIENTE (MMA). **O Bioma Cerrado**. Disponível em: <<https://www.mma.gov.br/biomas/cerrado>> Acesso em: 1.mai.2020.

- MORENO-ARMENDÁRIZ, M. A. et al. Deep green diagnostics: Urban green space analysis using deep learning and drone images. **Sensors**, v. 19, n. 23, 2019. ISSN 14248220.
- NIJHAWAN, R. et al. A deep learning hybrid CNN framework approach for vegetation cover mapping using deep features. **Proceedings - 13th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS)**, 2017.
- ORTEGA ADARME, M. et al. Evaluation of Deep Learning Techniques for Deforestation Detection in the Brazilian Amazon and Cerrado Biomes From Remote Sensing Imagery. **Remote Sensing**, v. 12, n. 6, p. 910, 2020.
- OSCO, L. P. et al. Semantic segmentation of citrus-orchard using deep neural networks and multispectral UAV-based imagery. **Precision Agriculture**, v. 22, n. 4, p. 1171–1188, 2021.
- PARANHOS FILHO, A. C. et al. Geotecnologias em aplicações ambientais. In: \_\_\_\_\_. **Classificação de Imagens**. Editora UFMS. Campo Grande, MS. 2016. cap. 6.
- PREFEITURA MUNICIPAL DE CAMPO GRANDE (PMCG). Unidade de Planejamento Urbano (PLANURB). **Carta Geotécnica de Campo Grande. Campo Grande**, 1991.
- PUISSANT, A.; ROUGIERA, S.; STUMPF, A. Object-oriented mapping of urban trees using random forest classifiers. **International Journal of Applied Earth Observation and Geoinformation**, v. 26, n. 1, p. 235–245, 2014.
- QGIS Development Team. 2009. QGIS Geographic Information System. Open Source Geospatial Foundation. Disponível em: <[https://qgis.org/pt\\_BR/site/](https://qgis.org/pt_BR/site/)> Acesso em: 7.abr.2020.

- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 9351, p. 234–241, 2015.
- SCHRADER, S.; POUNCEY, R. Erdas Field Guide. Erdas Inc. 4th ed. Atlanta. 1997. 656p.
- SCOTT, G. J. et al. Fusion of Deep Convolutional Neural Networks for Land Cover Classification of High-Resolution Imagery. **IEEE Geoscience and Remote Sensing Letters**, v. 14, n. 9, p. 1638–1642, 1 set. 2017.
- SISGRAN. **Sistema Municipal de Indicadores de Campo Grande – MS**. Disponível em: <<https://sites.google.com/view/sisgran-cg>> Acesso em: 16.mai.2020.
- SOUZA, C. A. SILVA, M. H. S. Análise da distribuição térmica da cidade de Campo Grande, Mato Grosso do Sul, no ano de 2015. **Revista Brasileira de Climatologia. Universidade Federal do Paraná**. Curitiba, PR. v. 21, jul-dez. 2017. Disponível em: <<https://revistas.ufpr.br/revistaabclima/article/download/45752/34046+&cd=3&hl=pt-BR&ct=clnk&gl=br>> Acesso em: 2.mai.2020.
- SUDRE, C. H. et al. Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 10553 LNCS, p. 240–248, 2017.
- TONG, X. Y. et al. Land-cover classification with high-resolution remote sensing images using transferable deep models. **Remote Sensing of Environment**, v. 237, 2020.

- TORRES, D. L. et al. Applying fully convolutional architectures for semantic segmentation of a single tree species in urban environment on high resolution UAV optical imagery. **Sensors (Switzerland)**, v. 20, n. 2, p. 563, 2020.
- TUCCI, C. E. M. **Hidrologia: ciência e aplicação**. 2 ed. Porto Alegre. Editora da Universidade Federal do Rio Grande do Sul/ABRH. 2001.
- WAGNER, F. H. et al. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. **Remote Sensing in Ecology and Conservation**, v. 5, n. 4, p. 360–375, 1 dez. 2019.
- WEN, D. et al. Semantic Classification of Urban Trees Using Very High-Resolution Satellite Imagery. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 10, n. 4, p. 1413–1424, 2017.
- WENG, Q. et al. Land-Use Classification via Extreme Learning Classifier Based on Deep Convolutional Features. **IEEE Geoscience and Remote Sensing Letters**, v. 14, n. 5, p. 704–708, 1 maio 2017.
- YU, X. 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.
- YUAN, Q. et al. Deep learning in environmental remote sensing: Achievements and challenges. **Remote Sensing of Environment**, n. 241, 2020.
- ZHANG, C. et al. Joint Deep Learning for land cover and land use classification. **Remote Sensing of Environment**, v. 221, p. 173–187, 2019.
- ZHANG, P. et al. Urban land use and land cover classification using novel deep learning models based on high spatial resolution satellite imagery. **Sensors (Switzerland)**, v. 18, n. 11, 2018.

ZHANG, X. et al. How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery?

**Remote Sensing**, v. 12, n. 3, 2020.

ZHAO, W.; DU, S.; EMERY, W. J. Object-Based Convolutional Neural Network for High-Resolution Imagery Classification. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 10, n. 7, p. 3386–3396, 2017.

ZHU, X. X. et al. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. **IEEE Geoscience and Remote Sensing Magazine Institute of Electrical and Electronics Engineers Inc.**, 2017.

# APÊNDICES

## Apêndice A. Treinamento de modelo *Deep Learning*, empregando a imagem aérea (ESRI *ArcGIS Pro*, 2021).

```
Train Deep Learning Model
=====
Parameters

Input Training Data      D:\AURIMAR\DADOS DE TREINAMENTO\BACIA AEREA\Dados_Classtyles_256_2
Output Model             D:\AURIMAR\DADOS DE TREINAMENTO\BACIA AEREA\Treinamento_256_ClassTiles_100_8_ResNet34_2
Max Epochs              100
Model Type              UNET
Batch Size              8
Model Arguments         class_balancing False;mixup False;focal_loss False;ignore_classes #;chip_size 224
Learning Rate           0,0001
Backbone Model          RESNET34
Pre-trained Model
Validation %            20
Stop when model stops improving  STOP_TRAINING
Output Model
Freeze Model            FREEZE_MODEL
=====
Environments

Processor Type          GPU
=====
Messages

Start Time: sexta-feira, 24 de dezembro de 2021 09:45:21
Learning Rate - 0.0001
epoch  training loss      validation loss      accuracy      Dice
0      0.902317464351654    0.9706555008888245 0.5881204605102539 0.3830822706222534
1      0.8979329466819763    0.8773415684700012 0.6387645602226257 0.4694981575012207
2      0.8188696503639221    0.7829930782318115 0.6890063881874084 0.5384387373924255
3      0.7990508079528809    0.7522072196006775 0.7058806419372559 0.5732413530349731
4      0.8287116289138794    0.7569169402122498 0.6993520855903625 0.5453937649726868
5      0.8069427609443665    0.72967529296875   0.711428701877594   0.5662848353385925
6      0.6815469861030579    0.6818142533302307 0.7345691919326782 0.6061285138130188
7      0.7162151336669922    0.6499798893928528 0.7436714172363281 0.6092227697372437
8      0.7154361605644226    0.6453841924667358 0.7566283941268921 0.6383335590362549
9      0.7156908512115479    0.6826753616333008 0.7280438542366028 0.5791572332382202
10     0.6805835366249084    0.6412366628646851 0.7611202001571655 0.6440766453742981
11     0.7326304316520691    0.65855473279953   0.7460695505142212 0.6274745464324951
12     0.6260421872138977    0.6057342290878296 0.7735263109207153 0.6532676815986633
13     0.6646179556846619    0.6513450145721436 0.755092442035675   0.614916205406189
14     0.5841274261474609    0.5936603546142578 0.7754295468330383 0.6595197319984436
15     0.6110982298851013    0.5735681056976318 0.7874378561973572 0.6853018999099731
16     0.5852247476577759    0.5525801181793213 0.7887307405471802 0.6841443777084351
17     0.5820899605751038    0.5423882603645325 0.7927442789077759 0.6971685290336609
18     0.6187533736228943    0.569369912147522  0.7872903943061829 0.673123300075531
19     0.5558943152427673    0.5694586634635925 0.7897565364837646 0.6751146912574768
20     0.5425260066986084    0.5362017154693604 0.7995509505271912 0.6957376003265381
21     0.6149815917015076    0.5887256860733032 0.7709823250770569 0.6599777936935425
22     0.553087592124939     0.5809904336929321 0.7906155586242676 0.6956160664558411
23     0.5633346438407898    0.5360633730888367 0.7971906065940857 0.6977291703224182
24     0.5509155988693237    0.5219136476516724 0.8071714043617249 0.7158123850822449
25     0.5224413871765137    0.5144686698913574 0.8095419406890869 0.7185149788856506
26     0.5135447978973389    0.5017135739326477 0.8123390674591064 0.7182783484458923
27     0.5206547975540161    0.5028279423713684 0.8130705952644348 0.7291247844696045
28     0.5906927585601807    0.561953604221344   0.796825110912323   0.6850603222846985
29     0.5070327520370483    0.5232239961624146 0.80829304445671082 0.7137256264686584
30     0.5078654885292053    0.5040712952613831 0.8171480894088745 0.7385346293449402
31     0.49370142817497253   0.4955061674118042 0.8141152858734131 0.7191703915596008
32     0.4929220378398895    0.4974876046180725 0.8205985426902771 0.7442463636398315
33     0.5153607726097107    0.4950143098831177 0.8162809014320374 0.7332454323768616
34     0.48508909344673157   0.47736385464668274 0.8236966133117676 0.7476930618286133
35     0.499822199344635     0.5136959552764893 0.8172943592071533 0.7345479130744934
36     0.4946441650390625    0.49078452587127686 0.8136066794395447 0.715328574180603
37     0.4687643349170685    0.4717901051044464 0.8262432813644409 0.7422630786895752
38     0.5017763376235962    0.4703581631183624 0.8277046084403992 0.7454290390014648
39     0.47622349858283997   0.466543376444577026 0.8263264894485474 0.7528554797172546
40     0.4909381568431854    0.4723946452140808 0.8259541988372803 0.7519589066505432
41     0.4665083587169647    0.4624060392379761 0.8315123915672302 0.7584150433540344
42     0.4796871840953827    0.45325639843940735 0.832145631313324   0.749838650226593
43     0.48495224118232727   0.44988545775413513 0.8349943161010742 0.76185542345047
44     0.48451563715934753   0.46117743849754333 0.8323835730552673 0.7550882697105408
45     0.4971890151500702    0.4545482099056244 0.8334781527519226 0.7605776190757751
46     0.4475092589855194    0.43558812141418457 0.837688148021698  0.7549735903739929
47     0.5187127590179443    0.4552404284477234 0.8321442604064941 0.7490283250808716
```

```

48      0.46329036355018616  0.4412502348423004  0.8377328515052795  0.7558998465538025
49      0.4684405028820038  0.4247075617313385  0.8429737091064453  0.7727974653244019
50      0.4788033664226532  0.4358670711517334  0.8401799201965332  0.768434464931488
51      0.45069289207458496  0.4371659457683563  0.838302493095398  0.7531997561454773
52      0.4549824297428131  0.4323031008243561  0.8413540124893188  0.7717405557632446
53      0.44143936038017273  0.43404000997543335  0.8394080400466919  0.7709444761276245
54      0.4436796009540558  0.4313890039920807  0.8411971926689148  0.7643742561340332
55      0.4448695182800293  0.4372527003288269  0.8409845232963562  0.7659639120101929
{'accuracy': '8.4297e-01'}

NoData  VEGETACAO_RASTEIRA  VEGETACAO_ARBOREA  SOLO_EXPOSTO
precision 0.908035      0.765096      0.818990      0.830409
recall   0.860870      0.879870      0.813881      0.729066
f1       0.883823      0.818479      0.816428      0.776445
Succeeded at sexta-feira, 24 de dezembro de 2021 13:59:52 (Elapsed Time: 4 hours 14 minutes 32 seconds)

```

## Apêndice B. Treinamento de modelo *Deep Learning*, empregando imagem orbital WorldView-2 (ESRI ArcGIS Pro, 2021).

Train Deep Learning Model

=====

Parameters

```

Input Training Data      D:\AURIMAR\DADOS DE TREINAMENTO\BACIA ORBITAL\Dados_Classsfiles_256
Output Model            D:\AURIMAR\DADOS DE TREINAMENTO\BACIA ORBITAL\Treinamento_256_ClassTiles_100_8_ResNet34
Max Epochs              100
Model Type              UNET
Batch Size              8
Model Arguments         class_balancing False;mixup False;focal_loss False;ignore_classes #;chip_size 224
Learning Rate           0,0001
Backbone Model         RESNET34
Pre-trained Model
Validation %            20
Stop when model stops improving      STOP_TRAINING
Output Model
Freeze Model           FREEZE_MODEL
=====

```

Environments

```

GPU ID      0
Processor Type  GPU
=====

```

Messages

Start Time: quinta-feira, 23 de dezembro de 2021 07:37:43

Learning Rate - 0.0001

epoch	training loss	validation loss	accuracy	Dice
0	1.347520112991333	1.0555588006973267	0.6302053332328796	0.008480609394609928
1	1.2531884908676147	1.066044569015503	0.6145857572555542	0.01565815880894661
2	1.2079488039016724	1.0450422763824463	0.62385493516922	0.023212281987071037
3	1.1730023622512817	0.9986497759819031	0.6382442712783813	0.03183569014072418
4	1.149364948272705	0.9323570728302002	0.6437902450561523	0.028042636811733246
5	1.1723483800888062	0.8915300965309143	0.6432865262031555	0.015384143218398094
6	1.1612178087234497	0.8771296143531799	0.6445747017860413	0.0197591632604599
7	1.1392393112182617	0.8719407916069031	0.6489320993423462	0.03330962359905243
8	1.1352884769439697	0.8650093078613281	0.6486548781394958	0.028599651530385017
9	1.1071850061416626	0.827504575252533	0.6499630212783813	0.028463168069720268
10	1.0737404823303223	0.8219538331031799	0.6538293957710266	0.042267806828022
11	0.9591581225395203	0.8124998211860657	0.7022045850753784	0.18774418532848358
12	0.9544088244438171	0.763522207736969	0.6964393854141235	0.15886721019476501
13	0.8825231194496155	0.7244096994400024	0.73687344789505	0.24256311357021332
14	0.8377359509468079	0.7259619832038879	0.761383593082428	0.3078227937221527
15	0.8690484762191772	0.7187948226928711	0.7421983480453491	0.2514141798019409
16	0.8252027630805969	0.6804251074790955	0.7678118348121643	0.2907835841178894
17	0.6708250641822815	0.7522407174110413	0.7362990975379944	0.4052352011203766
18	0.8381349444389343	0.6865430474281311	0.7335288524627686	0.24992795288562775
19	0.6809722781181335	0.6676608324050903	0.790269136428833	0.3936331570148468
20	0.7423624992370605	0.642569899559021	0.7820181250572205	0.33021482825279236
21	0.6432906985282898	0.5926344394683838	0.7962843179702759	0.39138853549957275
22	0.7861218452453613	0.6558099985122681	0.7616100311279297	0.30095601081848145
23	0.7042055726051331	0.6309838891029358	0.7855367064476013	0.3549128770828247
24	0.791009783744812	0.7190759181976318	0.7661123871803284	0.3663117587566376
25	0.9163873195648193	0.7021398544311523	0.7030978202819824	0.16796833276748657
26	0.7210731506347656	0.637231707572937	0.7914323210716248	0.35239216685295105
27	0.636667072729797	0.5975439548492432	0.7999604940414429	0.3948652148246765

{'accuracy': '7.9628e-01'}

```

NoData  VEGETACAO_ARBOREA  VEGETACAO_RASTEIRA  SOLO_EXPOSTO
precision 0.814724      0.811192      0.335745      0.090113
recall   0.914015      0.753756      0.132903      0.006519
f1       0.861518      0.781420      0.190427      0.012159
Succeeded at quinta-feira, 23 de dezembro de 2021 07:40:07 (Elapsed Time: 2 minutes 24 seconds)

```