



PÓS-GRADUAÇÃO EM GEOCIÊNCIAS APLICADAS E
GEODINÂMICA
INSTITUTO DE GEOCIÊNCIAS
UNIVERSIDADE DE BRASÍLIA

Projeto de Pesquisa de Mestrado
**EXTRAÇÃO AUTOMÁTICA DE EDIFICAÇÕES PARA A
PRODUÇÃO CARTOGRÁFICA UTILIZANDO
INTELIGÊNCIA ARTIFICIAL**

Philippe Borba

Orientador: Prof. Dr. Edilson de Souza Bias

Coorientador: Prof. Dr. Nilton Correia da Silva

Brasília, Agosto de 2020



PÓS-GRADUAÇÃO EM GEOCIÊNCIAS APLICADAS E
GEODINÂMICA
INSTITUTO DE GEOCIÊNCIAS
UNIVERSIDADE DE BRASÍLIA

PROJETO DE PESQUISA DE MESTRADO
EXTRAÇÃO AUTOMÁTICA DE EDIFICAÇÕES PARA A PRODUÇÃO
CARTOGRÁFICA UTILIZANDO INTELIGÊNCIA ARTIFICIAL

Área de Concentração: Geoprocessamento e Análise Ambiental

Linha de Pesquisa: Avaliação de Dados e Técnicas de Sensoriamento Remoto,
Geoprocessamento, Cartografia e Geodésia

Philippe Borba

Orientador: Prof. Dr. Edilson de Souza Bias

Coorientador: Prof. Dr. Nilton Correia da Silva

Proposta de Projeto de Mestrado apresentada
na Disciplina Seminário I, como pré-requisito
do Programa de Pós-Graduação em Geociên-
cias Aplicadas e Geodinâmica, do Instituto de
Geociências da Universidade de Brasília.

Professor: Dr. Henrique Llacer Roig

Brasília - DF

Agosto de 2020



Lista de Figuras

1	Distribuição espacial das 840 imagens brutas referentes aos estados de Santa Catarina e Rio Grande do Sul	9
2	Exemplo de imagem (a) e edificações adquiridas sobre a imagem (b) .	9
3	Fluxograma das etapas da pesquisa	13
4	Fluxograma dos passos para obtenção dos dados	15
5	Exemplo de imagens, rótulos e predições. Da esquerda pra direita, imagem de validação, imagem de rótulo (<i>ground truth</i>) e imagem predita. Os pixels amarelos correspondem aos de edificação. Tal diagrama é gerado durante o treinamento, para que se possa acompanhar o andamento do mesmo.	16
6	Exemplo de imagens e e predição. Os pixels amarelos correspondem aos de edificação.	17
7	Exemplo de generalização efetuada para melhorar a vetorização. Adaptado de https://github.com/Microsoft/USBuildingFootprints	18
8	Exemplo de generalização efetuada para melhorar a vetorização. Em amarelo os pixels correspondentes a edificações, em roxo os pixels que não correspondem a edificações. Na figura (a), os dados de pixels brutos, na figura (b), em azul, uma visualização dos dados vetoriais a serem obtidos após a generalização.	18



Lista de Tabelas

1	Tecnologias escolhidas para o Projeto de Pesquisa	10
2	Tecnologias escolhidas para o Projeto de Pesquisa	11
3	Cronograma da Pesquisa	20



Sumário

1	INTRODUÇÃO	4
2	OBJETIVO	7
2.1	Objetivos específicos	7
3	FUNDAMENTAÇÃO TEÓRICA	8
4	MATERIAIS E MÉTODOS	8
4.1	Conjunto de dados	8
4.2	Tecnologias escolhidas	10
4.3	Hardware	11
4.4	Descrição das métricas utilizadas no treinamento	11
4.4.1	Acurácia por pixel	11
4.4.2	Acurácia por classe	12
4.4.3	<i>Overall accuracy, precision, recall</i> , índice F_1 e o coeficiente de correlação de Matthews (<i>MCC</i>)	12
4.4.4	Índice de Jaccard (IoU) e índice médio de intersecção sobre união (mIoU)	13
4.5	Descrição do método utilizado para a pesquisa	13
4.6	Descrição do processo para obtenção dos polígonos de edificações . .	15
4.6.1	Pré-Processamento de dados	15
4.6.2	Treinamento da rede	16
4.6.3	Inferência utilizando a rede treinada	17
4.6.4	Vetorização dos rasters inferidos e pós-processamento vetorial .	17
4.6.5	Avaliação da qualidade dos dados vetoriais obtidos	18
5	RESULTADOS ESPERADOS	19
6	CRONOGRAMA	20
7	VIABILIDADE TÉCNICO-FINANCEIRA	20
8	REFERÊNCIAS BIBLIOGRÁFICAS	22



1 INTRODUÇÃO

Com o avanço da tecnologia, há cada vez mais satélites de imageamento modernos com alta resolução espacial, o que pode permitir um estudo mais acurado da superfície terrestre (BITTNER et al., 2018). As imagens dos referidos satélites permitem estudos de uso e cobertura do solo (LIU et al., 2020), de detecção de mudanças da paisagem (ZHANG et al., 2019), de expansão de mancha urbana (KHANAL et al., 2019), além de ser empregável insumos na produção cartográfica (HOBEL et al., 2015), dada a capacidade de se extrair informações de imagens, sendo dessa forma um dos principais insumos nas atividades de mapeamento.

Particularmente para a produção cartográfica, tal capacidade de se extrair informações é interessante pois, em tese, permite a automatização de processos e, por conseguinte, uma maior produção de Cartas Topográficas. O decreto-lei nº 243 de 28 de fevereiro de 1967 (BRASIL, 1967) define no artigo 2º o Sistema Cartográfico Nacional, que é constituído de entidades públicas e privadas, com a finalidade de realizar o mapeamento sistemático do território brasileiro, nas escalas 1:1.000.000 a 1:25.000. Dentre as entidades públicas responsáveis pelo mapeamento, pode-se citar a Diretoria de Serviço Geográfico (DSG) e o Instituto Brasileiro de Geografia e Estatística (IBGE).

O decreto-lei nº 6.666 de 27 de novembro de 2008 (BRASIL, 2008) criou a Comissão Nacional de Cartografia (CONCAR), a qual é responsável por normatizar a produção e disseminação de dados geoespaciais no âmbito do Governo Federal. A DSG e o IBGE, dentre outros órgãos, fazem parte da CONCAR. Para que dados produzidos se tornem dados de referência no SCN, é obrigatório que as informações obtidas sejam compatíveis com as normas e padrões nacionais definidos pela CONCAR, especificações estas abaixo relacionadas:

- Especificação Técnica para Produtos de Conjuntos de Dados Geoespaciais (ET-PCDG) (DSG, 2016a): especificação que especifica os produtos do SCN, bem como os metadados;
- Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais (ET-EDGV) (CONCAR, 2010): especificação que define o modelo de dados oficial brasileiro;
- Especificação Técnica para Aquisição de Dados Geoespaciais Vetoriais (ET-ADGV) (DSG, 2018): especificação que define as regras de aquisição da geometria dos dados, bem como os tamanhos mínimos dos objetos; e
- Especificação Técnica para Controle de Qualidade dos Produtos de Conjuntos



de Dados Geoespaciais (ET-CQDG) (DSG, 2016b): especificação que define os procedimentos para aferir a qualidade dos dados produzidos.

No SCN existe uma situação de falta de Cartas Topográficas na escala 1:25.000, dos 46.391 produtos dessa escala que recobrem o território brasileiro, há apenas 1.447 produtos disponíveis no Banco de Dados Geográficos do Exército (BDGEx) (DSG, 2010), representando 3,12% do total. Tal percentagem pode ser explicada pela dificuldade de mapeamento da escala 1:25.000, dado que de todas as escalas previstas para mapeamento no SCN (1:1.000.000 a 1:25.000), conforme descritos no decreto-lei nº243 (BRASIL, 1967) e especificado pela ET-PCDG (DSG, 2016a), a escala em questão é a que contém maior detalhamento, dado que as imagens adequadas utilizadas como insumos nessa escala devem ter maior resolução. Com isso, os produtos que requerem maior investimento de recursos humanos para a elaboração.

Atualmente, no âmbito dos projetos de mapeamento da DSG, parte das informações necessárias para a elaboração de produtos cartográficos é extraída manualmente, como as geometrias de edificações e de áreas edificadas. Portanto, a aplicação de métodos automáticos de extração de informação de imagens de satélite na produção cartográfica pode vir a acelerar algumas etapas do processo produtivo da DSG.

Para a extração de informações de imagens oriundas de sensoriamento remoto, tradicionalmente, utiliza-se de técnicas de segmentação baseadas em intervalos de histograma, em agrupamento de pixels, em regiões e em identificação de bordas (BLASCHKE, 2010; CHEN et al., 2018). Além disso, podem-se elencar sem exaurir todas as técnicas existentes, segmentação e classificação baseada em objeto, tradução do termo *object based image analysis* (OBIA), análise de textura, de reflectância e de formato (SU et al., 2008; WISEMAN et al., 2009; MA et al., 2017; KUCHARCZYK et al., 2020).

Como evolução dos conhecimentos supramencionados, pode-se encontrar na literatura, como em Li et al. (2014), em Damodaran et al. (2017), Thanh Noi e Kappas (2017), em Jozdani et al. (2019) e em Li et al. (2020), aquelas conhecidas como técnicas de aprendizado de máquina utilizadas no Sensoriamento Remoto, como *Support Vector Machines* (SVM), *Random Forests* (RF) e *Decision Trees* (DT).

Em adição, nos últimos anos também ocorreram avanços tecnológicos no campo da Computação, particularmente no campo da Inteligência Artificial (IA). Tais evoluções podem ser explicadas, em partes, pela melhoria do poder computacional e pelo advento de novos algoritmos de processamento de dados utilizando placas de vídeo (GPU) (CHEN, 2016; BALL et al., 2017). Nesse contexto, uma técnica que vem sendo largamente empregada em diversos estudos científicos é denominada *Deep Learning*



(DL).

Com o advento dos algoritmos de DL, a comunidade científica passou a se interessar nesse tipo de conhecimento dada a maior acurácia obtida ao utilizá-los, como pode ser verificado nos estudos Ortega Adarme et al. (2020), Mboga et al. (2019), Guirado et al. (2017), nos quais técnicas clássicas das geociências foram comparadas com aplicações utilizando o Estado da Arte de DL.

Segundo Lecun et al. (2015), Hoeser e Kuenzer (2020), ao contrário dos métodos que são baseados apenas nos valores de pixels para identificar padrões, o aspecto chave de DL é que este método se vale de múltiplas camadas de abstração obtidas por processos matemáticos não lineares que permitem a inferência de informações escondidas, conjunto de técnicas de IA conhecidas como redes convolucionais profundas (*deep convolutional neural networks*). Pode-se citar como características desse tipo de algoritmo, a necessidade de muitos dados para o treinamento e a exigência de alto poder computacional (NAJAFABADI et al., 2015).

No contexto das geociências, há diversos problemas que possuem soluções propostas utilizando *Deep Learning*, como detecção de objetos (MUSYAROFAN et al., 2020; LI et al., 2020; CHENG et al., 2016; HU et al., 2015; CHEN et al., 2019), classificação de imagens hiperespectrais (ZHU et al., 2017; DATA et al., 2015; AUDEBERT et al., 2019; PAN et al., 2020), estudo de uso e cobertura do solo (ALHASSAN et al., 2020; JOZDANI et al., 2019; CARRANZA-GARCÍA et al., 2019), superresolução (LANARAS et al., 2018; QIN et al., 2020; SALVETTI et al., 2020), detecção de mudanças (SONG; CHOI, 2020; KULKARNI; VENUGOPAL, 2020; LIU et al., 2019; WANG et al., 2018) e, finalmente, segmentação semântica (DIAKOIANNIS et al., 2019; ZHANG et al., 2016; KEMKER et al., 2018; YANG et al., 2018), que é uma técnica de visão computacional que permite ao computador reconhecer e extrair objetos de imagens (SHAPIRO, 2001), e se adequa ao caso de uso de extração automática de feições para a elaboração de produtos cartográficos.

Diante dos avanços elencados e dadas as necessidades de mapeamento supra-mencionadas, uma pesquisa cujo resultado fosse aplicar o Estado da Arte de DL em atividades de mapeamento seria justificável. Porém, para que se possa utilizar na produção cartográfica as informações inferidas utilizando-se uma rede convolucional profunda, algumas operações de generalização cartográfica são necessárias para pós-processar os dados inferidos devido aos serrilhados, excesso de vértices e ao formato das linhas nem sempre serem compatíveis com as regras de aquisição definidas pelo SCN. Autores como Partovi et al. (2017), Li et al. (2017), Lokhat e Touya (2016) tratam de algumas formas de generalizar os dados obtidos, logo, pesquisas sobre esses procedimentos também são necessárias.



Sendo assim, identifica-se um problema científico de encontrar uma metodologia em que se possa utilizar arquitetura de rede neural convolucional profunda, treinada com um conjunto consistente de dados, para inferir informações geográficas no formato vetorial que possam representar as formas poligonais de feições no terreno como por exemplo, corpos d'água, vegetação, terreno exposto, áreas densamente edificadas e edificações, com acurácia aderente à escala 1:25.000, seguindo as normas do SCN. Logo, uma pesquisa cujo objetivo seja investigar formas de utilizar inteligência artificial para realizar tais extrações para a cartografia se justifica.

Portanto, considerando-se apenas o problema de extração automática de edificações, o presente projeto de pesquisa tem como proposta de hipótese científica que é possível treinar uma rede convolucional profunda que seja capaz de segmentar e extrair feições vetoriais no formato poligonal de edificações, geometricamente consistentes, com acurácia compatível com a escala 1:25.000, em conformidade com as normas definidas pelo SCN.

2 OBJETIVO

Essa pesquisa tem por objetivo determinar uma metodologia de treinamento de rede neural convolucional profunda, cuja rede treinada seja capaz de segmentar imagens de satélite de alta resolução com a finalidade de extrair geometrias de edificação com características e níveis de qualidade aderentes aos padrões definidos pelo SCN, na escala 1:25.000.

2.1 Objetivos específicos

Como objetivos específicos da pesquisa, pretende-se:

- (a) Identificar arquiteturas de redes neurais para realizar segmentação semântica;
- (b) Verificar qual é a melhor arquitetura para ser utilizada em atividade de mapeamento, levando-se em consideração o tempo de processamento em relação a acurácia obtida;
- (c) Verificar se os dados segmentados utilizando as arquiteturas identificadas tem acurácia compatível com o SCN;
- (d) Identificar o melhor método, ou combinação de métodos, para simplificar (generalizar) os polígonos obtidos; e



- (e) Avaliar se os polígonos obtidos do processo de generalização possuem acurácia compatível para a elaboração de produtos cartográficos na escala 1:25.000.

3 FUNDAMENTAÇÃO TEÓRICA

Será elaborada na disciplina Seminário II.

4 MATERIAIS E MÉTODOS

A presente seção apresentará como a pesquisa será desenvolvida. Para tal, a seção 4.1 descreverá os dados que serão utilizados para os treinamentos de redes neurais, a seção 4.2 descreverá as tecnologias escolhidas para o desenvolvimento do projeto, a seção 4.3 descreverá os equipamentos utilizados na pesquisa, a seção 4.4 elencará as métricas utilizadas no treinamento, a seção 4.5 descreverá o método utilizado para verificar qual rede é a mais adequada para o uso em atividades de mapeamento, e, por fim, a seção 4.6 detalhará os passos para o treinamento e avaliação.

4.1 Conjunto de dados

O 1º Centro de Geoinformação (1º CGEO), organização militar diretamente subordinada à DSG, elaborou um conjunto de dados (*dataset*) utilizando imagens obtidas por meio de voo fotogramétrico, dos Estados do Rio Grande do Sul e Santa Catarina, as quais foram cedidas ao 1º CGEO, respectivamente, pela Secretaria de Planejamento, Orçamento e Gestão do Estado (SPGG) do Rio Grande do Sul e pela Secretaria de Estado do Desenvolvimento Econômico Sustentável (SDS) do Estado de Santa Catarina.

Há 840 imagens no formato ECW e elas ocupam um espaço em disco total de 196Gb. As resoluções espaciais são de 35 cm para as imagens referentes ao Estado do Rio Grande do Sul e de 39 cm para o Estado de Santa Catarina, ambas em RGB. A resolução radiométrica é de 8 bits para ambas. A distribuição espacial das imagens pode ser visualizada na figura 1.

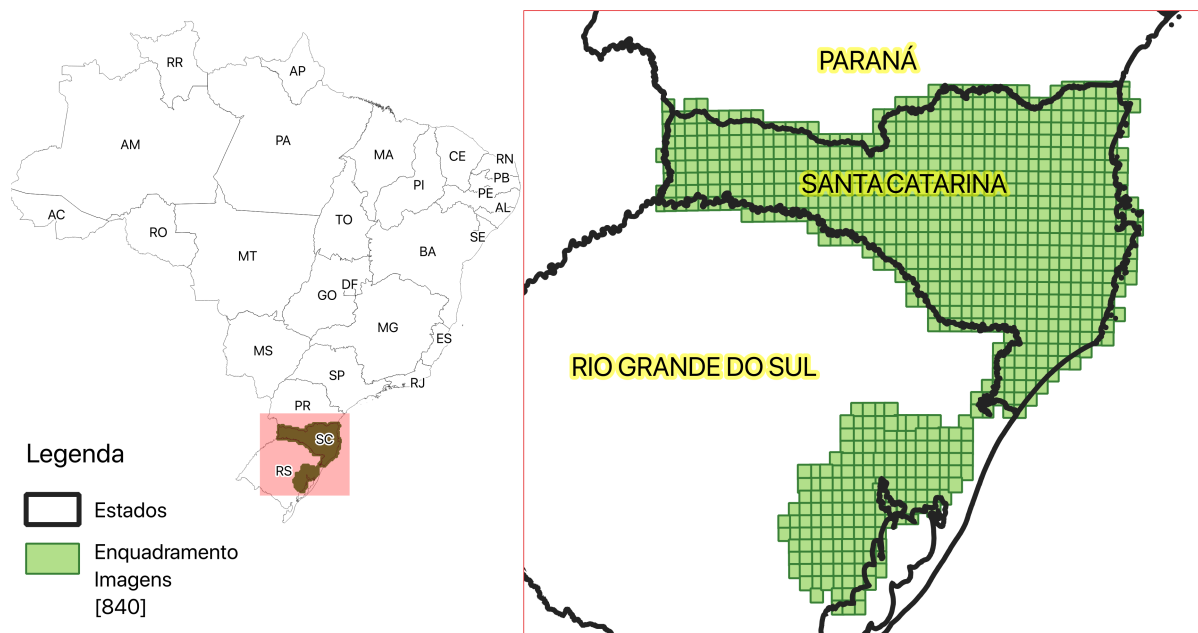


Figura 1: Distribuição espacial das 840 imagens brutas referentes aos estados de Santa Catarina e Rio Grande do Sul

Das ortoimagens supradescritas e apresentadas na figura 1, o 1º CGEO extraiu os polígonos que correspondem às edificações, obtendo-se dessa forma uma base com mais de 1 milhão e seiscentas mil edificações, no formato vetorial, com um detalhamento maior que o necessário para a escala 1:25.000. Um exemplo de extrato de imagem e das edificações extraídas (polígonos azuis) pode ser visto na figura 2.



Figura 2: Exemplo de imagem (a) e edificações adquiridas sobre a imagem (b)



4.2 Tecnologias escolhidas

Como critério de escolha das tecnologias, procurou-se utilizar todas com licença de software livre e de código aberto, bem como as soluções com melhores documentações, maiores comunidades e, na medida do possível, melhor desempenho. A tabela 1 elenca as escolhas, bem como as descrições e justificativas.

Tecnologia	Descrição e Justificativa
Kubuntu 18.04	Sistema operacional Linux, baseado na distribuição ubuntu, com interface KDE.
Docker (MERKEL, 2014)	O docker é um software livre, tendência de mercado, que serve para replicar o ambiente de desenvolvimento e de produção em qualquer máquina, padronizando a infraestrutura e encapsulando as dependências.
Tensorflow 2.2 (ABADI et al., 2015) e Keras (CHOLLET et al., 2015)	Framework para implementação e treinamento de redes neurais. Juntamente com o Pytorch, é líder de mercado em termos de performance e de facilidade de desenvolvimento. Foi escolhido por conta do alto desempenho, da rica documentação e da crescente popularidade. A biblioteca Keras funciona juntamente com a biblioteca Tensorflow e serve para escrever de modo mais fácil as arquiteturas de redes neurais.
Python 3.8 (ROSSUM; DRAKE, 2009)	Linguagem de programação utilizada juntamente com o framework Tensorflow para treinamento de redes neurais. No escopo desse trabalho, também será utilizada para implementação de scripts de pré-processamento, de plugins para o QGIS, de rotinas para generalização e rotinas de avaliação da qualidade dos dados.
Jupyter (KLUYVER et al., 2016)	Ferramenta de prototipação rápida e visualização de dados utilizando <i>scripts</i> Python.
QGIS 3.12 (QGIS Development Team, 2009)	Sistema de Informações Geográficas (SIG) escrito em software livre e de código aberto, baseado em C++. Possui a possibilidade de ser estendido por meio de plugins escritos na linguagem de programação Python. Nesse trabalho, servirá para a visualização dos dados, preparação das máscaras de treinamento, geoprocessamento e avaliação da qualidade.
PostgreSQL 12 (The PostgreSQL Global Development Group, 1996) e PostGIS 3.0 (Refractons Research Inc, 2005)	PostgreSQL é um Sistema Gerenciador de Banco de Dados (SGBD) escrito em software livre e de código aberto. PostGIS é uma extensão do PostgreSQL que permite o armazenamento e a manipulação de dados geográficos. Os dados fornecidos pelo 1º CGEO foram extraídos utilizando essas duas tecnologias.

Tabela 1: Tecnologias escolhidas para o Projeto de Pesquisa



4.3 Hardware

Conforme elencado na Introdução, os métodos de *deep learning* necessitam de hardware especializado denominado GPU (CHEN, 2016; BALL et al., 2017). Sendo assim, a pesquisa contará com duas infraestruturas, uma do Laboratório de Inteligência Artificial (Lab IA) do curso de Ciência da Computação da UnB, campus Gama e uma pessoal do aluno, ambas descritas na tabela 2.

Hardware Pessoal	Hardware do Lab IA
<ol style="list-style-type: none">1. Processador Core i7-4770;2. 32 Gb de memória RAM;3. HD SSD de 256 Gb para o sistema e HD SSD de 1Tb para armazenamento dos dados; e4. Placa de vídeo dedicada NVIDIA GeForce RTX 2080 Ti com 11Gb de memória de vídeo.	<ol style="list-style-type: none">1. Processador Core i5;2. 14 Gb de memória RAM;3. HD de 1Tb para armazenamento dos dados; e4. Placa de vídeo dedicada NVIDIA TITAN Xp com 12Gb de memória de vídeo.

Tabela 2: Tecnologias escolhidas para o Projeto de Pesquisa

4.4 Descrição das métricas utilizadas no treinamento

Para cada tile inferido, pode-se calcular as métricas: acurácia por pixel (*Pixel Accuracy*), acurácia por classe (*class-average accuracy*), *overall accuracy* (*OA*), *precision*, *recall*, índice F_1 (F_1 score), o coeficiente de correlação de Matthews (*MCC*), índice de Jaccard (*IoU*), e índice médio de intersecção sobre união (mIoU). As subseções a seguir descreverão cada item acima relacionado.

4.4.1 Acurácia por pixel

A acurácia por pixel, segundo Yu et al. (2018) é uma avaliação de acurácia por pixel. Sejam \hat{l}_i e l_i , respectivamente o rótulo predito e o rótulo da verdade de campo (*ground truth*) e N o número total de pixels de todas as imagens de teste. A acurácia por pixel (PA) pode ser calculada por:

$$PA = \frac{1}{N} \sum_i^N \delta(\hat{l}_i, l_i) \quad (1)$$



onde $\delta(x, y)$ é uma função definida por

$$\delta(x, y) = \begin{cases} 1, & \text{se } x = y \\ 0, & \text{caso contrário} \end{cases} \quad (2)$$

4.4.2 Acurácia por classe

Ainda segundo Yu et al. (2018), para cada classe k , a acurácia da predição ($ACC(k)$) é definida por:

$$ACC(k) = \frac{\sum_i^N (\delta(l_i, k) \& \delta(\hat{l}_i, k))}{\sum_i^N \delta(l_i, k)} \quad (3)$$

onde o denominador representa o total de pixels da classe k do conjunto de verdade de campo, N é o número total de pixels de todas as imagens de teste e $\delta(x, y)$ é a função definida na equação 2. Dado o total de classes L , a acurácia média da classe k é dada por

$$C_{ACC} = \frac{\sum_k^L ACC(k)}{L} \quad (4)$$

4.4.3 Overall accuracy, precision, recall, índice F_1 e o coeficiente de correlação de Matthews (MCC)

Segundo Diakogiannis et al. (2019), *overall accuracy*, *precision*, *recall*, índice F_1 e o coeficiente de correlação de Matthews (MCC) são definidos por:

$$OA = \frac{TP + TN}{FP + FN} \quad (5)$$

$$precision = \frac{TP}{TP + FP} \quad (6)$$

$$recall = \frac{TP}{TP + FN} \quad (7)$$

$$F_1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (8)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (9)$$

onde TP , FP , FN e TN são os verdadeiros positivos, falso positivos, falsos negativos e verdadeiros negativos da inferência respectivamente.



4.4.4 Índice de Jaccard (IoU) e índice médio de intersecção sobre união (mIoU)

Segundo Yu et al. (2018), o índice de Jaccard (IoU) pode ser calculado por:

$$IoU = J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (10)$$

onde A e B são, respectivamente, a verdade de campo e os dados preditos.

Ainda segundo Yu et al. (2018), o índice médio de intersecção sobre união (mIoU) pode ser calculado por:

$$mIoU = \frac{1}{m} \sum \frac{A_{pred} \cap A_{true}}{A_{pred} \cup A_{true}} \quad (11)$$

onde m é o número de classes esperadas, A_{pred} é o conjunto de predições e A_{true} é o conjunto correspondente à verdade de campo.

4.5 Descrição do método utilizado para a pesquisa

A pesquisa proposta neste documento seguirá o fluxograma descrito na figura 3, cujas explicações podem ser visualizados nos itens imediatamente após à referida ilustração.

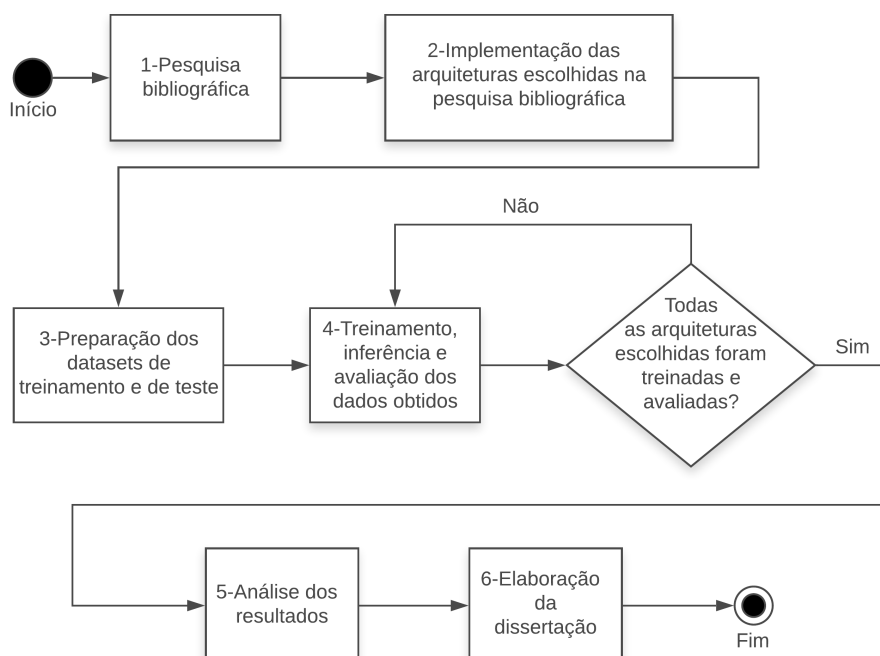


Figura 3: Fluxograma das etapas da pesquisa



1. Pesquisa bibliográfica: etapa que será realizado o levantamento do Estado da Arte das arquiteturas de redes convolucionais profundas tratadas no presente projeto de pesquisa, levando-se em consideração a acurácia de cada método descrito;
2. Implementação das arquiteturas escolhidas na pesquisa bibliográfica: de posse das arquiteturas selecionadas no item 1, nesta etapa as mesmas serão implementadas em Python, utilizando os *frameworks* Tensorflow e Keras;
3. Preparação dos datasets de treinamento e de teste: com base nos dados descritos na seção 4.1, as imagens deverão ser recortadas para um formato que a rede possa entender, como por exemplo 512x512 pixels. Os recortes deverão ser divididos em três conjuntos, um de treino, um de teste e um de validação, na proporção de, aproximadamente, 60% para o primeiro e 20% para o segundo, e 20% para o terceiro. Tal separação tem por objetivo impedir que a rede memorize os casos do dataset de teste e que exista um conjunto diferente de imagens para validação, cuja acurácia obtida nesse conjunto servirá para verificar a qualidade do treinamento;
4. Treinamento, inferência e avaliação dos dados obtidos: Para cada arquitetura implementada, será realizado um treinamento utilizando-se os datasets de treinamento e teste e métricas específicas que serão descritas na seção 4.4. Após o treinamento, será feita a inferência no dataset de teste, obtendo-se imagens binárias em que os pixels correspondentes ao valor 1 serão aqueles que pertencem a uma edificação e os de valor 0 que não são de edificações. Em seguida, será executada a operação de vetorização, para transformar os dados do formato raster para o formato vetorial, seguido de uma operação de generalização, para ajustar o formato dos vetores obtidos. Por fim, será realizada a avaliação dos dados vetoriais;
5. Análise dos resultados: com base nos dados obtidos por cada treinamento, as métricas definidas na seção 4.4 serão calculadas e consolidadas em tabelas. Em seguida, será realizada uma comparação dos dados obtidos e conclusões poderão ser tiradas dessa análise; e
6. Elaboração da dissertação: De posse de todos os procedimentos executados nas etapas anteriores e, levando-se em consideração a melhor arquitetura para o problema proposto, será escrita a dissertação de mestrado, na qual será proposta uma metodologia para o treinamento da mesma e a inferência dos dados.



4.6 Descrição do processo para obtenção dos polígonos de edificações

Para a obtenção dos polígonos é necessário executar os procedimentos descritos no fluxograma apresentado na figura 4, cujas explicações estão dispostas nas subseções a seguir.

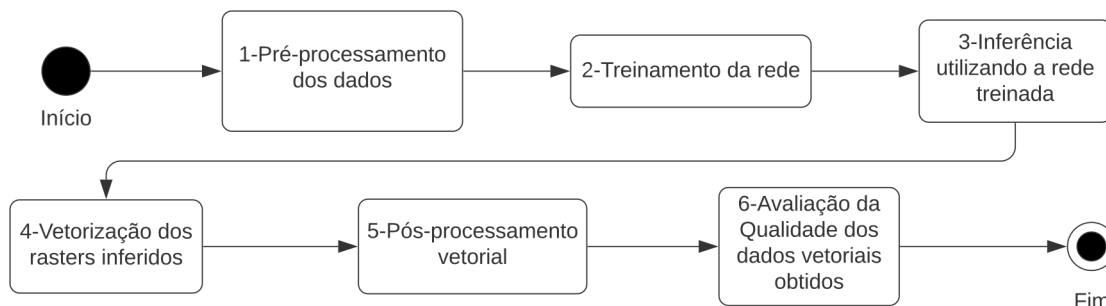


Figura 4: Fluxograma dos passos para obtenção dos dados

4.6.1 Pré-Processamento de dados

Primeiramente, é necessário realizar a etapa de Pré-Processamento de Dados, descrita pela etapa 1 da figura 4. Nesta etapa, as imagens são divididas em tamanhos menores, de acordo com a arquitetura de rede escolhida. Em seguida, com base nos polígonos já adquiridos, é gerada uma imagem binária que determina a localização e a ausência de edificações na cena. Esse procedimento é semelhante ao descrito no item (c) da seção 4.5, pois de posse dos dados obtidos no processamento do referido item, dependendo da arquitetura de rede escolhida, será necessário pegar novamente um recorte menor.

Nessa etapa é feito um processo chamado *data augmentation*, no qual operações de histograma, operações de escala e rotações são aplicados às amostras para gerar imagens sintéticas, aumentando assim a quantidade de imagens do dataset (KRIZHEVSKY et al., 2017).

Na etapa de pré-processamento de dados, existem as seguintes oportunidades de pesquisa:

- (a) Estudo da influência do tamanho das imagens recortadas e da sobreposição dos recortes no resultado do treinamento (REINA et al., 2020);
- (b) Estudo da influência do tamanho do *dataset* no resultado do treinamento (SO-EKHOE et al., 2016); e



- (c) Estudo da influência da geração de imagens sintéticas para o treinamento, também chamado *data augmentation* (GARCIA-GARCIA et al., 2018).

4.6.2 Treinamento da rede

Após o recorte, de posse das imagens geradas na etapa anterior, é realizado o treinamento da rede, cuja teoria foi descrita na seção 3. É importante que durante essa etapa os gráficos de treinamento sejam monitorados a fim de identificar a ocorrência de *underfit* ou *overfit* (GOODFELLOW et al., 2016), bem como monitorar os resultados das predições. A figura 5 exemplifica os resultados parciais da etapa de validação obtidos durante o treinamento.



Figura 5: Exemplo de imagens, rótulos e predições. Da esquerda pra direita, imagem de validação, imagem de rótulo (*ground truth*) e imagem predita. Os pixels amarelos correspondem aos de edificação. Tal diagrama é gerado durante o treinamento, para que se possa acompanhar o andamento do mesmo.

Na etapa de treinamento da rede, existem as seguintes oportunidades de pesquisa:

- (a) Estudo da influência da arquitetura escolhida nas métricas de treinamento (MA et al., 2019);
- (b) Estudo da influência da inicialização da rede nas métricas de treinamento (MA et al., 2019);
- (c) Estudo da influência das funções de perda (*loss functions*) nas métricas de treinamento, como estudado em Guo et al. (2020) e Diakogiannis et al. (2020); e



- (d) Estudo da influência da escolha dos hiperparâmetros na convergência e nas métricas de treinamento (JÖNSSON et al., 2020).

4.6.3 Inferência utilizando a rede treinada

Após o processo de treinamento, é realizada a inferência citada no fluxograma da figura 4, na qual são utilizados os pesos exportados da rede treinada para que se possa extrair informações de imagens. O resultado da inferência é uma imagem binária, onde os pixels de valor 1 correspondem a pixels identificados como pertencentes a edificação e pixels com valor 0 não pertencentes a uma edificação. Um exemplo de resultado de tal procedimento pode ser visualizado na figura 6.

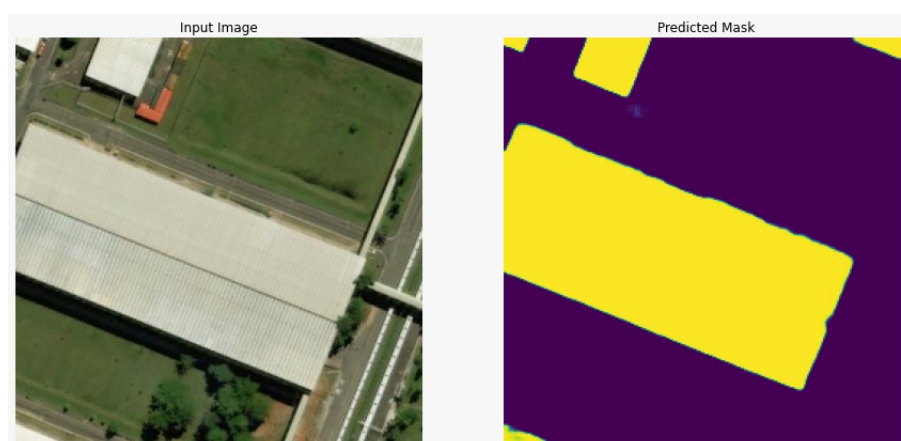


Figura 6: Exemplo de imagens e e predição. Os pixels amarelos correspondem aos de edificação.

Como oportunidade de pesquisa, pode-se citar o estudo de inferência simples versus influência com sobreposição e posterior média (DIAKOIANNIS et al., 2020).

4.6.4 Vetorização dos rasters inferidos e pós-processamento vetorial

Os próximos processos descritos no fluxograma da figura 4 são os de Vetorização dos rasters inferidos e o de Pós-processamento vetorial. O processo de vetorização nada mais é do que transformar os agrupamentos de pixels em polígonos no formato vetorial. Devido ao resultado irregular descrito na figura 6, os polígonos obtidos também serão irregulares. Sendo assim, é necessário que sejam realizados pós processamentos para obter polígonos simples que se assemelham a edificações, como os descritos na figura 2. A figura 7 mostra um exemplo de procedimento aplicado em cada região para generalizar os polígonos obtidos da vetorização.

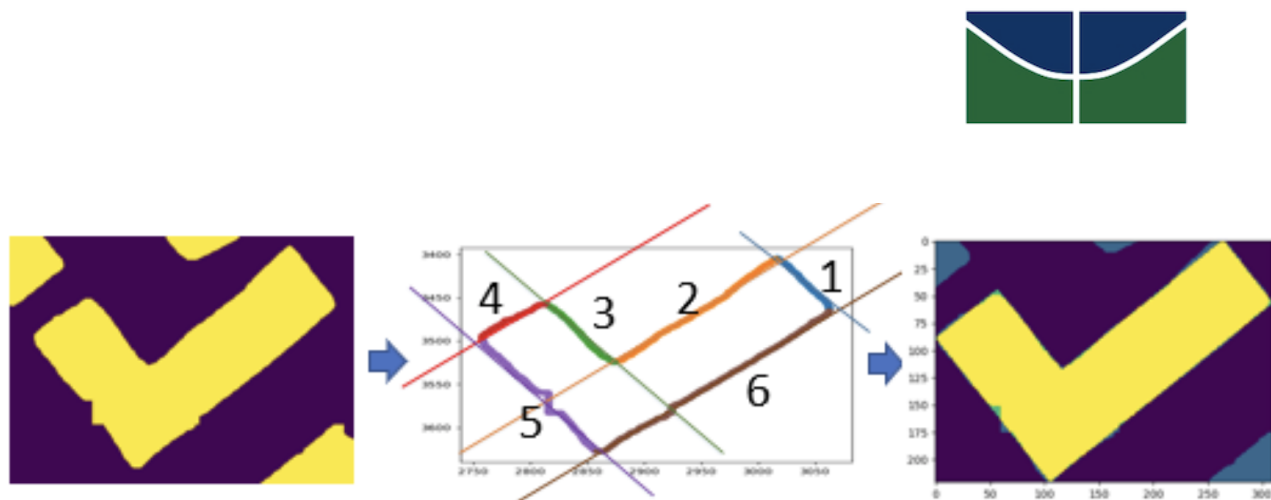


Figura 7: Exemplo de generalização efetuada para melhorar a vetorização. Adaptado de <https://github.com/Microsoft/USBuildingFootprints>

A figura 8 mostra os dados originais (a) e os dados generalizados (linhas em azul) (b).

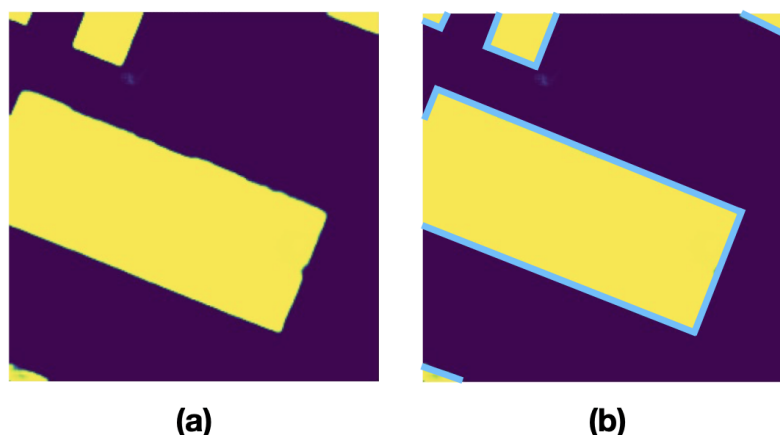


Figura 8: Exemplo de generalização efetuada para melhorar a vetorização. Em amarelo os pixels correspondentes a edificações, em roxo os pixels que não correspondem a edificações. Na figura (a), os dados de pixels brutos, na figura (b), em azul, uma visualização dos dados vetoriais a serem obtidos após a generalização.

Como oportunidade de pesquisa para a etapa de vetorização dos rasters inferidos, pode-se citar o estudo da influência da vetorização baseada em *threshold* de valor versus a influência da filtragem a priori no resultado da acurácia final. Com relação à etapa de pós-processamento vetorial, pode-se citar o estudo da influência de diferentes métodos de generalização e simplificação, para obtenção de polígonos com menos vértices, na acurácia final do produto vetorial (PARTOVI et al., 2017);

4.6.5 Avaliação da qualidade dos dados vetoriais obtidos

Por fim, deve ser realizada uma avaliação da qualidade dos dados vetoriais obtidos, com a finalidade de determinar qual é a escala de mapeamento adequada para que



os dados inferidos possam ser utilizados.

Nessa etapa, existe a oportunidade de estudo de métodos de pareamento de objetos (*matching methods*) para a avaliação vértice a vértice da acurácia (AVBELJ et al., 2015).

5 RESULTADOS ESPERADOS

Como resultados da pesquisa em questão, pretende-se:

1. Obter uma metodologia de treinamento de rede neural que seja capaz de receber imagens de alta resolução e segmentar as regiões que contém edificações, de forma que seja possível extrair a geometria desses objetos em formato vetorial, incluindo a arquitetura
2. Analisar comparativamente a qualidade dos resultados em relação a um conjunto de dados de referência com o objetivo de concluir se os dados obtidos são adequados para a elaboração de produtos cartográficos na escala 1:25.000;
3. Desenvolver rotinas utilizando-se tecnologias livres e de código aberto para:
 - (a) automatizar o treinamento da rede neural;
 - (b) segmentar conjunto de imagens utilizando a rede treinada;
 - (c) avaliar a acurácia dos dados obtidos em relação ao conjunto de dados de referência;
 - (d) obter os polígonos em formato vetorial compatível com os padrões da *Open Geospatial Consortium* (OGC) e;
 - (e) avaliar a qualidade dos dados vetoriais obtidos sob a óptica do Mapeamento Sistemático Brasileiro.

Vale salientar que, devido o fato das implementações efetuadas neste trabalho serem escritas utilizando tecnologias livres e de código aberto, a aplicação do resultados seja em outros trabalhos acadêmicos, seja em diferentes projetos de mapeamento, pode ser verificada utilizando-se as rotinas desenvolvidas e, modificando-nas quando for o caso.



6 CRONOGRAMA

As atividades da pesquisa foram planejadas de acordo com os dados da Tabela 3. Os meses do item 1- Cursar Disciplinas da tabela 3 que estão em "?" se referem a possibilidade de extensão do período letivo tendo em vista a pandemia do COVID-19.

Atividades	Meses																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1- Cursar Disciplinas	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	?	?	?	?								
2- Revisão Bibliográfica				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓						
3- Implementação de Algoritmos				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓						
4- Treinamento da Rede Neural					✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓					
5- Análise dos Resultados															✓	✓	✓	✓	✓	✓	✓			
6- Elaboração de artigos									✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
7- Elaboração da Dissertação												✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
8- Entrega da Dissertação																							✓	
9- Defesa																								✓

Tabela 3: Cronograma da Pesquisa

7 VIABILIDADE TÉCNICO-FINANCEIRA

Com respeito à viabilidade técnica, conforme exposto na seção 4.1, o dados já estão disponíveis, ou seja, um tempo considerável foi economizado para a pesquisa, dado que extrair por conta própria os dados de um conjunto de imagens, da mesma forma que o 1º CGEO fez para montar o *dataset* descrito na referida seção, seria inviável no tempo disponível para a pesquisa.

Ademais, conforme exemplificado na introdução teórica, há diversos estudos atuais no âmbito acadêmico relativos ao uso de técnicas de *Deep Learning* para segmentação semântica, aplicadas ao Sensoriamento Remoto, como Diakogiannis et al. (2019), Zhang et al. (2016), Kemker et al. (2018) e Yang et al. (2018), com bons resultados, mostrando assim que trata-se de uma técnica viável para o presente projeto.

Quanto à viabilidade financeira, o aluno não recebe bolsa por ser militar de carreira do Exército. Não há necessidade de trabalhos de campo, nem haverá custos com aquisição de dados, uma vez que os dados já foram obtidos em projetos de mapeamento anteriores, nem haverá aquisição de equipamentos para a pesquisa, dado que o aluno utilizará a infraestrutura própria, bem como do Laboratório de Inteligência



Artificial da UnB Campus Gama. Como serão utilizadas somente tecnologias livres de código aberto, não haverá custo adicional ao laboratório com aquisição de licenças.

Portanto, como não há custos extras ao orçamento da UnB com a pesquisa, já que os equipamentos disponíveis atendem ao projeto em questão, como há dados suficientes para os treinamentos e uma vez que há literatura científica suficiente embasando a esta proposta, pode-se concluir que ela é viável sob o ponto de vista técnico-financeiro.



8 REFERÊNCIAS BIBLIOGRÁFICAS

ABADI, M.; AGARWAL, A.; BARHAM, P.; BREVDO, E.; CHEN, Z.; CITRO, C.; CORRADO, G. S.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; GOODFELLOW, I.; HARP, A.; IRVING, G.; ISARD, M.; JIA, Y.; JOZEFOWICZ, R.; KAISER, L.; KUDLUR, M.; LEVENBERG, J.; MANÉ, D.; MONGA, R.; MOORE, S.; MURRAY, D.; OLAH, C.; SCHUSTER, M.; SHLENS, J.; STEINER, B.; SUTSKEVER, I.; TALWAR, K.; TUCKER, P.; VANHOUCKE, V.; VASUDEVAN, V.; VIÉGAS, F.; VINYALS, O.; WARDEN, P.; WATTENBERG, M.; WICKE, M.; YU, Y.; ZHENG, X. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. [S.l.], 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>.

ALHASSAN, V.; HENRY, C.; RAMANNA, S.; STORIE, C. A deep learning framework for land-use/land-cover mapping and analysis using multispectral satellite imagery. *Neural Computing and Applications*, v. 32, n. 12, p. 8529–8544, 2020. ISSN 14333058. Disponível em: <<https://doi.org/10.1007/s00521-019-04349-9>>.

AUDEBERT, N.; Le Saux, B.; LEFEVRE, S. Deep learning for classification of hyperspectral data: A comparative review. *IEEE Geoscience and Remote Sensing Magazine*, v. 7, n. 2, p. 159–173, 2019. ISSN 21686831.

AVBELJ, J.; MULLER, R.; BAMLER, R. A metric for polygon comparison and building extraction evaluation. *IEEE Geoscience and Remote Sensing Letters*, Institute of Electrical and Electronics Engineers Inc., v. 12, n. 1, p. 170–174, 2015. ISSN 1545598X.

BALL, J. E.; ANDERSON, D. T.; CHAN, C. S. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. *Journal of Applied Remote Sensing*, v. 11, n. 04, p. 1, 2017. ISSN 1931-3195.

BITTNER, K.; ADAM, F.; CUI, S.; KÖRNER, M.; REINARTZ, P. Building Footprint Extraction From VHR Remote Sensing Images Combined With Normalized DSMs Using Fused Fully Convolutional Networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 11, n. 8, p. 2615–2629, 2018. ISSN 21511535.

BLASCHKE, T. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier B.V., v. 65, n. 1, p. 2–16, 2010. ISSN 09242716. Disponível em: <<http://dx.doi.org/10.1016/j.isprsjprs.2009.06.004>>.

BRASIL. *Decreto-Lei nº 243, de 28 de fevereiro de 1967*. 1967. 2438 p. Disponível em: <<https://www2.camara.leg.br/legin/fed/declei/1960-1969/decreto-lei-243-28-fevereiro-1967-376132-publicacaooriginal-1-pe.html>><https://www2.camara.leg.br/legin/fed/emecon/2016/emendaconstitucional-95-15-dezembro-2016-784029-publicacaooriginal-151558-pl.html><<https://www2.camara.leg.br/legin/fed/decret/2006/decreto-5800-8-junho-2006-543167-publicacaooriginal-53181-pe.html>>.



BRASIL. *Decreto nº 6.666, de 27 de novembro de 2008*. 2008. Disponível em: <http://www.planalto.gov.br/ccivil/_03/_Ato2007-2010/2008/Decreto/D6666.htm><http://www.inde.gov.br/images/inde/20@Decreto6666_27112008.pdf>.

CARRANZA-GARCÍA, M.; GARCÍA-GUTIÉRREZ, J.; RIQUELME, J. C. A framework for evaluating land use and land cover classification using convolutional neural networks. *Remote Sensing*, v. 11, n. 3, 2019. ISSN 20724292.

CHEN, G.; WENG, Q.; HAY, G. J.; HE, Y. Geographic object-based image analysis (GEOBIA): emerging trends and future opportunities. *GIScience and Remote Sensing*, Taylor & Francis, v. 55, n. 2, p. 159–182, 2018. ISSN 15481603. Disponível em: <<https://doi.org/10.1080/15481603.2018.1426092>>.

CHEN, J. X. The Evolution of Computing: AlphaGo. *Computing in Science and Engineering*, IEEE, v. 18, n. 4, p. 4–7, 2016. ISSN 15219615.

CHEN, Q.; WANG, L.; WU, Y.; WU, G.; GUO, Z.; WASLANDER, S. L. Aerial imagery for roof segmentation: A large-scale dataset towards automatic mapping of buildings. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier B.V., v. 147, p. 42–55, jan 2019. ISSN 09242716.

CHENG, G.; ZHOU, P.; HAN, J. Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 54, n. 12, p. 7405–7415, 2016. ISSN 01962892.

CHOLLET, F. et al. *Keras*. [S.l.]: GitHub, 2015. <<https://github.com/fchollet/keras>>.

CONCAR. *Especificação Técnica para a Estruturação de Dados Geoespaciais Vetoriais (ET-EDGV)*. Brasília, DF: [s.n.], 2010. 246 p. Disponível em: <http://www.geoportal.eb.mil.br/images/PDF/ET_EDGV_Vs_2_1_3.pdf>.

DAMODARAN, B. B.; HÖHLE, J.; LEFÈVRE, S. Attribute profiles on derived features for urban land cover classification. *Photogrammetric Engineering and Remote Sensing*, v. 83, n. 3, p. 183–193, 2017. ISSN 00991112.

DATA, H.; CHEN, Y.; LIN, Z.; CHEN, Y.; LIN, Z.; ZHAO, X.; MEMBER, S. Deep Learning-Based Classification of Hyperspectral Data. *IEEE*, v. 7, n. June 2014, p. 1–14, 2015.

DIAKOGIANNIS, F. I.; WALDNER, F.; CACCETTA, P.; WU, C. ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. 2019. Disponível em: <<http://arxiv.org/abs/1904.00592>>.

DIAKOGIANNIS, F. I.; WALDNER, F.; CACCETTA, P.; WU, C. ISPRS Journal of Photogrammetry and Remote Sensing ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. 2020. Disponível em: <<https://doi.org/10.1016/j.isprsjprs.2020.01.013>>.

DSG. *Banco de Dados Geográficos do Exército*. Brasília, DF: [s.n.], 2010. Disponível em: <www.bdgex.eb.mil.br>.



DSG. *Especificação Técnica para Controle de Qualidade de Dados Geoespaciais (ET-CQDG)*. Brasília: [s.n.], 2016. 1–94 p. Disponível em: <<https://www.geoportal.eb.mil.br/portal/inde2>>.

DSG. *Especificação Técnica para Controle de Qualidade de Dados Geoespaciais (ET-CQDG)*. Brasília: [s.n.], 2016. 1–94 p. Disponível em: <<https://www.geoportal.eb.mil.br/portal/inde2>>.

DSG. *Especificação Técnica para Aquisição de Dados Geoespaciais Vetoriais (ET-ADGV) Versão 3.0, 1ª Edição*. Brasília, DF: [s.n.], 2018.

GARCIA-GARCIA, A.; ORTS-ESCOLANO, S.; OPREA, S.; VILLENA-MARTINEZ, V.; MARTINEZ-GONZALEZ, P.; GARCIA-RODRIGUEZ, J. A survey on deep learning techniques for image and video semantic segmentation. *Applied Soft Computing Journal*, Elsevier B.V., v. 70, p. 41–65, 2018. ISSN 15684946. Disponível em: <<https://doi.org/10.1016/j.asoc.2018.05.018>>.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press, 2016.

GUIRADO, E.; TABIK, S.; ALCARAZ-SEGURA, D.; CABELLO, J.; HERRERA, F. Deep-learning Versus OBIA for scattered shrub detection with Google Earth Imagery: *Ziziphus lotus* as case study. *Remote Sensing*, v. 9, n. 12, p. 1–22, 2017. ISSN 20724292.

GUO, W.; LI, W.; GONG, W.; CUI, J. Extended Feature Pyramid Network with Adaptive Scale Training Strategy and Anchors for Object Detection in Aerial Images. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 12, n. 5, p. 784, mar 2020. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/12/5/784>>.

HOBEL, H.; ABDALLA, A.; FOGLIARONI, P.; FRANK, A. U. A semantic region growing algorithm: Extraction of urban settings. In: *Lecture Notes in Geoinformation and Cartography*. [S.l.: s.n.], 2015. ISBN 9783319167862. ISSN 18632351.

HOESER, T.; KUENZER, C. Object detection and image segmentation with deep learning on Earth observation data: A review-part I: Evolution and recent trends. *Remote Sensing*, v. 12, n. 10, 2020. ISSN 20724292.

HU, F.; XIA, G. S.; HU, J.; ZHANG, L. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sensing*, v. 7, n. 11, p. 14680–14707, 2015. ISSN 20724292.

JÖNSSON, D.; EILERTSEN, G.; ZHENG, J.; YNNERMAN, A.; UNGER, J. Visual Analysis of the Impact of Neural Network Hyper-Parameters. In: *Machine Learning Methods in Visualisation for Big Data*. [S.l.: s.n.], 2020. ISBN 978-3-03868-113-7.

JOZDANI, S. E.; JOHNSON, B. A.; CHEN, D. Comparing Deep Neural Networks , Ensemble Classifiers , and Support Vector Machine Algorithms. *Remote Sensing*, v. 11, n. 1, p. 1–24, 2019.



KEMKER, R.; SALVAGGIO, C.; KANAN, C. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 145, p. 60–77, 2018. ISSN 09242716.

KHANAL, N.; UDDIN, K.; MATIN, M. A.; TENNESON, K. Automatic detection of spatiotemporal urban expansion patterns by fusing OSM and Landsat data in Kathmandu. *Remote Sensing*, v. 11, n. 19, 2019. ISSN 20724292.

KLUYVER, T.; RAGAN-KELLEY, B.; PÉREZ, F.; GRANGER, B.; BUSSONNIER, M.; FREDERIC, J.; KELLEY, K.; HAMRICK, J.; GROUT, J.; CORLAY, S.; IVANOV, P.; AVILA, D.; ABDALLA, S.; WILLING, C. Jupyter notebooks – a publishing format for reproducible computational workflows. In: LOIZIDES, F.; SCHMIDT, B. (Ed.). *Positioning and Power in Academic Publishing: Players, Agents and Agendas*. [S.l.], 2016. p. 87 – 90.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *COMMUNICATIONS OF THE ACM*, v. 60, n. 6, 2017. Disponível em: <<http://code.google.com/p/cuda-convnet/>>

KUCHARCZYK, M.; HAY, G. J.; GHAFARIAN, S.; HUGENHOLTZ, C. H. Geographic Object-Based Image Analysis: A Primer and Future Directions. *Remote Sensing*, v. 12, n. 12, p. 2012, 2020. ISSN 2072-4292.

KULKARNI, T.; VENUGOPAL, N. Automatic semantic segmentation for change detection in remote sensing images. *Advances in Intelligent Systems and Computing*, Springer US, v. 705, n. 0123456789, p. 337–344, 2020. ISSN 21945357. Disponível em: <<https://doi.org/10.1007/s11063-019-10174-x>>.

LANARAS, C.; BIOUCAS-DIAS, J.; GALLIANI, S.; BALTSAVIAS, E.; SCHINDLER, K.; SENSING, R.; ZURICH, E. Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. 2018. Disponível em: <<https://doi.org/10.1016/j.isprsjprs.2018.09.018>>.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, 2015. ISSN 14764687.

LI, J.; LI, X.; XIE, T. Morphing of building footprints using a turning angle function. *ISPRS International Journal of Geo-Information*, MDPI AG, v. 6, n. 6, jun 2017. ISSN 22209964.

LI, K.; HU, X.; JIANG, H.; SHU, Z.; ZHANG, M. Attention-Guided Multi-Scale Segmentation Neural Network for Interactive Extraction of Region Objects from High-Resolution Satellite Imagery. 2020.

LI, M.; ZANG, S.; ZHANG, B.; LI, S.; WU, C. A review of remote sensing image classification techniques: The role of Spatio-contextual information. *European Journal of Remote Sensing*, Associazione Italiana di Telerilevamento, v. 47, n. 1, p. 389–411, jun 2014. ISSN 22797254.



LIU, Q.; KAMPFFMEYER, M.; JESSEN, R.; SALBERG, A.-B. Dense Dilated Convolutions Merging Network for Land Cover Classification. n. March, 2020. Disponível em: <<http://arxiv.org/abs/2003.04027>{\%}0A<http://dx.doi.org/10.1109/TGRS.2020.2976658>>.

LIU, Y.; PANG, C.; ZHAN, Z.; ZHANG, X.; YANG, X. Building Change Detection for Remote Sensing Images Using a Dual Task Constrained Deep Siamese Convolutional Network Model. 2019. Disponível em: <<http://arxiv.org/abs/1909.07726>>.

LOKHAT, I.; TOUYA, G. Enhancing building footprints with squaring operations. *Journal of Spatial Information Science*, v. 13, n. 2016, p. 33–60, 2016. ISSN 1948660X.

MA, L.; LI, M.; MA, X.; CHENG, L.; DU, P.; LIU, Y. *A review of supervised object-based land-cover image classification*. 2017. 277–293 p.

MA, L.; LIU, Y.; ZHANG, X.; YE, Y.; YIN, G.; JOHNSON, B. A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 152, n. March, p. 166–177, 2019. ISSN 09242716. Disponível em: <<https://doi.org/10.1016/j.isprsjprs.2019.04.015>>.

MBOGA, N.; GEORGANOS, S.; GRIPPA, T.; LENNERT, M.; VANHUYSSE, S.; WOLFF, E. Fully convolutional networks and geographic object-based image analysis for the classification of VHR imagery. *Remote Sensing*, v. 11, n. 5, 2019. ISSN 20724292.

MERKEL, D. Docker: lightweight linux containers for consistent development and deployment. *Linux journal*, v. 2014, n. 239, p. 2, 2014.

MUSYAROFAH; SCHMIDT, V.; KADA, M. Object detection of aerial image using mask-region convolutional neural network (mask R-CNN). In: *IOP Conference Series: Earth and Environmental Science*. [S.l.]: Institute of Physics Publishing, 2020. v. 500, n. 1. ISSN 17551315.

NAJAFABADI, M. M.; VILLANUSTRE, F.; KHOSHGOFTAAR, T. M.; SELIYA, N.; WALD, R.; MUHAREMAGIC, E. Deep learning applications and challenges in big data analytics. *Journal of Big Data*, v. 2, p. 1, 2015.

Ortega Adarme, M.; Queiroz Feitosa, R.; Nigri Happ, P. N.; Aparecido De Almeida, C.; Rodrigues Gomes, A. Evaluation of Deep Learning Techniques for Deforestation Detection in the Brazilian Amazon and Cerrado Biomes From Remote Sensing Imagery. *Remote Sensing*, MDPI AG, v. 12, n. 6, p. 910, mar 2020.

PAN, B.; XU, X.; SHI, Z.; ZHANG, N.; LUO, H.; LAN, X. DSSNet: A Simple Dilated Semantic Segmentation Network for Hyperspectral Imagery Classification. *IEEE Geoscience and Remote Sensing Letters*, IEEE, p. 1–5, 2020. ISSN 1545-598X.

PARTOVI, T.; BAHMANYAR, R.; KRAUS, T.; REINARTZ, P. Building Outline Extraction Using a Heuristic Approach Based on Generalization of Line Segments. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 10, n. 3, p. 933–947, 2017. ISSN 21511535.



QGIS Development Team. *QGIS Geographic Information System*. [S.I.], 2009. Disponível em: <<http://qgis.org>>.

QIN, M.; MAVROMATIS, S.; HU, L.; ZHANG, F.; LIU, R.; SEQUEIRA, J.; DU, Z. Remote Sensing Single-Image Resolution Improvement Using A Deep Gradient-Aware Network with Image-Specific Enhancement. *Remote Sensing*, MDPI AG, v. 12, n. 5, p. 758, feb 2020. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/12/5/758>>.

Refractions Research Inc. *PostGIS*. [S.I.], 2005. Disponível em: <<http://postgis.net>>.

REINA, G. A.; PANCHUMARTHY, R.; THAKUR, S. P.; BASTIDAS, A.; BAKAS, S. Systematic Evaluation of Image Tiling Adverse Effects on Deep Learning Semantic Segmentation. *Frontiers in Neuroscience*, v. 14, n. February, p. 1–14, 2020. ISSN 1662453X.

ROSSUM, G. V.; DRAKE, F. L. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009. ISBN 1441412697.

SALVETTI, F.; MAZZIA, V.; KHALIQ, A.; CHIABERGE, M. Multi-Image Super Resolution of Remotely Sensed Images Using Residual Attention Deep Neural Networks. *Remote Sensing*, v. 12, n. 14, p. 2207, jul 2020. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/12/14/2207>>.

SHAPIRO, L. *Computer vision*. Upper Saddle River, NJ: Prentice Hall, 2001. ISBN 0130307963.

SOEKHOE, D.; PUTTEN, P. van der; PLAAT, A. On the impact of data set size in transfer learning using deep neural networks. In: BOSTRÖM, H.; KNOBBE, A.; SOARES, C.; PAPAPETROU, P. (Ed.). *Advances in Intelligent Data Analysis XV*. Cham: Springer International Publishing, 2016. p. 50–60. ISBN 978-3-319-46349-0.

SONG, A.; CHOI, J. Fully Convolutional Networks with Multiscale 3D Filters and Transfer Learning for Change Detection in High Spatial Resolution Satellite Images. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 12, n. 5, p. 799, mar 2020. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/12/5/799>>.

SU, W.; LI, J.; CHEN, Y.; LIU, Z.; ZHANG, J.; LOW, T. M.; SUPPIAH, I.; HASHIM, S. A. M. Textural and local spatial statistics for the object-oriented classification of urban areas using high resolution imagery. *International Journal of Remote Sensing*, Taylor Francis, v. 29, n. 11, p. 3105–3117, 2008. Disponível em: <<https://doi.org/10.1080/01431160701469016>>.

Thanh Noi, P.; KAPPAS, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors (Basel, Switzerland)*, v. 18, n. 1, 2017. ISSN 14248220.

The PostgreSQL Global Development Group. *PostgreSQL*. [S.I.], 1996. Disponível em: <www.postgresql.org>.



WANG, Q.; ZHANG, X.; CHEN, G.; DAI, F.; GONG, Y.; ZHU, K. Change detection based on Faster R-CNN for high-resolution remote sensing images. *Remote Sensing Letters*, Taylor and Francis Ltd., v. 9, n. 10, p. 923–932, oct 2018. ISSN 21507058.

WISEMAN, G.; KORT, J.; WALKER, D. Quantification of shelterbelt characteristics using high-resolution imagery. *Agriculture, ecosystems & environment*, Elsevier, v. 131, n. 1-2, p. 111–117, 2009.

YANG, H.; WU, P.; YAO, X.; WU, Y.; WANG, B.; XU, Y. Building extraction in very high resolution imagery by dense-attention networks. *Remote Sensing*, v. 10, n. 11, p. 1–16, 2018. ISSN 20724292.

YU, H.; YANG, Z.; TAN, L.; WANG, Y.; SUN, W.; SUN, M.; TANG, Y. Methods and datasets on semantic segmentation: A review. *Neurocomputing*, v. 304, p. 82–103, 2018. Disponível em: <<https://doi.org/10.1016/j.neucom.2018.03.037>>.

ZHANG, C.; WEI, S.; JI, S.; LU, M. Detecting large-scale urban land cover changes from very high resolution remote sensing images using CNN-based classification. *ISPRS International Journal of Geo-Information*, v. 8, n. 4, 2019. ISSN 22209964.

ZHANG, L.; ZHANG, L.; DU, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, v. 4, n. 2, p. 22–40, 2016. ISSN 21686831.

ZHU, X. X.; TUIA, D.; MOU, L.; XIA, G. S.; ZHANG, L.; XU, F.; FRAUNDORFER, F. *Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources*. [S.l.]: Institute of Electrical and Electronics Engineers Inc., dec 2017. 8–36 p.