

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

**SEGMENTAÇÃO SEMÂNTICA TEMPORAL POR *DEEP LEARNING* PARA DETECÇÃO
DE MUDANÇA DE STATUS EM GARIMPOS NA REGIÃO AMAZÔNICA, A PARTIR DE
IMAGENS DE SENSORES REMOTOS**

PAULO ROBERTO BATISTA DE OLIVEIRA JÚNIOR

DISSERTAÇÃO DE MESTRADO Nº 206

Orientador:

Prof. Dr. Henrique Llacer Roig

Brasília – DF

2022

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

SEGMENTAÇÃO SEMÂNTICA TEMPORAL POR *DEEP LEARNING* PARA DETECÇÃO DE
MUDANÇA DE STATUS EM GARIMPOS NA REGIÃO AMAZÔNICA, A PARTIR DE
IMAGENS DE SENSORES REMOTOS

Área de Concentração: Geoprocessamento e Análise Ambiental

Linha de Pesquisa: Detecção de Mudanças e Evolução do Meio Ambiente

Paulo Roberto Batista de Oliveira Júnior

Dissertação de Mestrado submetida ao Instituto
de Geociências da Universidade de Brasília,
como parte dos requisitos para obtenção do grau
de Mestre em Geociências Aplicadas.

Orientador: Prof. Dr. Henrique Llacer Roig

Brasília – DF

2022

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

B048s Batista de Oliveira Junior, Paulo Roberto
Segmentação Semântica Temporal por Deep Learning para
Detecção de Mudança de status em Garimpos na Região Amazônica,
a partir de Imagens de Sensores Remotos / Paulo Roberto
Batista de Oliveira Junior; orientador Henrique Llacer
Roig. -- Brasília, 2023.
103 p.

Dissertação(Mestrado em Geociências Aplicadas) --
Universidade de Brasília, 2023.

1. Deep Learning. 2. Segmentação Semântica. 3. Detecção de
Mudança. 4. Sensoriamento Remoto. 5. Garimpo. I. Llacer
Roig, Henrique, orient. II. Título.

BANCA EXAMINADORA

Dr. Henrique Llacer Roig
Universidade de Brasília

Dra. Tati de Almeida
Universidade de Brasília

Dr. Weeberb João Réquia Júnior
Fundação Getúlio Vargas

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Código de Financiamento 001, do Exército Brasileiro (EB) e do Centro Gestor e Operacional do Sistema de Proteção da Amazônia (CENSIPAM).

DEDICATÓRIA

Dedico este trabalho:

*À minha esposa **Ana Carolina** e aos meus filhos **Arthur e Theo**, por todo amor, carinho, cumplicidade e paciência. Por acreditarem em mim e muitas vezes suportarem minha ausência em apoio a esta empreitada.*

*Aos meus pais **Paulo e Brasilene**, aos meus avós **Erlis e Aldair**, à minha irmã **Nayane** e aos meus sogros **Valdetaro e Ana Lúcia**, por todo apoio moral e pelos conselhos sempre assertivos.*

*Aos meus amigos **Azeredo, Luiz Cláudio, Marco Aurélio** e todos os outros que me apoiaram psicologicamente e/ou intelectualmente no decorrer do curso.*

*Aos Doutores e Mestres, em especial ao Professor Doutor **Henrique Llacer Roig**, pelos conselhos e pela ajuda quando decidi trilhar o caminho do conhecimento.*

“Nas grandes batalhas da vida, o primeiro passo para a vitória é o desejo de vencer” (Mahatma Gandhi)

RESUMO

A atividade garimpeira na Região Amazônica foi responsável pelo desmatamento de centenas de quilômetros quadrados de floresta ao longo dos anos. Existem, atualmente, mais de 2750 garimpos legalizados no Brasil e mais de 4400 pontos de garimpos ilegais na pan-amazônia. Com intuito de monitorar o desmatamento na região, órgãos ambientais e de segurança trabalham para dar suporte às autoridades nos processos de fiscalização e controle de atividades ambientais, inclusive no que se refere ao desmatamento provocado pela ação garimpeira. Para o garimpo, em específico, o monitoramento constante das áreas é fundamental, tendo em vista a necessidade de se verificar a atividade ou a inatividade do mesmo. Neste cenário, uma solução promissora para a otimização dos processos de classificação das imagens e reavaliação do estágio do garimpo seria a utilização das Redes Neurais do Aprendizado Profundo, uma vez que estas, em especial as Convolucionais, são eficazes para o reconhecimento de imagens, a segmentação semântica e a detecção de mudança. Neste contexto, algoritmos de Aprendizado Profundo foram avaliados, selecionados e aplicados para monitoramento e análise da evolução dos status dos garimpos nas regiões de Tapajós e Kayapó, na Floresta Amazônica Brasileira, a partir de imagens oriundas do satélite CBERS 04A. Foram testadas as arquiteturas de redes neurais Unet, Unet++, Manet, Linknet, FPN, PSPNet, PAN, DeepLabV3 e DeepLabV3+. A arquitetura de melhor resultado foi a Unet++, em conjunto com o *backbone EfficientNet*, que obteve Acurácia de 98,87%, *F1 score* de 89,20% e *Intersection over Union* de 84,16%, o que atesta a eficiência da rede para segmentação de garimpos na região Amazônica. Para a análise temporal da evolução dos status do garimpo, utilizou-se a detecção de mudança a partir da Estrutura de Pós Classificação do método Fluxo Duplo. Para quantificar as mudanças, foi aplicado o Índice de Similaridade Estrutural. Como resultado, concluiu-se que o modelo apresentado é uma ferramenta viável para o monitoramento de áreas garimpeiras, uma vez que o mesmo foi capaz de identificar mudanças ocorridas entre os anos de 2020 e de 2021, bem como é um instrumento capaz de amparar às tomadas de decisão de Órgãos de Proteção e Fiscalização, de Órgãos Ambientais e de Órgãos de Segurança e Controle.

Palavras Chave: Garimpo, Amazônia, Monitoramento, Redes Neurais, Segmentação Semântica, Detecção de Mudança.

ABSTRACT

The mining activity in the Amazon Region has been responsible for the deforestation of hundreds of kilometers of forest over the years. Currently, there are more than 2750 legalized gold mines in Brazil and more than 4400 illegal gold mines in the Pan-amazon. In order to monitor the deforestation in the region, environmental and security agencies work to support the authorities in the processes of inspection and control of environmental activities, including the deforestation caused by mining activities. For mining, in particular, constant monitoring of the areas is essential to verify their activity or inactivity. In this scenario, a promising solution for optimizing image classification processes and reassessing the mining stage would be the use of Deep Learning Neural Networks, especially Convolutional Networks, which are effective for image recognition, semantic segmentation and change detection. In this context, Deep Learning algorithms were evaluated, selected and applied for monitoring and analyzing the evolution of the mining status in the regions of Tapajós and Kayapó, in the Brazilian Amazon Forest, based on images from CBERS 04A satellite. The Neural network architectures Unet, Unet++, Manet, Linknet, FPN, PSPNet, PAN, DeepLabV3 and DeepLabV3+ were tested. The architecture with the best result was Unet++, together with the EfficientNet backbone, which obtained an Accuracy of 98.87%, a F1 score of 89.20% and an Intersection over Union of 84.16%. For the temporal analysis of the evolution of the mining status, the Post Classification Structure of the Double Flow method was used. To quantify the changes, the Structural Similarity Index was applied. As a result, it was concluded that the presented model is a viable tool for monitoring mining areas, since it was able to identify changes that occurred between 2020 and 2021. The model can also be an instrument capable of supporting decision-making by Protection, Environmental and Security institutions.

Key Words: Mining, Amazon, Monitoring, Neural Networks, Semantic Segmentation, Change Detection.

Lista de Abreviaturas

NOME	ABREVIATURA
Aprendizado de Máquina	ML
Aprendizado Profundo	DL
Banco de Dados de Aplicação	BDA
Banco de Dados de Estudo	BDE
Banco de Dados de Teste	BDT
Câmera imageadora de Campo Largo	WFI
Câmera Multiespectral	MUX
Câmera Multiespectral e Pancromática de Ampla Varredura	WPM
Centro Gestor e Operacional do Sistema de Proteção da Amazônia	CENSIPAM
Enfoque Meta Analítico	EMA
Erro Quadrático Médio	EQM
Fully Convolution Networks	FCNs
Fundação Nacional do Índio	FUNAI
Gravador de Dados Digital	DDR
Ground Truth	GT
Índice de Similaridade Estrutural	ISE
Instituto Chico Mendes de Conservação da Biodiversidade	ICMBio
Instituto Nacional de Pesquisas Espaciais	INPE
Inteligência Artificial	IA
Interface de Programação de Aplicação	API
Intersection over Union	IoU
Monitor do Ambiente Espacial	SEM
Randon Forest	RF
Rede Amazônica de Informação Socioambiental Georreferenciada	RAISG
Redes Neurais Convolucionais	CNN
Sistema de Coleta de Dados	DCS
Sistema de Informação Geográfica	SIG
Support Vector Machine	SVM
Teoria do Enfoque Meta Analítico Consolidado	TEMAC
Transmissor de Dados de Imagem	DTS
Web of Science	WoS

Lista de Figuras

Figura 1: Modelo TEMAC.....	21
Figura 2: Segmentação Semântica – Nuvem de palavras construídas a partir das palavras-chaves dos artigos publicados.....	25
Figura 3: Segmentação Semântica – Mapa de calor dos pesquisadores de maior destaque.....	27
Figura 4: Segmentação Semântica – Mapa de calor de co-citation.....	28
Figura 5: Detecção de Mudança – Nuvem de palavras construídas a partir das palavras-chaves dos artigos publicados.....	32
Figura 6: Detecção de Mudança – Mapa de calor dos pesquisadores de maior destaque.....	34
Figura 7: Detecção de Mudança – Mapa de calor de co-citation.....	35
Figura 8: Garimpagem de ouro – Nuvem de palavras construídas a partir das palavras-chaves dos artigos publicados.....	40
Figura 9: Garimpagem de ouro – Mapa de calor de co-citation.....	42
Figura 10: Exemplo de Rede Neural Profunda.....	45
Figura 11: Campo de IA e seus subconjuntos.....	46
Figura 12: Exemplo de propagação de redes neurais profundas.....	47
Figura 13: Funções de ativação mais comuns.....	48
Figura 14: Camadas de Redes Neurais Convolucionais.....	50
Figura 15: Exemplo de operação matemática de Convolução.....	51
Figura 16: Exemplo de operações estatísticas de Max Pooling e Average Pooling.....	51
Figura 17: Etapas do processo de aprendizagem das Redes Neurais Convolucionais.....	52
Figura 18: Exemplo de segmentação semântica.....	53
Figura 19: Exemplo de segmentação semântica.....	54
Figura 20: a) estrutura siamesa; b) estrutura baseada em transferência de aprendizado; c) estrutura de pós classificação.....	55
Figura 21: Esquematização das métricas precisão e recall.....	58
Figura 22: Métrica Intersection over Union (IoU).....	58
Figura 23: Influência da escolha da taxa de aprendizagem na função de perda.....	60
Figura 24: Exemplo de <i>Data Augmentation</i>	61
Figura 25: Área de estudos de Tapajós, na Amazônia Brasileira.....	64
Figura 26: Área de estudos de Kayapó, na Amazônia Brasileira.....	65
Figura 27: Fluxograma do projeto.....	66

Figura 28: Sumários das características das câmeras do CBERS 04A.....	68
Figura 29: a) banda pancromática; b) banda vermelha; c) composição RGB; d) fusão PAN e composição colorida.....	69
Figura 30: Processo de <i>Data Augmentation</i>	70
Figura 31: Esquematização do processo de criação dos bancos de dados de estudo e aplicação.....	71
Figura 32: Esquematização para definição da melhor arquitetura.....	74
Figura 33: Esquematização para definição do melhor backbone.....	75
Figura 34: Esquematização para definição do melhor modelo treinado e resultado de suas métricas de avaliação.....	76
Figura 35: Esquematização do teste na área de Kayapó e da detecção de mudança.....	76
Figura 36: Resultados da escolha da arquitetura mais adequada.....	78
Figura 37: Resultados da escolha da arquitetura mais adequada.....	79
Figura 38: Gráficos da evolução da perda e das métricas de avaliação durante a fase de treinamento e validação.....	80
Figura 39: Resultado final do modelo treinado.....	81
Figura 40: a) imagem original, b)máscara, c) máscara predita e d) diferença entre as máscaras das áreas garimpeiras.....	82
Figura 41: a) imagem original, b)máscara, c) máscara predita e d) diferença entre as máscaras das áreas garimpeiras em Kayapó.....	83
Figura 42: a) imagem original, b)máscara, c) máscara predita e d) diferença entre as máscaras das áreas garimpeiras em Kayapó.....	84
Figura 43: Exemplos de garimpos em Kayapó.....	85
Figura 44: Imagens SEM mudanças, a) imagem 2020, b) imagem 2021, c) diferença.....	86
Figura 45: Imagens COM mudanças: a) imagem 2020, b) imagem 2021 e c) diferença.....	87

Lista de Gráficos

Gráfico 1: Revistas mais relevantes na área de Sensoriamento Remoto, conforme valor do Fator de Impacto.....	23
Gráfico 2: Segmentação Semântica – Evolução do número de publicações e citações ano a ano.....	24
Gráfico 3: Segmentação Semântica – Autores que mais publicaram.....	27
Gráfico 4: Arquiteturas de CNNs mais utilizadas para segmentação semântica.....	30
Gráfico 5: Revistas com maior número de publicações sobre Segmentação Semântica.....	30
Gráfico 6: Detecção de Mudança – Evolução do número de publicações e citações ano a ano.....	31
Gráfico 7: Detecção de Mudança - Autores que mais publicaram.....	34
Gráfico 8: Detecção de Mudança – Revistas com maior número de publicações.....	37
Gráfico 9: Garimpagem de ouro – Evolução do número de publicações e citações ano a ano.....	40
Gráfico 10: Garimpagem de ouro – Visão geral dos países que mais publicaram.....	43
Gráfico 11: Garimpagem de ouro – Revistas com maior número de publicações.....	43
Gráfico 12: Quantidade de imagens com e sem mudança após decorrido um ano.....	88

Lista de Tabelas

Tabela 1: Segmentação Semântica – Artigos mais citados.....	26
Tabela 2: Detecção de Mudança – Artigos mais citados.....	33
Tabela 3: Garimpagem de ouro – Artigos mais citados.....	41
Tabela 4: Matriz de Confusão para dados binários.....	57
Tabela 5: Composição dos bancos de dados.....	67
Tabela 6: Divisão dos dados em treinamento, validação e teste.....	71
Tabela 7: Hiperparâmetros utilizados para definição da arquitetura mais adequada.....	73
Tabela 8: Resultado final das etapas de treinamento e validação.....	80
Tabela 9: Resultado final do teste em Kayapó.....	83
Tabela 10: Erro percentual do teste em Kayapó.....	84
Tabela 11: Extrato da tabela de resultados do ISE.....	88

Sumário

CAPÍTULO 01 – INTRODUÇÃO.....	17
1.1 Contextualização.....	17
1.2 Objetivos.....	19
1.2.1 Geral.....	19
1.2.2 Específicos.....	19
1.3 Justificativas.....	20
CAPÍTULO 2 – ESTADO DA ARTE.....	21
2.1 Teoria do Enfoque Meta Analítico Consolidado – TEMAC.....	21
2.2 Preparação da Pesquisa.....	22
2.3 Apresentação e inter-relação dos dados, detalhamento, modelo integrador e validação por evidências.....	22
2.3.1 Segmentação Semântica no Sensoriamento Remoto.....	24
2.3.2 Detecção de Mudança no Sensoriamento Remoto.....	31
2.3.3 CNN x Técnicas Clássicas de Aprendizado de Máquina.....	37
2.3.4 Garimpagem de ouro e Sensoriamento Remoto.....	39
2.3.5 Considerações Finais acerca do Estado da Arte.....	44
CAPÍTULO 3 – FUNDAMENTAÇÃO TEÓRICA.....	45
3.1 Aprendizado Profundo.....	45
3.1.1 Componentes de uma rede neural profunda.....	47
3.1.2 <i>Backpropagation</i>	49
3.2 Redes Neurais Convolucionais.....	50
3.3 Segmentação Semântica e Detecção de Mudança.....	53

3.3.1 Segmentação Semântica.....	53
3.3.2 Detecção de Mudança.....	54
3.4 Métricas de Avaliação.....	56
3.5 Hiperparâmetros.....	59
3.6 <i>Data Augmentation</i>	61
3.7 Métodos de Comparação de Imagens.....	62
4 METODOLOGIA.....	64
4.1 Áreas de Estudo.....	64
4.2 Etapas do Projeto.....	65
4.3 Banco de dados.....	66
4.3.1 Satélite CBERS 04A.....	67
4.3.2 Preparação do banco de dados.....	70
4.4 Processamento dos dados.....	71
4.4.1 Treinamento, validação e teste.....	72
4.4.2 Definição da melhor arquitetura.....	73
4.4.3 Definição do <i>backbone</i> mais adequado.....	74
4.4.4 Estabelecimento do melhor modelo.....	75
4.4.5 Teste e detecção de mudança.....	76
5 RESULTADOS E DISCUSSÕES.....	78
5.1 Arquitetura com melhor desempenho.....	78
5.2 <i>Backbone</i> de melhor performance.....	79
5.3 Modelo Final.....	80
5.4 Teste em Kayapó.....	83

5.5 Detecção de Mudança.....	86
6 CONCLUSÃO.....	89
7 SUGESTÕES PARA TRABALHOS FUTUROS.....	91
8 REFERÊNCIAS BIBLIOGRÁFICAS.....	92

CAPÍTULO 01 – INTRODUÇÃO

Neste capítulo será apresentado uma contextualização acerca do tema, o objetivo geral e os objetivos específicos da pesquisa, bem como a justificativa para o objeto proposto.

1.1 Contextualização

Conforme dados compilados do Sistema de Detecção de Desmatamento em Tempo Real – Deter (INPE, 2021a), entre agosto de 2015 e julho de 2022, a atividade garimpeira na Região Amazônica foi responsável pelo desmatamento de aproximadamente 600 km² de floresta.

Esta atividade, cuja característica envolve a extração de minérios com potencial econômico, pode causar impactos socioambientais e econômicos relevantes (FERNANDES; ALAMINO; ARAUJO, 2014). Desde a década de 90, RODRIGUES et al. (1994) já apontava diversos problemas apresentados pela expansão da atividade garimpeira, tais como: destruição de nichos ecológicos, contaminação dos recursos hídricos e biota, comprometimento de atividades econômicas, geração de focos de doenças endêmicas, exposição da população à contaminação por mercúrio, desorganização social, desrespeito aos direitos do cidadão e ao direito ambiental, perda de recurso natural não renovável, além de outros problemas ocasionados por esta prática.

Conforme dados atualizados até janeiro de 2022, da Secretaria de Geologia, Mineração e Transformação Mineral do Ministério de Minas e Energia (MME, 2022), existem mais de 2750 garimpos legalizados no Brasil. Em termos de garimpos ilegais, segundo a Rede Amazônica de Informação Socioambiental Georreferenciada (RAISG, 2020), estima-se a existência de mais de 4400 pontos de garimpos na região da pan-amazônia.

Tendo em vista o apelo nacional e internacional para conservação da Amazônia, diversos órgãos, públicos e privados, trabalham no monitoramento do desmatamento da região, inclusive no que se refere ao desmatamento provocado pela atividade garimpeira. Dentre eles, pode-se destacar o Instituto Nacional de Pesquisas Espaciais (INPE), a Rede Amazônica de Informação Socioambiental Georreferenciada (RAISG) e o Centro Gestor e Operacional do Sistema de Proteção da Amazônia (CENSIPAM).

Estas instituições, em termos gerais, buscam um monitoramento efetivo, rápido e constante da região com intuito de dar suporte às autoridades nos processos de fiscalização e controle de atividades ambientais. Este monitoramento, na maior parte dos casos, consiste em uma análise

continua de imagens de sensoriamento remoto. Isto ocorre, por exemplo, com os sistemas DETER (INPE, 2021b) e PRODES (INPE, 2021c), que foram desenvolvidos para supervisionar a Amazônia Legal Brasileira. Segundo (SOUZA et al., 2019), ambos sistemas realizam o mapeamento dos incrementos de desmatamento através de fotointerpretação por especialistas.

Em que pesem os esforços, verifica-se que existem dificuldades impostas para os processos de monitoramento do desmatamento no bioma amazônico. Além dos mesmos demandarem a utilização de pessoas capacitadas e experientes em analisar imagens de sistemas remotos, há os obstáculos naturais inerentes à região, como o regime de nebulosidade, os vapores de água sazonais, a concentração de ozônio e a alta concentração de aerossóis na região (MARTINS et al., 2018).

No contexto da exploração aurífera, o monitoramento constante das áreas é fundamental tendo em vista a necessidade de se verificar a atividade ou inatividade do garimpo. Em virtude da variação do preço de mercado do bem mineral, muitas áreas podem ter uma dinâmica intermitente. Por este motivo, com intuito de verificar a dinâmica da vegetação, é essencial que os polígonos identificados como garimpo sejam frequentemente reanalisados para o cálculo efetivo e o mais acurado possível da atividade garimpeira. Caso seja verificado que determinada área esteja em processo de recuperação da vegetação, faz-se necessário enquadrar o garimpo como inativo e recalcular a taxa de atividade.

Considerando o exposto, uma solução plausível e promissora para a otimização dos processos de classificação das imagens e reavaliação do estágio do garimpo seria o Aprendizado de Máquina (*Machine Learning*), mais precisamente o Aprendizado Profundo (*Deep Learning*). O aprendizado de máquina é uma área da Inteligência Artificial que tem por objetivo desenvolver técnicas computacionais sobre o aprendizado e construir sistemas capazes de adquirir conhecimento de forma automática (MONARD; BARANAUSKAS, 2003). O Aprendizado Profundo é uma técnica de Aprendizado de Máquina caracterizada por Redes Neurais cuja arquitetura envolve múltiplas camadas com capacidade de explorar e representar informações a partir do aprendizado de um conjunto de dados (ZHU et al., 2017).

HOESER, BACHOFER e KUENZER (2020), em uma ampla revisão da literatura científica, verificaram que a técnica de *Deep Learning* tem sido aplicada em diversas áreas, tais como: setores de transporte, assentamentos, uso e cobertura da terra, detecção de navios, extração de feições, entre outras. No contexto do sensoriamento remoto, a automatização na classificação do Uso e Cobertura Terra (*Automating Land Use and Land Cover – LULC*) por métodos de *Deep Learning* é efetiva,

tanto no que se refere ao mapeamento dos alvos, quanto à detecção de mudança (MA et al., 2019). O reconhecimento de imagens, a detecção de objetos e a segmentação semântica são recursos aprendidos de maneira eficaz por Redes Neurais (ZHU et al., 2017).

Dado o exposto, considerando o Aprendizado Profundo como um caminho promissor para avaliar a evolução do status de garimpos, esta dissertação utilizou a Segmentação Semântica Temporal com intuito de criar algoritmos para monitoramento de garimpos na Amazônia Brasileira, a partir de imagens oriundas de sensores remotos.

1.2 Objetivos

1.2.1 Geral

Esta pesquisa tem por objetivo avaliar, selecionar e aplicar algoritmos de *Deep Learning* para monitoramento e análise temporal da evolução dos status dos garimpos na Floresta Amazônica Brasileira, a partir de imagens oriundas do satélite CBERS 04A, utilizando como estudo de caso áreas garimpeiras nas regiões de Tapajós e Kayapó, no estado do Pará, entre os anos de 2020 e 2021.

1.2.2 Específicos

Como objetivos específicos da pesquisa, pretende-se:

- (a) Testar arquiteturas de redes neurais para a extração automática de garimpos;
- (b) Selecionar e aplicar a melhor arquitetura e o melhor *backbone* para identificação de desmatamentos provocados por garimpos;
- (c) Avaliar a performance da arquitetura e do *backbone* selecionados;
- (d) Identificar métodos de detecção de mudança;
- (e) Aplicar a detecção de mudança em áreas garimpeiras; e
- (f) Avaliar os resultados obtidos pela detecção de mudança.

1.3 Justificativas

A Amazônia é maior floresta tropical do mundo (ocupando, só no Brasil, um território com cerca de 4.200.000 km²), possui a maior bacia hidrográfica do planeta (aproximadamente 6.000.000 km² de área coberta), bem como possui a maior biodiversidade dentre as florestas tropicais (MMA, 2022). Além disso, ela conserva uma fauna e flora com milhões de espécies e é fundamental para o controle e manutenção climática no país (MOSS; MOSS, 2014). Isto posto, devido sua importância nacional e mundial, há um debate constante para sua conservação.

Dentre as diversas alterações a respeito da preservação da Amazônia, destacam-se os impactos socioambientais e econômicos provocados pelas atividades garimpeiras na floresta. Esta atividade, por exemplo, provocou, segundo dados da Comissão Pastoral da Terra (CPT), a morte de 109 indígenas no ano de 2021 (PASSOS et al., 2021). Outro exemplo do impacto dessa atividade pode ser verificada no relatório da Hutukara Associação Yanomami, que aponta um crescimento de 3350% na área garimpeira na Terra Indígena Yanomami, entre os anos de 2016 e 2020 (HUTUKARA; YE'KWANA, 2022).

Com intuito de monitorar a atividade garimpeira, órgãos ambientais e de segurança buscam analisar e fotointerpretar continuamente imagens da região oriundas de sensores remotos. Neste contexto, esta pesquisa se justifica pela importância da automatização da verificação dos status dos garimpos para aplicações estratégicas, tanto no meio civil, quanto no âmbito do Ministério da Defesa. Esta automatização possibilita o levantamento da evolução de áreas garimpeiras de modo ágil e com custos reduzidos, proporcionando ferramentas fundamentais para monitoramento das áreas, bem como para amparar as tomadas de decisão de Órgãos de Proteção e Fiscalização, como o Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio), a Fundação Nacional do Índio (FUNAI) e o CENSIPAM, e de Órgãos de Segurança e Controle, como as Forças Armadas e a Polícia Federal.

CAPÍTULO 2 – ESTADO DA ARTE

Tendo em vista o objetivo do projeto apresentado na seção 1.2 e com o propósito de nortear a fundamentação teórica a ser apresentada no Capítulo 3, este capítulo tem por finalidade apresentar o Artigo remetido a Revista Brasileira de Cartografia acerca do Estado da Arte dos assuntos Segmentação Semântica (*Semantic Segmentation*) e Detecção de Mudança (*Change Detection*), inseridos no contexto do Aprendizado Profundo (*Deep Learning*), ambos associados a área de Sensoriamento Remoto. Será apresentado, ainda, o Estado da Arte relativo à detecção de áreas de garimpo a partir de sensores remotos.

2.1 Teoria do Enfoque Meta Analítico Consolidado – TEMAC

No intuito de analisar e apresentar uma revisão da literatura de modo detalhado, foi empregada, no período de julho 2021 a julho de 2022, a Teoria do Enfoque Meta Analítico Consolidado – TEMAC (MARIANO; ROCHA, 2017). Conforme os autores, esta teoria é a unificação e consolidação dos métodos do Enfoque Meta Analítico – EMA, o qual utiliza critérios relacionados ao impacto de revistas, citações de autores, citações de artigos e frequência de palavras-chaves para seleção dos materiais a serem utilizados na pesquisa metodológica. Tendo em vista que o método do EMA foi abordado de diferentes maneiras ao longo dos anos 2000, a TEMAC surgiu com o objetivo de unificar e consolidar suas abordagens. Esta teoria é dividida em três etapas: (1) preparação da pesquisa, (2) apresentação e inter-relação dos dados e (3) detalhamento, modelo integrador e validação por evidências. A figura 1 apresenta estas etapas.

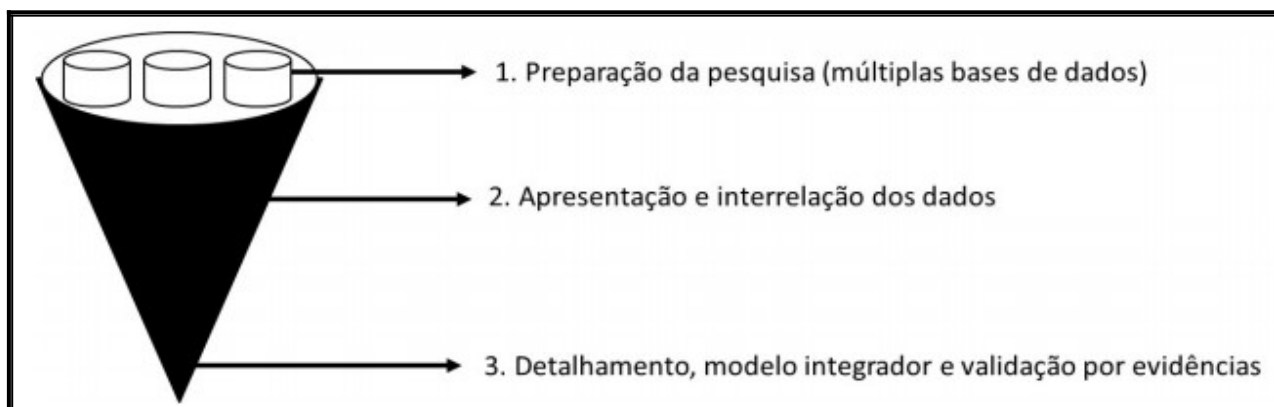


Figura 1: Modelo TEMAC

Fonte: Mariano e Rocha (2017)

Na **preparação da pesquisa** foram analisadas e definidas as questões sobre os termos de pesquisa, o espaço temporal e a base de dados. Na etapa **apresentação e inter-relação dos dados** foram feitas análises em cima dos seguintes filtros bibliométricos: revistas mais relevantes; evolução dos temas ano a ano; artigos mais citados; autores que mais publicaram; países que mais pesquisaram; e revistas que mais publicaram. Na etapa **detalhamento, modelo integrador e validação por evidências** foram criados *clusters* de aproximação de dados e foram efetuadas análises de *co-citation* e *coupling*. A *co-citation* é um índice bibliométrico utilizado para verificar os autores que frequentemente são citados juntos e o *coupling* usa o princípio de que trabalhos que referenciam um mesmo artigo possuem similaridades em seus conteúdos (MARIANO; ROCHA, 2017). Neste artigo, especificamente, as etapas apresentação e inter-relação dos dados e detalhamento, modelo integrador e validação por evidências foram realizadas concomitantemente.

2.2 Preparação da Pesquisa

Para esta etapa, definiu-se pela utilização da base de dados do *Web of Science* (WoS). O WoS é uma das bases de dados de publicações e citações de pesquisa mais antigo, mais amplamente utilizado e mais confiável do mundo (BIRKLE et al., 2020). A partir da mesma, é possível utilizar filtros automáticos de pesquisa que podem ser divididos em categorias, tais como: área de conhecimento, instituições, países, revistas, autores de relevância, entre outros. O espaço-tempo das análises foi delimitado entre os anos de 1945 e 2021, que compreende todo o período disponibilizado na base de dados. Tendo em vista a essência deste artigo englobar, principalmente, as áreas de Aprendizado de Máquina e Sensoriamento Remoto, as pesquisas foram definidas nos tópicos referentes as técnicas de Aprendizado Profundo, Segmentação Semântica e Detecção de Mudança. Para esses tópicos, utilizou-se das palavras-chave “*Semantic Segmentation*” e “*Change Detection*” associadas a expressão “*Deep Learning*” e refinadas na categoria “*Remote Sensing*”. Nestes casos, a quantidade de registros encontrados foram de 500 e 219, respectivamente. Para o tópico “*Gold Mining*”, foi aplicada na pesquisa esta palavra-chave e a mesma foi refinada na categoria “*Remote Sensing*”. Nesta análise, foram encontrados 25 artigos registrados.

2.3 Apresentação e inter-relação dos dados, detalhamento, modelo integrador e validação por evidências

Nesta fase foram feitas as análises dos filtros bibliométricos para cada pesquisa realizada no sítio do *Web of Science* (<https://apps-webofknowledge.ez54.periodicos.capes.gov.br>), acessado

através do Portal de Periódico da CAPES (<https://www-periodicos-capes-gov-br.ez54.periodicos.capes.gov.br>), conforme definido na preparação da pesquisa.

Em um primeiro momento, foi feita uma análise das revistas mais relevantes sobre o tema, uma vez que para todas as pesquisas foi utilizada como refinamento a categoria *Remote Sensing*. Para este estudo recorreu-se ao *InCites Journal Citation Reports* – JCR para definir a relevância de cada periódico. Para isto, utilizou-se como critério de classificação o Fator de Impacto (FI). Este fator “estima o grau de relevância de artigos, cientistas e periódicos científicos, em determinada área do conhecimento” (MARIANO; ROCHA, 2017) e “é obtido dividindo-se o número total de citações dos artigos acumulados nos últimos dois anos pelo total acumulado de artigos publicados pela revista no referido período” (MARZIALE; MENDES, 2002). Ao todo foram indicadas 50 revistas. O gráfico 1 aponta as 10 revistas mais relevantes, posicionadas conforme valor do Fator de Impacto na área de Sensoriamento Remoto.

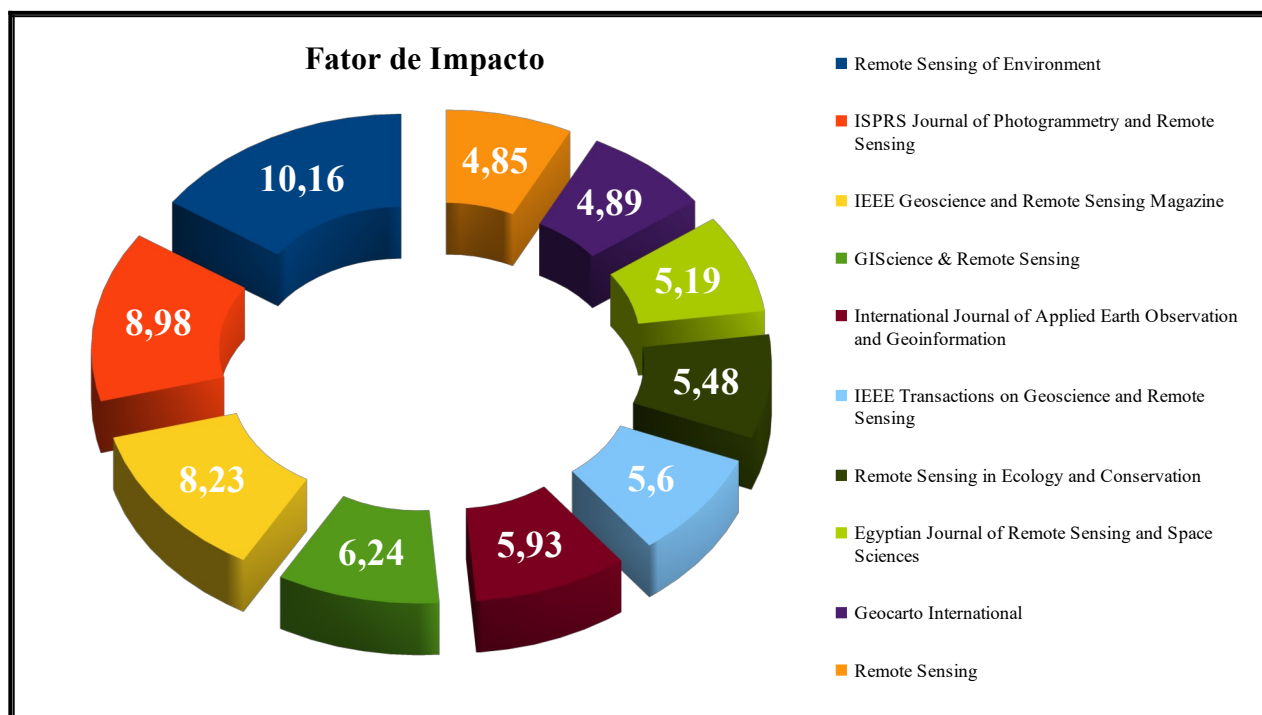


Gráfico 1: Revistas mais relevantes na área de Sensoriamento Remoto, conforme valor do Fator de Impacto

Fonte: Próprio autor. Adaptado de *Journal Citation Reports*, 2021

Após este estudo inicial, foram realizados as análises dos resultados dos demais filtros bibliométricos, conforme as palavras-chave definidas na fase anterior.

2.3.1 Segmentação Semântica no Sensoriamento Remoto

Conforme apresentado no item 2.1, a pesquisa pelo termo “*Semantic Segmentation*”, associado a expressão “*Deep Learning*” e refinado na categoria “*Remote Sensing*”, apresentou 421 resultados.

Com intuito de compreender as principais tônicas relacionadas a esta área, inicialmente analisou-se a evolução do tema ano a ano. Em 2016, primeira data com registro deste assunto, foi publicado um único artigo – *Semantic Segmentation of Aerial Images with an Ensemble of CNNs* (MARMANIS et al., 2016). No decorrer dos anos, verificou-se um aumento expressivo do número de publicações e citações a respeito do tópico. Em 2020 e 2021, por exemplo, mais de 140 artigos foram publicados e chegou-se a um total de 2304 citações. A partir destes dados pode-se notar a atualidade do assunto e deduzir que o mesmo vem ganhando uma relevância significativa na produção científica. O gráfico 2 apresenta essa evolução temporal.

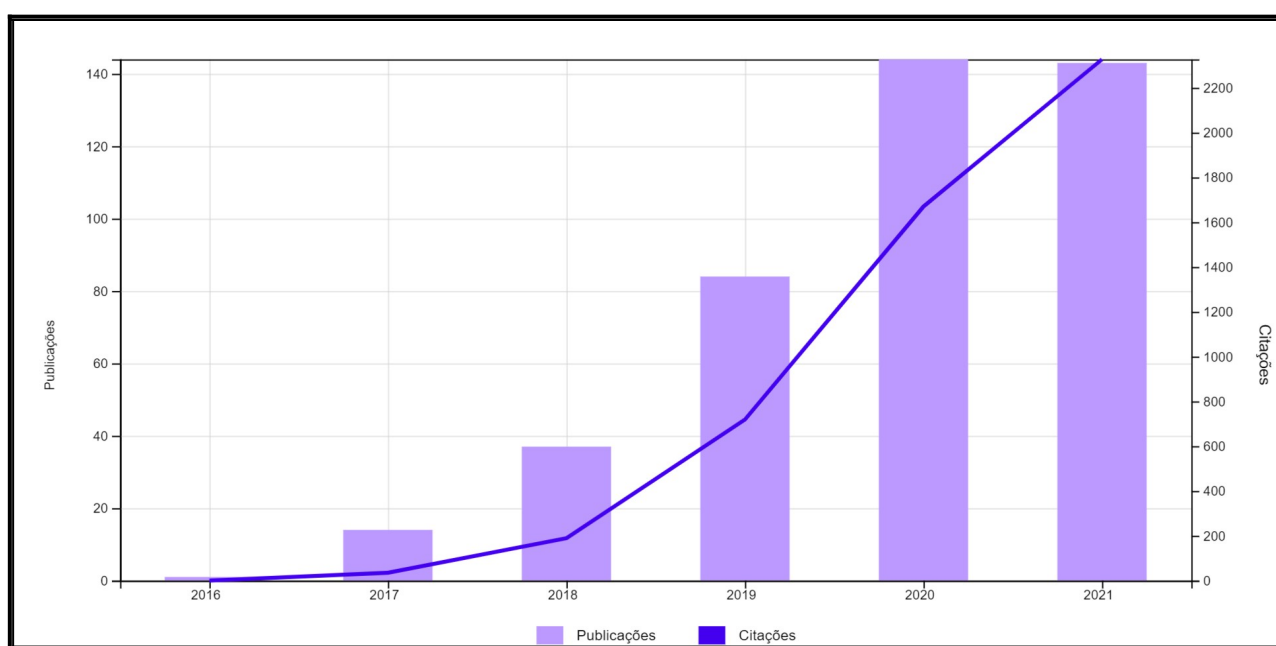


Gráfico 2: Segmentação Semântica – Evolução do número de publicações e citações ano a ano

Fonte: Web of Science, 2021

Com o objetivo de entender os principais conteúdos apresentados nestas publicações e verificar a tendência das mesmas, foi construída uma nuvem de palavras a partir das palavras-chaves dos artigos publicados, excetuando as palavras utilizadas no processo da pesquisa, conforme apresentado na figura 2.

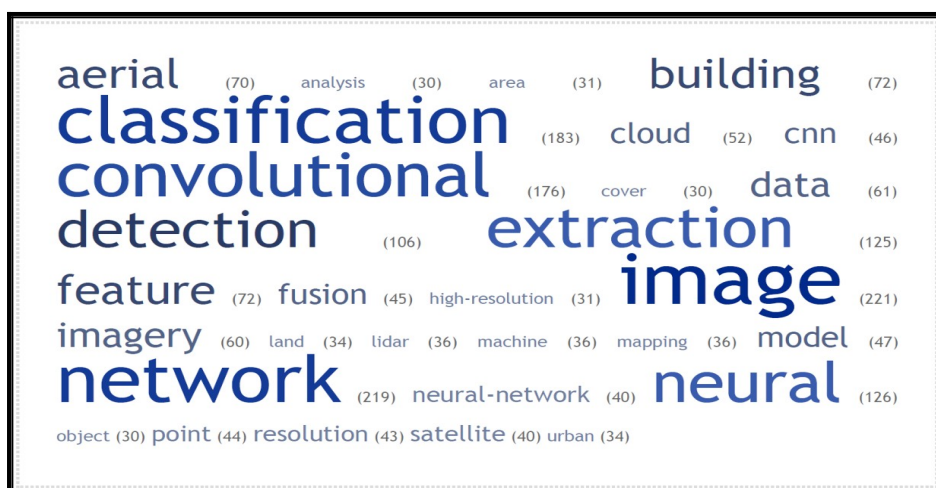


Figura 2: Segmentação Semântica – Nuvem de palavras construídas a partir das palavras-chaves dos artigos publicados

Fonte: Próprio autor

Na análise da figura 2, destacam-se os termos *image*, *network*, *classification*, *convolucional*, *neural*, *extration* e *detection*. Deste conjunto de palavras, pode-se inferir que as pesquisas realizadas mostram a propensão do uso de Redes Neurais Convolucionais para classificação de imagens. Esta tendência inicia-se a partir do artigo de Marmanis et al. (2016). Neste artigo é utilizado um conjunto de CNN, especialmente as *Fully Convolution Networks* (FCNs), para segmentação semântica pixel a pixel em imagens de altíssima resolução. Na análise dos 15 artigos mais citados, apresentados na tabela 1, observa-se a mesma tendência.

Artigo	Ano de Publicação	Nr Citações
1 Road Extraction by Deep Residual U-Net (ZHANG; LIU; WANG, 2018)	2018	390
2 Deep learning in remote sensing applications: A meta-analysis and review (MA et al., 2019)	2019	362
3 Classification for High Resolution Remote Sensing Imagery Using a Fully Convolutional Network (FU et al., 2017)	2017	164
4 Semantic Segmentation of Aerial Images with an Ensemble of CNNs (MARMANIS et al., 2016)	2016	152
5 Building Extraction in Very High Resolution Remote Sensing Imagery Using Deep Learning and Guided Filters (XU et al., 2018a)	2018	151

6 <i>Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning (KEMKER; SALVAGGIO; KANAN, 2018)</i>	2018	138
7 <i>Gated Convolutional Neural Network for Semantic Segmentation in High-Resolution Images (WANG et al., 2017a)</i>	2017	118
8 <i>Segment-before-Detect: Vehicle Detection and Classification through Semantic Segmentation of Aerial Images (AUDEBERT; LE SAUX; LEFÈVRE, 2017)</i>	2017	114
9 <i>Learning Aerial Image Segmentation From Online Maps (KAISER et al., 2017)</i>	2017	95
10 <i>Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks (WURM et al., 2019)</i>	2019	86
11 <i>End-to-End Change Detection for High Resolution Satellite Images Using Improved UNet plus (PENG; ZHANG; GUAN, 2019)</i>	2019	80
12 <i>ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data (DIAKOGLIANNIS et al., 2020)</i>	2020	79
13 <i>Road Extraction from High-Resolution Remote Sensing Imagery Using Deep Learning (XU et al., 2018b)</i>	2018	79
14 <i>Fully Convolutional Networks for Semantic Segmentation of Very High Resolution Remotely Sensed Images Combined With DSM (SUN; WANG, 2018)</i>	2020	75
15 <i>Hourglass-ShapeNetwork Based Semantic Segmentation for High Resolution Aerial Imagery (LIU et al., 2017)</i>	2018	68

Tabela 1: Segmentação Semântica – Artigos mais citados

Fonte: Próprio autor. Adaptado do *Web of Science* (2021)

No que concerne à análise dos autores, a mesma foi realizada, em um primeiro momento, a partir do levantamento dos pesquisadores que mais publicaram sobre o tema, conforme destacado no gráfico 3. Após esta etapa, foi observada a importância dos autores, a partir da análise de *coupling*, e foi realizada a análise de *co-citation*. Estes resultados foram apresentados nas figuras 3 e 4, respectivamente.

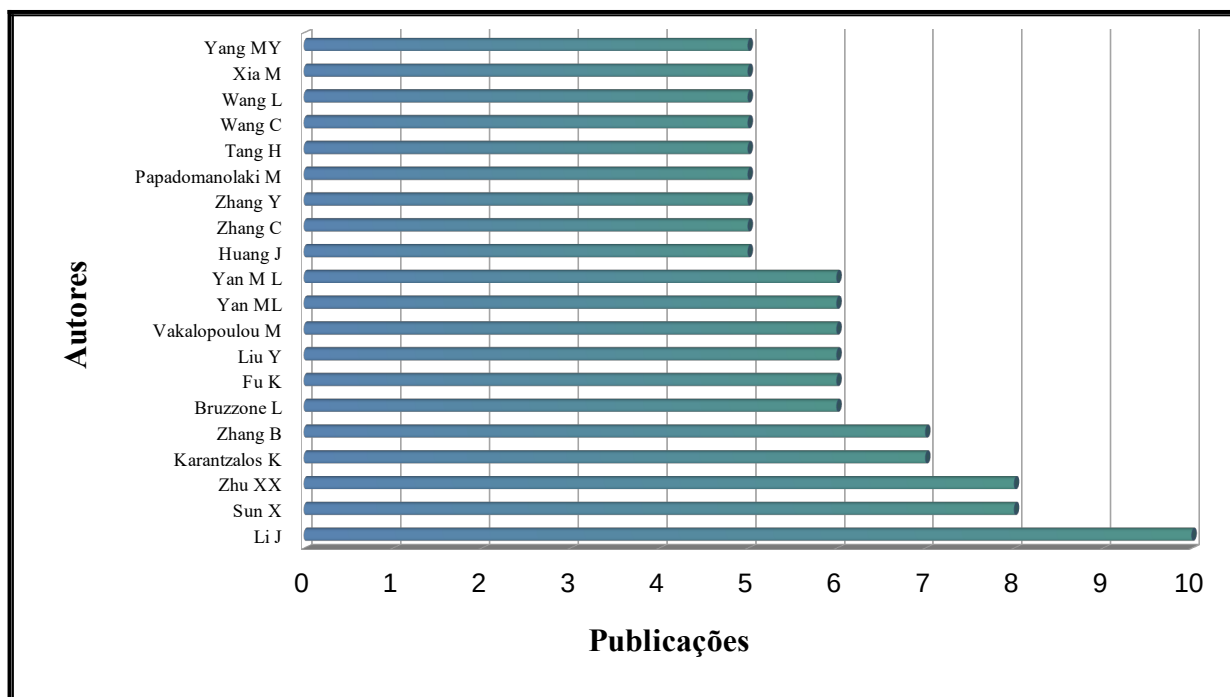


Gráfico 3: Segmentação Semântica – Autores que mais publicaram

Fonte: Próprio autor. Adaptado do *Web os Science* (2021)

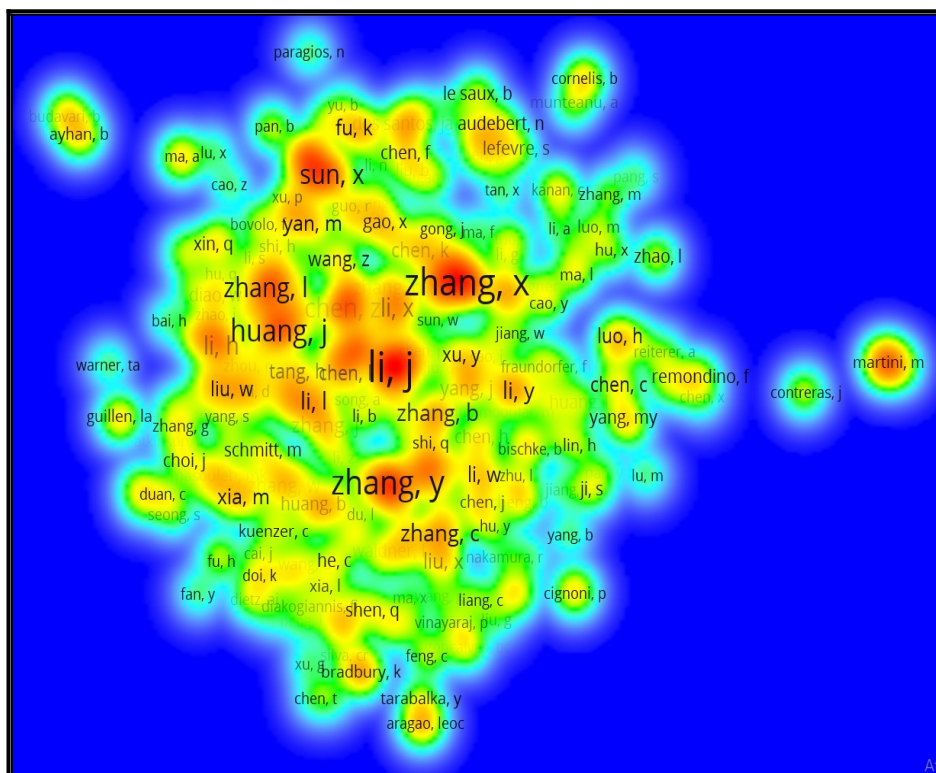


Figura 3: Segmentação Semântica – Mapa de calor representando os pesquisadores de maior destaque

Fonte: Próprio autor

Do mapa de calor apresentado na figura 3, destacaram-se os seguintes autores: Jonathan Li, Xinchang Zhang, Yi Zhang.

Jonathan Li se destaca por seus artigos voltados para extração e detecção de feições, como prédios, vegetações, navios, etc. Seu artigo mais citado nestes temas é *Attention-Gate-Based Encoder-Decoder Network for Automatic Building Extraction*, publicado em 2021, com 29 citações.

Os artigos de Xinchang Zhang referem-se, principalmente, à utilização de técnicas de deep learning para análise de imagens aéreas. Seu artigo mais citado foi *Research Progress on Few-Shot Learning for Remote Sensing Image Interpretation*, de 2021, com 55 citações.

O principal artigo de Yi Zhang, *SPMF-Net: Weakly Supervised Building Segmentation by Combining Superpixel Pooling and Multi-Scale Feature Fusion*, evidencia uma abordagem diferente das segmentações tradicionais ao propor uma segmentação supervisionada fraca. Este artigo foi publicado em 2020 e possui 20 citações.

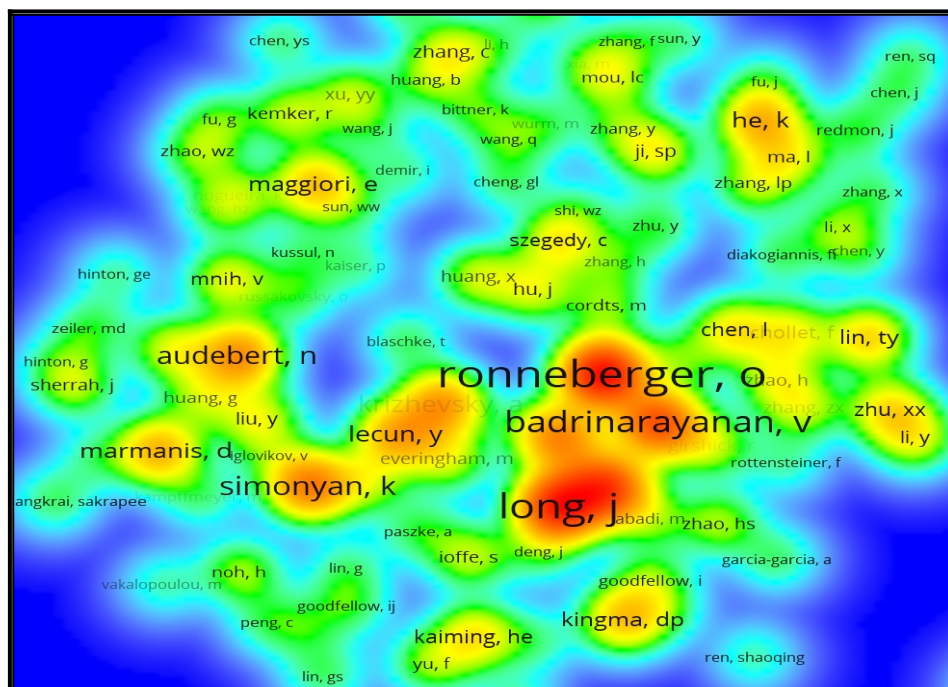


Figura 4: Segmentação Semântica – Mapa de calor de co-citation

Fonte: Próprio autor

No mapa de calor apresentado na figura 4 verifica-se três pesquisadores em destaque: Olaf Ronneberger, Jonathan Long e Vijay Badrinarayanan.

Na pesquisa de Olaf Ronneberger, destaca-se o artigo *U-Net: Convolutional Networks for Biomedical Image Segmentation* (RONNEBERGER; FISCHER; BROX, 2015). Este artigo apresenta a *U-Net* como uma arquitetura de rede neural convolucional capaz de capturar o contexto dos pixels da imagem para segmentar a mesmas de maneira mais precisa e com uma quantidade menor de exemplos de treinamento.

Na pesquisa de Jonathan Long, ressalta-se o trabalho *Fully Convolutional Networks for Semantic Segmentation* (LONG; SHELHAMER; DARRELL, 2015). Este artigo objetiva-se a construir redes convolucionais totais para segmentação de imagens de maneira precisa, a partir de dados de entrada de tamanhos arbitrários.

Na pesquisa de Vijay Badrinarayanan, destaca-se o artigo *SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation* (BADRINARAYANAN; HANDA; CIPOLLA, 2015). Este artigo apresenta a SegNet como uma rede neural convolucional para segmentação semântica, demonstrando suas peculiaridades frente as demais redes existentes.

É notório que os três trabalhos supracitados apresentam diferentes arquiteturas de redes neurais convolucionais para aplicação à segmentação semântica, o que corrobora o potencial de tais ferramentas para esta atividade.

Tendo em vista a análise acerca dos artigos publicados e do *co-citation* apontarem as Redes Neurais Convolucionais como sendo as principais ferramentas para a segmentação semântica de imagens, foi feita um estudo específico no artigo *Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review—Part II: Applications* (HOESER; BACHOFER; KUENZER 2020), publicado na revista *Remote Sensing*, com intuito de verificar, especificamente, as arquiteturas de redes neurais mais utilizadas. Este artigo busca apresentar o Estado da Arte para Segmentação Semântica aplicada ao Sensoriamento Remoto entre os anos de 2010 e 2020.

Dentre as diversas categorias disponíveis naquele trabalho, foi explorada, especialmente, a seção 2.5.2. Nesta, foram realizadas análises em cima de 261 artigos para averiguação das arquiteturas mais empregadas para segmentação semântica. Como resultado, verificou-se que as principais arquiteturas usadas foram: U-Net, com aproximadamente 33% dos resultados; Arquiteturas Customizadas, com 22%; *Fully Convolutional Network*, com 17%; SegNet, com 7%; e ResNet, com 5%. O gráfico 4 apresenta os resultados encontrados.

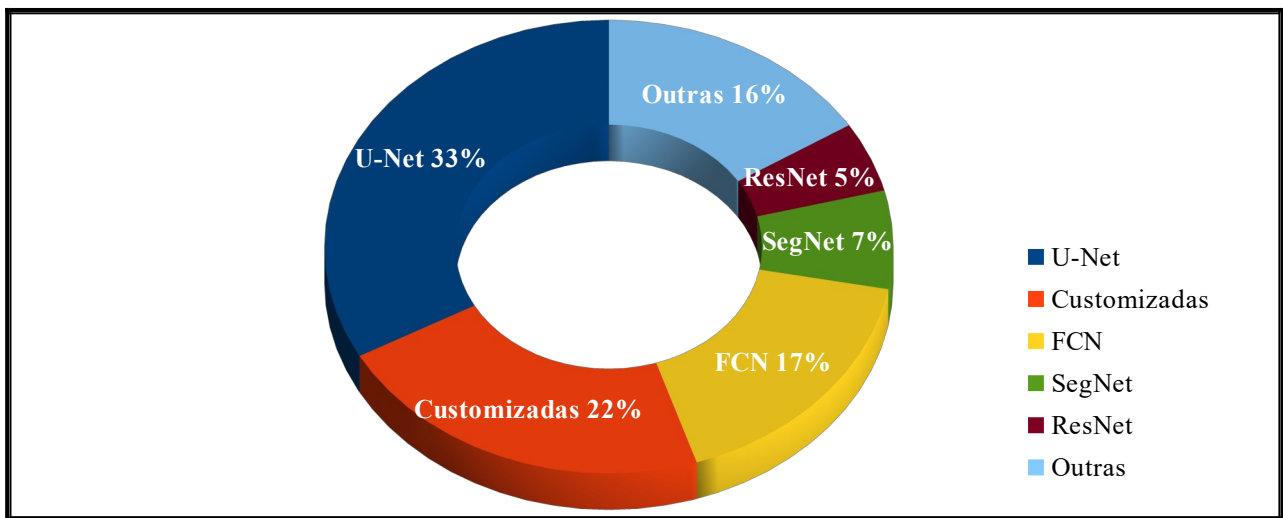


Gráfico 4: Arquiteturas de CNNs mais utilizadas para segmentação semântica

Fonte: Próprio autor. Adaptada de Hoeser, Bachofer e Kuenzer (2020)

No que concerne aos países que mais pesquisaram o tema, destacaram-se a China (209 registros), os Estados Unidos (44 registros) e a Alemanha (41 registros). O Brasil, é o décimo em número de publicações, com 13 registros (*Web of Science*, 2021).

Por fim, no tocante as revistas com maior número de publicações sobre o tema, as dez de maior destaque são apresentadas no gráfico 5.

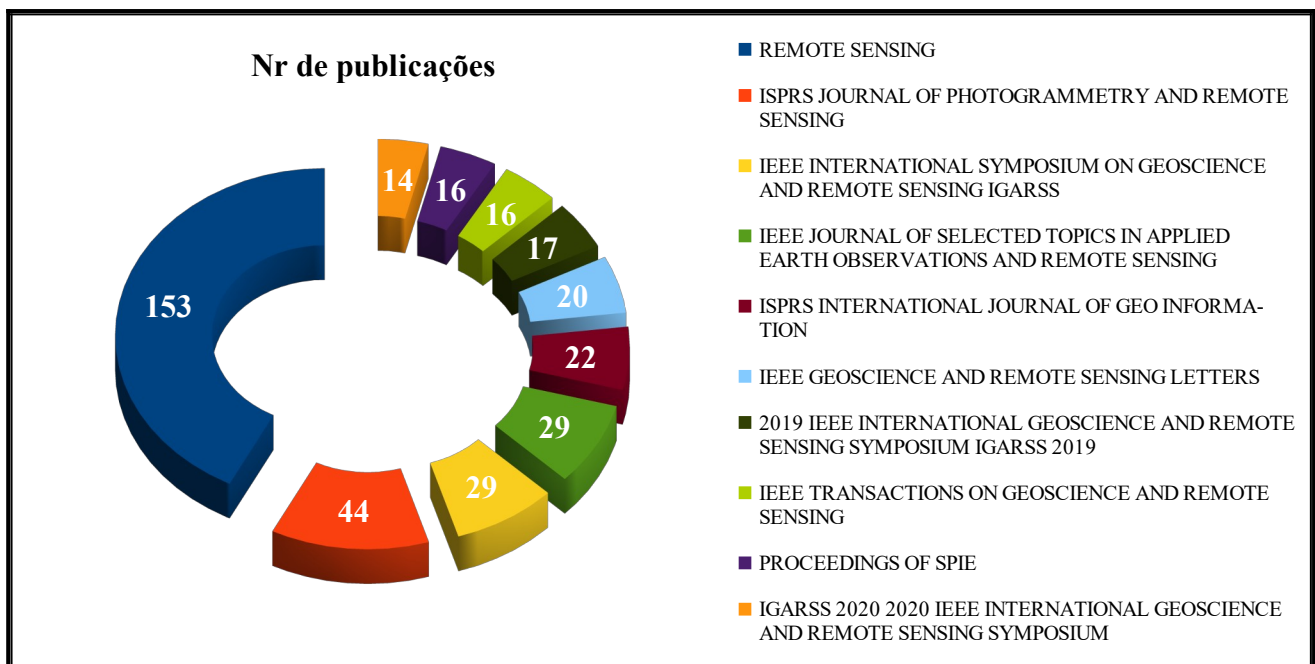


Gráfico 5: Revistas com maior número de publicações sobre Segmentação Semântica

Fonte: Próprio autor. Adaptado do *Web of Science* (2021)

2.3.2 Detecção de Mudança no Sensoriamento Remoto

Para esta análise seguiu-se os mesmos procedimentos apresentados na seção 2.2.1. Conforme apresentado no item 2.1, a pesquisa pelo termo “*Change Detection*”, associado a expressão “*Deep Learning*” e refinado na categoria “*Remote Sensing*”, apresentou 170 resultados.

A respeito da evolução do tema ano a ano, verificou-se que o número de publicações cresceram no decorrer do tempo. Em 2016, primeira data com registro deste tema, foram publicados 4 artigos. Em 2017 e 2018 foram publicados 9 e 10 artigos respectivamente. Já em 2019 houve um crescimento exponencial, sendo publicados 44 artigos. Uma alta de 440% em relação ao ano anterior. Em 2020 foram publicados 48 artigos e em 2021 o número chegou a 55. Além disso, chegou-se a um total superior a 800 citações. Esta expansão é apresentada no gráfico 6.

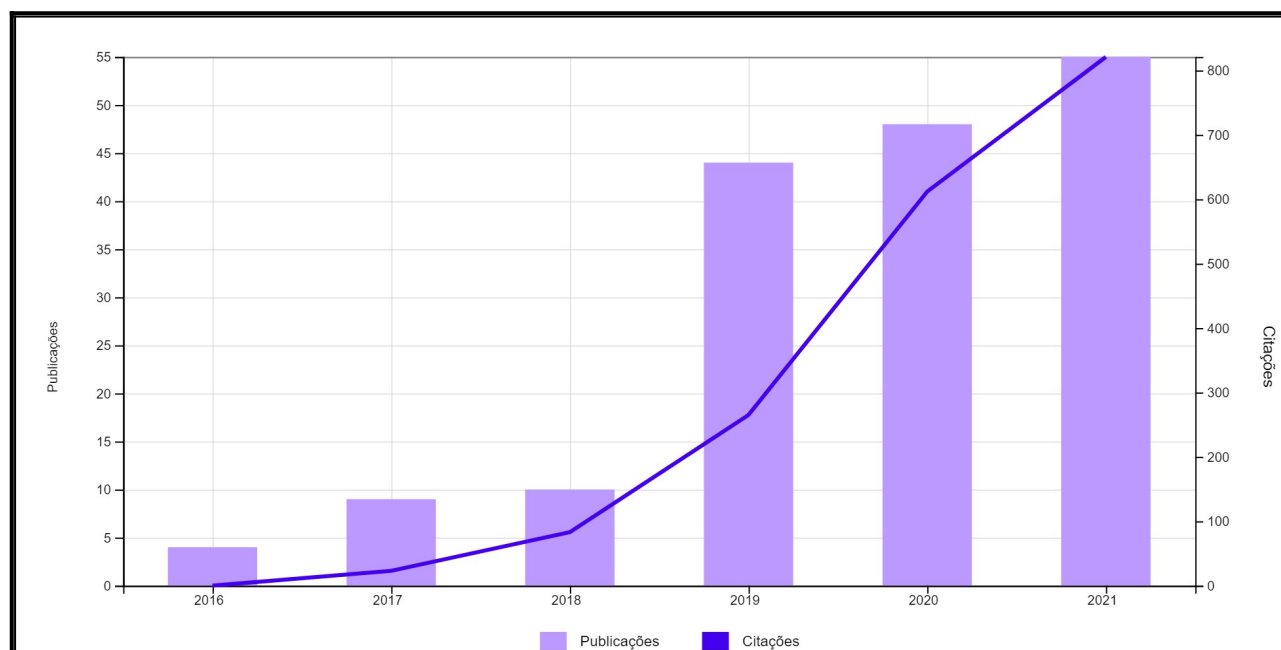


Gráfico 6: Detecção de Mudança – Evolução do número de publicações e citações ano a ano

Fonte: Web of Science, 2021

Para verificar os principais tópicos apresentados nestas publicações e analisar as tendências das mesmas, foi construída a nuvem de palavras a partir das palavras-chaves dos artigos publicados, excetuando as palavras utilizadas no processo da pesquisa, conforme apresentado na figura 5.

5 <i>End-to-End Change Detection for High Resolution Satellite Images Using Improved UNet plus (PENG; ZHANG; GUAN, 2019)</i>	2019	80
6 <i>Unsupervised Deep Slow Feature Analysis for Change Detection in Multi-Temporal Remote Sensing Images (DU et al., 2019)</i>	2019	59
7 <i>Feature-Level Change Detection Using Deep Representation and Feature Change Analysis for Multispectral Imagery (ZHANG et al., 2016)</i>	2016	57
8 <i>Forest Change Detection in Incomplete Satellite Images With Deep Neural Networks (KHAN et al., 2017)</i>	2017	55
9 <i>Change detection based on Faster R-CNN for high-resolution remote sensing images (WANG et al., 2018)</i>	2018	41
10 <i>Change Detection in Hyperspectral Images Using Recurrent 3D Fully Convolutional Networks (SONG et al., 2018)</i>	2018	40
11 <i>Landslide mapping from multi-sensor data through improved change detection-based Markov random field (LU et al., 2019)</i>	2019	38
12 <i>Multitemporal Very High Resolution From Space: Outcome of the 2016 IEEE GRSS Data Fusion Contest (MOU et al., 2017)</i>	2017	37
13 <i>Detecting Large-Scale Urban Land Cover Changes from Very High Resolution Remote Sensing Images Using CNN-Based Classification (ZHANG et al., 2019)</i>	2019	35
14 <i>Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges (SHI et al., 2020)</i>	2020	33
15 <i>Forest Damage Assessment Using Deep Learning on High Resolution Remote Sensing Data (HAMDI; BRANDMEIER; STRAUB, 2019)</i>	2019	32

Tabela 2: Detecção de Mudança – Artigos mais citados
Fonte: Próprio autor. Adaptado do *Web of Science* (2021)

No tocante aos autores, o gráfico 7 e a figuras 6 e 7 apresentam, respectivamente, os pesquisadores que mais publicaram, quais são os mais relevantes em suas respectivas áreas (análise de *coupling*) e a análise de *co-citation*.

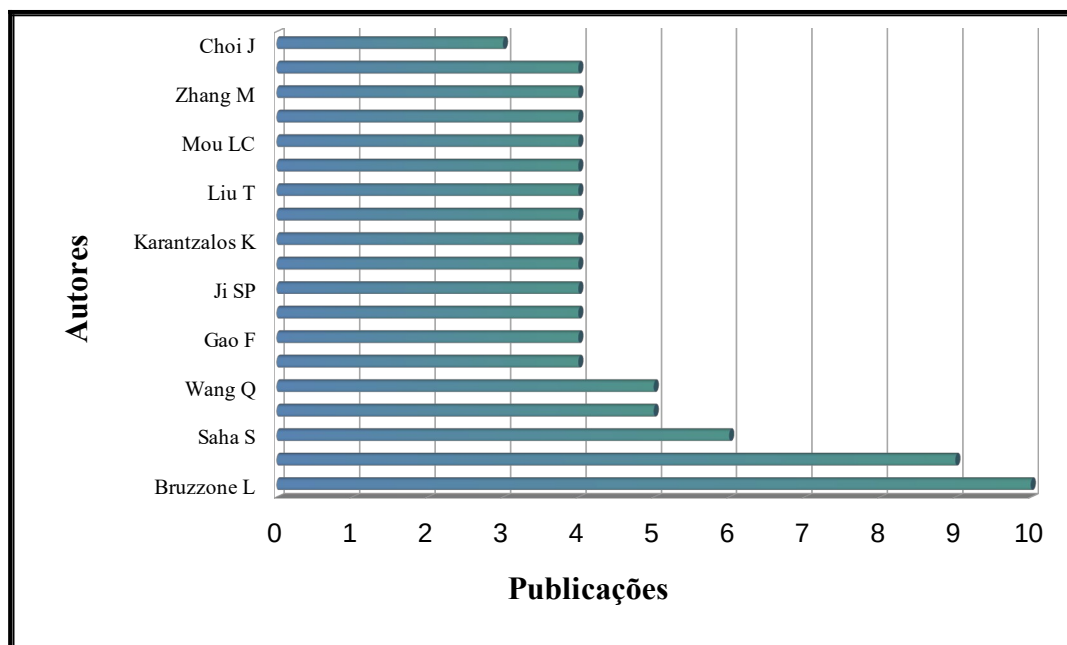


Gráfico 7: Detecção de Mudança - Autores que mais publicaram

Fonte: Próprio autor. Adaptado do *Web os Science* (2021)

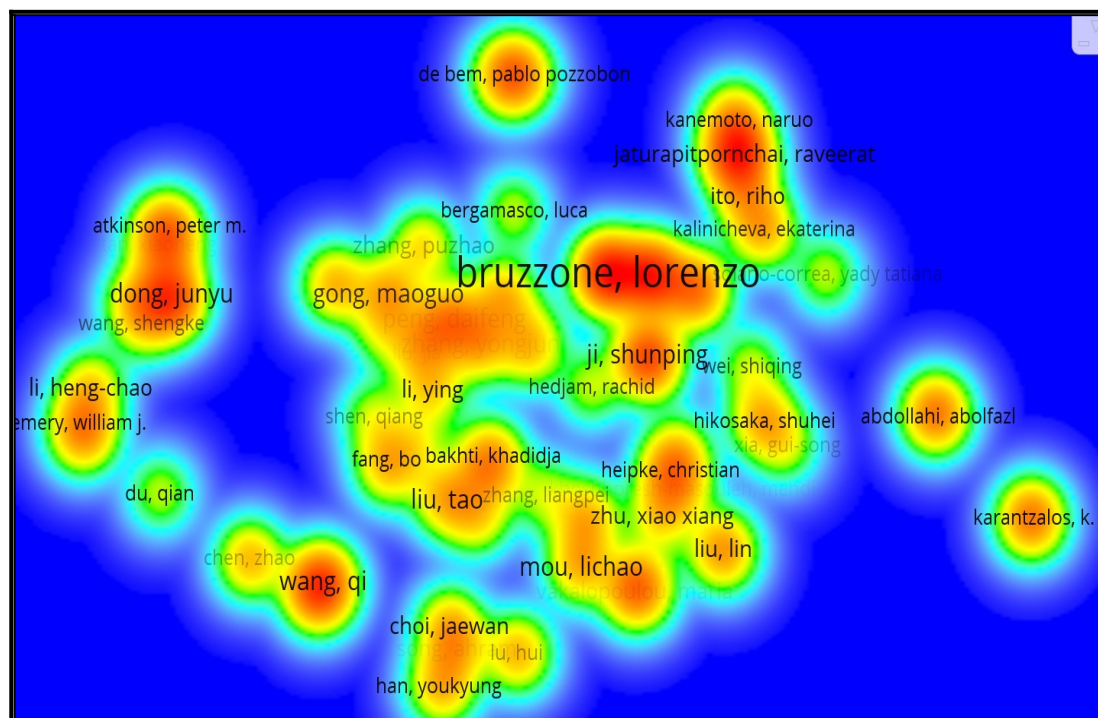


Figura 6: Detecção de Mudança – Mapa de calor representando os pesquisadores de maior destaque

Fonte: Próprio autor

Do mapa de calor apresentado na figura 6, destacaram-se os seguintes autores: Lorenzo Bruzzone, Qi Wang, Junyu Dong, entre outros.

O trabalho de Lorenzo Bruzzone se destaca pela utilização de redes convolucionais para detecção de mudanças semi supervisionadas. Seu principal artigo, *SemiCDNet: A Semisupervised Convolutional Neural Network for Change Detection in High Resolution Remote-Sensing Images*, publicado em 2021, foi citado 45 vezes.

Qi Wang se sobressai por suas pesquisas de detecção de mudança em imagens hiperespectrais. Seu artigo mais citado, *GETNET: A General End-to-End 2-D CNN Framework for Hyperspectral Image Change Detection*, publicado em 2019, possui 258 citações.

Junyu Dong se destaca por seus artigos voltados para análise da detecção de mudança a partir de imagens SAR. O seu principal artigo, *Transferred Deep Learning for Sea Ice Change Detection From Synthetic-Aperture Radar Images*, publicado em 2019, foi citado 32 vezes.

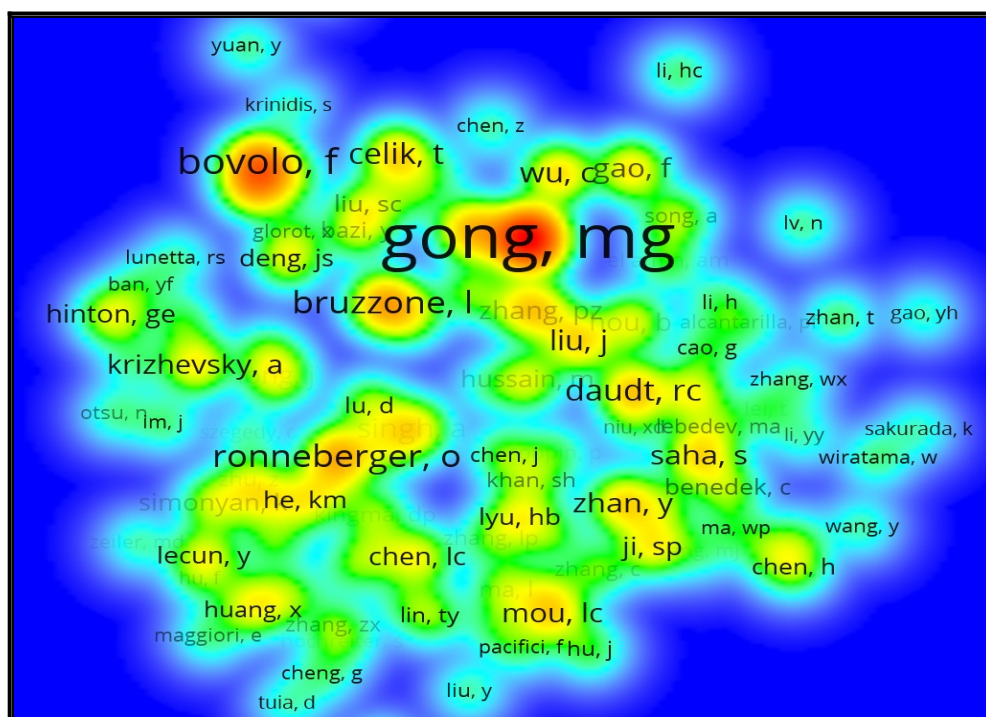


Figura 7: Detecção de Mudança – Mapa de calor de co-citation

Fonte: Próprio autor

No mapa de calor apresentado na figura 7 verifica-se 2 autores em destaque: Maoguo Gong e Francesca Bovolo.

Na pesquisa de Maoguo Gong, destacam-se artigos voltados para detecção de mudança a partir de imagens SAR, como por exemplo os artigos: *Change detection in synthetic aperture radar images based on deep neural networks* (GONG et al., 2016), publicado na revista *IEEE transactions on neural networks and learning systems*; e *Feature learning and change feature classification based on deep learning for ternary change detection in SAR images* (GONG; YANG; ZHANG, 2017), publicado na revista *ISPRS Journal of Photogrammetry and Remote Sensing*.

Na pesquisa de Francesca Bolovo, ressaltam-se trabalhos de detecção de mudança não supervisionada em imagens multitemporais. Os artigos de maior destaque são: *A detail-preserving scale-driven approach to change detection in multitemporal SAR images* (BOVOLO; BRUZZONE, 2006); *A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain* (BOVOLO; BRUZZONE, 2007a); e *A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment* (BOVOLO; BRUZZONE, 2007b). Todos os artigos foram publicados na revista *IEEE Transactions on Geoscience and Remote Sensing*.

Para o estudo das principais arquiteturas utilizadas na detecção de mudança, foi realizado o estudo no artigo *Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges* (SHI et al., 2020), publicado na revista *Remote Sensing*. Este artigo busca apresentar o Estado da Arte para Detecção de Mudança baseada em inteligência artificial e aplicada ao Sensoriamento Remoto. Dentre tópicos disponíveis no artigo, o item 5 - *Mainstream Networks in AI*, trata exclusivamente das redes de Inteligência Artificial utilizadas na detecção de mudança. No tópico são citadas e apresentadas as 6 principais: *Autoencoder* (AE); *Deep Belief Network* (DBN); *Convolutional Neural Network* (CNN); *Recurrent Neural Network* (RNN); *Pulse Couple Neural Network* (PCNN); e *Generative Adversarial Networks* (GANs). No caso das redes convolucionais, SHI et al. (2020) destacou sua alta capacidade de aprendizagem, bem como o desempenho satisfatório em diferentes tarefas de processamento de imagem. Dentre as arquiteturas existentes, foram destacadas as seguintes: *VGGNet*, *CaffeNet*, *SegNet*, *UNet*, *InceptionNet* e *ResNet*.

No que concerne aos países que mais pesquisaram o tema destacaram-se a China (95 registros), Itália (17 registros), Alemanha (15 registros) e Estados Unidos (15 registros). O Brasil ocupa a décima quarta posição, com 3 registros (*Web of Science*, 2021).

No tocante as revistas com maior número de publicações sobre o tema, as 10 de maior destaque são apresentadas no gráfico 8.

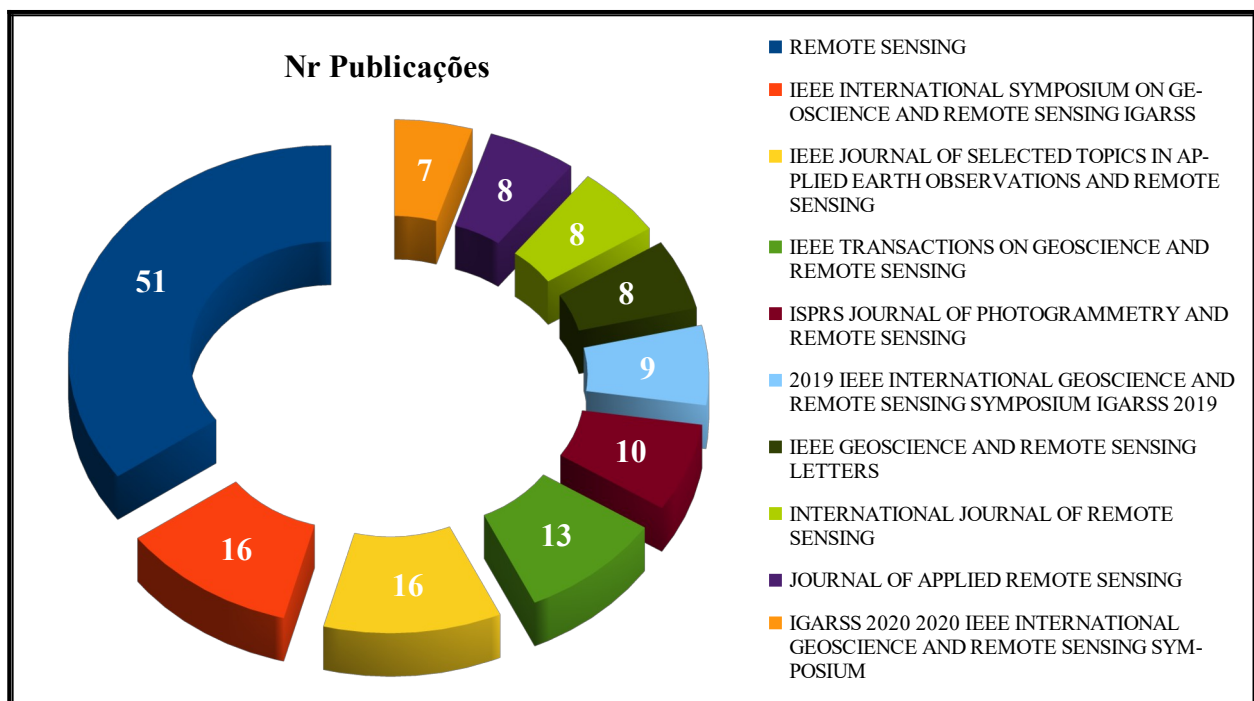


Gráfico 8: Detecção de Mudança – Revistas com maior número de publicações

Fonte: Próprio autor. Adaptado do *Web of Science* (2021)

2.3.3 CNN x Técnicas Clássicas de Aprendizado de Máquina

Considerando o que foi averiguado nos itens 2.2.1 e 2.2.2, em que verificou-se uma tendência em se utilizar redes neurais convolucionais para segmentação semântica e detecção de mudança, constatou-se a necessidade de investigar artigos que comparassem a utilização de CNNs e Técnicas Clássicas de Aprendizado de Máquina. Conforme MA et al. (2019), no contexto do aprendizado de máquina, antes do desenvolvimento de técnicas de aprendizado profundo, os pesquisadores envolvidos com o sensoriamento remoto utilizavam o *Support Vector Machine* (SVM) ou classificadores de conjunto, como o *Random Forest* (RF), para classificação de imagens.

Neste contexto, foi feita uma análise no WoS dos termos “Deep Learning and Forest”, refinados na categoria *Remote Sensing*. Foram encontrados ao todo 407 resultados. Explorando os artigos mais citados que tratam deste tema, destacaram-se:

- *Deep learning classification of land cover and crop types using remote sensing data* (KUSSUL et al., 2017), citado 443 vezes. Este artigo traz uma comparação entre o uso de CNN e RF para cobertura do solo e classificação de tipo de cultivo a partir de imagens multitemporais dos satélites *Landsat-8* e *Sentinel-1A*. Nos resultados apresentados no artigo verificou-se que as CNNs produziram classificações mais precisas se comparadas às RF. Enquanto a 1-D e 2-D CNN

alcançaram 92.5% e 94.6%, respectivamente, a acurácia da RF foi de 88.7%. Além disso, foi apresentado que as CNNs produziram melhores resultados para determinação das principais culturas analisadas (trigo, milho, girassol, soja e beterraba), atingindo uma acurácia de 85%;

- *Deep learning based multi-temporal crop classification* (ZHONG; HU; ZHOU, 2019), citado 158 vezes. Este artigo tem como objetivo desenvolver uma estrutura de classificação baseada em aprendizagem profunda para séries temporais utilizando o *Landsat Enhanced Vegetation Index* (EVI), cujo índice se baseia pela quantificação dos verdes da vegetação. Foram projetados dois modelos de aprendizado profundo: *Long Short-Term Memory* (LSTM) e *one-dimensional convolutional* (Conv1D). Além disso, o artigo traz uma comparação entre os modelos projetados e outros três classificadores: *Gradient Boosting Machine* (XGBoost), *Random Forest* e *Support Vector Machine*. Como conclusão, o estudo mostrou que o modelo do Conv1D forneceu uma eficiência maior para representação de séries temporais em tarefas de classificação multitemporais, atingindo 85.54% de acurácia, enquanto RF e SVM atingiram 82.38% e 82.45% respectivamente;

- *Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery* (MAHDIANPARI et al., 2018), citado 123 vezes. Este artigo apresenta uma investigação das ferramentas de aprendizagem profunda para a classificação de classes de áreas úmidas a partir de imagens multiespectrais *RapidEye*. No artigo foram examinados sete Redes Neurais Convolucionais (DenseNet121, InceptionV3, VGG16, VGG19, Xception, ResNet50 e InceptionResNetV2). Os resultados da classificação obtidos das CNNs foram comparados as ferramentas de aprendizado de máquina *Random Forest* e *Support Vector Machine* (SVM). O artigo ilustra que a precisão de classificação das CNNs (InceptionResNetV2 – 96.17%, ResNet50 – 94.81%, Xception – 92.57%) são superiores as obtidas por *Random Forest* (76.08%) e SVM (74.89%).

- *Comparing fully convolutional networks, random forest, support vector machine and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system* (LIU et al., 2018), citado 102 vezes. Neste artigo é realizada uma comparação entre *Fully Convolutional Networks* (FCN), *Patch-based Deep Convolutional Neural Networks* (DCNN), *Random Forest* e *Support Vector Machine* para classificar sete tipos de cobertura natural do solo. Os resultados alcançados no artigo indicam que a DCNN pode produzir um desempenho inferior em comparação com classificadores convencionais quando o tamanho da amostra de treinamento é pequeno, porém tende a mostrar uma precisão maior para amostras de

treinamento grandes. Os resultados também mostram que a FCN é a mais eficiente na utilização das informações na amostra de treinamento e que a mesma possui uma precisão maior do que todos os outros métodos de classificação, independentemente do número de amostras de treinamento.

- *Deep Convolutional Neural Network for Complex Wetland Classification Using Optical Remote Sensing Imagery (REZAEE et al., 2018)*, citado 76 vezes. Este artigo propõe uma estrutura de classificação baseada em rede neural convolucional (CNN) para mapeamento de zonas úmidas, a partir de dados de sensoriamento remoto óptico. Foi utilizada como base a CNN AlexNet. Para verificar a eficiência da rede, foi realizada a comparação com o classificador *Randon Forest*.

Em uma análise sucinta destes artigos, nota-se que os mesmos convergem para a superioridade das Redes Neurais Convolucionais quando comparadas com as Técnicas Clássicas de Aprendizado de Máquina, como SVM e *Randon Forest*. Isto fica evidente quando da acareação dos resultados apresentados pelos autores considerados. Ressalta-se, entretanto, que não há uma uniformidade das CNN utilizadas nas comparações, ou seja, as arquiteturas de redes utilizadas nos trabalhos foram adaptadas para o problema proposto no respectivo artigo. Isto posto, pode-se inferir que, se a rede convolucional não for adequada ao caso em estudo, a mesma pode ter resultados inferiores aos apresentados.

2.3.4 Garimpagem de ouro e Sensoriamento Remoto

Conforme apresentado no item 2.1, a pesquisa pelo termo “*Gold Mining*”, refinado na categoria “*Remote Sensing*”, apresentou 25 resultados.

A respeito da evolução do tema ano a ano, verificou-se que a primeira publicação data do ano 2000, com o artigo *Detecting areas disturbed by gold mining activities through JERS-1 SAR images, Roraima State, Brazilian Amazon (ALMEIDA-FILHO; SHIMABUKURO, 2000)*. Após hiatos de publicações no início dos anos 2000, com artigos publicados apenas em 2006 (2 artigos), 2008 (1 artigo) e 2010 (1 artigo), a partir de 2014 as publicações acerca do tema foram recorrentes. Com relação ao número de citações, o mesmo aumentou consideravelmente a partir de 2015. O gráfico 9 apresenta a evolução temporal deste tema.

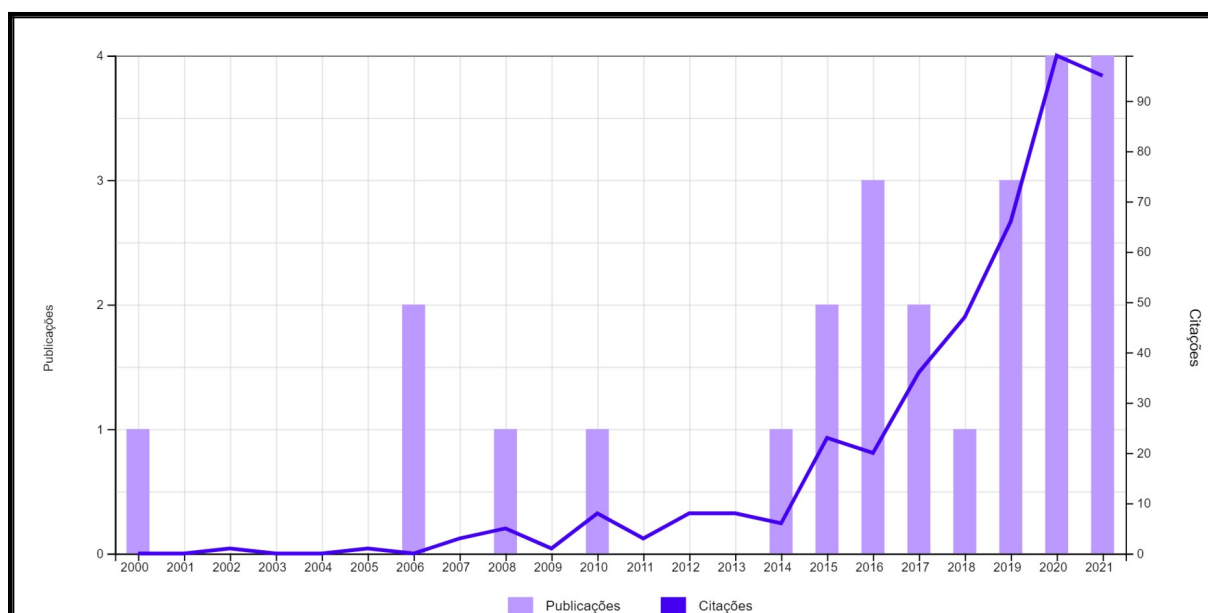


Gráfico 9: Garimpagem de ouro – Evolução do número de publicações e citações ano a ano

Fonte: Web of Science, 2021

Na verificação dos principais tópicos foi construída a nuvem de palavras a partir das palavras-chaves dos artigos, excetuando a palavra utilizada no processo da pesquisa, conforme figura 8.

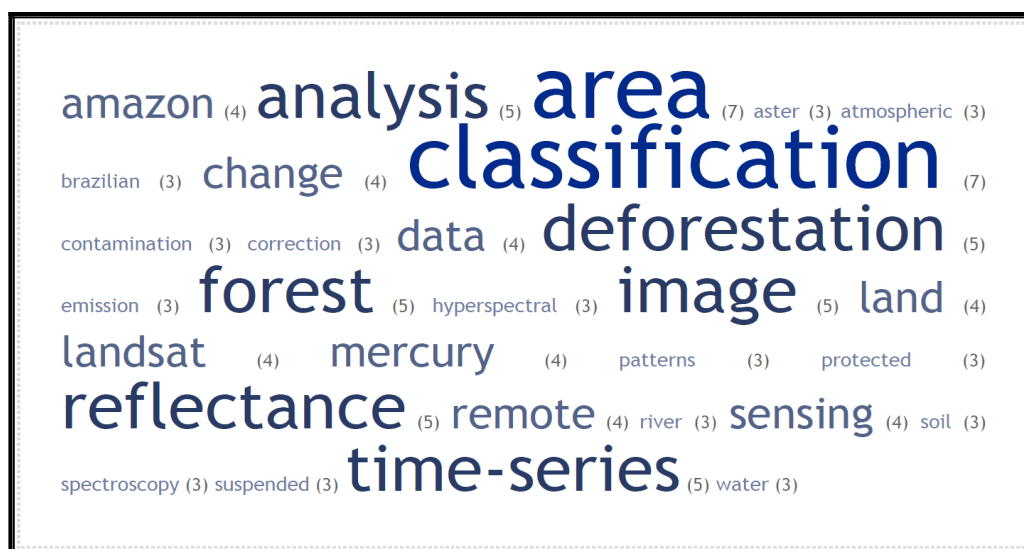


Figura 8: Garimpagem de ouro – Nuvem de palavras construídas a partir das palavras-chaves dos artigos publicados

Fonte: Próprio autor

Neste contexto, os termos que mais se destacaram foram: *area*, *classification*, *analysis*, *deforestation*, *forest*, *image*, *reflectance* e *time-series*. Destes dados, pode-se inferir que para classificar áreas de mineração aurífera é necessário analisar imagens das áreas florestadas e

desmatadas a partir de uma série temporal. A análise dos artigos mais citados reforça essa propensão, conforme apresentado na tabela 3.

Artigo	Ano de Publicação	Nr Citações
1 <i>Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities (LOBO; COSTA; NOVO, 2015)</i>	2015	81
2 <i>Deforestation and Forest Degradation Due to Gold Mining in the Peruvian Amazon: A 34-Year Perspective (CABALLERO ESPEJO et al., 2018)</i>	2018	66
3 <i>Mapping mineralogical alteration using principal-component analysis and matched filter processing in the Takab area, north-west Iran, from ASTER data (MOORE et al., 2008)</i>	2008	50
4 <i>Spaceborne Mine Waste Mineralogy Monitoring in South Africa, Applications for Modern Push-Broom Missions: Hyperion/OLI and EnMAP/Sentinel-2 (MIELKE et al., 2014)</i>	2014	48
5 <i>Distribution of Artisanal and Small-Scale Gold Mining in the Tapajós River Basin (Brazilian Amazon) over the Past 40 Years and Relationship with Water Siltation (LOBO et al., 2016)</i>	2016	37
6 <i>Channel and Floodplain Change Analysis over a 100-Year Period: Lower Yuba River, California (GHOSHAL et al., 2010)</i>	2010	28
7 <i>Indicator kriging applied to absorption band analysis in hyperspectral imagery: A case study from the Rodalquilar epithermal gold mining area, SE Spain (MEER, 2006)</i>	2006	27
8 <i>Mapping the expansion of galamsey gold mines in the cocoa growing area of Ghana using optical remote sensing (SNAPIR; SIMMS; WAINE, 2017)</i>	2017	23
9 <i>Land use/land cover dynamics using landsat data in a gold mining basin-the Ankobra, Ghana (OBODAI et al., 2019)</i>	2019	16
10 <i>A statistical model for spatial patterns of Buruli ulcer in the Amansie West district, Ghana (DUKER; STEIN; HALE, 2006)</i>	2006	13

Tabela 3: Garimpagem de ouro – Artigos mais citados
Fonte: Próprio autor. Adaptado do *Web of Science* (2021)

No tocante aos autores, os que mais publicaram e os mais relevantes são: Felipe de Lucia Lobo, Maycira Costa, Kelvin Telmer e Evelyn Marcia Leao de Moraes Novo, todos com 2 publicações. A análise de *co-citation* é apresentada na figura 9.

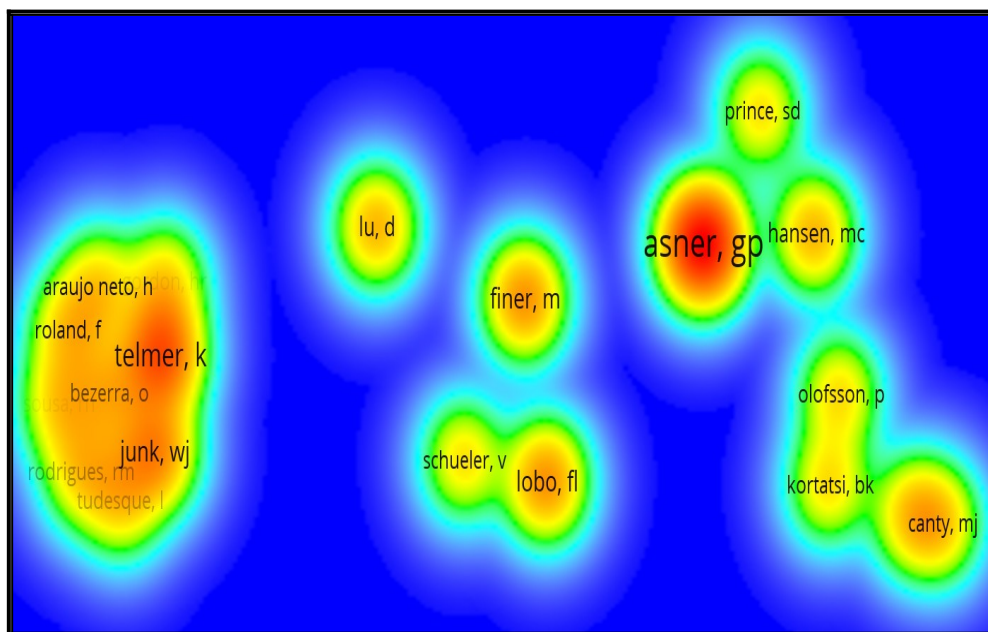


Figura 9: Garimpagem de ouro – Mapa de calor de co-citation

Fonte: Próprio autor

No mapa de calor da figura 2.10 verifica-se 3 autores em destaque: Gregory P. Asner, Kevin Telmer e Felipe de Lucia Lobo.

Na pesquisa de Gregory P. Asner, considerando a área de mineração, destacam-se os artigos *Elevated rates of gold mining in the Amazon revealed through high-resolution monitoring* (ASNER et al., 2013), que avalia a mineração de ouro em estradas e rios na Amazônia Peruana entre os anos de 1999 e 2012, e *Accelerated losses of protected forests from gold mining in the Peruvian Amazon* (ASNER; TUPAYACHI, 2016), que utiliza imagens de alta resolução, oriundas de sensores remotos, para avaliar as mudanças anuais na extensão da mineração no período de 1999 e 2016.

Na pesquisa de Kevin Telmer, sobressaem os trabalhos voltados para a emissão de mercúrio durante a mineração de ouro artesanal. O artigo de maior destaque é: *World emissions of mercury from artisanal and small scale gold mining* (TELMER; VEIGA, 2009).

Na pesquisa de Felipe de Lucia Lobo, voltada à área de mineração, destacam-se artigos voltados para análise dos impactos da atividade garimpeiras nas águas amazônicas. os seguintes artigos: *Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities* (LOBO; COSTA; NOVO, 2015), que utiliza imagens *Landsat* para avaliar a

qualidade da água impactada pela atividade garimpeira no período de 1973 e 2013; e *Mapping mining areas in the brazilian amazon using msi/sentinel-2 imagery* (LOBO et al., 2018), que propõe um método de detecção e mapeamento de áreas de garimpo, a partir de imagens de do Sentinel-2, para posterior integração dos mesmos em um ambiente de Sistema de Informação Geográfica (SIG).

No que concerne aos países que mais pesquisaram o tema destacaram-se os países apresentados no gráfico 10.

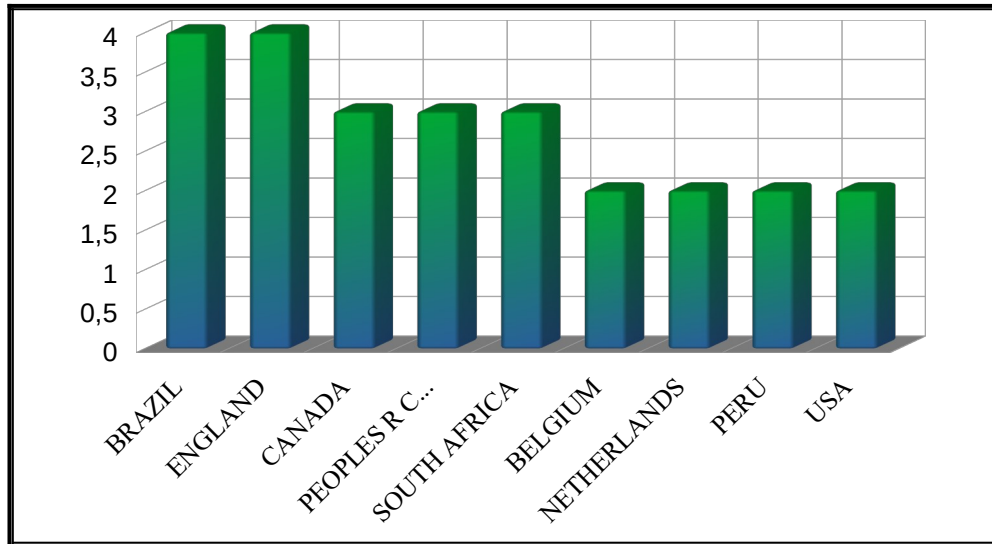


Gráfico 10: Garimpagem de ouro – Visão geral dos países que mais publicaram

Fonte: Próprio autor

No tocante as revistas com maior número de publicações sobre o tema, destacam-se as apresentadas no gráfico 11:

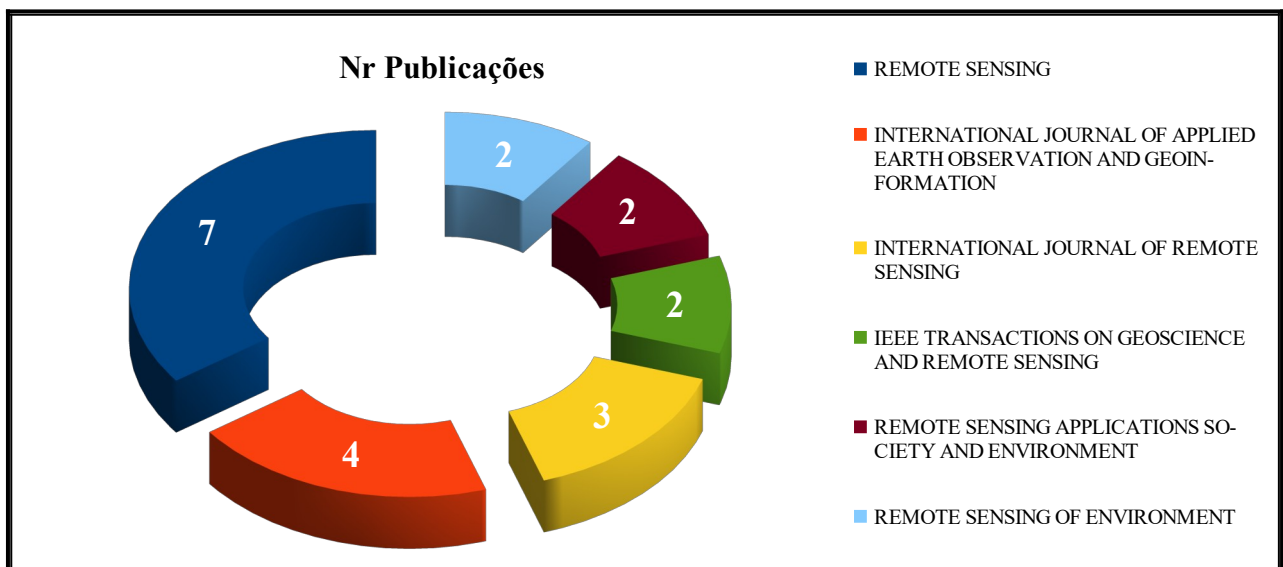


Gráfico 11: Garimpagem de ouro – Revistas com maior número de publicações

Fonte: Próprio autor. Adaptado do *Web of Science* (2021)

Após a revisão sistêmica da literatura apresentada, verificou-se que a utilização de imagens de sensores remotos possibilita a incorporação de técnicas de *deep learning*, especialmente as Redes Neurais Convolucionais, para mapeamento da atividade de mineração na Amazônia. Além disso, notou-se que a relação entre aprendizado profundo e mineração carece de pesquisas e que ainda há áreas de conhecimento a serem exploradas.

2.3.5 Considerações Finais acerca do Estado da Arte

A revisão sistemática da literatura apresentada apontou, dentro das técnicas de Aprendizado Profundo de Segmentação Semântica e Detecção de Mudança, as revistas mais relevantes sobre os temas, a evolução dos mesmos ano a ano, quais os artigos mais citados e quais autores, países e revistas que mais publicaram. Além disso, a utilização dos índices de *co-citation* e *coupling* auxiliou nos estudos acerca dos autores mais importantes em cada um dos assuntos. Outrossim, a aplicação da Teoria do Enfoque Meta Analítico Consolidado foi fundamental para que os artigos primordiais fossem assinalados e apresentados.

O estudo confirmou, ainda, a percepção intuitiva da notoriedade que os temas expostos vêm ganhando frente a comunidade científica ao longo dos anos e que países com grande desenvolvimento tecnológico como China, Estados Unidos e Alemanha, lideram as pesquisas nessas áreas. Ademais, pôde-se inferir que não há uma convergência na identificação das melhores CNN para classificar as imagens de sensores remotos. Notou-se que, para cada atividade desenvolvida pelos pesquisadores, as redes foram diferentemente escolhidas e/ou adaptadas. Isto posto, verifica-se que o ramo das Redes Neurais Convolucionais ainda está em evolução, com diversos algoritmos, arquiteturas e ferramentas sendo utilizadas e aperfeiçoadas para soluções de problemas.

Por fim, o Estado da Arte destacou, ainda, que a implementação de técnicas de *Deep Learning* para análise de imagens advindas de sensoriamento remoto é eficaz, principalmente quando comparadas as técnicas clássicas de Aprendizado de Máquina, como *Random Forest* e *Support Vector Machine*.

CAPÍTULO 3 – FUNDAMENTAÇÃO TEÓRICA

Levando em consideração o apresentado no Cap 2 – Estado da Arte, nota-se que as técnicas de Aprendizado Profundo para Segmentação Semântica e Detecção de Mudança de imagens oriundas de sensores remotos são agudamente eficazes quando utilizam-se CNN. Isto posto, o objetivo deste capítulo é apresentar os principais conceitos de Aprendizado Profundo, Redes Neurais Convolucionais, Segmentação Semântica e Detecção de Mudança, bem como exemplificar Arquiteturas Computacionais, Métricas de Avaliação e Hiperparâmetros utilizadas em CNN. Além disso, será exposto dois métodos de análise de similaridade de imagens (Erro Médio Quadrático e Índice de Similaridade Estrutural) e a conceptualização de *data augmentation*, que é um recurso fundamental para o treinamento de redes neurais profundas.

3.1 Aprendizado Profundo

O Aprendizado Profundo (*Deep Learning – DL*) é uma técnica caracterizada pela utilização de redes neurais profundas cuja arquitetura envolve múltiplas camadas (de entrada, ocultas e de saída) com capacidade de explorar e representar informações a partir do aprendizado de um conjunto de dados (ZHU et al., 2017). Tais redes consistem de neurônios, nós ou células dispostas e interconectadas (ZAJMI; AHMED; JAHARADAK, 2018). A figura 10 mostra um exemplo de uma rede neural profunda.

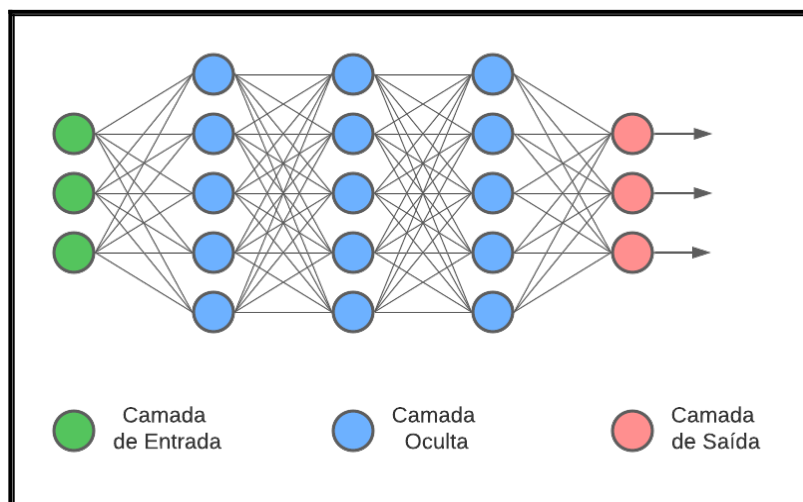


Figura 10: Exemplo de Rede Neural Profunda

Fonte: Próprio autor

O DL é um subconjunto do aprendizado de máquina (*Machine Learning – ML*). O ramo ML tem por finalidade obter uma representação matemática para modelar, a partir de processos de

treinamento, o comportamento de uma função (BOCHIE et al., 2020). O ML, por sua vez, é um subcampo da Inteligência Artificial (IA). A IA é uma área da Ciência da Computação que busca automatizar a aprendizagem para execução de tarefas, aprimorando-se iterativamente a partir de dados e informações coletadas. A figura 11 apresenta o campo da IA e seus subconjuntos ML e DL.

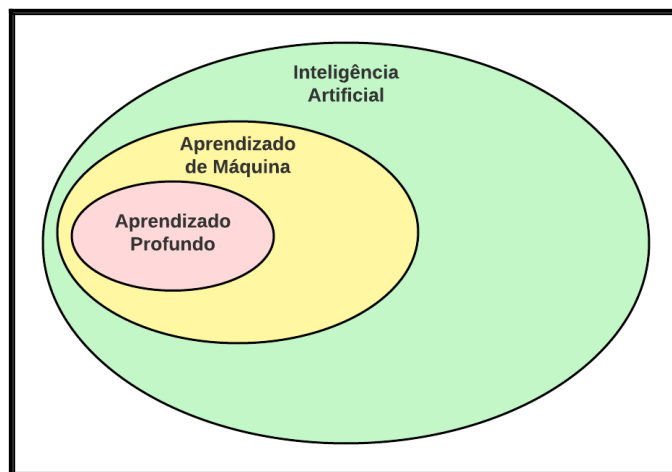


Figura 11: Campo de IA e seus subconjuntos

Fonte: Próprio autor

Conforme exposto no capítulo 2, foi confirmada a percepção intuitiva da notoriedade que o aprendizado profundo vêm alcançando junto à comunidade científica ao longo dos anos. Segundo DENG (2014), os fatores que impulsionam a utilização do aprendizado profundo são o desenvolvimento da capacidade de processamento de sistemas computacionais, a diminuição do custo de *hardwares* e os avanços nos algoritmos de aprendizado de máquina. Para ZHAO et al. (2019), a popularidade dessa técnica cresceu, a partir de 2006, devido a grande capacidade de aprendizado que a mesma demonstra na utilização de dados de treinamento de larga escala, o rápido desenvolvimento de sistemas computacionais de alta performance e a evolução das arquiteturas das redes e das estratégias de treinamento.

No que se refere a aplicação dos modelos de Aprendizado Profundo, o cerne é encontrar, a partir de algoritmos de retropropagação (*backpropagation*), estruturas intrincadas em um determinado conjunto de dados. O intuito é indicar à máquina, a partir das características de uma camada anterior, quais parâmetros internos devem ser alterados para calcular a representação de uma camada em análise. Para esse fim, suas redes neurais hierarquizam as características do dado, de modo que as representações de níveis mais altos sejam formadas pela combinação das representações de níveis mais baixos (LECUN; BENGIO; HINTON, 2015). Para que esse processo ocorra, as redes neurais profundas utilizam-se de diversos componentes, tais como Pesos, Bias e Funções de Ativação, os quais serão trabalhados a seguir.

3.1.1 Componentes de uma rede neural profunda

As redes neurais profundas são compostas por camada de entrada (*input layer*), camada oculta (*hidden layer*) e camada de saída (*output layer*). A camada de entrada e seus nós representam os recursos preditivos. A camada de saída são as previsões de destino e a camada oculta agrega as informações para capturar as interações dos dados de entrada (TORKY; ABURAWWASH, 2018).

Para propagar pela rede, os dados de entrada sofrem transformações lineares e não lineares ao longo de todo percurso. As transformações lineares são realizadas por pesos e bias, conforme apresentado na figura 12

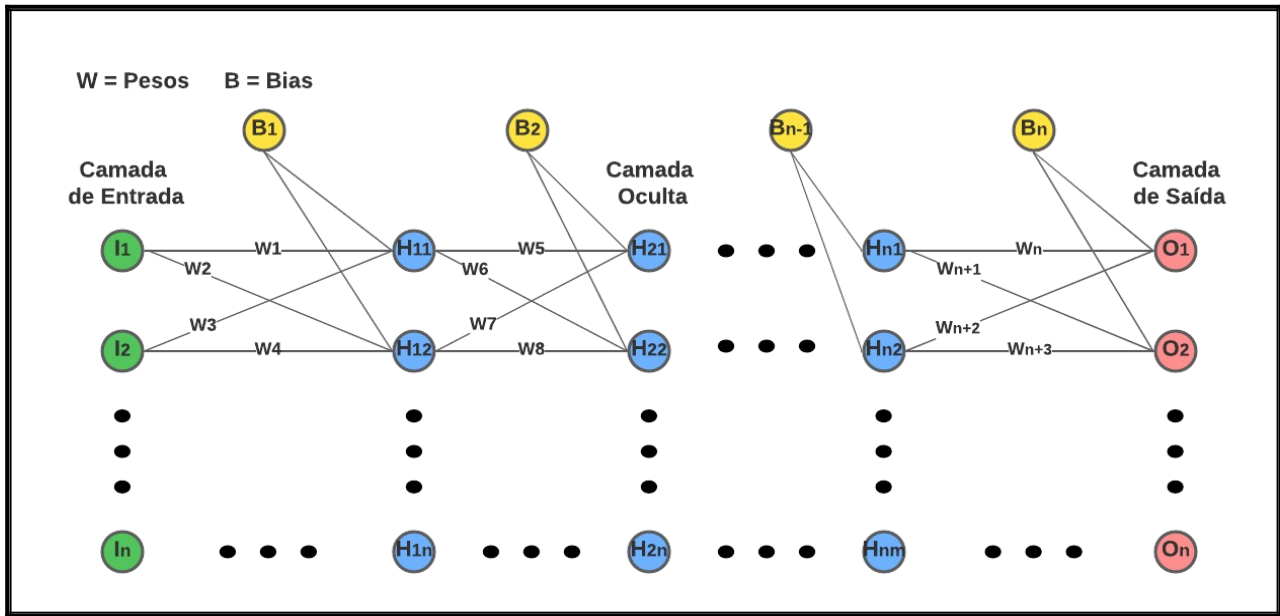


Figura 12: Exemplo de propagação de redes neurais profundas

Fonte: Próprio autor. Adaptado de ABIODUN et al. (2018)

Na figura, “I” representa os nós da camada de entrada, “H” os nós das camadas ocultas, “O” os nós da camada de saída, “B” as bias e “W” os pesos. A equação 1 representa um exemplo de transformação linear que ocorre entre a camada de entrada e a camada oculta, conforme segue:

$$H_h = \sum_{n=1}^n w_n I_n + B_h, \text{ onde } 1 \leq h \leq n \quad (1)$$

Levando em consideração o início de propagação da rede, nas fórmulas apresentadas tem-se que H_h é o valor de entrada do nó da camada oculta, W_n são os pesos utilizados entre as saídas da camada de entrada e o nó considerado da camada oculta, I_n é valor de saída da camada de entrada, B_h é a bia. O raciocínio utilizado na transição da camada de entrada para a primeira camada oculta é análogo às transformações lineares das camadas posteriores.

Para as transformações não lineares, utiliza-se a função de ativação. Esta função é um recurso matemático que introduz não linearidade à rede neural e permite que a mesma expresse uma função mais complexa (LIU et al., 2017). Utilizando a fórmula 1 como base, pode-se representar matematicamente a função de ativação conforme as equações 2 e 3:

$$S_h = Activation(H_h) \quad (2)$$

$$S_h = Activation\left(\sum_h^n w_n I_n + B_h\right) \quad (3)$$

Nas equações, S_h representa o valor de saída da camada oculta considerada.

As funções de ativação mais comuns são *Sigmoid*, *Tanh* e *Rectified Linear Unit* (ReLU). A figura 13 apresenta essas funções e as equações 4, 5 e 6 suas respectivas fórmulas matemáticas.

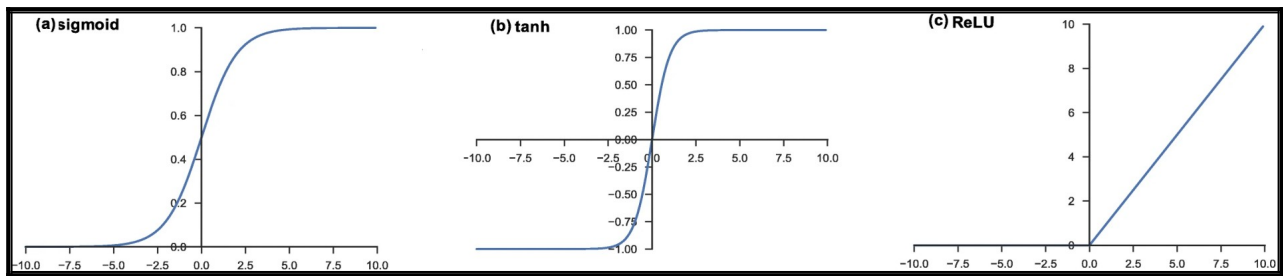


Figura 13: Funções de ativação mais comuns

Fonte: Adaptado de YAMASHITA et al., 2018

$$sigmoid(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (5)$$

$$ReLU(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } x \geq 0 \end{cases} \quad (6)$$

Unindo a equação 1, 2 e 4, tem-se um exemplo de saída da camada oculta. Este exemplo é apresentado nas fórmulas 7 e 8.

$$S_h = Activation(H_h) = \frac{1}{1 + e^{-H_h}} \quad (7)$$

$$S_h = \frac{1}{1 + e^{-\sum_h^n w_n I_n + B_h}} \quad (8)$$

3.1.2 Backpropagation

Backpropagation é o algoritmo utilizado para treinamento de redes neurais profundas. Seu objetivo é otimizar os pesos para que a rede neural aprenda a mapear as entradas para as saídas (DATA SCIENCE ACADEMY, 2022). Para isto, o algoritmo indica como a máquina deve alterar os parâmetros internos de uma camada anterior para calcular a representação da camada vigente (LECUN; BENGIO; HINTON, 2015). O algoritmo de *backpropagation* constitui-se de duas etapas principais: *forward pass* e *backward pass*.

O *forward pass* é o movimento das informações da camada de entrada para as camadas ocultas intermediárias e, posteriormente, para a camada de saída (TORKY; ABURAWWASH, 2018). Nesta etapa, a imagem de entrada é processada com pesos iniciais para calcular uma previsão de saída. Esta previsão é comparada com a saída esperada (*Ground Truth*) e suas compatibilidades são medidas a partir das funções de perda (*loss function*) (TAGHANAKI et al., 2019), como *Cross Entropy*, *Weighted Cross Entropy*, *Focal Loss*, *Dice Loss*, *Boundary Loss*, entre outras.

O *backward pass* é a fase na qual calcula-se o gradiente da função de perda da camada de saída para que o mesmo possa ser aplicado para atualizar os pesos da rede (CHRISTIANSEN et al., 2014). Nesta etapa é utilizada a regra da cadeia (*chain rule*) para atualizar os pesos dos parâmetros iniciais e reiniciar o processo (GUO et al., 2016). A otimização dos pesos é realizada conforme pode-se observar nas equações 7 e 8.

$$W_n = W_a + \Delta W \quad (7)$$

$$\Delta W = -\eta \frac{\delta E(W)}{\delta W} \quad (8)$$

Nestas fórmulas, W_n é o novo peso, W_a é o peso a ser atualizado, E é a função de perda e η é a taxa de aprendizagem. Esta taxa representa o tamanho do passo para que se minimize a função de perda.

3.2 Redes Neurais Convolucionais

LECUN, BENGIO e HINTON (2015) relatam que as CNN são os principais modelos representativos do Aprendizado Profundo.

As Redes Neurais Convolucionais são uma classe de redes neurais especializadas no processamento de dados a partir do emprego de operações matemáticas de convolução (GOODFELLOW; BENGIO; COURVILLE, 2016). Essas redes foram projetadas para, automaticamente e de forma adaptativa, aprenderem hierarquias espaciais de classes por meio da retropropagação. Para isto, as mesmas utilizam múltiplas camadas de construção, como as camadas de convolução (*convolution layers*), as camadas de *pooling* (*pooling layers*) e as camadas totalmente conectadas (*fully connected layers*) (YAMASHITA et al., 2018). A figura 14 apresenta estas camadas.

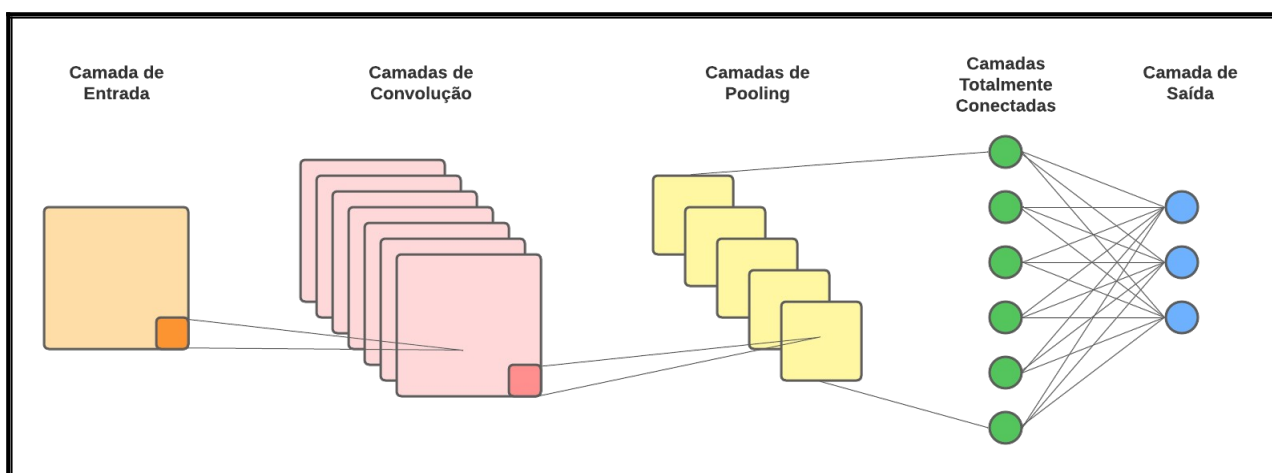


Figura 14: Camadas de Redes Neurais Convolucionais

Fonte: Próprio autor. Adaptado de PHUNG e RHEE (2019)

A camada *convolution* é a responsável por executar a extração de recursos, que consiste em uma combinação de operações lineares (convolução) e não lineares (função de ativação) (YAMASHITA et al., 2018).

A convolução é uma operação matemática na qual uma matriz de números (*kernel*) é aplicada em todas as posições de uma matriz de entrada (*tensor*). O resultado de cada aplicação é associado a uma posição específica em um tensor de saída, gerando mapas de característica (*feature map*) (GUO et al., 2016). Um exemplo desta operação é apresentado na figura 15.

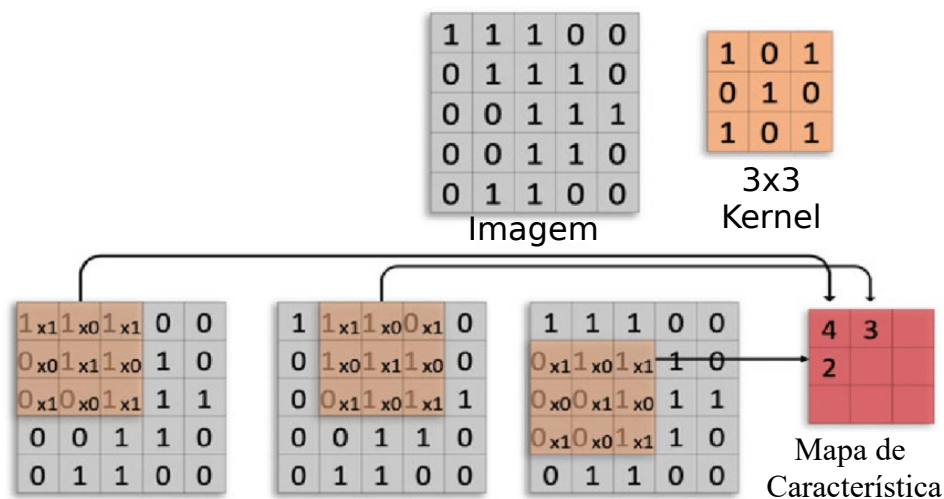


Figura 15: Exemplo de operação matemática de Convolução
Fonte: Adaptado de NG e FENG (2020)

As camadas de *pooling* são utilizadas para reduzir a dimensão dos mapas de característica. Seus objetivos são extrair informações a partir de um contexto espacial mais amplo (SUN; WANG, 2018) utilizando, para isto, a mesclagem de recursos semanticamente semelhantes (LECUN; BENGIO; HINTON, 2015). Desta forma, o *pooling* introduz uma inalterabilidade de translação e promove a diminuição do número de parâmetros a serem aprendidos posteriormente (YAMASHITA et al., 2018). As operações estatísticas de *pooling* mais comuns são as de Máximo (*Max Pooling*), que gera a ativação máxima da região de entrada, e de Média (*Average Pooling*), que gera a ativação a partir da média aritmética. Estas operações são apresentadas na figura 16.

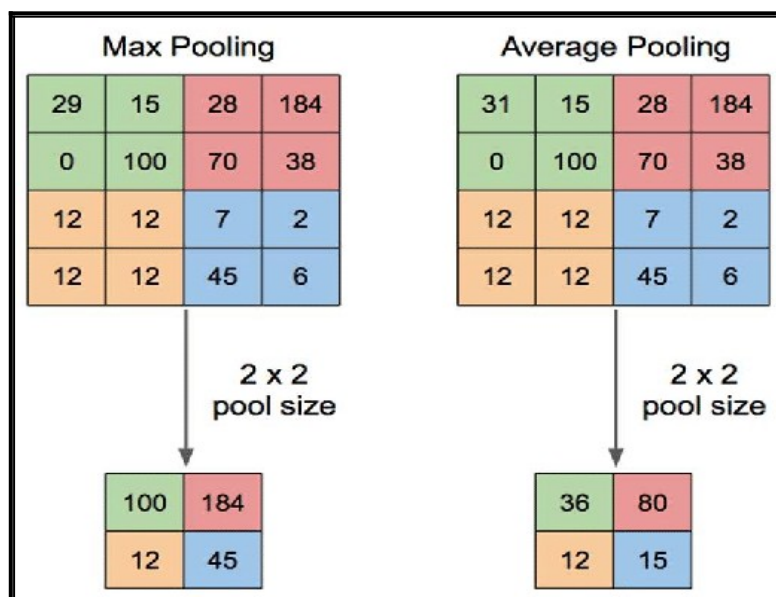


Figura 16: Exemplo de operações estatísticas de Max Pooling e Average Pooling

Fonte: YANI, BUDHI IRAWAN e CASI SETININGSIH (2019)

As *Fully Connected Layers* são camadas localizadas ao final do processo de convolução. Uma vez que os recursos extraídos pelas camadas de convolução e a resolução reduzida pelas camadas de *pooling* são criados, eles são mapeados por um subconjunto de *Fully Connected Layers* (YAMASHITA et al., 2018). Estas camadas utilizam as características extraídas no processo de convolução para categorizá-las em classes, fornecendo suas probabilidades (HOESER; KUENZER, 2020).

Quanto ao processo de aprendizagem das Redes Neurais Convolucionais, o mesmo é dividido em três etapas: Treinamento, Validação e Teste. O treinamento consiste em um processo para encontrar *kernels* (nas *convolution layers*) e pesos (nas *fully connected layers*) que minimizem as diferenças entre as previsões de saídas e os resultados esperados (YAMASHITA et al., 2018). É nesta etapa que ocorrem as fases *forward stage* e *backward stage* (GUO et al., 2016).

A validação é utilizada para avaliar o ajuste do treinamento enquanto ajusta os parâmetros do mesmo. Para seleção do modelo, o processo de avaliação torna-se mais tendencioso conforme o conjunto de dados de validação é incorporado à sua configuração (ELGENDY, 2020).

A etapa de teste consiste na avaliação do modelo final ajustado no conjunto de dados de treinamento e validação (ELGENDY, 2020). A figura 17 retrata as três etapas do processo de aprendizagem das Redes Neurais Convolucionais.

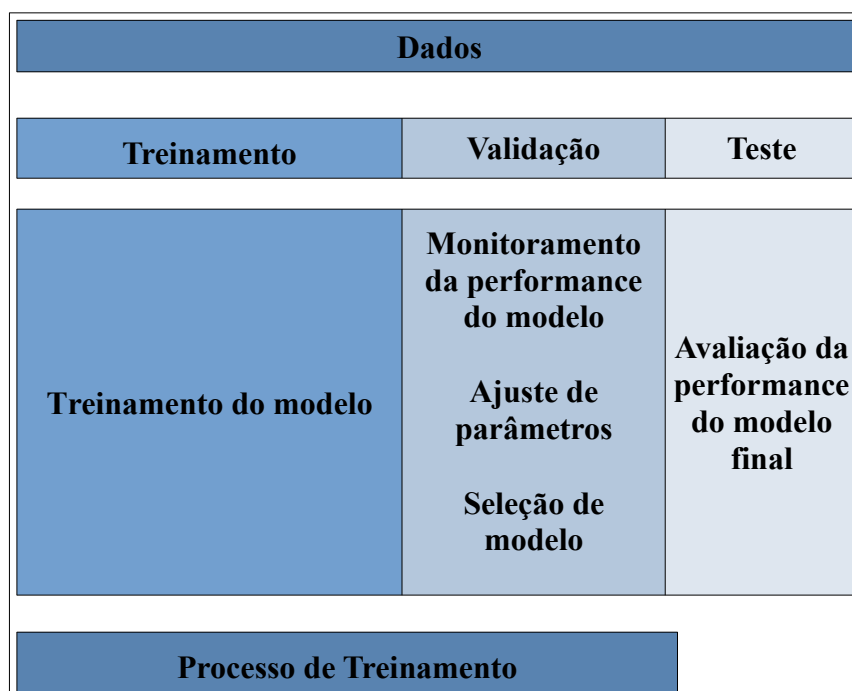


Figura 17: Etapas do processo de aprendizagem das Redes Neurais Convolucionais
Fonte: Próprio autor. Adaptado de YAMASHITA et al. (2018)

3.3 Segmentação Semântica e Detecção de Mudança

3.3.1 Segmentação Semântica

A segmentação semântica tem por finalidade a divisão da imagem em classes semânticas significativas, onde o menor segmento criado pode ser do tamanho de um pixel (KEMKER; SALVAGGIO; KANAN, 2018). Nesta técnica, os algoritmos utilizados são formulados por processos de clusterização de dados (ILEA; WHELAN, 2011), de modo que os mesmos se agrupem conforme aprendizado da rede. A figura 18 apresenta um exemplo de segmentação semântica.

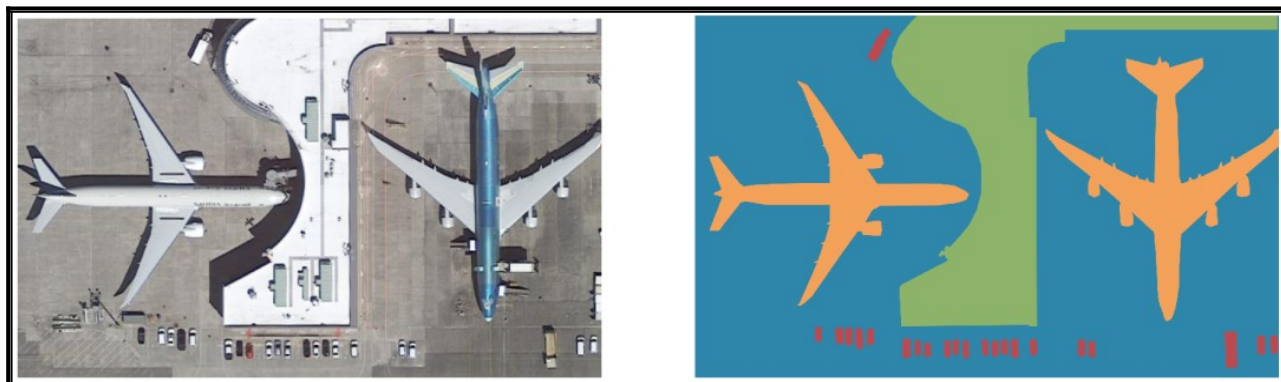


Figura 18: Exemplo de segmentação semântica

Fonte: Adaptado de HOESER e KUENZER (2020)

Para GUO et al. (2016), a principal razão de sucesso da segmentação semântica advém dos modelos de redes neurais convolucionais, uma vez que essa técnica possui a capacidade de lidar com as previsões em nível de pixel, a partir de redes pré treinadas, em conjuntos de dados de grande volume.

Dentre os principais benefícios dessa técnica, GUO et al. (2018) destaca a melhora da eficiência computacional, a eliminação de ruídos de fundo, fornecimento de *insights* teóricos sobre funcionamento de sistemas visuais e a abrangência da técnica quando comparada a detecção e reconhecimento de objetos.

Em termos de aplicação, pode-se incluir direção autônoma, interação homem-máquina, fotografia computacional, mecanismos de busca de imagens, realidade aumentada (GARCIA-GARCIA et al., 2018), transporte, cobertura e uso da terra, agricultura, vegetação natural, entre outros (HOESER; BACHOFER; KUENZER, 2020).

3.3.2 Detecção de Mudança

A detecção de mudança é o processo de verificar e identificar as diferenças entre objetos e fenômenos em períodos de tempo distintos (SHI et al., 2020). Seu principal objetivo é detectar o conjunto de pixels que são “significativamente diferentes” entre a última imagem da sequência e as imagens anteriores (RADKE et al., 2005). A figura 19 apresenta um exemplo de detecção de mudança.

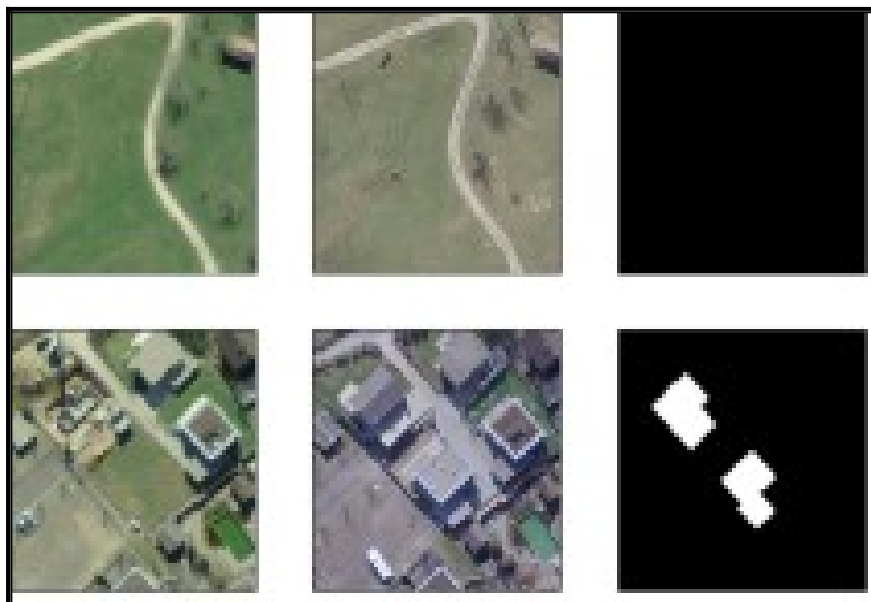


Figura 19: Exemplo de segmentação semântica

Fonte: Adaptado de RŮŽIČKA et al. (2020)

Como resultado da detecção de mudança, pode-se encontrar o conjunto de diferenças entre os dados temporais analisados, tais como aparecimento ou desaparecimento de objetos, movimento de objetos em relação ao fundo e mudanças da forma de objetos (RADKE et al., 2005).

Em termos de aplicação, pode-se citar análise de mudança de uso da terra, monitoramento de cultivo itinerante, avaliação de desmatamento, avaliação de danos, monitoramento de desastres, medições de derretimento de neve, entre outras (SINGH, 1989).

Com relação as estruturas da detecção de mudança, as mesmas podem ser divididas em três categorias: Fluxo Único, Fluxo Duplo e Modelo Multi Integrado.

Tendo em vista que a detecção de mudança se baseia na comparação de dois períodos diferentes, é usualmente utilizada a Estrutura de Fluxo Duplo (SHI et al., 2020). Esta estrutura pode ser categorizada em três tipos: Estrutura Siamesa, Estrutura Baseada em Transferência de Aprendizado e Estrutura de Pós Classificação. A figura 20 esquematiza essas três estruturas.

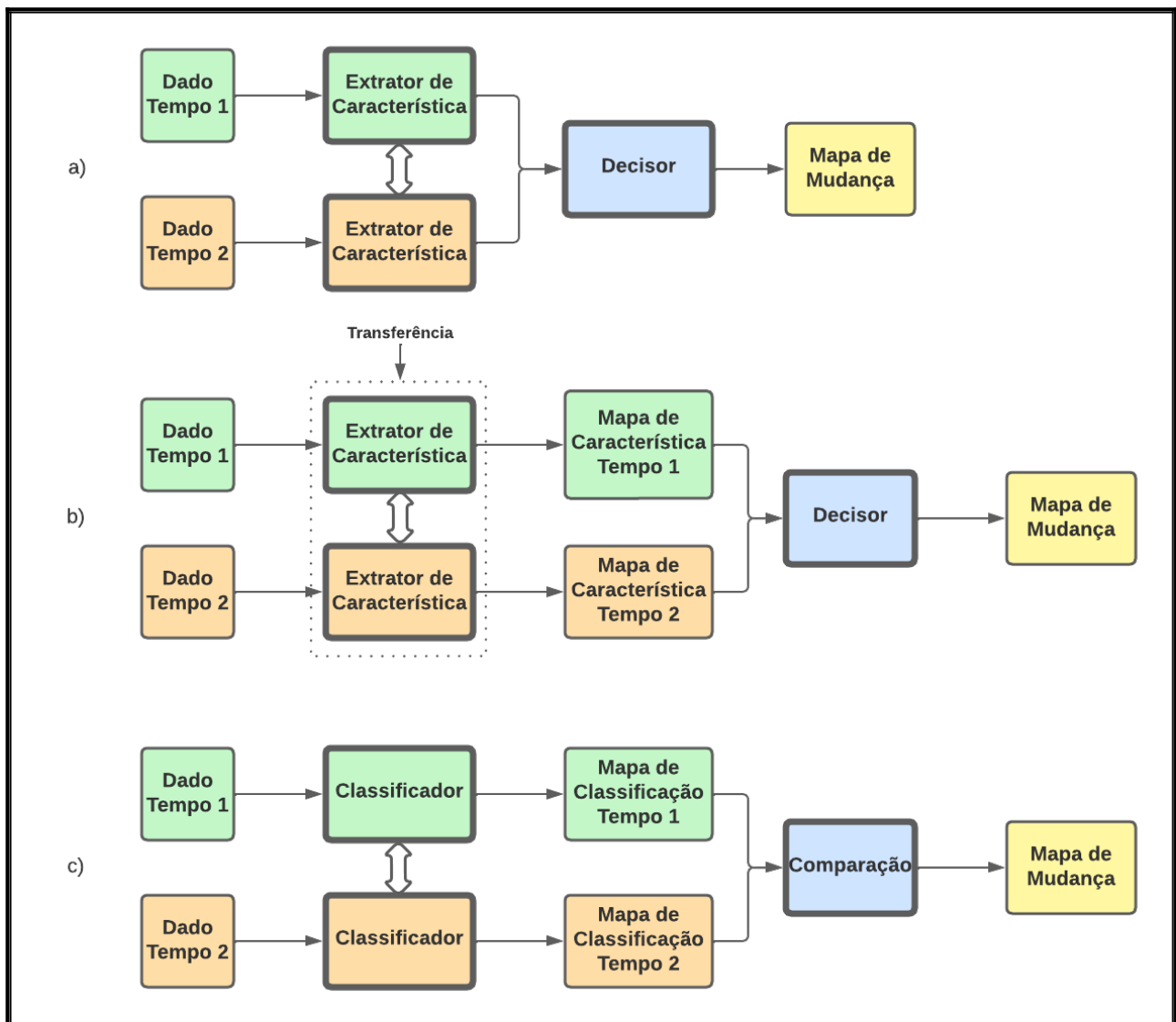


Figura 20: a) estrutura siamesa; b) estrutura baseada em transferência de aprendizado; c) estrutura de pós-classificação

Fonte: Próprio autor. Adaptado de SHI et al. (2020)

A estrutura siamesa é formada por duas sub-redes de estruturas idênticas, ou seja, com a mesma configuração e com os mesmos parâmetros e pesos. Estas sub-redes são treinadas simultaneamente e a atualização dos parâmetros é realizado de modo espelhado em ambas. Nesta estrutura, após extração das características, os decisores realizam análises de mudança para obter o Mapa de Mudança (SHI et al., 2020)

A estrutura baseada em transferência utiliza-se do treinamento em um domínio para obter melhores resultados em um outro domínio. Para isto, as características de nível inferior aprendidas na origem são transferidas para o novo domínio (BALL; ANDERSON; CHAN, 2017). Neste tipo de estrutura são utilizados extratores de característica, os quais podem ser os mesmos, para gerar mapas de características. A partir destes, os decisores analisam as mudanças ocorridas.

A estrutura de pós classificação consiste em obter a classificação das imagens individualmente para depois comparar os resultados obtidos. Neste processo, os classificadores podem ser treinados de forma conjunta ou não. Nesta estrutura, o desempenho do modelo treinado influencia diretamente o resultado da detecção de mudança (SHI et al., 2020).

Conforme apresentado no capítulo 2, tanto para segmentação semântica quanto para detecção de mudança, diversas arquiteturas podem ser utilizadas. Dentre as mais recorrentes, pode-se citar: U-Net (RONNEBERGER; FISCHER; BROX, 2015), U-Net++ (ZHOU et al., 2018), SegNet (BADRINARAYANAN; HANDA; CIPOLLA, 2015), LinkNet (CHAURASIA; CULURCIELLO, 2017), VGG (SIMONYAN; ZISSERMAN, 2015), entre outras.

Essas arquiteturas, em termos gerais, são formadas por redes neurais totalmente convolucionais (*fully convolutional neural network*) e podem utilizar de dados de outras redes previamente treinadas. Estas redes pré treinadas, com milhares de parâmetros extraídos de diferentes modelos de DL, são denominadas *Backbones* (ELHARROUSS et al., 2022).

3.4 Métricas de Avaliação

As métricas de avaliação de ML são utilizadas para qualificar os resultados performados por um algoritmo. Podem ser classificadas em Métricas de Regressão, como erro médio quadrático e R^2 , ou Métricas de Classificação, como acurácia e *Intersection over Union*).

Para segmentação semântica, especificamente, os algoritmos de DL normalmente utilizam as Métricas de Classificação para avaliar os resultados alcançados por suas redes neurais durante as fases de treinamento e teste. Segundo MINAEE et al. (2020), as métricas mais utilizadas para avaliar a performance de modelos de segmentação são: *Pixel Accuracy*, *Mean Pixel Accuracy* (MPA), *Intersection over Union* (IoU), *F1 score* e *Dice coefficient*.

As técnicas apresentadas por MINAEE et al. (2020) são baseadas no conceito de Matriz de Confusão. Esta matriz é uma ferramenta para avaliação de desempenho, uma vez que a mesma quantifica a sobreposição de uma classificação descrevendo o desempenho do modelo (HEYDARIAN; DOYLE; SAMAVI, 2022).

Para classificações binárias, a Matriz de Confusão apresenta em cada linha o *Ground Truth* e em cada coluna a predição do algoritmo. Cada elemento da matriz é definido pela comparação dos dados de entrada e de saída. Para entrada e saída verdadeiras e iguais, dá-se o nome de *True Positive* (TP). Para entrada e saída falsas e iguais, dá-se o nome de *True Negative* (TN). Para entradas e

saídas diferentes, dão-se os nomes *False Positive* (FP) e *False Negative* (FN). A tabela 4 apresenta um exemplo da Matriz de Confusão.

	Predição – Verdade	Predição – Falso
Ground Truth – Verdade	TP	FP
Ground Truth – Falso	FN	TN

Tabela 4: Matriz de Confusão para dados binários

Fonte: Próprio autor

TP significa que as predições verdadeiras alcançadas pela rede neural foram corretas (dados do *ground truth* = dados de predição = 1). TN significa que as predições falsas também foram alcançadas corretamente pelo algoritmo (dados do *ground truth* = dados de predição = 0). FN significa que a rede previu incorretamente as classes positivas e FP significa que a rede previu incorretamente as classes negativas.

Isto posto, pode-se calcular os valores das métricas de avaliação a partir da matriz apresentada na tabela 3.1. As fórmulas 9 a 11, apresentam, respectivamente, as métricas *Pixel Accuracy*, *F1 score* e *Intersection over Union* (IoU).

$$Pixel\ Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (09)$$

$$F1\ score = 2 \times \frac{(precision \times recall)}{(precision + recall)} = 2 \times \frac{(\frac{TP}{TP+FP}) * (\frac{TP}{TP+FN})}{(\frac{TP}{TP+FP}) + (\frac{TP}{TP+FN})} \quad (10)$$

$$IoU = \frac{TP}{(TP+FP+FN)} \quad (11)$$

Pixel Accuracy é a métrica que avalia os pixels que foram classificados corretamente, a partir de uma comparação direta entre os dados preditos e os dados do *Ground Truth*. Tendo em vista que esta comparação é realizada tanto entre o objeto predito quanto o *background* da imagem, pode ocorrer um desbalanceamento de classe, haja vista que, normalmente, o fundo possui quantidades de pixels expressivamente maiores. Esse desbalanceamento pode ocasionar a falsa impressão da qualidade do modelo.

F1 score é a combinação das métricas de precisão e *recall* (SINGH; RANI, 2020). Precisão é a medida de qualidade das predições e é obtida pela divisão dos TP com a soma dos FP e TP. Nesta

métrica, quantos menos previsões forem feitas, menos erros serão cometidos, uma vez que os dados de TN e FN não são levados em consideração. *Recall* é a métrica que mede a quantidade de elementos classificados como positivos entre o todas as predições tidas como verdadeiras, ou seja, é a divisão de TP com a soma dos TP e FN. Nesta métrica as predições tidas como falsas (FP e TN) não são consideradas. A figura 21 esquematiza essas duas métricas.

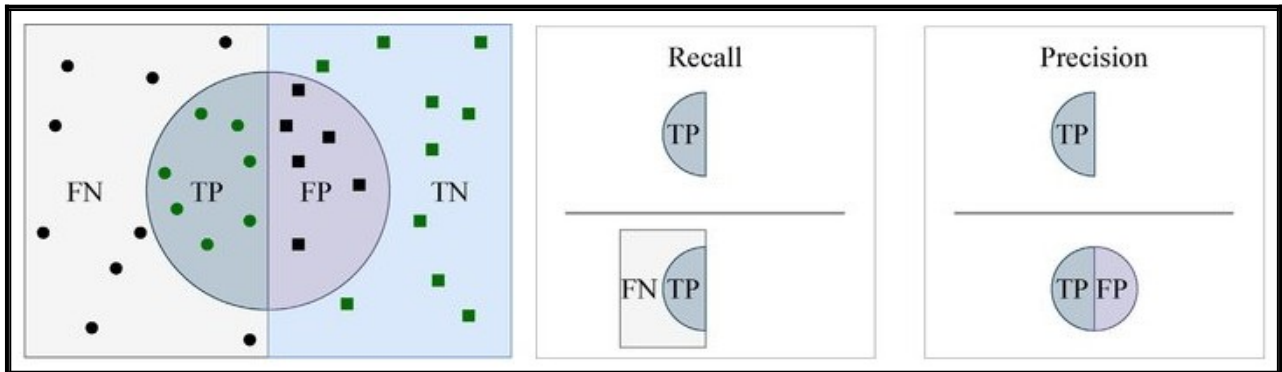


Figura 21: Esquematização das métricas precisão e recall

Fonte: adaptado de MALEKI et al. (2020)

No contexto apresentado, o *F1 score* utiliza-se de um balanceamento da precisão e do *recall* a partir de uma média harmônica entre ambos, obtendo, desta forma, as qualidades apresentadas por cada uma delas.

IoU é a métrica de avaliação que compara a área obtida pela interseção dos dados preditos e do *Ground Truth* com a área obtida pela união entre ambos, conforme apresentado na figura 22.

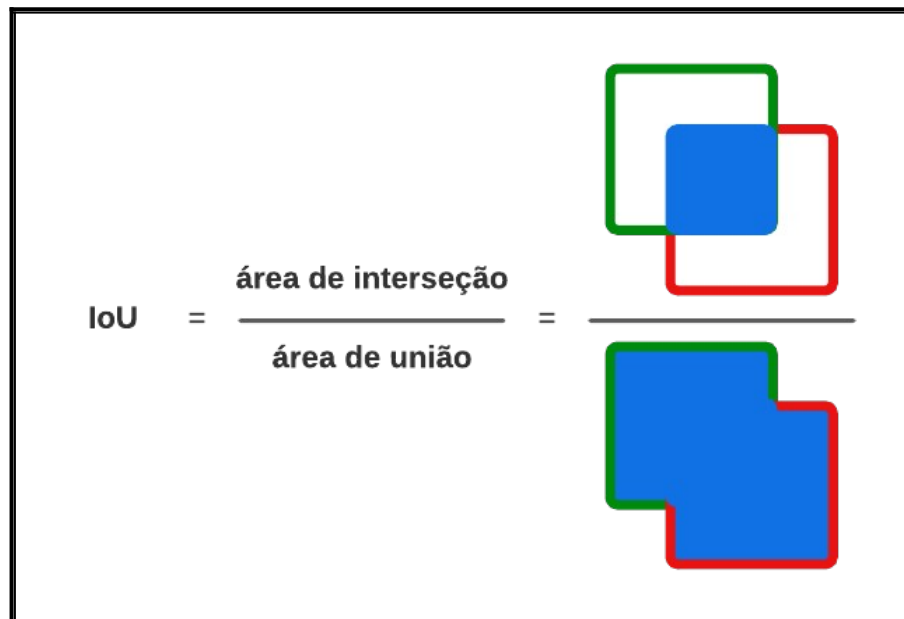


Figura 22: Métrica Intersection over Union (IoU)

Fonte: PADILLA, NETTO e DA SILVA (2020)

Nessa métrica os valores variam entre 0 e 1, sendo o valor igual a 1 considerado uma sobreposição perfeita da segmentação. Ademais, a IoU tende a medir a performance para os piores cenários, penalizando instâncias únicas de má classificação. Em comparação, por exemplo, com o *F1 score*, que busca uma avaliação do desempenho médio, o IoU propende a ter resultados com pior avaliação. Ainda assim, segundo MINAE et al. (2020), essa métrica é a mais comumente utilizada para segmentação semântica. A equação 12 apresenta a fórmula geral da IoU.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (12)$$

Conforme apresentado, a métrica IoU é a mais utilizada para verificar o desempenho de uma segmentação semântica, porém, existem trabalhos que priorizam a utilização da F1 Score, como, por exemplo, os artigos “Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks” (KAMPFFMEYER; SALBERG; JENSSEN, 2016), “DenseU-Net-Based Semantic Segmentation of Small Objects in Urban Remote Sensing Images” (DONG; PAN; LI, 2019), “Semantic Segmentation of Crop Type in Africa: A Novel Dataset and Analysis of Deep Learning Methods” (M RUSTOWICZ et al., 2019), “Semantic Segmentation Using Deep Learning with Vegetation Indices for Rice Lodging Identification in Multi-date UAV Visible Images” (YANG et al., 2020), entre outros. Isto posto, verifica-se que a métrica a ser utilizada para avaliação dos modelos, em geral, depende da arbitrariedade do pesquisador.

3.5 Hiperparâmetros

Hiperparâmetros são atributos que impactam diretamente na performance do modelo de aprendizado de máquina (YANG; SHAMI, 2020). Enquanto os parâmetros são adaptados durante a fase de treinamento, os hiperparâmetros são especificados antes da fase de aprendizagem e influenciam na complexidade, no comportamento, na velocidade e no aprendizado dos modelos (BISCHL et al., 2021). São exemplos de hiperparâmetros os otimizadores, a taxa de aprendizagem, o tamanho do *batch* e o número de épocas de treinamento.

Os otimizadores tem por objetivo diminuir o erro entre os resultados obtidos por uma rede em comparação com os resultados desejados (SILVA, 2018). Dentre os diversos métodos de otimização de redes neurais, pode-se citar, como exemplo, Adam (KINGMA; BA, 2017), SGD com *momentum* (SUTSKEVER et al., 2013), Adagrad (DUCHI; HAZAN; SINGER, 2011) e Adabound (LUO et al., 2019).

A taxa de aprendizagem é o hiperparâmetro utilizado para determinar o tamanho do passo de aprendizado da rede, ou seja, ela determina a velocidade em que os pesos se ajustarão aos pesos ideais. Uma taxa de aprendizado baixa torna o aprendizado da rede lento. O oposto, ou seja, uma taxa muito elevada, pode impedir a convergência do processo de aprendizado (APAYDIN et al., 2020). A figura 3.14 apresenta um exemplo de mudanças na função de perda, no decorrer das épocas de treinamento, com base na taxa de aprendizagem escolhida inicialmente.

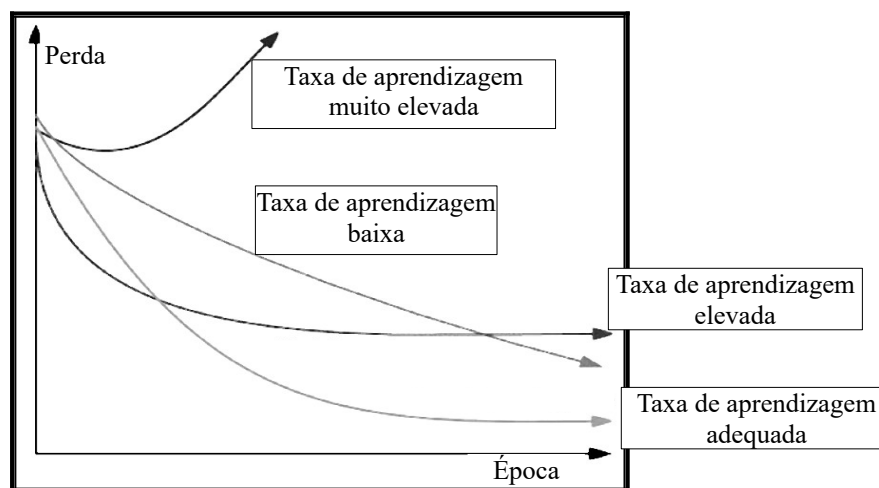


Figura 23: Influência da escolha da taxa de aprendizagem na função de perda

Fonte: Próprio autor. Adaptado de APAYDIN et al. (2020).

O tamanho do *batch* define o número de exemplos de treinamento que serão utilizados durante uma iteração da rede. Esse hiperparâmetro pode ser utilizado em três opções: *Batch Mode*, no qual o tamanho do lote é igual ao conjunto de dados total; *Mini-Batch Mode*: no qual o tamanho do lote é maior que um e menor que o tamanho total do conjunto de dados; e *Stochastic Mode*: no qual o tamanho do lote é igual a um. Tamanhos de *batch* menores proporcionam treinamentos mais rápidos e requerem menos memória, porém a estimativa do gradiente será menos precisa. Tamanhos de *batch* grandes garantem a convergência para o treinamento, entretanto o custo computacional é elevado e há a possibilidade de degradar o desempenho da generalização (GAO; ZHONG, 2020).

O número de épocas de treinamento determina a quantidade de vezes que o conjunto de dados completo será treinado. É ao final de cada época de treinamento que os parâmetros dos modelos são atualizados (BROWNLEE, 2018). Isto posto, poucas épocas de treinamento podem ocasionar o *underfitting*, que é uma aprendizagem aquém da capacidade de treinamento e que não se adaptou aos dados treinados, ao passo que muitas épocas podem proporcionar o *overfitting*, que é o excesso de aprendizagem, na qual a rede se adapta perfeitamente aos dados treinados, porém não generaliza para novos dados.

3.6 Data Augmentation

Conforme apresentado na seção 3.2, as redes neurais convolucionais são modelos hierárquicos de aprendizagem com vários níveis de representações, onde os níveis mais altos capturam conceitos mais abstratos (TAYLOR; NITSCHKE, 2017). Neste contexto, é de conhecimento comum que quanto maior a quantidade de dados que um algoritmo tem acesso, mais eficaz ele pode ser (PEREZ; WANG, 2017). Nesta perspectiva, um dos problemas enfrentados pelas redes neurais é a dificuldade de coletar dados de qualidade e que possuam equilíbrio entre suas classes (MIKOŁAJCZYK; GROCHOWSKI, 2018).

Para resolver este problema, uma ferramenta preponderante é o *Data Augmentation*. Esta técnica é utilizada para, a partir de transformações que preservam as características fundamentais do dado, inflar artificialmente um conjunto de dados (TAYLOR; NITSCHKE, 2017).

Dentre os diversos procedimentos que se pode adotar para aumentar o conjunto de dados, pode-se citar: *crop*, *rotate*, *transpose*, *brightness*, entre outros. A figura 24 apresenta um exemplo de aplicação de *Data Augmentation*, onde a imagem a) da figura é a original, e as imagens de b) a f) são derivações da mesma.

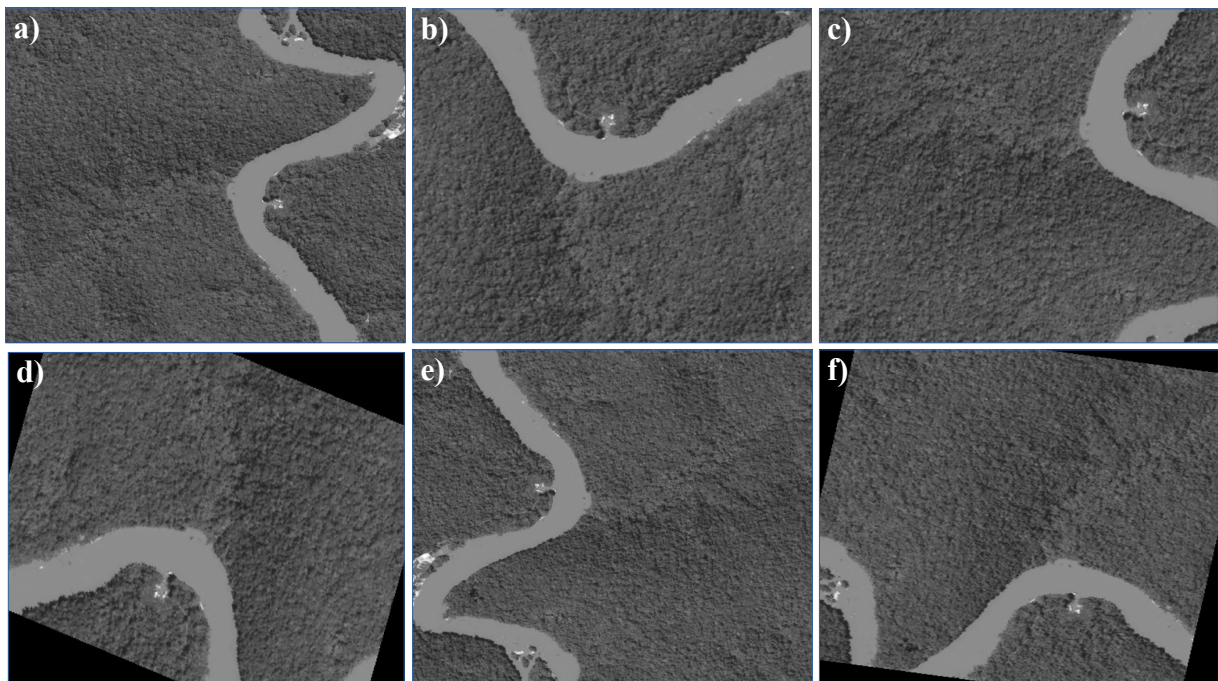


Figura 24: Exemplo de *Data Augmentation*

Fonte: Próprio autor

3.7 Métodos de Comparação de Imagens

A comparação de imagens é um dos processos essenciais na área de processamento de imagens. A estimativa de similaridade e dissimilaridade são fundamentais para comparar algoritmos dedicados à redução de ruído, imagens correspondência, codificação de imagem e restauração (KATUKAM, 2015). Dentre os variados métodos de comparação de imagens, pode-se citar o Erro Quadrático Médio (EQM) e o Índice de Similaridade Estrutural (ISE). Ambos métodos buscam quantificar a similaridade entre imagens a partir da comparação de seus pixels.

O EQM é considerado o mais comum estimador de qualidade de imagens (SARA; AKTER; UDDIN, 2019) e calcula a soma das diferenças ao quadrado de cada pixel das imagens, dividido pelo tamanho das mesmas. Neste método, quanto mais próximo a zero for o resultado alcançado, mais similares são as imagens. A equação do EQM é apresentada na fórmula 13.

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [f(i, j) - g(i, j)]^2 \quad (13)$$

Na equação, m representa o número de linhas de pixels das imagens, n o número de colunas de pixels, i o índice da linha, j o índice da coluna, f representa os dados da matriz da imagem original e g os dados da matriz da imagem a ser comparada.

O ISE é um método proposto por (WANG et al., 2004) e busca analisar padrões locais de intensidade de pixels a partir da comparação da luminância, do contraste e da estrutura das imagens. Seus resultados variam entre 0 e 1, sendo 1 o valor alcançado para imagens iguais. As equações do ISE são apresentadas nas fórmulas 14 a 18.

$$SSIM(x, y) = l(x, y) * c(x, y) * s(x, y) \quad (14)$$

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x + \mu_y + C_1} \quad (15)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (16)$$

$$s(x, y) = \frac{2\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (17)$$

$$SSIM(x, y) = \frac{(2\mu_x2\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (18)$$

Nestas fórmulas, $l(x,y)$ é a função de luminância, $c(x,y)$ a de contraste e $s(x,y)$ a de estrutura. μ_x e μ_y são a média de brilho, σ_x e σ_y são os desvios padrão do contraste, σ_{xy} é a covariância e os C são constantes incluídas para evitar instabilidade quando os denominadores se aproximam de zero.

Em uma comparação sucinta entre esses dois métodos, (GANDHI; KULKARNI, 2013) cita que na utilização do EQM algumas distorções não são claramente visíveis e outras, que estão presentes nas imagens, não mudam a qualidade das mesmas. Isto ocorre pois este método não se correlaciona com a percepção visual humana, na qual o erro está relacionado com a perda de qualidade. Em contrapartida, o ISE visa melhorar a eficácia das avaliações de qualidade das imagens conforme o Sistema Visual Humano, a partir das semelhanças existentes entre as mesmas (WANG et al., 2017b).

4 METODOLOGIA

O objetivo desta seção é apresentar as áreas de estudo utilizadas, as etapas de que foram seguidas até alcançar os resultados finais, os Bancos de Dados elaborados, os processamentos de dados realizados, os processos para definições de arquiteturas e *backbones*, bem como os procedimentos realizados para estabelecer o modelo final do trabalho.

4.1 Áreas de Estudo

O projeto utilizará as regiões de Tapajós (área aproximada de 20.000km²) e Kayapó (área aproximada de 2.000 km²), no estado do Pará, uma vez que estas são duas das maiores e mais antigas áreas de desmatamento provocado pela atividade garimpeira na amazônia brasileira.

As etapas treinamento, validação e teste das redes neurais serão distribuídas para a área de Tapajós, tendo em vista o tamanho da área e a quantidade de garimpos existentes na mesma. A área de Kayapó será utilizada para verificar a eficácia do algoritmo em regiões diferentes da de treinamento. As figuras 25 e 26 apresentam a localização e a área garimpeira da região do Tapajós e de Kayapó, respectivamente.

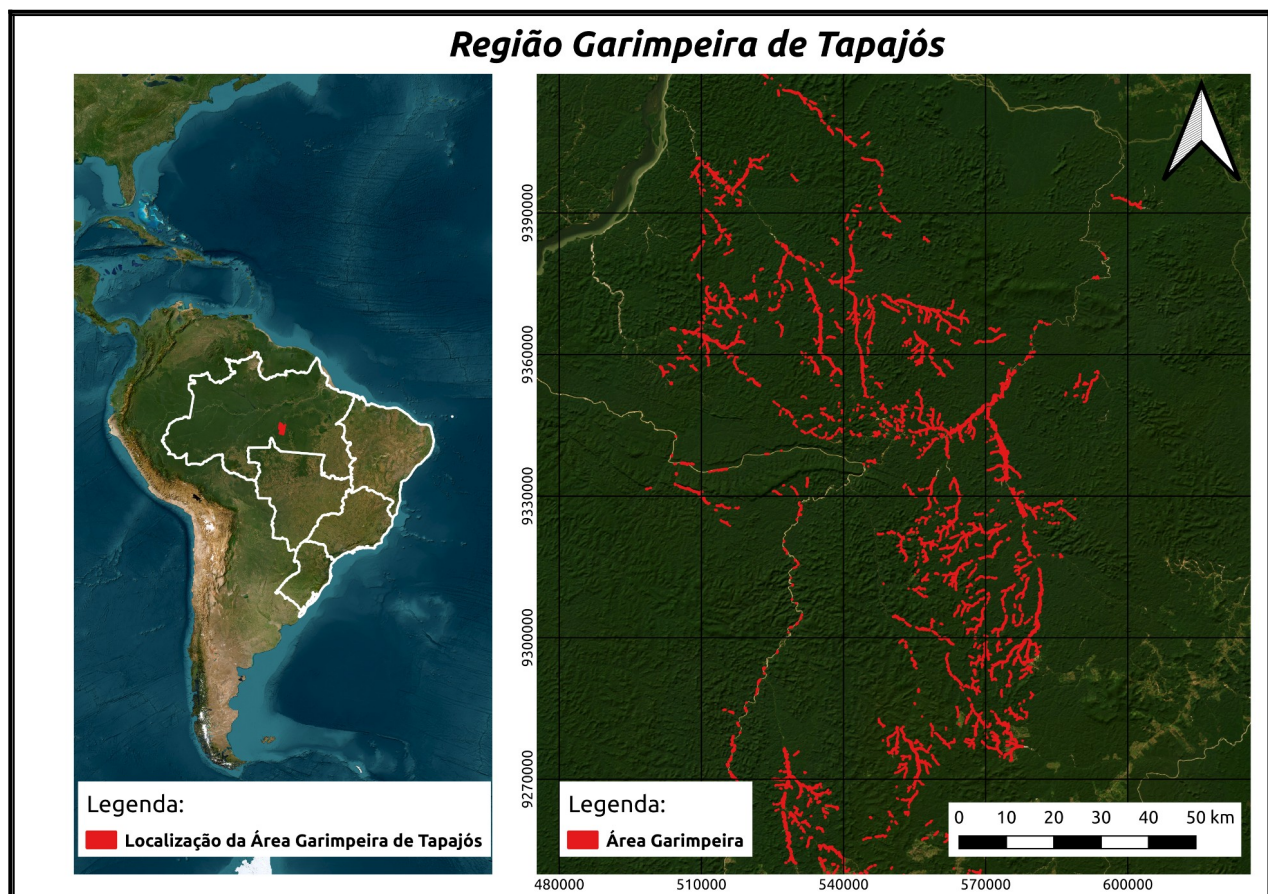


Figura 25: Área de estudos de Tapajós, na Amazônia Brasileira

Fonte: Próprio autor

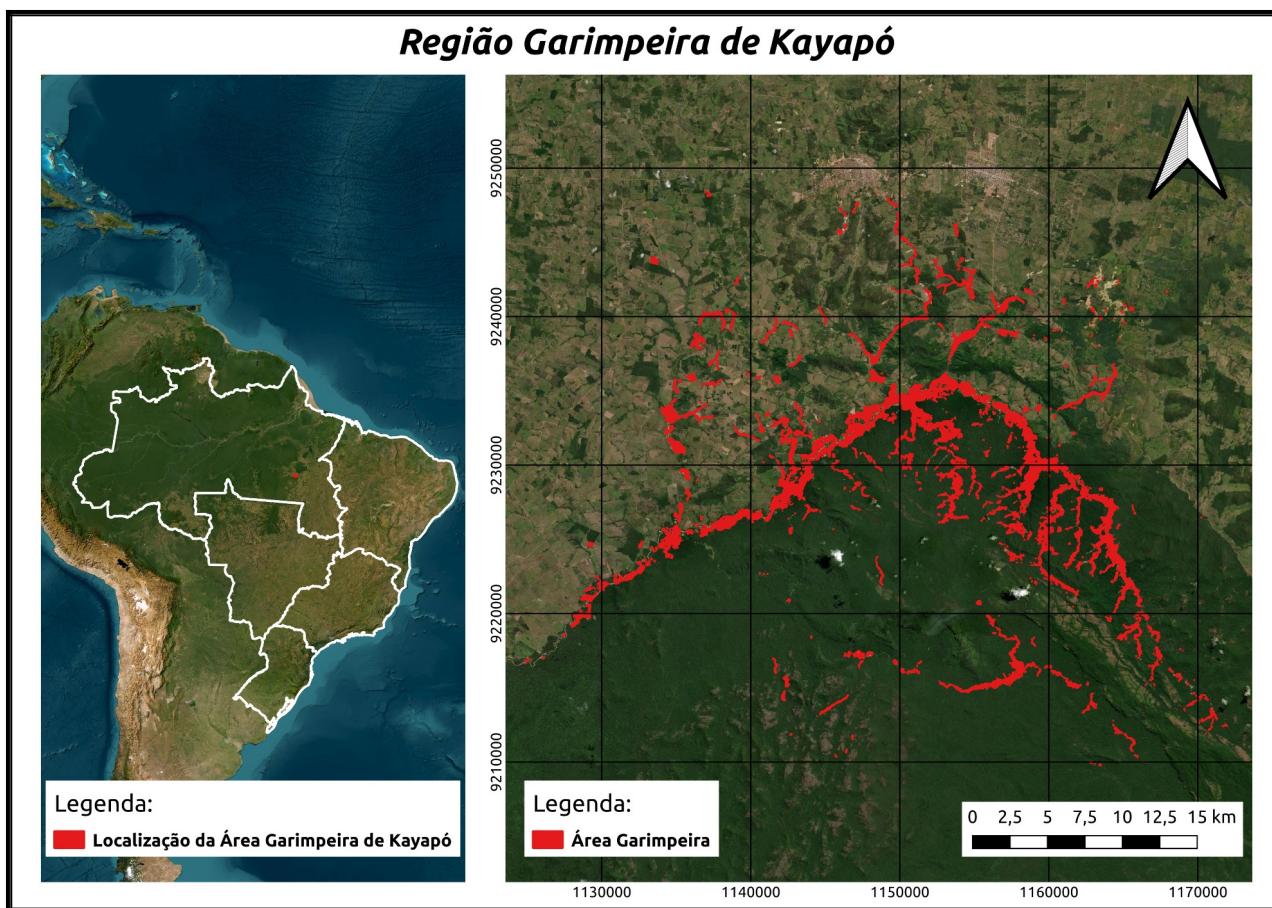


Figura 26: Área de estudos de Kayapó, na Amazônia Brasileira

Fonte: Próprio autor

4.2 Etapas do Projeto

A seguir serão apresentadas as etapas realizadas neste projeto. Ao final desta seção será apresentado um fluxograma geral.

1. Revisão bibliográfica acerca dos temas Segmentação Semântica, Detecção de Mudança e Mineração Aurífera aplicados ao Sensoriamento Remoto.
2. Definição das áreas de estudo.
3. Criação do Banco de Dados do projeto, composto por imagens dos satélites CBERS 4A, dos anos de 2020 e 2021, e por polígonos de áreas garimpeiras.
4. Estudo e análise dos principais conceitos relacionados à Redes Neurais Convolucionais.
5. Estudo de arquiteturas de CNNs voltadas para o emprego em segmentação semântica e detecção de mudança.
6. Aplicação de estruturas de redes neurais capazes de detectar os desmatamentos provocados por garimpos e suas mudanças ao longo do tempo.

7. Teste da rede neural com melhor desempenho nas etapas de treinamento e validação.
8. Apresentação dos resultados:
 - Resultado 1: detecção do garimpo.
 - Resultado 2: mudança da área garimpeira com o passar do tempo.

A figura 27 apresenta o fluxograma geral do projeto.

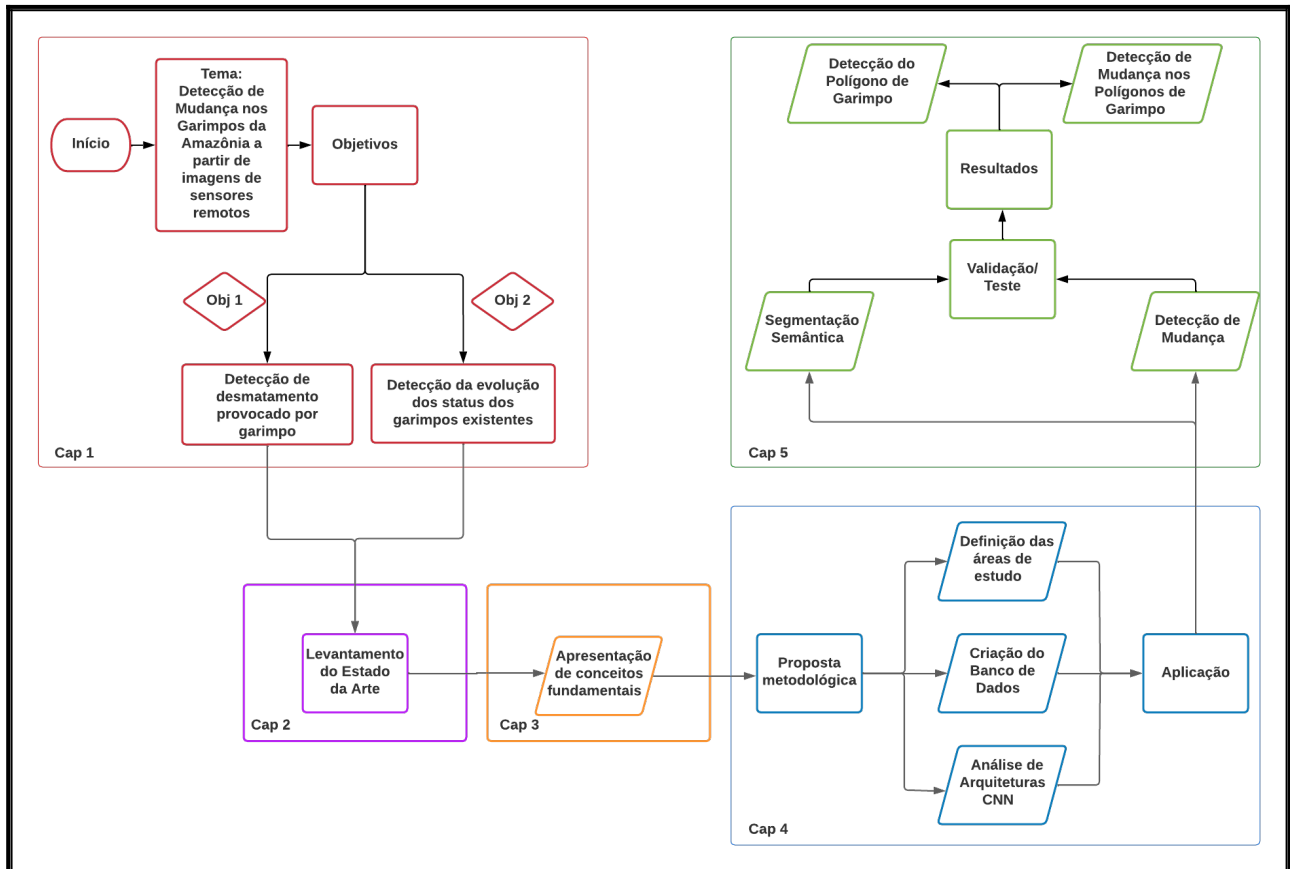


Figura 27: Fluxograma do projeto
Fonte: Próprio autor

4.3 Banco de dados

O projeto foi composto por três bancos de dados: Banco de Dados de Estudo (BDE), Banco de Dados de Aplicação (BDA) e Banco de Dados de Teste (BDT).

O BDE é composto por imagens e máscaras de áreas de, aproximadamente, 5km², retiradas de uma única cena do satélite CBERS 04A, da região de Tapajós, referente ao ano de 2020, e submetidas a processos de *data augmentation*. Este banco foi utilizado para analisar e verificar os modelos e os parâmetros mais adequados para atender o objetivo proposto, uma vez que o mesmo possui uma quantidade menor de dados.

O BDA é formado por imagens e máscaras retiradas de duas cenas do satélite CBERS 04A, da região de Tapajós, referente ao ano de 2020, e submetidas a processos de *data augmentation*. Este banco foi utilizado para avaliar o desempenho do melhor modelo testado no BDE e para determinar os polígonos referentes as áreas garimpeiras de Tapajós.

O BDT é formado por imagens retiradas de duas cenas do satélite CBERS 04A, sendo uma atinente às áreas garimpeiras da região de Kayapó, em 2020, e a outra referente as áreas de garimpo na região de Tapajós, em 2021. Este banco foi utilizado para testar o modelo adquirido em uma região diferente da qual foi realizado os treinamentos e para detectar a mudança dos garimpos na região de Tapajós. A tabela 5 apresenta esses três bancos:

Banco	Região	Ano	Dados	Quantidade de insumos (Nr imagens 5km ²)
BDE	Tapajós	2020	Imagens retiradas de uma única cena do satélite CBERS 04A, após <i>data augmentation</i>	13200
			Máscaras das imagens	13200
BDA	Tapajós	2020	Imagens retiradas de duas cenas do satélite CBERS 04A, após <i>data augmentation</i>	38000
			Máscaras das Imagens	38000
BDT	Tapajós	2021	Imagens retiradas de uma única cena do satélite CBERS 04A	438
	Kayapó	2020	Imagens retiradas de uma única cena do satélite CBERS 04A	30
			Máscaras das Imagens	30

Tabela 5: Composição dos bancos de dados

Fonte: Próprio autor

4.3.1 Satélite CBERS 04A

O CBERS 04A é o satélite mais moderno desenvolvido a partir do Programa CBERS. Este programa é um parceria entre Brasil e China, iniciada em 1988, envolvendo o Instituto Nacional de Pesquisas Espaciais (INPE) e a Academia Chinesa de Tecnologia Espacial (CAST). Seu objetivo é desenvolver um programa de construção de satélites de sensoriamento remoto. O programa já foi responsável pelo desenvolvimento dos satélites CBERS-1, 2, 2B, 3 e 4, além do 04A.

O CBERS 04A foi lançado em dezembro de 2019 e é equipado com Câmera Multiespectral e Pancromática de Ampla Varredura (WPM), Câmera Multiespectral (MUX), Câmera imageadora de

Campo Largo (WFI), Transmissor de Dados de Imagem (DTS), Gravador de Dados Digital (DDR), Sistema de Coleta de Dados (DCS) e Monitor do Ambiente Espacial (SEM) (INPE/CBERS, 2019). A figura 28 apresenta um sumário das características das câmeras do CBERS 04A.

Características das câmeras do CBERS 04A			
Característica	WPM	MUX	WFI
Bandas Espectrais	0,45-0,52 μ m (B) 0,52-0,59 μ m (G) 0,63-0,69 μ m (R) 0,77-0,89 μ m (NIR) 0,45-0,90 μ m (PAN)	0,45-0,52 μ m (B) 0,52-0,59 μ m (G) 0,63-0,69 μ m (R) 0,77-0,89 μ m (NIR)	0,45-0,52 μ m (B) 0,52-0,59 μ m (G) 0,63-0,69 μ m (R) 0,77-0,89 μ m (NIR)
Resolução	2 m 8 m	16,5 m	55 m
Largura da Faixa Imageada	92 km	95 km	684 km
Visada Lateral de Espelho	não	não	não
Revisita	31 dias	31 dias	5 dias
Quantização	10 bits	8 bits	10 bits
Taxa de Dados Bruta	1800.8 Mbps 450.2 Mbps	65 Mbps	50 Mbps

Figura 28: Sumários das características das câmeras do CBERS 04A

Fonte: INPE/CBERS (2019)

Tendo em vista a necessidade de que o levantamento das áreas garimpeiras, relativas as regiões de estudo, corresponderem ao *Ground Truth* da forma mais precisa possível, os levantamentos foram realizados na escala 1:5000. Para isto, utilizou-se das imagens disponibilizadas pela câmera WPM, uma vez que esta é a que possui a melhor resolução se comparada às demais. Isto posto, para se atingir a resolução espacial de 2 metros, foi realizado, em um primeiro momento, a composição de cores das imagens correspondentes as bandas RGB (multiespectrais – 8 metros de resolução) e posteriormente, uma fusão entre esta composição e a imagem da banda pancromática (2 metros de resolução). A figura 29 apresenta o extrato de uma imagem adquirida pela câmera WPM do satélite CBERS 04A, de uma da região garimpeira de

Tapajós, na escala 1:5000, contendo: na letra a) a banda pancromática; na letra b) a banda vermelha; na letra c) a composição RGB; e na letra d) a fusão da banda pancromática e da composição RGB.

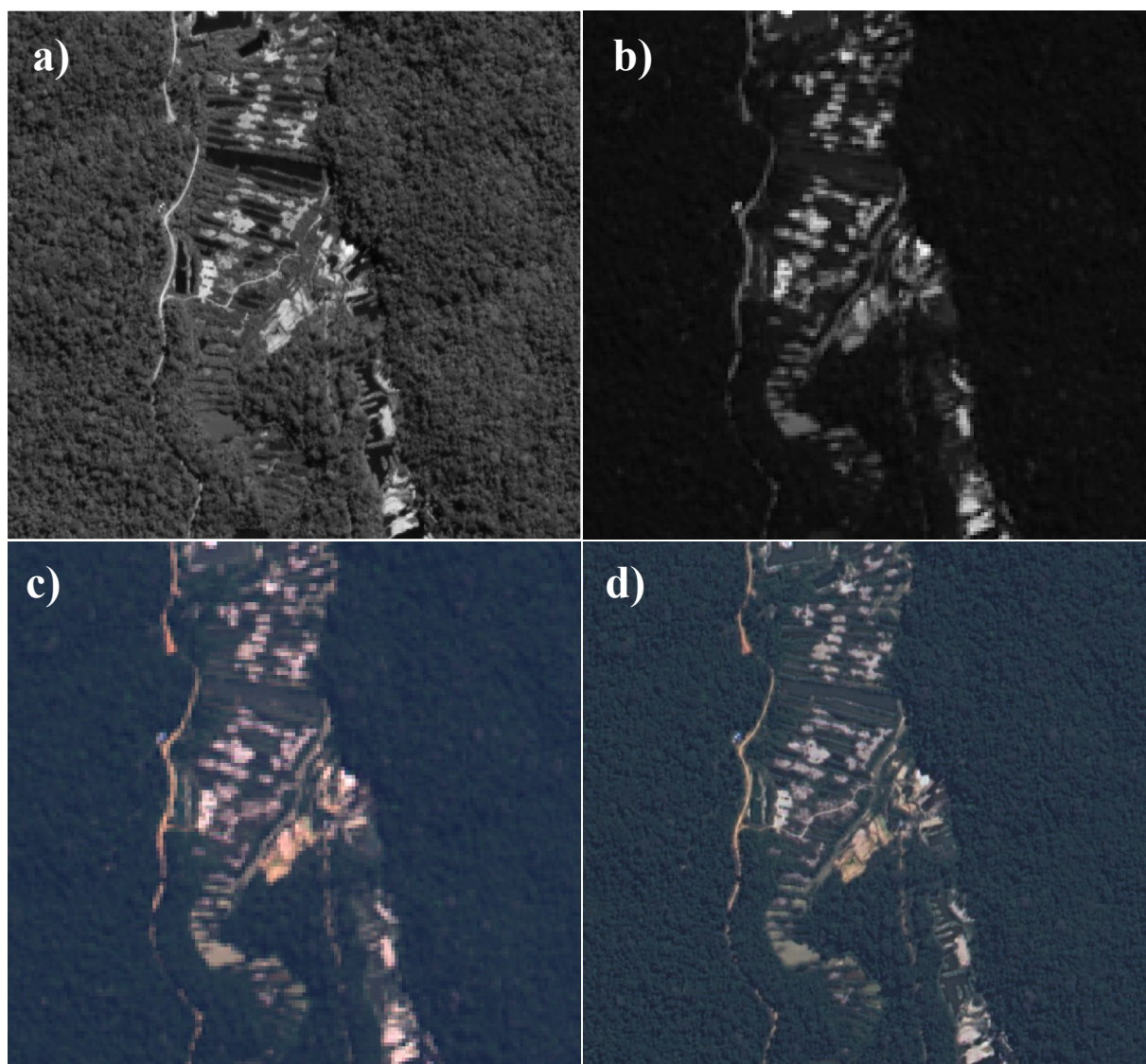


Figura 29: a) banda pancromática; b) banda vermelha; c) composição RGB; d) fusão PAN e composição colorida
Fonte: Próprio autor

No que se refere ao tamanho da cena do CBERS 04A, as mesmas possuem, aproximadamente, 9000km². Desta forma, para trabalhar os dados, a cena foi dividida em uma grade de quadrados de, aproximadamente, 5km² cada. Com esta divisão, cada um dos quadrados formados foi trabalhado como uma imagem única e, em cada um delas, foi realizados todos os processos necessários para formar os dados de entrada utilizados pelas redes neurais, quais sejam: definição de polígonos garimpeiros, criação de máscaras e *data augmentation*. Ressalta-se, ainda, que as cenas selecionadas não continham nuvens, uma vez que estas poderiam influenciar no treinamento das redes neurais e que a extração das mesmas não é objetivo deste trabalho.

4.3.2 Preparação do banco de dados

Para preparação do banco de dados, inicialmente foram adquiridos, a partir de fotointerpretação, os polígonos de garimpo referentes a cada uma das grades criadas, na escala 1:5000. Estes foram adquiridos tendo como base os polígonos de desmatamentos obtidos do sistema DETER. Após obtenção dos polígonos, foi realizada uma união vetorial entre os polígonos de garimpo e as grades. Posteriormente, esta união foi rasterizada para criação de imagens no formato TIFF. Para realização de todos estes procedimentos, utilizou-se do programa QGIS, na versão 3.16.11-Hannover (disponível em https://qgis.org/pt_BR/site/).

A partir das imagens rasterizadas, deu-se início ao processamento das mesmas para criação de máscaras, cujo o objetivo foi formar o *Ground Truth* das redes neurais. Neste processamento, definiu-se que os valores dos pixels seriam iguais a zero nos locais que não eram garimpo e iguais a 255 nos locais de garimpo (dados binários). Para este processo utilizou-se do programa ImageJ (disponível em <https://imagej.net/>).

Definidos as imagens e suas máscaras, utilizou-se o processo de *data augmentation* para aumentar o número de dados de entrada. Este procedimento, programado em *python 3.8*, realizou, aleatoriamente, diversas transformações nas imagens, conforme apresentado na figura 30.

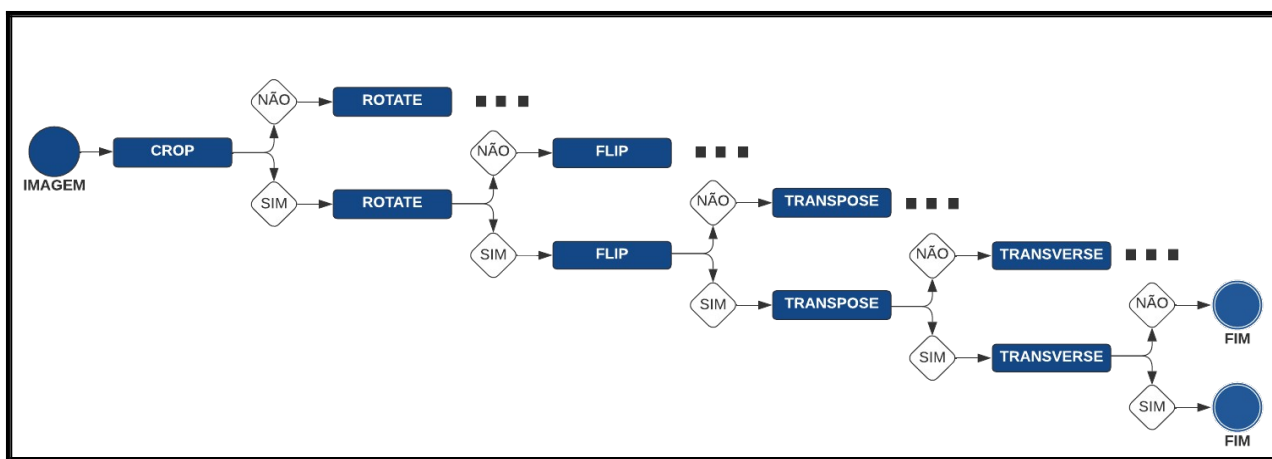


Figura 30: Processo de *Data Augmentation*

Fonte: Próprio Autor.

A partir do *data augmentation*, foram gerados o conjunto de dados do BDE e BDA, os quais foram divididos em treinamento (70% dos dados), validação (15%) e teste (15%). A divisão foi realizada nesta proporção uma vez que segue o que foi apresentado nos artigos do Estado da Arte, bem como atende a necessidade de se ter uma quantidade de dados de treinamento adequada. A tabela 6 apresenta os bancos e seus respectivos conjuntos de dados e a figura 31 esquematiza todo o processo de criação dos bancos de estudo e aplicação.

Banco	Região	Dados Originais				
		Nr imagens (grades)	Nr polígonos de garimpo	Nr imagens Treinamento	Nr imagens Validação	Nr imagens Teste
BDE	Tapajós (2020)	13200	49380	9240	1980	1980
BDA	Tapajós (2020)	38000	137640	26600	5700	5700

Tabela 6: Divisão dos dados em treinamento, validação e teste

Fonte: Próprio autor.

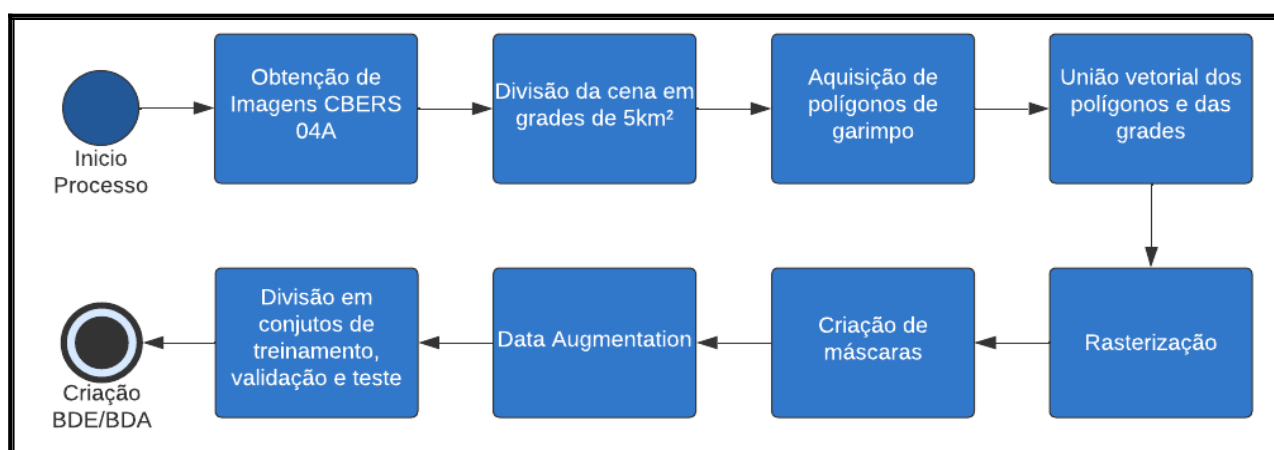


Figura 31: Esquemática do processo de criação dos bancos de dados de estudo e aplicação

Fonte: Próprio Autor.

Para criação do BDT utilizou-se os mesmos procedimentos. Porém, neste banco, o processo foi finalizado após a criação de máscaras para as imagens de Kayapó 2020 e após a divisão da cena em grades para a imagem de Tapajós 2021.

4.4 Processamento dos dados

Com os dados adquiridos e adequados em seus respectivos grupos (treinamento, validação e teste), utilizou-se o *framework* de aprendizado de máquina *Pytorch* e o serviço de nuvem *Google Colaboratory (Colab)* para processamento dos mesmos.

O *Pytorch* é uma biblioteca de tensores, otimizada para aprendizado profundo, que pode utilizar tanto Unidades de Processamento Gráfico – GPUs, quanto Unidades de Processamento Central – CPUs (PYTORCH, 2022). Dentre as vantagens do *Pytorch*, STEVENS (2020) elenca que o mesmo minimiza as sobrecargas de memória enquanto foca na flexibilidade e velocidade,

padroniza a execução imediata de operações, oferece inúmeras bibliotecas, além de seus *scripts* poderem ser pré compilados e importados de linguagens *python* e C++.

O *Colab* é um serviço de nuvem gratuito, disponibilizado pelo Google, baseado no código aberto *Jupyter Notebook*. É voltado, especialmente, para aprendizado de máquina e análise de dados. Dentre suas vantagens, ressalta-se a possibilidade de processamento de dados em GPUs.

4.4.1 Treinamento, validação e teste

Para os processos de treinamento, validação e teste aplicados neste projeto, utilizou-se da biblioteca *Segmentation Models Pytorch* – SMP (YAKUBOVSKIY, 2022), em virtude desta possuir uma elevada Interface de Programação de Aplicação (API), além de conter 9 modelos de arquiteturas, centenas de *backbones* com dados pré treinados, diversas métricas de avaliação e variadas funções de perda.

Apesar dos levantamentos realizados para os capítulos 2 e 3, não foi possível determinar o melhor modelo e os melhores hiperparâmetros a serem utilizados para atingir o objetivo proposto deste trabalho. Isto posto, resolveu-se por testar todas as arquiteturas disponibilizadas no SMP, bem como avaliar cinco diferentes *backbones* disponíveis na biblioteca (ResNet, RegNet, GERNet, EfficientNet e VGG). Além disso, definiu-se a utilização das métricas apresentadas no capítulo 3 (Acurácia, *F1Score* e *IoU*) para avaliação do modelo final.

Inicialmente, os treinamentos preliminares foram realizados no BDE. Os resultados destes definiram os parâmetros utilizados no conjunto de dados do BDA. O modelo estabelecido pelos dados deste conjunto foi aplicado nos dados do BDT. Para facilitar o entendimento deste processo, os procedimentos foram divididos conforme segue:

- a) Avaliação e definição da arquitetura mais adequada para detecção de garimpo;
- b) Verificação do *backbone* de melhor desempenho;
- c) Utilização da arquitetura e do *backbone* nos dados do BDA e estabelecimento do melhor modelo; e
- d) Utilização do melhor modelo nos dados do Banco de Dados de Teste.

4.4.2 Definição da melhor arquitetura

Para definir a arquitetura mais adequada para detecção de garimpo, foram testadas, com os dados do BDE, as 9 arquiteturas disponíveis no SMP (Unet, Unet++, Manet, Linknet, FPN, PSPNet, PAN, DeepLabV3, DeepLabV3+). Para esses teste foram selecionados os hiperparâmetros apresentados na tabela 7.

Hiperparâmetro	Valor
Otimizador	Adam
Função de Perda	<i>Dice Loss</i>
Backbone	ResNet
Nr Épocas	10
Batch Size	16
Métrica	IoU

Tabela 7: Hiperparâmetros utilizados para definição da arquitetura mais adequada

Fonte: Próprio autor.

O otimizador Adam foi o escolhido tendo em vista que o mesmo é eficiente para se trabalhar com grande volume de dados, requer menos memória e menos parâmetros para ajuste, além de já estar implementado na biblioteca do Pytorch.

A Função de Perda *Dice Loss* foi selecionada pois a maior parte das imagens são formadas pelo *background* e esta função é utilizada para maximizar as áreas de interseção do *foreground*, enquanto minimiza a união das mesmas.

O número de épocas igual a 10 foi utilizado pois o objetivo principal dos testes eram verificar a arquitetura e o *backbone* que se destacavam no treinamento das CNN. Para o treinamento final, no BDA, foram selecionadas 60 épocas. Esta quantidade foi decidida após teste com 20, 40, 60 e 80 épocas. Com 20 e 40 épocas verificou-se que as perdas e as métricas de avaliação ainda estavam em evolução (figura 5.3). Com 60 épocas os dados tenderam a estabilização e não apresentaram características de *overfitting*. Para 80 épocas, os dados não foram analisados devido a limitações computacionais.

O tamanho do *Batch* igual a 16 foi escolhido por ser o maior *batch* possível dentro das limitações impostas pelo Google Colab.

A métrica de avaliação IoU foi priorizada por sua característica de penalizar os erros com um peso maior, conforme apresentado no item 3.4.

Determinados os hiperparâmetros, as arquiteturas foram testadas conforme apresentado na figura 32.

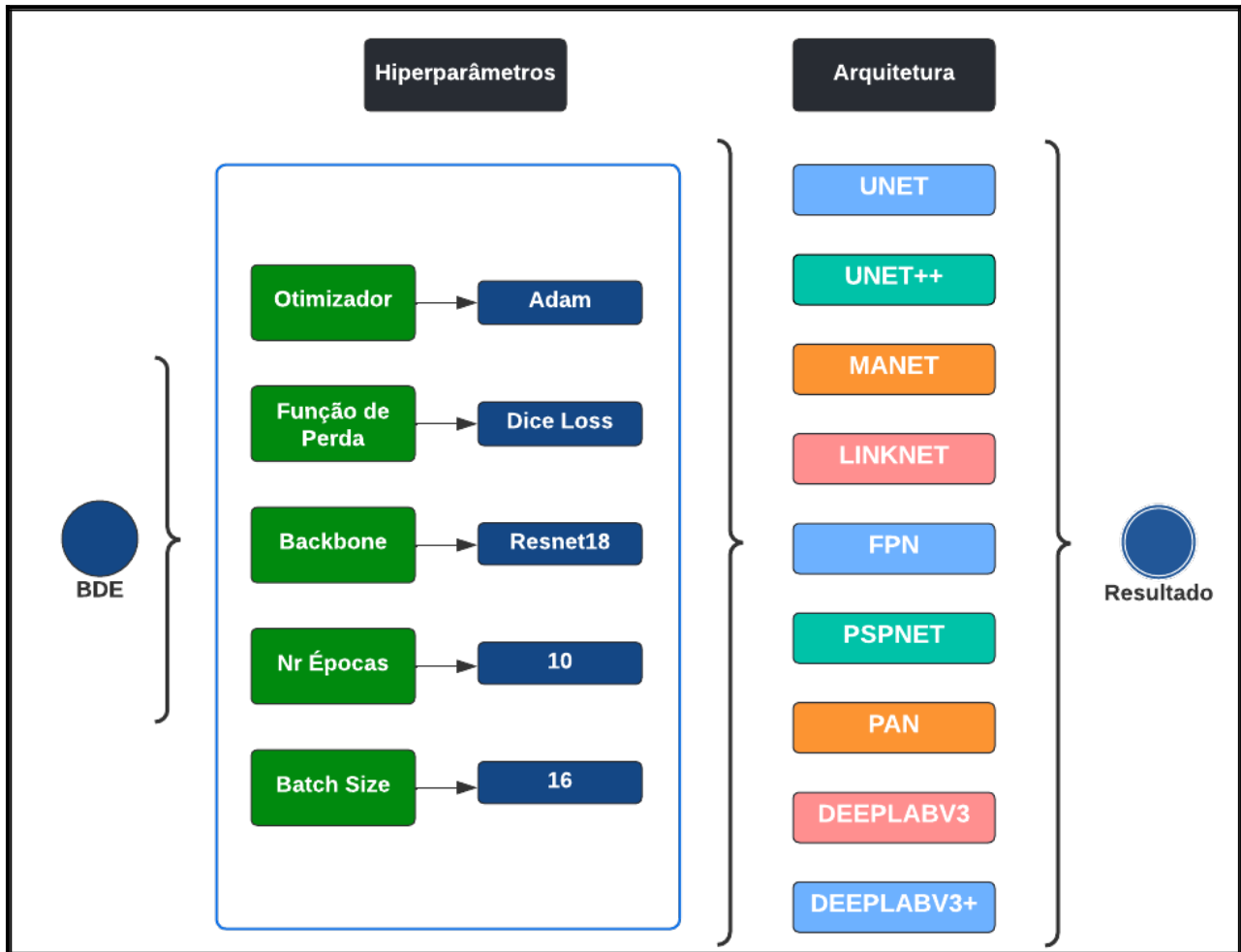


Figura 32: Esquematização para definição da melhor arquitetura

Fonte: Próprio autor.

Definida a arquitetura que apresentou melhor performance, a mesma foi selecionada para processar novamente os dados e determinar o *backbone* que provesse a maior precisão.

4.4.3 Definição do *backbone* mais adequado

Nesta fase foram testados os *backbones* ResNet18, RegNet_16, GERNet_s, EfficientNet-b3 e VGG13, cujos pesos originaram-se do banco de dados ImageNet, que é uma base que fornece dados de imagem para o treinamento de modelos de reconhecimento de objetos em larga escala. Estes

backbones contém, respectivamente, 11, 8, 6, 10 e 9 milhões de parâmetros pré treinados. A figura 33 apresenta a esquematização desta etapa.

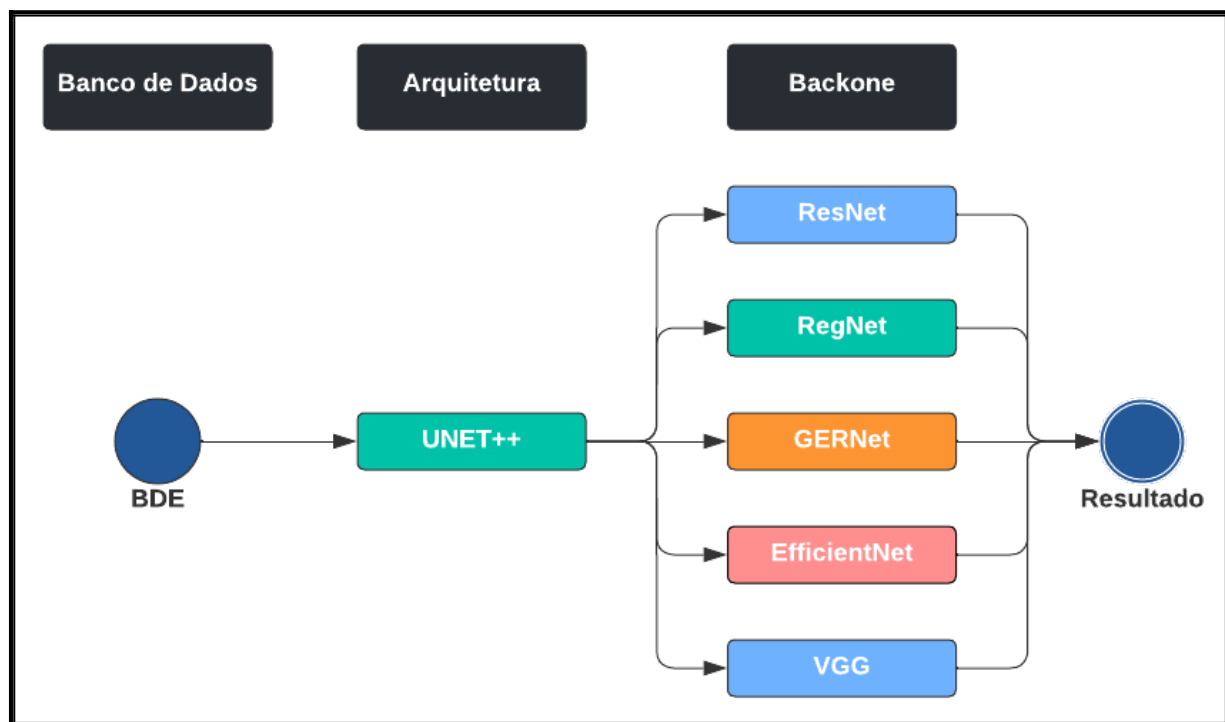


Figura 33: Esquematização para definição do melhor backbone

Fonte: Próprio autor.

4.4.4 Estabelecimento do melhor modelo

Definida a arquitetura e o *backbone*, utilizou-se os dados do BDA para estabelecer o modelo final do trabalho.

Nesta fase, aplicou-se a UNET++ com 20, 40 e 60 épocas de processamento. Após análise, verificou-se que 60 épocas foram adequadas para que os dados de treinamento e validação se estabilizassem, conforme será apresentado no capítulo 5.

O *backbone* selecionado foi o EfficientNet-b6 (40 milhões de parâmetros pré treinados). As métricas de avaliação, a função de perda, o otimizador e o tamanho do *batch* foram mantidos. A figura 34 apresenta a esquematização desta etapa.

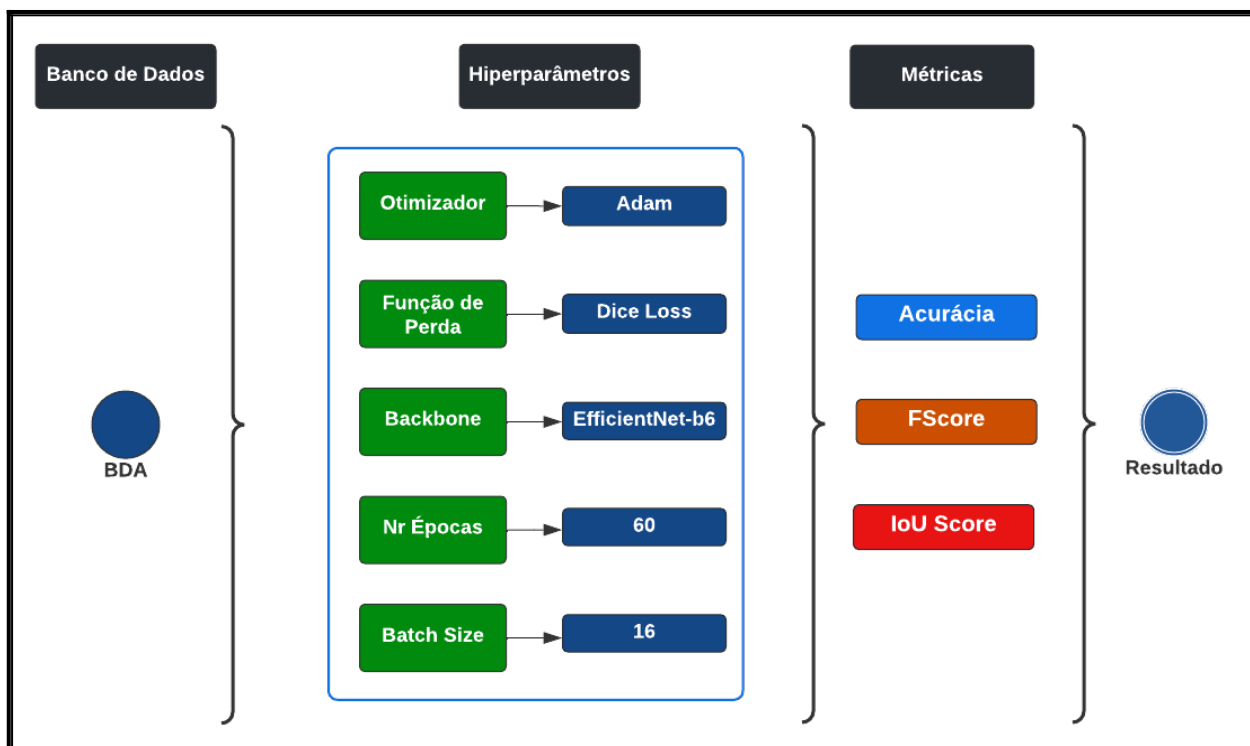


Figura 34: Esquematização para definição do melhor modelo treinado e resultado de suas métricas de avaliação
Fonte: Próprio autor.

4.4.5 Teste e detecção de mudança

Após definição do modelo final, o mesmo foi utilizado no BDT com intuito de testar o resultado em uma área diferente da treinada (região de Kayapó) e para detectar mudanças ocorridas no período de um ano (2020 – 2021). A figura 35 esquematiza esta etapa.

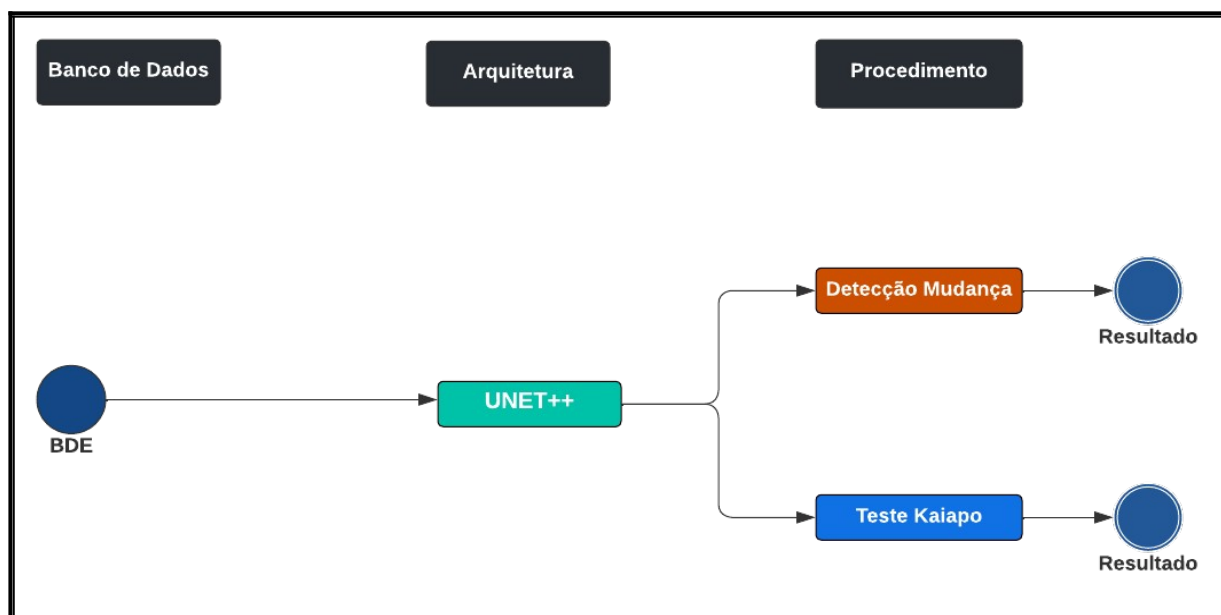


Figura 35: Esquematização do teste na área de Kayapó e da detecção de mudança
Fonte: Próprio autor

Para o teste em Kayapó, uma amostra de 30 imagens aleatórias da região, cada uma com área de 5km², retiradas de uma única cena do CBERS 04A, contendo polígonos de garimpo, foi selecionada. Estas imagens foram submetidas ao modelo para verificação da perda e dos valores das métricas de avaliação. Foi averiguado, ainda, o erro percentual entre os resultados alcançados com os dados do BDA e do BDE.

Para a detecção de mudança, dentre as estruturas apresentadas na seção 3.3.2, foi utilizada a Estrutura de Fluxo Duplo com Classificação Posterior, com o modelo de treinamento adquirido em apenas uma das épocas. Esta estrutura foi a escolhida uma vez que foram levantadas as máscaras de *Ground Truth* apenas das imagens de Tapajós no ano de 2020, não havendo, desta forma, a possibilidade de utilizar as imagens de 2021 para treinamento da rede. Sendo assim, empregou-se o algoritmo treinado nas imagens de 2020 para levantar as previsões do ano 2021.

Com as máscaras de previsão de ambas as épocas, recorreu-se a operações aritméticas para subtrair os valores dos pixels em cada uma das imagens, mostrando graficamente a mudança entre as mesmas. Para determinar a similaridade entre as imagens, utilizou-se do Índice de Similaridade Estrutural, uma vez que o mesmo possui consistência à percepção do olho humano e apresenta resultados percentuais de similaridade. Ademais, foi considerado imagens com mudanças as que obtiveram similaridade menores que 95%. Com isto, pôde-se separar as imagens conforme a manutenção ou mudança das áreas garimpeiras.

5 RESULTADOS E DISCUSSÕES

Neste capítulo serão apresentados os resultados e as discussões dos testes realizados para encontrar a arquitetura e o *backbone* que melhor atenderam ao objetivo da dissertação, o modelo final, os resultados dos testes na área de Kayapó e a detecção de mudança do estágio do garimpo na região de Tapajós, após decorrido um ano.

5.1 Arquitetura com melhor desempenho

Conforme apresentado na seção 4.4.2, foram realizados testes para definir a arquitetura mais adequada para detecção de garimpo. Nesta etapa verificou-se que todas as arquiteturas obtiveram avaliações de precisão de aproximadamente 68%, com exceção da arquitetura PSPNet, que obteve um desempenho de 63,96%, bem aquém das demais. No geral, as arquiteturas que mais se destacaram foram: UNET++, que apresentou um IoU Score de 68,93%; DeepLabV3, com 68,72%; e UNET, com 68,60%.

Tendo em vista o objetivo desta etapa, a arquitetura selecionada para desenvolvimento do trabalho foi a UNET++. Os resultados de todas as arquiteturas são apresentados na figura 36.

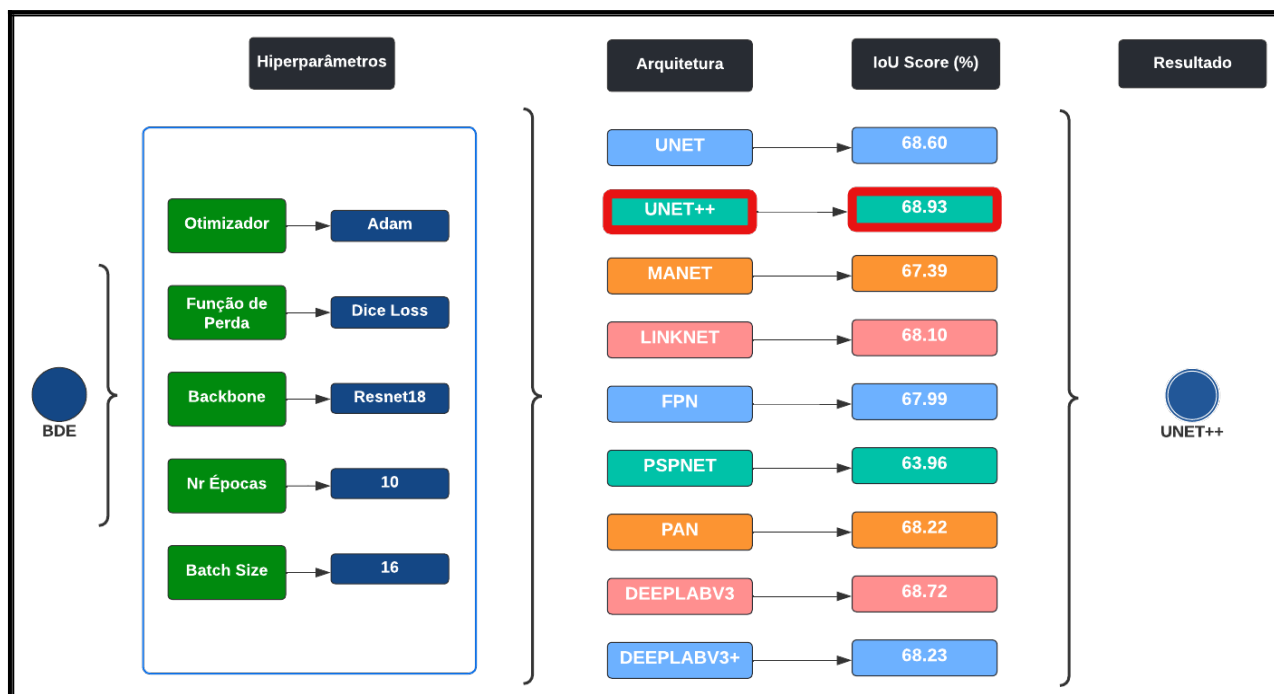


Figura 36: Resultados da escolha da arquitetura mais adequada

Fonte: Próprio autor

Em comparação ao apresentado no Estado da Arte, os resultados alcançados nesta etapa foram inferiores. Isto ocorreu pelo fato de que a quantidade de dados de entrada, o número de épocas de treinamento e o número de algoritmos pré treinados foram reduzidos. Esta redução foi proposital, uma vez que o objetivo desta fase foi testar as arquiteturas, não definir o modelo de treinamento. Além disso, um número menor de dados de entrada e épocas de treinamento propiciaram encontrar, celeremente, a arquitetura com maior potencial para segmentar as áreas garimpeiras analisadas.

5.2 Backbone de melhor performance

De acordo com o relatado na seção 4.4.3, foram realizados testes para definir o *backbone* nos quais os dados pré treinados mais se adaptaram à detecção de garimpos. Nesta fase, todos os *backbones* utilizados resultaram em performances com F1 *score* variando entre 75% e 83% e IoU entre 67% e 75%. Nesta etapa, destacou-se o *backbone EfficientNet*, que proporcionou um F1 *score* de 82,44% e um IoU de 74,06%. Os resultados do processamento com os *backbones* são apresentados na figura 37.

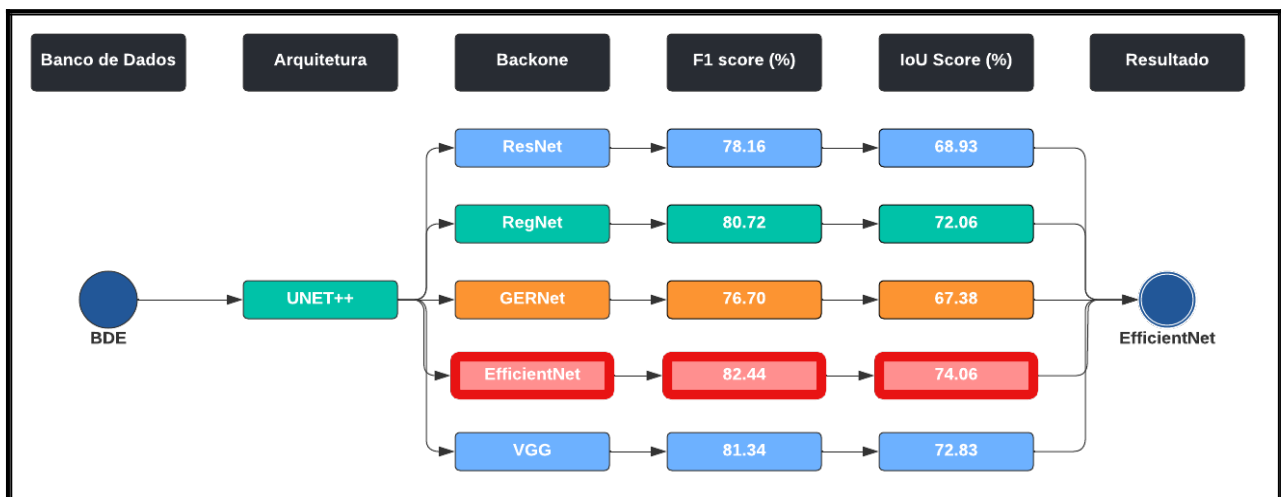


Figura 37: Resultados da escolha da arquitetura mais adequada

Fonte: Próprio autor

Apesar da melhora dos resultados com a utilização de *backbones* mais robustos, os mesmos ainda se mantiveram aquém do apresentado no Estado da Arte. Assim como explicado no item 5.1, a quantidade de dados de entrada e o número de épocas de treinamento foram preponderantes nesses resultados.

5.3 Modelo Final

Definida a arquitetura e o *backbone*, conforme determinado na seção 4.4.4, foi treinado o modelo final no BDA. O resultado final desta fase é apresentado na tabela 8. Os gráficos da figura 38 apresentam a evolução da perda e das métricas de avaliação durante a fase de treinamento e validação.

Época	Etapa	Dice Loss	Acurácia	F Score	IoU
20	Treinamento	0,06055	98,86%	93,95%	88,61%
	Validação	0,07676	98,72%	92,33%	85,78%
40	Treinamento	0,04290	99,00%	95,71%	91,79%
	Validação	0,06201	98,84%	93,80%	88,35%
60	Treinamento	0,03624	99,04%	96,38%	93,02%
	Validação	0,05641	98,89%	94,36%	89,34%

Tabela 8: Resultado final das etapas de treinamento e validação

Fonte: Próprio autor

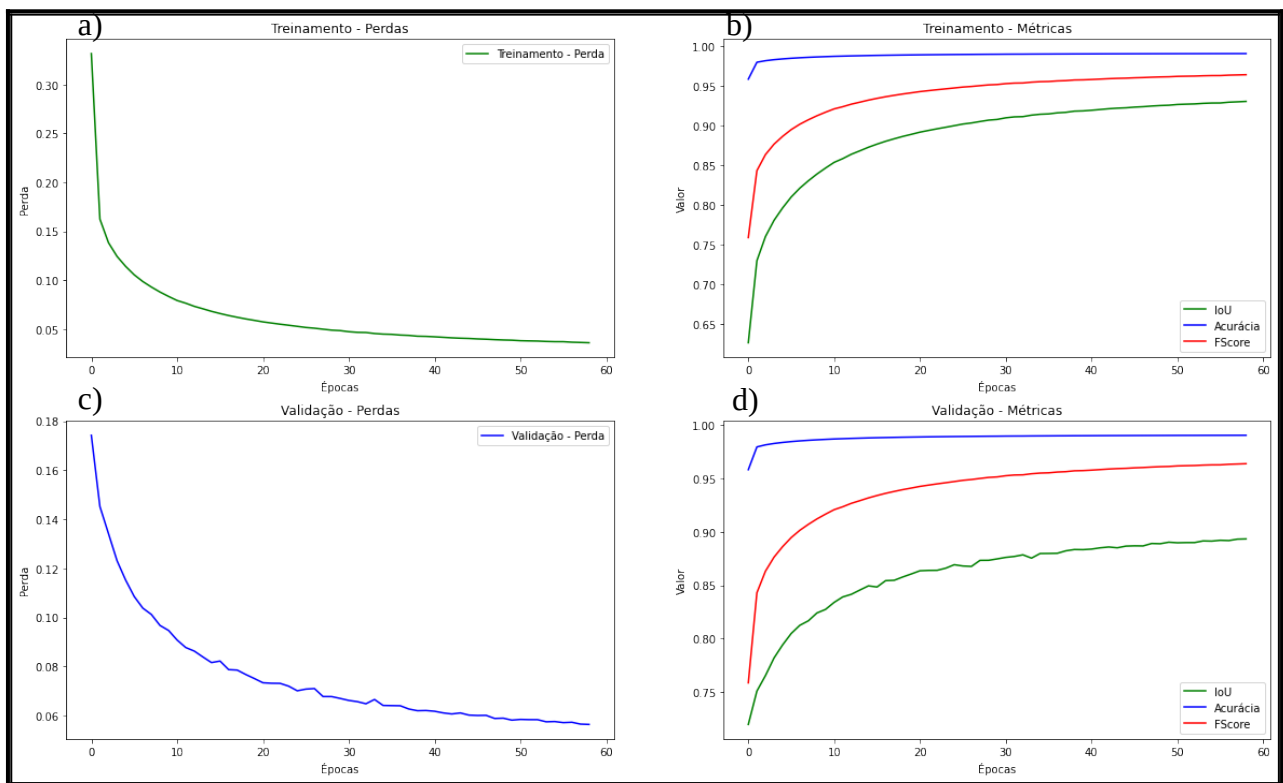


Figura 38: Gráficos da evolução da perda e das métricas de avaliação durante a fase de treinamento e validação

Fonte: Próprio autor

Pelos gráficos da figura 38, verifica-se que a medida que a CNN foi realizando seu treinamento, os pesos foram continuamente ajustados de forma que as áreas garimpeiras pudessem ser melhor delimitadas. A curva decrescente da perda demonstra que os pesos foram corrigidos, ao longo das épocas, em uma escala cada vez menor. Nota-se que entre as épocas 50 e 60 as funções de perda tenderam à estabilização.

Para as métricas de avaliação, nota-se uma curva crescente, demonstrando que a rede alcançava, a cada época, uma precisão maior. Verifica-se que a acurácia, desde o início do treinamento, apresentava valores elevados. Isto deve-se ao fato de que, nas imagens utilizadas, as áreas não garimpeira serem consideravelmente maiores que as áreas garimpeiras. Como consequência disto, a maioria dos pixels foram previamente posicionados nos locais “corretos”, impactando no resultado apresentado por esta métrica.

Para as métricas F1 *score* e IoU, verificou-se um resultado muito bom, se comparado aos apresentados na literatura do capítulo 2. A diferença de performance entre ambas ocorre, conforme explicado no capítulo 3, uma vez que a métrica F1 *score* tende a medir o desempenho médio das instâncias classificadas e a IoU tende a penalizar com maior peso as instâncias que foram erroneamente classificadas.

Com o treinamento e a validação da CNN, encontrou-se o modelo final do trabalho. Este modelo foi testado no conjunto de imagens de teste do BDA. Os resultados obtidos estão disponibilizados na figura 39 e um extrato das imagens são apresentados nas figuras 40:

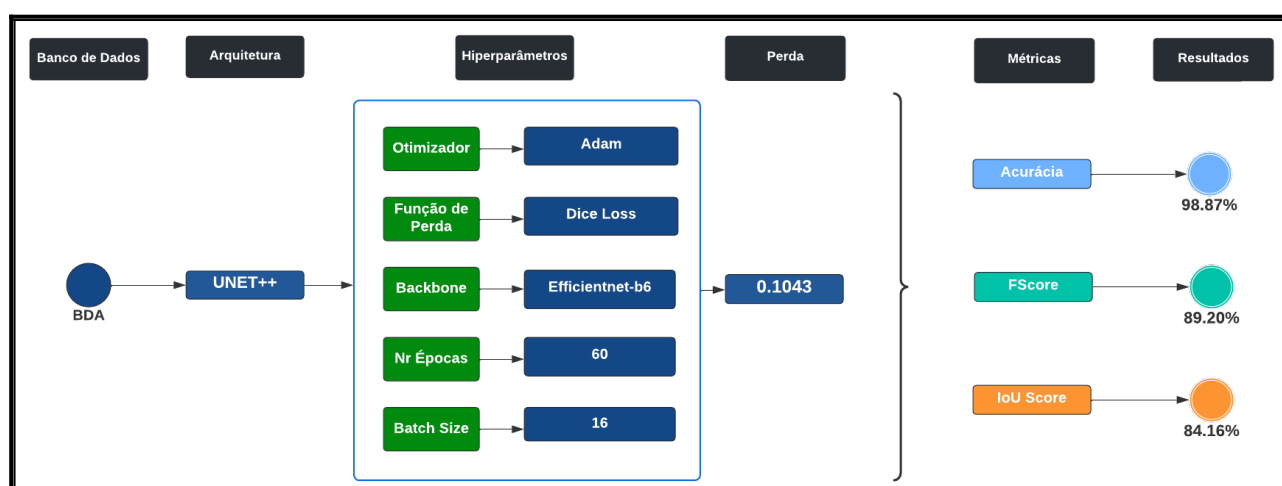


Figura 39: Resultado final do modelo treinado

Fonte: Próprio autor

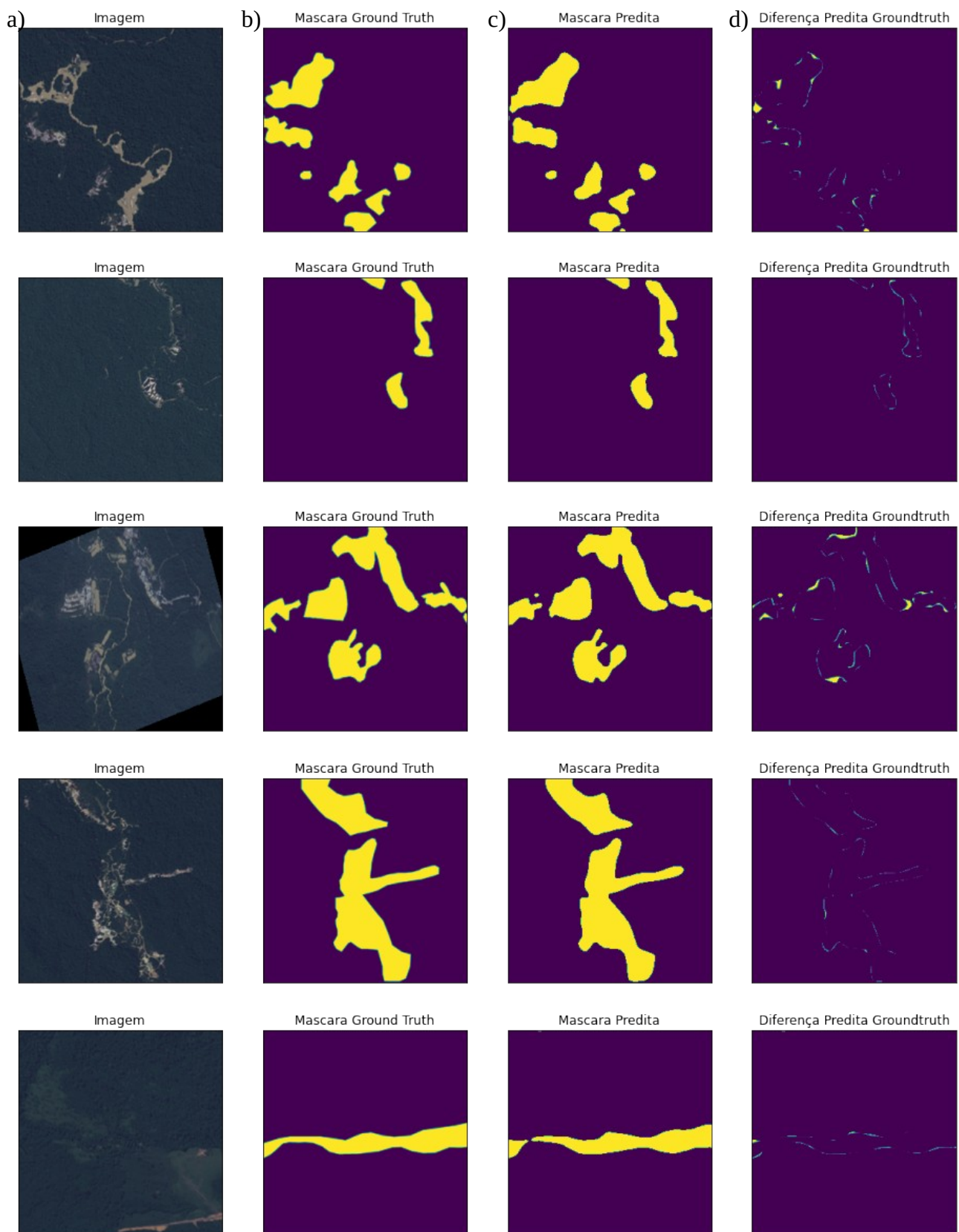


Figura 40: a) imagem original, b)máscara, c) máscara predita e d) diferença entre as máscaras das áreas garimpeiras

Fonte: Próprio autor.

A diminuição dos valores das métricas de avaliação eram esperados para a fase de testes, uma vez que esta etapa ocorreu em áreas nas quais a rede estava sendo utilizada pela primeira vez. Apesar dessa diminuição, os valores encontrados foram ótimos, se comparados aos resultados apresentados nas literaturas do capítulo 2. Sobre a diferença de resultados das métricas de avaliação, o mesmo raciocínio lógico utilizado nos conjuntos de Treinamento e Validação pode ser utilizado nesta etapa.

5.4 Teste em Kayapó

Com intuito de verificar o desempenho do algoritmo em uma área garimpeira diferente de Tapajós, onde o mesmo foi treinado, o modelo foi testado na área de Kayapó, conforme descrito na seção 4.4.5. Ao utilizar o modelo final nas 30 imagens pré levantadas, chegou-se aos resultados apresentados na tabela 9. Um extrato das imagens são demonstrados nas figuras 41 e 42.

Etapa	Acurácia	F Score	IoU
Teste	96,10%	82,67%	71,00%

Tabela 9: Resultado final do teste em Kayapó

Fonte: Próprio autor.

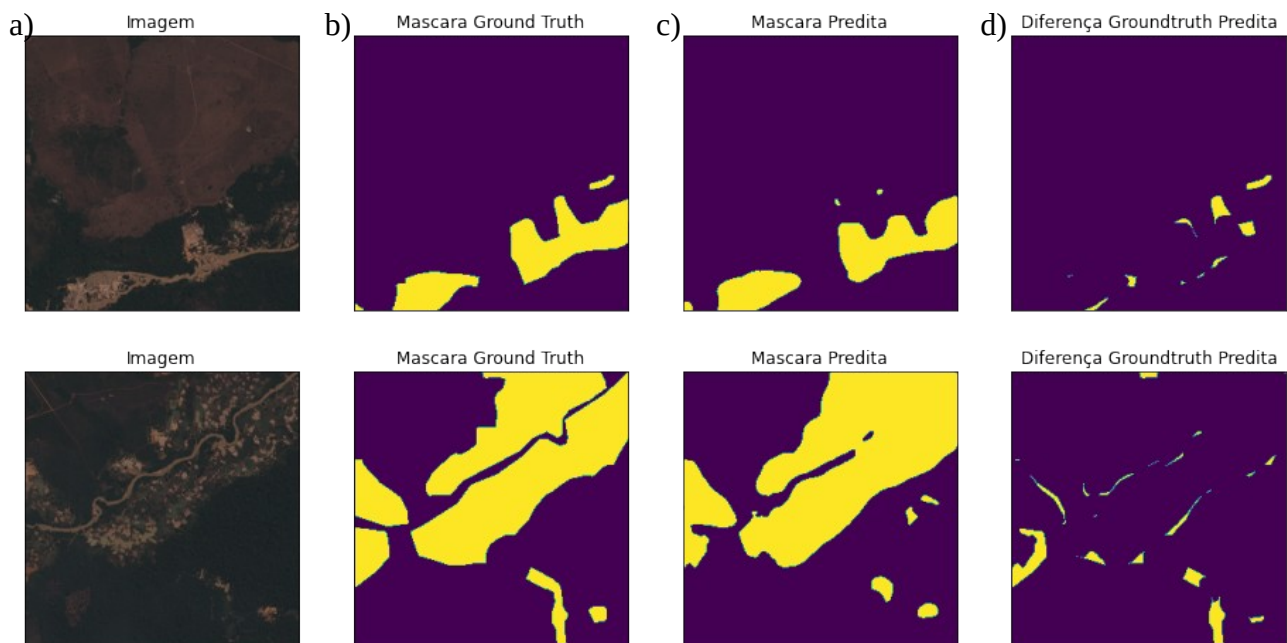


Figura 41: a) imagem original, b)máscara, c) máscara predita e d) diferença entre as máscaras das áreas garimpeiras em Kayapó

Fonte: Próprio autor

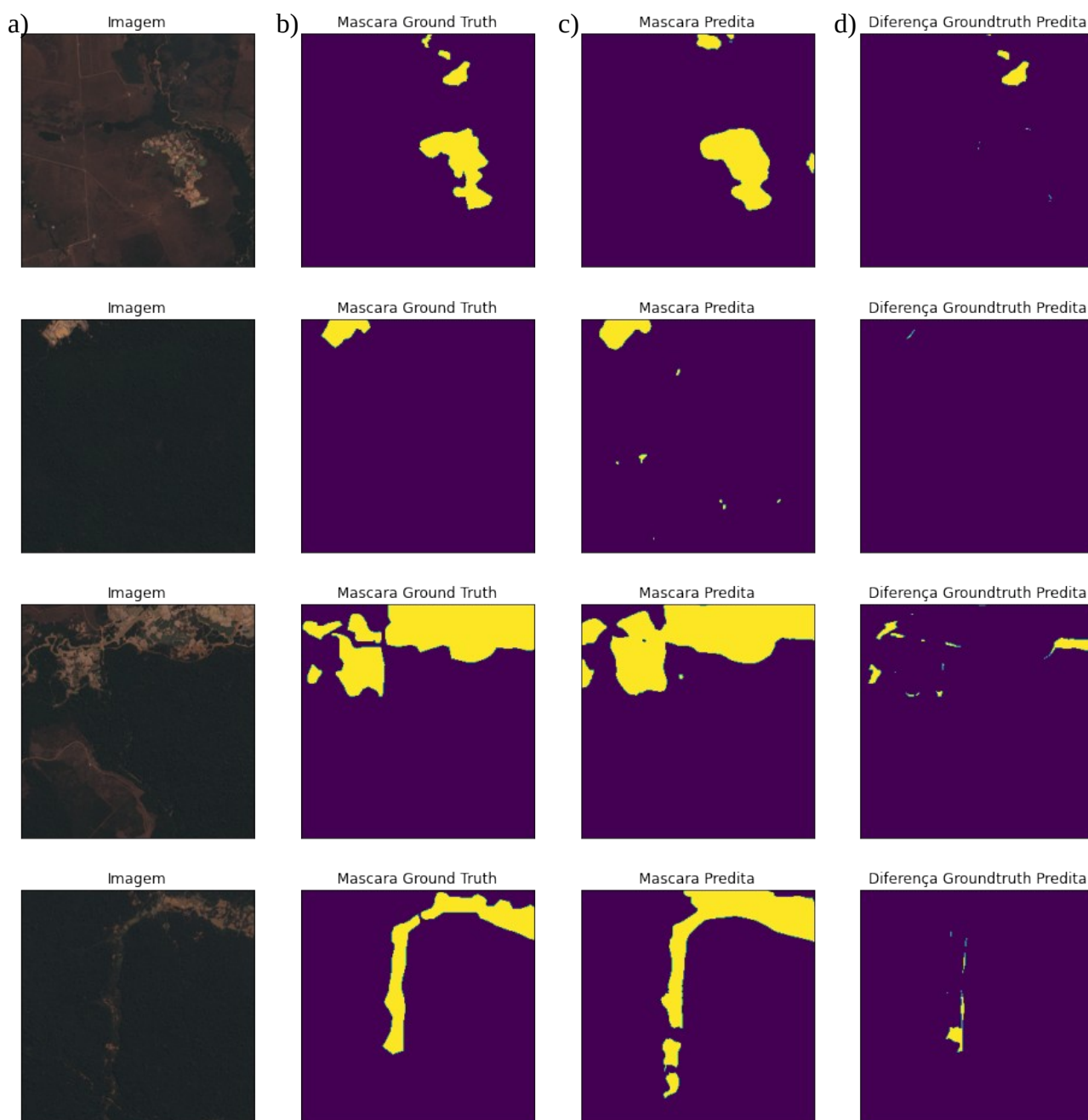


Figura 42: a) imagem original, b)máscara, c) máscara predita e d) diferença entre as máscaras das áreas garimpeiras em Kayapó

Fonte: Próprio autor

Neste teste, os valores teóricos esperados eram os encontrados no teste do BDA, conforme figura 39. Neste contexto, a tabela 10 apresenta o erro percentual para as métricas de avaliação.

Métrica	Acurácia	F Score	IoU
Erro Percentual	2,88%	7,90%	18,54%

Tabela 10: Erro percentual do teste em Kayapó

Fonte: Próprio autor

Os resultados alcançados pelo teste em Kayapó foram inferiores aos resultados em Tapajós. Este fato era esperado, uma vez que o treinamento ocorreu em sua totalidade na região de Tapajós. Nota-se, ainda, uma diferença considerável no erro percentual entre *F Score* e IoU. Este fato é explicado pelas características de cada uma das métricas em relação aos erros, conforme explicado no item 5.3.

Além disso, pode-se inferir que características específicas dos garimpos da região de Kayapó, como, por exemplo, o tamanho da área ocupada ao redor dos rios, suas extensões, a quantidade de piscinas de lavagem, entre outros, não foram constantemente treinadas nos garimpos da região de Tapajós, o que influenciou no entendimento do algoritmo quando da análise desta área. A figura 43 apresenta exemplos de garimpos específicos da região de Kayapó.

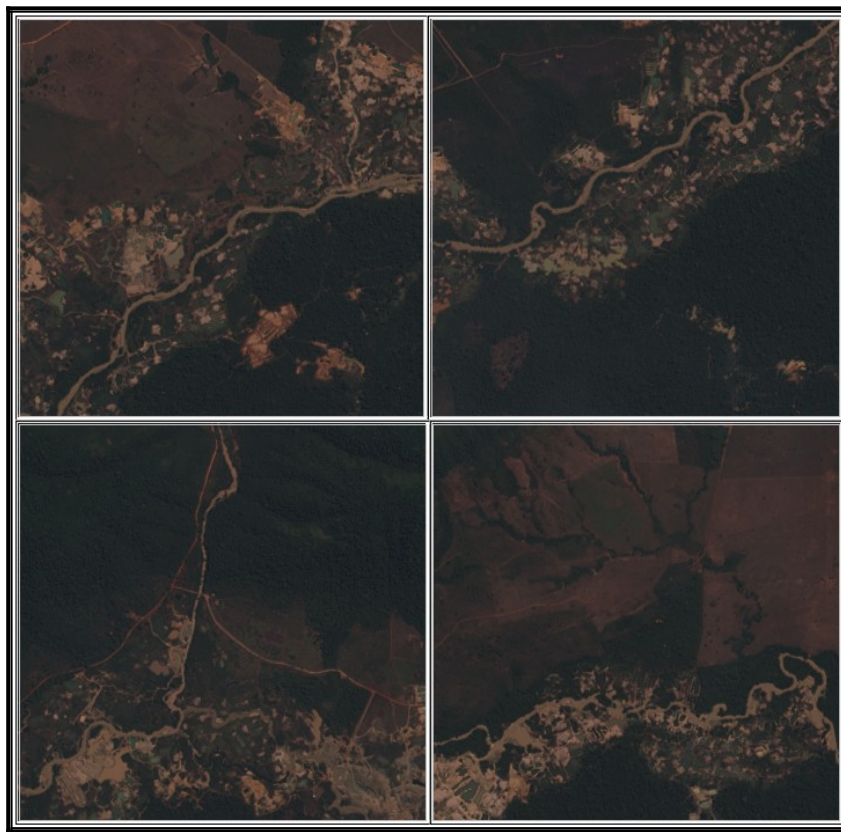


Figura 43: Exemplos de garimpos em Kayapó
Fonte: Próprio autor

Com os resultados obtidos em Kayapó, verifica-se que para melhores efeitos nesta região, ou em demais áreas da Amazônia, há a necessidade de que se obtenha uma maior quantidade de dados de entrada, bem como requer-se a realização de treinamentos em diferentes ambientes garimpeiros. Estes treinamentos, com dados de diferentes áreas, aperfeiçoariam a rede para que a mesma fosse capaz de aprender a segmentar os variados garimpos existentes na região amazônica de modo mais eficaz.

5.5 Detecção de Mudança

A fim de detectar as mudanças ocorridas nas áreas garimpeiras de Tapajós entre os anos de 2020 e 2021, o modelo foi utilizado no BDT, nas imagens referentes ao ano de 2021. Para esta etapa não houve teste do modelo, apenas sua aplicação prática. Na primeira etapa, utilizou-se do esquema da Estrutura de Fluxo Duplo com Classificação Posterior para levantar as máscaras de previsão dos anos de 2020 e 2021. Conforme apresentado na seção 4.4.5, com o modelo de treinamento adquirido em apenas uma das épocas, utilizou-se de operações aritméticas de subtração para determinar a diferença entre os valores dos pixels em cada uma das imagens, com objetivo de apresentar os resultados graficamente. Um extrato dos resultados das imagens sem mudança é apresentado nas figuras 44 e com mudanças na figura 45.

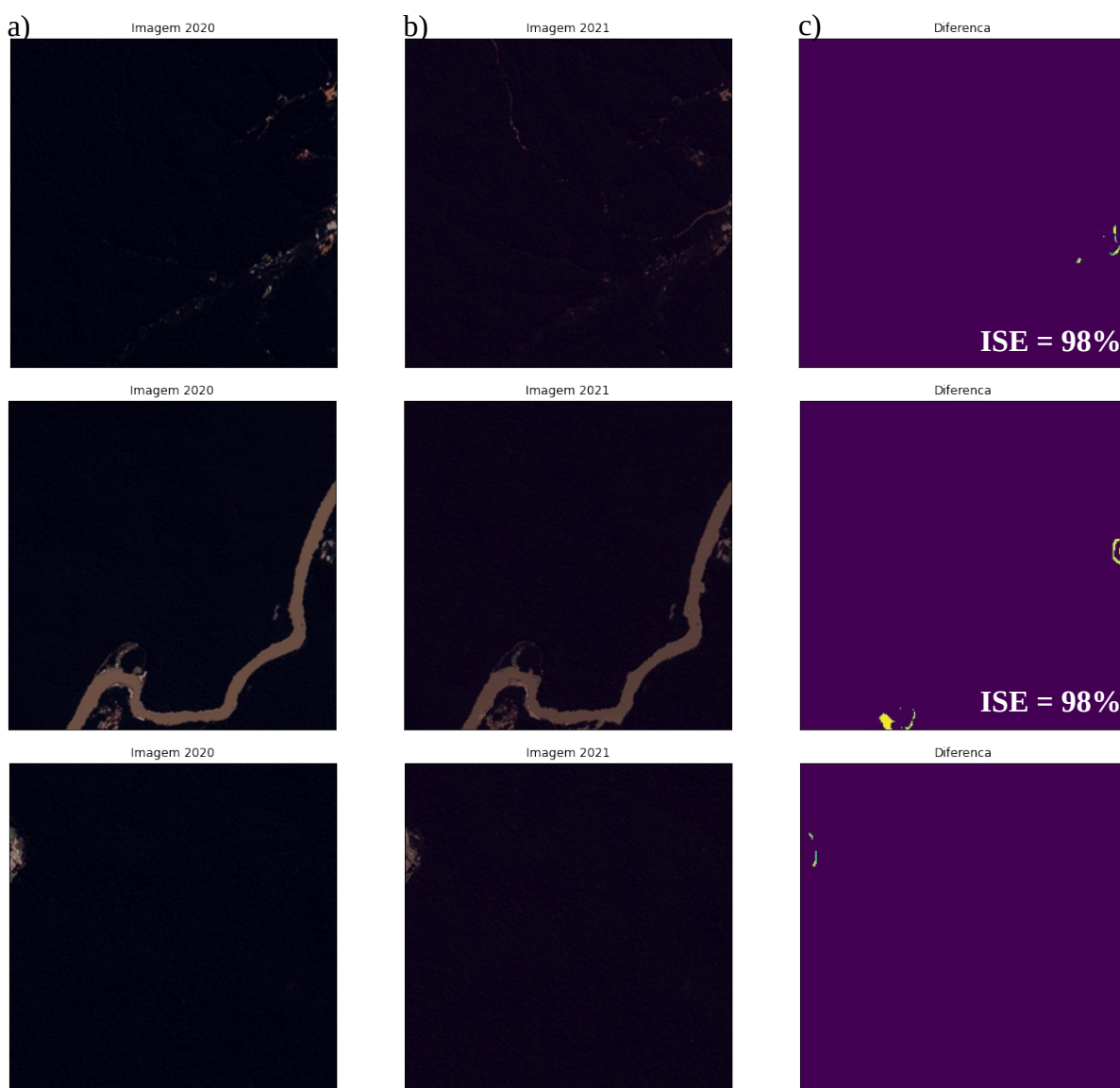


Figura 44: Imagens SEM mudanças: a) imagem 2020, b) imagem 2021 e c) diferença entre as mesmas
Fonte: Próprio autor.

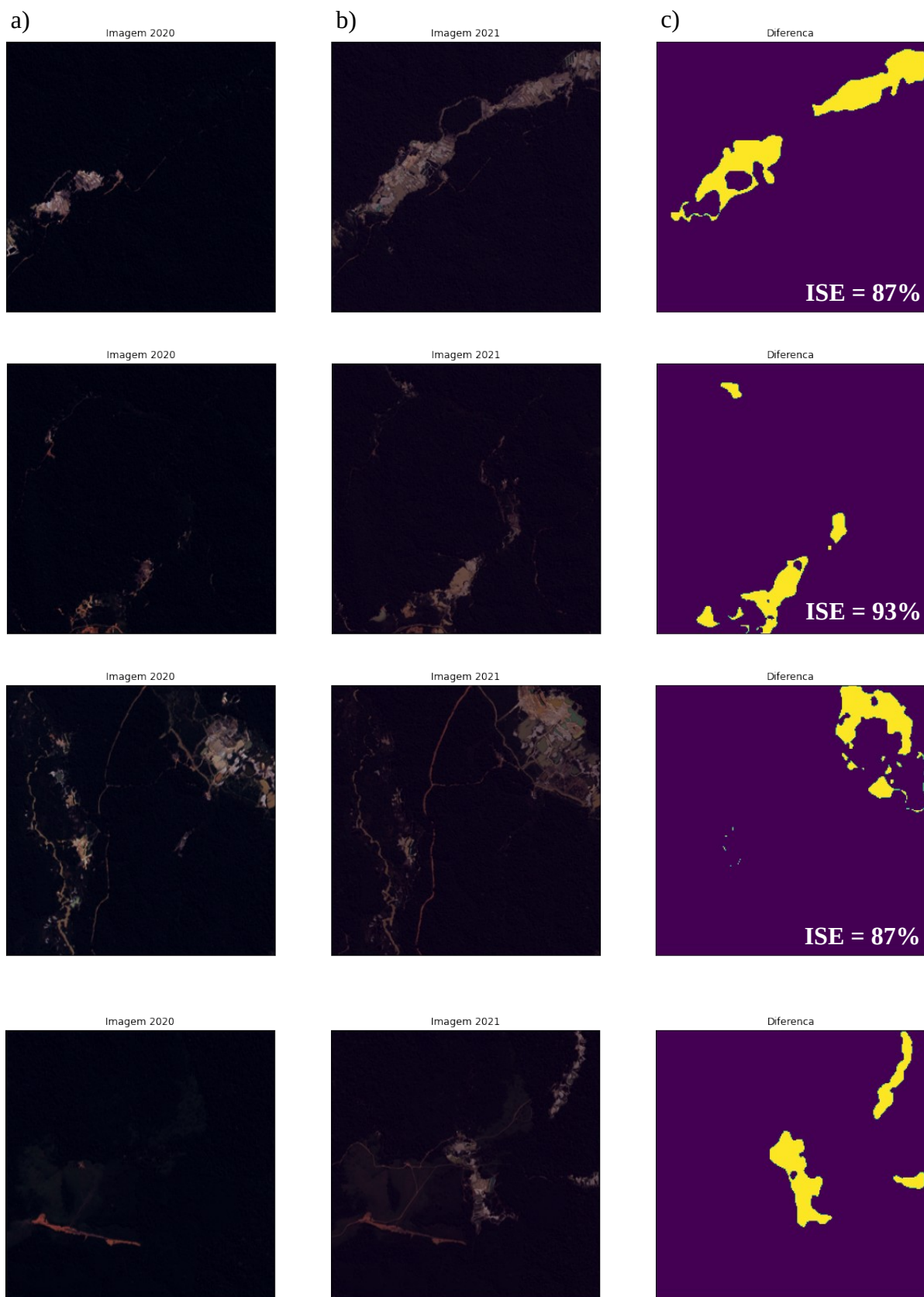


Figura 45: Imagens COM mudanças: a) imagem 2020, b) imagem 2021 e c) diferença entre as mesmas

Fonte: Próprio autor.

Conforme apresentado na seção 4.4.5, utilizou-se do Índice de Similaridade Estrutural para quantificar as similaridades entre as imagens. Índices superiores a 95% resultaram em imagens sem mudanças. Índices inferiores, resultaram em imagens com mudanças. Ao aplicar o ISE no conjunto de dados do BDT de Tapajós, obteve-se 327 imagens sem mudanças e 111 imagens com mudanças, ou seja, houveram alterações em, aproximadamente, 25% das imagens levantadas, conforme apresentado no gráfico 12. A partir desses dados, pode-se analisar o novo impacto ocasionado pela alteração das áreas garimpeiras nessas 111 imagens, bem como controlar a evolução das mesmas. A tabela 11 apresenta um extrato dos resultados de similaridade pelo ISE.

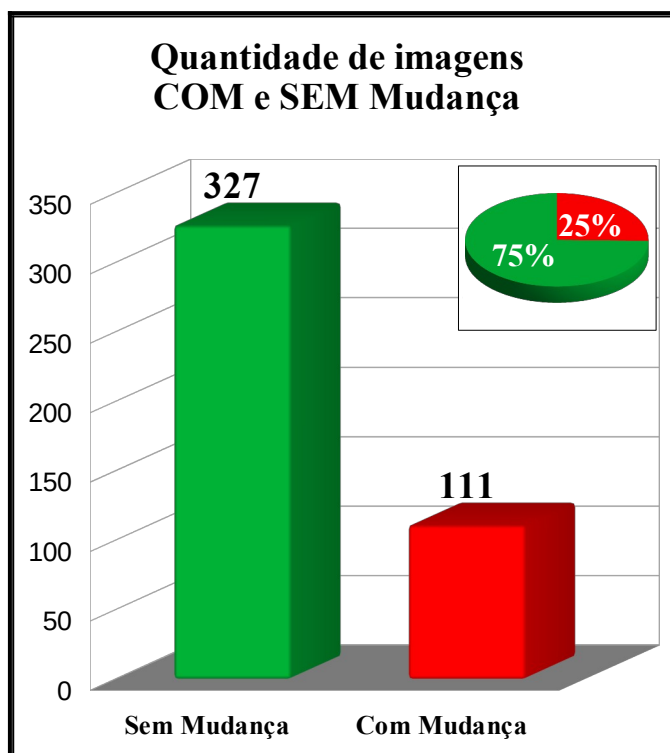


Gráfico 12: Quantidade de imagens com e sem mudança após decorrido um ano

Fonte: Próprio autor

Id Imagem	ISE	Mudança
1	92.05%	Sim
2	99.76%	Não
3	99.92%	Não
4	91.69%	Sim
5	97.03%	Não
6	98.50%	Não

Id Imagem	ISE	Mudança
7	97.49%	Não
8	93.05%	Sim
9	91.50%	Sim
10	100.00%	Não
11	100.00%	Não
12	90.28%	Sim

Tabela 11: Extrato da tabela de resultados do ISE

Fonte: Próprio autor

6 CONCLUSÃO

O presente trabalho teve por objetivo avaliar, selecionar e aplicar algoritmos de Aprendizado Profundo para monitoramento e análise temporal da evolução dos status dos garimpos nas regiões de Tapajós e Kayapó, na Floresta Amazônica Brasileira, a partir de imagens oriundas do satélite CBERS 04A, entre os anos de 2020 e 2021.

Para cumprir tal objetivo, inicialmente evidenciou-se a eficiência da utilização de Redes Neurais Convolucionais para segmentação semântica e detecção de mudança, em imagens advindas de sensores remotos, a partir do apresentado no Estado da Arte.

Posteriormente, foram realizados diversos experimentos para identificar a melhor arquitetura e o melhor *backbone* para o problema proposto. Como resultados destes experimentos, definiu-se a arquitetura UNET++ com o *backbone EfficientNet* como modelo mais adequado para o trabalho.

Após realização dos experimentos iniciais, foram feitos treinamentos, validações e testes em 38000 imagens advindas de duas cenas do satélite CBERS 04A, na região de Tapajós. Ao final desta etapa, as métricas de avaliação apresentaram como resultados uma acurácia de 98,87%, um *F1 score* de 89,20% e um IoU de 84,16%. Tais resultados são condizentes com os encontrados no Estado da Arte, corroborando a eficiência da técnica e do modelo utilizado.

Em seguida, para verificar a performance do algoritmo em uma área diferente da área de treinamento, o modelo foi avaliado na região de Kayapó. Nesta, os resultados alcançados foram acurácia de 96,10%, *F1 score* de 82,67% e IoU de 71,00%. Tais resultados, inferiores aos encontrados na região de Tapajós, eram esperados. As diferentes características existentes entre as áreas garimpeiras foi fator determinante na performance do algoritmo nessa região.

Na última etapa da pesquisa foi realizada a análise temporal, a partir da detecção de mudança, ocorrida nas áreas garimpeiras de Tapajós entre os anos de 2020 e 2021. Para isto, foi utilizada o esquema da Estrutura de Fluxo Duplo com Classificação Posterior e o Índice de Similaridade Estrutural com intuito de, respectivamente, detectar e quantificar a mudança entre as imagens. No conjunto de 438 imagens analisadas, obteve-se 327 imagens sem mudanças e 111 imagens com mudanças.

Diante do exposto, ratificou-se que, após a avaliação, seleção e análise de algoritmos, tanto a segmentação semântica quanto a detecção de mudança foram métodos eficientes para segmentar e

verificar a evolução temporal do status de garimpos na região amazônica, atingindo, desta forma, o objetivo geral e os objetivos específicos propostos.

Além disso, verificou-se que a pesquisa realizada tem potencial para ser utilizada como ferramenta de monitoramento de áreas garimpeiras, bem como um instrumento de amparo às tomadas de decisão de Órgãos de Proteção e Fiscalização, de Órgãos Ambientais e de Órgãos de Segurança e Controle. Para isto, no entanto, há a necessidade de se realizar o treinamento com uma base de dados de entrada que englobe áreas garimpeiras em toda amazônia, tendo em vista que uma base completa propiciará ao modelo uma maior quantidade de instâncias para aprendizagem, o que o aperfeiçoaria a segmentação das áreas garimpeiras como um todo.

Ademais, os resultados obtidos pelo modelo mostraram-se adequados para serem utilizados na identificação de novas áreas garimpeiras, na análise de locais de crescimento de garimpos, na verificação de possíveis impactos gerados por esta expansão, na evolução temporal das áreas, entre outros.

Por fim, ressalta-se que as limitações computacionais impostas pelo *Google Colab*, como tempo de processamento (máximo de 12 horas), HD (107,77 GB) e memória RAM (12,72 GB), foram determinantes nas escolhas dos hiperparâmetros de entrada para treinamento do modelo. Apesar dos resultados expressivos alcançados, a utilização de parâmetros mais robustecidos como, por exemplo, *batch* 32 e *encoder* efficientnet-b7, poderia aumentar, ainda mais, a precisão encontrada.

7 SUGESTÕES PARA TRABALHOS FUTUROS

Há, na comunidade científica, um entendimento geral de que o Aprendizado Profundo está em constante evolução. As arquiteturas, redes, *backbones*, métricas de avaliação, técnicas de processamento, entre outras atividades, surgiram e evoluíram ao longo dos anos e é notório a tendência da continuidade desse desenvolvimento. Isto posto, diante do que foi apresentado nesta dissertação, sugere-se, em complemento à pesquisa realizada, os seguintes trabalhos futuros:

a) Utilização de algoritmos de máscara de nuvens para as imagens advindas da Câmera Multiespectral e Pancromática de Ampla Varredura do CBERS 04A;

b) Treinamento, validação e teste da CNN apresentada a partir de dados de entrada de garimpos de diversas regiões da Amazônia;

c) Adição de outros índices para análise e segmentação de imagens, como o *Normalized Difference Vegetation Index* (NDVI);

d) Detecção de mudanças com a utilização de redes siamesas, tendo em vista o potencial de sua utilização em técnicas de DL; e

e) Testes com outros hiperparâmetros de entrada, a partir da utilização de computadores e/ou servidores com alta capacidade de processamento.

8 REFERÊNCIAS BIBLIOGRÁFICAS

ABIODUN, O. I. et al. State-of-the-art in artificial neural network applications: A survey. **Heliyon**, v. 4, n. 11, p. e00938, 1 nov. 2018.

ALMEIDA-FILHO, R.; SHIMABUKURO, Y. Detecting areas disturbed by gold mining activities through JERS-1 SAR images Roraima State Brazilian Amazon. **International Journal of Remote Sensing**, v. 21, 1 nov. 2000.

APAYDIN, H. et al. Comparative Analysis of Recurrent Neural Network Architectures for Reservoir Inflow Forecasting. **Water**, 24 maio 2020.

ASNER, G. P. et al. Elevated rates of gold mining in the Amazon revealed through high-resolution monitoring. **Proceedings of the National Academy of Sciences of the United States of America**, v. 110, n. 46, p. 18454–18459, 12 nov. 2013.

ASNER, G. P.; TUPAYACHI, R. Accelerated losses of protected forests from gold mining in the Peruvian Amazon. v. 12, n. 9, p. 094004, set. 2016.

AUDEBERT, N.; LE SAUX, B.; LEFÈVRE, S. Segment-before-Detect: Vehicle Detection and Classification through Semantic Segmentation of Aerial Images. **Remote Sensing**, v. 9, n. 4, p. 368, 13 abr. 2017.

BADRINARAYANAN, V.; HANDA, A.; CIPOLLA, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling. **arXiv:1505.07293 [cs]**, 27 maio 2015.

BALL, J. E.; ANDERSON, D. T.; CHAN, C. S. A 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, 23 set. 2017.

BIRKLE, C. et al. Web of Science as a data source for research on scientific and scholarly activity. **Quantitative Science Studies**, v. 1, p. 1–14, 23 jan. 2020.

BISCHL, B. et al. **Hyperparameter Optimization: Foundations, Algorithms, Best Practices and Open Challenges**. arXiv, , 24 nov. 2021. Disponível em: <<http://arxiv.org/abs/2107.05847>>. Acesso em: 24 dez. 2022

BOCHIE, K. et al. **Aprendizado Profundo em Redes Desafiadoras: Conceitos e Aplicações**. [s.l.: s.n.].

BOVOLO, F.; BRUZZONE, L. A detail-preserving scale-driven approach to change detection in multitemporal SAR images. **Geoscience and Remote Sensing, IEEE Transactions on**, v. 43, p. 2963–2972, 1 jan. 2006.

BOVOLO, F.; BRUZZONE, L. A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in the Polar Domain. **Geoscience and Remote Sensing, IEEE Transactions on**, v. 45, p. 218–236, 1 fev. 2007a.

BOVOLO, F.; BRUZZONE, L. A Split-Based Approach to Unsupervised Change Detection in Large-Size Multitemporal Images: Application to Tsunami-Damage Assessment. **Geoscience and Remote Sensing, IEEE Transactions on**, v. 45, p. 1658–1670, 1 jul. 2007b.

BROWNLEE, J. What is the Difference Between a Batch and an Epoch in a Neural Network. **Machine Learning Mastery**, v. 20, 2018.

CABALLERO ESPEJO, J. et al. Deforestation and Forest Degradation Due to Gold Mining in the Peruvian Amazon: A 34-Year Perspective. **Remote Sensing**, v. 10, n. 12, p. 1903, 29 nov. 2018.

CHAURASIA, A.; CULURCIELLO, E. **LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation**. 2017 IEEE Visual Communications and Image Processing (VCIP). **Anais...dez.** 2017. Disponível em: <<http://arxiv.org/abs/1707.03718>>. Acesso em: 13 dez. 2022

CHRISTIANSEN, N. H. et al. Comparison of Neural Network Error Measures for Simulation of Slender Marine Structures. **Journal of Applied Mathematics**, v. 2014, p. e759834, 2 mar. 2014.

DATA SCIENCE ACADEMY. **Deep Learning Book, 2022**. Disponível em: <<https://www.deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/>>. Acesso em: 15 ago. 2022.

DENG, L. A tutorial survey of architectures, algorithms, and applications for deep learning. **APSIPA Transactions on Signal and Information Processing**, v. 3, p. e2, ed 2014.

DIAKOGIANNIS, F. I. et al. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 162, p. 94–114, abr. 2020.

DONG, R.; PAN, X.; LI, F. DenseU-Net-Based Semantic Segmentation of Small Objects in Urban Remote Sensing Images. **IEEE Access**, v. 7, p. 65347–65356, 2019.

DU, B. et al. Unsupervised Deep Slow Feature Analysis for Change Detection in Multi-Temporal Remote Sensing Images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 57, n. 12, p. 9976–9992, dez. 2019.

DUCHI, J.; HAZAN, E.; SINGER, Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. **Journal of Machine Learning Research**, v. 12, n. 61, p. 2121–2159, 2011.

DUKER, A. A.; STEIN, A.; HALE, M. A statistical model for spatial patterns of Buruli ulcer in the Amansie West district, Ghana. **International Journal of Applied Earth Observation and Geoinformation**, v. 8, n. 2, p. 126–136, jun. 2006.

ELGENDY, M. **Deep Learning for Vision Systems**. [s.l.] Manning, 2020.

ELHARROUSS, O. et al. **Backbones-Review: Feature Extraction Networks for Deep Learning and Deep Reinforcement Learning Approaches**. arXiv, , 16 jun. 2022. Disponível em: <<http://arxiv.org/abs/2206.08016>>. Acesso em: 8 out. 2022

FERNANDES, F. R. C.; ALAMINO, R. DE C. J.; ARAUJO, E. R. **Recursos minerais e comunidade: impactos humanos, socioambientais e econômicos**. [s.l.] CETEM/MCTI, 2014.

FU, G. et al. Classification for High Resolution Remote Sensing Imagery Using a Fully Convolutional Network. **Remote Sensing**, v. 9, n. 5, p. 498, 18 maio 2017.

GANDHI, S. A.; KULKARNI, C. V. **MSE Vs SSIM**. Disponível em: <<https://www.ijser.org/paper/MSE-Vs-SSIM.html>>. Acesso em: 25 dez. 2022.

GAO, F.; ZHONG, H. **Study on the Large Batch Size Training of Neural Networks Based on the Second Order Gradient**. arXiv, , 16 dez. 2020. Disponível em: <<http://arxiv.org/abs/2012.08795>>. Acesso em: 24 dez. 2022

GARCIA-GARCIA, A. et al. A survey on deep learning techniques for image and video semantic segmentation. **Applied Soft Computing**, v. 70, p. 41–65, 1 set. 2018.

GHOSHAL, S. et al. Channel and Floodplain Change Analysis over a 100-Year Period: Lower Yuba River, California. **Remote Sensing**, v. 2, n. 7, p. 1797–1825, 19 jul. 2010.

GONG, M. et al. Change Detection in Synthetic Aperture Radar Images Based on Deep Neural Networks. **IEEE Transactions on Neural Networks and Learning Systems**, v. 27, n. 1, p. 125–138, jan. 2016.

GONG, M.; YANG, H.; ZHANG, P. Feature learning and change feature classification based on deep learning for ternary change detection in SAR images. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 129, p. 212–225, jul. 2017.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Disponível em: <<https://www.deeplearningbook.org/>>. Acesso em: 8 ago. 2022.

GUO, Y. et al. Deep learning for visual understanding: A review. **Recent Developments on Deep Big Vision**, v. 187, p. 27–48, 26 abr. 2016.

GUO, Y. et al. A review of semantic segmentation using deep neural networks. **International Journal of Multimedia Information Retrieval**, v. 7, n. 2, p. 87–93, 1 jun. 2018.

HAMDI, Z. M.; BRANDMEIER, M.; STRAUB, C. Forest Damage Assessment Using Deep Learning on High Resolution Remote Sensing Data. **Remote Sensing**, v. 11, n. 17, p. 1976, 22 ago. 2019.

HEYDARIAN, M.; DOYLE, T. E.; SAMAVI, R. MLCM: Multi-Label Confusion Matrix. **IEEE Access**, v. 10, p. 19083–19095, 2022.

HOESER, T.; BACHOFER, F.; KUENZER, C. Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review—Part II: Applications. **Remote Sensing**, v. 12, n. 18, p. 3053, 18 set. 2020.

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, p. 1667, 22 maio 2020.

HUTUKARA, A. Y.; YE'KWANA, A. W. **Yanomami sob ataque: garimpo ilegal na Terra Indígena Yanomami e propostas para combatê-lo. | Acervo | ISA**. [s.l.: s.n.]. Disponível em: <<https://acervo.socioambiental.org/acervo/documentos/yanomami-sob-ataque-garimpo-ilegal-na-terra-indigena-yanomami-e-propostas-para>>. Acesso em: 30 nov. 2022.

ILEA, D. E.; WHELAN, P. F. Image segmentation based on the integration of colour–texture descriptors—A review. **Semi-Supervised Learning for Visual Content Analysis and Understanding**, v. 44, n. 10, p. 2479–2501, 1 out. 2011.

INPE. **DETER**. Disponível em: <<http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/deter/deter>>. Acesso em: 17 out. 2021a.

INPE. **PRODES**. Disponível em: <<http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>>. Acesso em: 17 out. 2021b.

INPE/CBERS. **Câmeras Imageadoras CBERS 04A**. Disponível em: <<http://www.cbers.inpe.br/sobre/cameras/cbers04a.php>>. Acesso em: 19 set. 2022.

Ji, S.; WEI, S.; LU, M. Fully Convolutional Networks for Multisource Building Extraction From an Open Aerial and Satellite Imagery Data Set. **IEEE Transactions on Geoscience and Remote Sensing**, v. 57, n. 1, p. 574–586, jan. 2019.

KAISER, P. et al. Learning Aerial Image Segmentation From Online Maps. **IEEE Transactions on Geoscience and Remote Sensing**, v. 55, n. 11, p. 6054–6068, nov. 2017.

KAMPFFMEYER, M.; SALBERG, A.-B.; JENSSEN, R. **Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks**. 2016. Disponível em: <https://www.cv-foundation.org/openaccess/content_cvpr_2016_workshops/w19/html/Kampffmeyer_Semantic_Segmentation_of_CVPR_2016_paper.html>. Acesso em: 13 dez. 2022

KATUKAM, R. Image comparison Methods & Tools: A Review. 2015.

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, nov. 2018.

KHAN, S. H. et al. Forest Change Detection in Incomplete Satellite Images With Deep Neural Networks. **IEEE Transactions on Geoscience and Remote Sensing**, v. 55, n. 9, p. 5407–5423, set. 2017.

KINGMA, D. P.; BA, J. **Adam: A Method for Stochastic Optimization**. arXiv, , 29 jan. 2017. Disponível em: <<http://arxiv.org/abs/1412.6980>>. Acesso em: 24 dez. 2022

KUSSUL, N. et al. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. **IEEE Geoscience and Remote Sensing Letters**, v. PP, p. 1–5, 31 mar. 2017.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 28 maio 2015.

LIU, T. et al. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. **GIScience & Remote Sensing**, v. 55, n. 2, p. 243–264, 4 mar. 2018.

LIU, Y. et al. Hourglass-ShapeNetwork Based Semantic Segmentation for High Resolution Aerial Imagery. **Remote Sensing**, v. 9, n. 6, p. 522, 25 maio 2017.

LOBO, F. et al. Distribution of Artisanal and Small-Scale Gold Mining in the Tapajós River Basin (Brazilian Amazon) over the Past 40 Years and Relationship with Water Siltation. **Remote Sensing**, v. 8, n. 7, p. 579, 9 jul. 2016.

LOBO, F. D. L. et al. Mapping Mining Areas in the Brazilian Amazon Using MSI/Sentinel-2 Imagery (2017). **Remote Sensing**, v. 10, n. 8, p. 1178, ago. 2018.

LOBO, F. L.; COSTA, M. P. F.; NOVO, E. M. L. M. Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities. **Remote Sensing of Environment**, Special Issue: Remote Sensing of Inland Waters. v. 157, p. 170–184, 1 fev. 2015.

LONG, J.; SHELHAMER, E.; DARRELL, T. Fully Convolutional Networks for Semantic Segmentation. **arXiv:1411.4038 [cs]**, 8 mar. 2015.

LU, P. et al. Landslide mapping from multi-sensor data through improved change detection-based Markov random field. **Remote Sensing of Environment**, v. 231, p. 111235, set. 2019.

LUO, L. et al. **Adaptive Gradient Methods with Dynamic Bound of Learning Rate**. arXiv, , 26 fev. 2019. Disponível em: <<http://arxiv.org/abs/1902.09843>>. Acesso em: 24 dez. 2022

LYU, H.; LU, H.; MOU, L. Learning a Transferable Change Rule from a Recurrent Neural Network for Land Cover Change Detection. **Remote Sensing**, v. 8, n. 6, p. 506, 16 jun. 2016.

M RUSTOWICZ, R. et al. **Semantic Segmentation of Crop Type in Africa: A Novel Dataset and Analysis of Deep Learning Methods**. 2019. Disponível em: <https://openaccess.thecvf.com/content_CVPRW_2019/html/cv4gc/Rustowicz_Semantic_Segmentation_of_Crop_Type_in_Africa_A_Novel_Dataset_CVPRW_2019_paper.html>. Acesso em: 13 dez. 2022

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, jun. 2019.

MAHDIANPARI, M. et al. Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery. **Remote Sensing**, v. 10, n. 7, p. 1119, jul. 2018.

MALEKI, F. et al. Overview of Machine Learning Part 1. **Neuroimaging Clinics of North America**, v. 30, p. e17–e32, 1 nov. 2020.

MARIANO, A.; ROCHA, M. **Revisão da Literatura: Apresentação de uma Abordagem Integradora**. 3 set. 2017.

MARMANIS, D. et al. Semantic Segmentation of Aerial Images with an Ensemble of CNNs. **ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. III–3, p. 473–480, 6 jun. 2016.

MARTINS, V. S. et al. Seasonal and interannual assessment of cloud cover and atmospheric constituents across the Amazon (2000–2015): Insights for remote sensing and climate analysis. **ISPRS Journal of Photogrammetry and Remote Sensing**, SI: Latin America Issue. v. 145, p. 309–327, 1 nov. 2018.

MARZIALE, M. H. P.; MENDES, I. A. C. O fator de impacto das publicações científicas. **Revista Latino-Americana de Enfermagem**, v. 10, p. 466–467, jul. 2002.

MEER, F. VAN DER. Indicator kriging applied to absorption band analysis in hyperspectral imagery: A case study from the Rodalquilar epithermal gold mining area, SE Spain. **International Journal of Applied Earth Observation and Geoinformation**, v. 8, n. 1, p. 61–72, jan. 2006.

MIELKE, C. et al. Spaceborne Mine Waste Mineralogy Monitoring in South Africa, Applications for Modern Push-Broom Missions: Hyperion/OLI and EnMAP/Sentinel-2. **Remote Sensing**, v. 6, n. 8, p. 6790–6816, 25 jul. 2014.

MIKOŁAJCZYK, A.; GROCHOWSKI, M. **Data augmentation for improving deep learning in image classification problem**. [s.l: s.n.]. p. 122

MINAEE, S. et al. **Image Segmentation Using Deep Learning: A Survey**. arXiv, , 14 nov. 2020. Disponível em: <<http://arxiv.org/abs/2001.05566>>. Acesso em: 10 set. 2022

MMA. **No Dia da Amazônia, conheça curiosidades da maior floresta tropical do mundo e ações que unem preservação e desenvolvimento**. Disponível em: <<https://www.gov.br/mma/pt-br/assuntos/noticias/no-dia-da-amazonia-conheca-curiosidades-da-maior-floresta-tropical-do-mundo-e-acoes-que-unem-preservacao-e-desenvolvimento>>. Acesso em: 6 dez. 2022.

MME. **Boletim do Setor Mineral 2021**. , 3 mar. 2022. Disponível em: <<https://www.gov.br/mme/pt-br/assuntos/secretarias/geologia-mineracao-e-transformacao-mineral/publicacoes-1/boletim-do-setor-mineral/boletim-do-setor-mineral-2013-4o-trim-2020.pdf/view>>

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos Sobre Aprendizado de Máquina. **Sistemas Inteligentes Fundamentos e Aplicações**, n. 1, p. 89--114, 2003.

MOORE, F. et al. Mapping mineralogical alteration using principal-component analysis and matched filter processing in the Takab area, north-west Iran, from ASTER data. **International Journal of Remote Sensing**, v. 29, n. 10, p. 2851–2867, 20 maio 2008.

MOSS, G.; MOSS, M. **Os rios voadores, a Amazônia e o clima brasileiro**. São Paulo: Horizonte, 2014.

MOU, L. et al. Multitemporal Very High Resolution From Space: Outcome of the 2016 IEEE GRSS Data Fusion Contest. **IEEE Journal of Selected Topics in Applied Earth Observations and**

Remote Sensing, v. 10, n. 8, p. 3435–3447, ago. 2017.

NG, D.; FENG, M. Medical Image Recognition: An Explanation and Hands-On Example of Convolutional Networks. Em: CELI, L. A. et al. (Eds.). **Leveraging Data Science for Global Health**. Cham: Springer International Publishing, 2020. p. 263–284.

OBODAI, J. et al. Land use/land cover dynamics using landsat data in a gold mining basin-the Ankobra, Ghana. **Remote Sensing Applications: Society and Environment**, v. 13, p. 247–256, jan. 2019.

PADILLA, R.; NETTO, S.; DA SILVA, E. **A Survey on Performance Metrics for Object-Detection Algorithms**. [s.l: s.n.].

PASSOS, C. et al. **Garimpo ilegal: mortes em consequência e massacre na Terra Indígena Yanomami**. [s.l: s.n.]. Disponível em: <<https://cptnacional.org.br/downloads?task=download.send&id=14249&catid=89&m=0>>. Acesso em: 30 nov. 2022.

PENG, D.; ZHANG, Y.; GUAN, H. End-to-End Change Detection for High Resolution Satellite Images Using Improved UNet++. **Remote Sensing**, v. 11, n. 11, p. 1382, 10 jun. 2019.

PEREZ, L.; WANG, J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning. 13 dez. 2017.

PHUNG; RHEE. A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets. **Applied Sciences**, v. 9, n. 21, p. 4500, 23 out. 2019.

PYTORCH. **PyTorch documentation — PyTorch 1.12 documentation**. Disponível em: <<https://pytorch.org/docs/stable/index.html>>. Acesso em: 21 set. 2022.

RADKE, R. J. et al. Image change detection algorithms: a systematic survey. **IEEE Transactions on Image Processing**, v. 14, n. 3, p. 294–307, mar. 2005.

RAISG. **Amazônia saqueada**. Disponível em: <<https://saqueada.amazoniasocioambiental.org/story>>. Acesso em: 17 out. 2021.

REZAEI, M. et al. Deep Convolutional Neural Network for Complex Wetland Classification Using Optical Remote Sensing Imagery. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, 2018.

RODRIGUES, R. M. et al. **Estudo dos impactos ambientais decorrentes do extrativismo mineral e poluição mercurial no Tapajós: pré-diagnóstico**. [s.l.] CETEM, 1994.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. **arXiv:1505.04597 [cs]**, 18 maio 2015.

RŮŽIČKA, V. et al. **Deep Active Learning in Remote Sensing for data efficient Change Detection**. arXiv, , 25 ago. 2020. Disponível em: <<http://arxiv.org/abs/2008.11201>>. Acesso em: 9 out. 2022

SARA, U.; AKTER, M.; UDDIN, M. S. Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. **Journal of Computer and Communications**, v. 7, n. 3, p. 8–18, 4 mar. 2019.

SHI, W. et al. Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges. **Remote Sensing**, v. 12, n. 10, p. 1688, 25 maio 2020.

SILVA, R. E. V. DA. Um estudo comparativo entre redes neurais convolucionais para a classificação de imagens. 2018.

SIMONYAN, K.; ZISSERMAN, A. **Very Deep Convolutional Networks for Large-Scale Image Recognition**. arXiv, , 10 abr. 2015. Disponível em: <<http://arxiv.org/abs/1409.1556>>. Acesso em: 13 dez. 2022

SINGH, A. Review Article Digital change detection techniques using remotely-sensed data. **International Journal of Remote Sensing**, v. 10, n. 6, p. 989–1003, jun. 1989.

SINGH, R.; RANI, R. **Semantic Segmentation using Deep Convolutional Neural Network: A Review**. Rochester, NY, 1 abr. 2020. Disponível em: <<https://papers.ssrn.com/abstract=3565919>>. Acesso em: 13 dez. 2022

SNAPIR, B.; SIMMS, D. M.; WAINE, T. W. Mapping the expansion of galamsey gold mines in the cocoa growing area of Ghana using optical remote sensing. **International Journal of Applied Earth Observation and Geoinformation**, v. 58, p. 225–233, jun. 2017.

SONG, A. et al. Change Detection in Hyperspectral Images Using Recurrent 3D Fully Convolutional Networks. **Remote Sensing**, v. 10, n. 11, p. 1827, 17 nov. 2018.

SOUZA, A. et al. Metodologia Utilizada nos Projetos PRODES e DETER. p. 33, 2019.

STEVENS, E.; ANTIGA, L.; VIEHMANN, T. **Deep Learning with PyTorch**. [s.l.] Simon and Schuster, 2020.

SUN, W.; WANG, R. Fully Convolutional Networks for Semantic Segmentation of Very High Resolution Remotely Sensed Images Combined With DSM. **IEEE Geoscience and Remote**

Sensing Letters, v. 15, n. 3, p. 474–478, mar. 2018.

SUTSKEVER, I. et al. **On the importance of initialization and momentum in deep learning**. Proceedings of the 30th International Conference on Machine Learning. **Anais...PMLR**, 26 maio 2013. Disponível em: <<https://proceedings.mlr.press/v28/sutskever13.html>>. Acesso em: 24 dez. 2022

TAGHANAKI, S. et al. **Deep Semantic Segmentation of Natural and Medical Images: A Review**. [s.l: s.n.].

TAYLOR, L.; NITSCHKE, G. Improving Deep Learning using Generic Data Augmentation. 20 ago. 2017.

TELMER, K.; VEIGA, M. World emissions of mercury from artisanal and small scale gold mining. Em: **Mercury Fate and Transport in the Global Atmosphere**. [s.l: s.n.]. p. 131–172.

TORKY, A.; ABURAWWASH, A. A Deep Learning Approach to Automated Structural Engineering of Prestressed Members. **International Journal of Structural and Civil Engineering Research**, p. 347–352, 1 jan. 2018.

WANG, H. et al. Gated Convolutional Neural Network for Semantic Segmentation in High-Resolution Images. **Remote Sensing**, v. 9, n. 5, p. 446, 5 maio 2017a.

WANG, J. et al. Associations among Image Assessments as Cost Functions in Linear Decomposition: MSE, SSIM, and Correlation Coefficient. 4 ago. 2017b.

WANG, Q. et al. Change detection based on Faster R-CNN for high-resolution remote sensing images. **Remote Sensing Letters**, v. 9, n. 10, p. 923–932, 3 out. 2018.

WANG, Q. et al. GETNET: A General End-to-End 2-D CNN Framework for Hyperspectral Image Change Detection. **IEEE Transactions on Geoscience and Remote Sensing**, v. 57, n. 1, p. 3–13, jan. 2019.

WANG, Z. et al. Image Quality Assessment: From Error Visibility to Structural Similarity. **IEEE Transactions on Image Processing**, v. 13, n. 4, p. 600–612, abr. 2004.

WURM, M. et al. Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 150, p. 59–69, abr. 2019.

XU, Y. et al. Building Extraction in Very High Resolution Remote Sensing Imagery Using Deep Learning and Guided Filters. **Remote Sensing**, v. 10, n. 1, p. 144, 19 jan. 2018a.

XU, Y. et al. Road Extraction from High-Resolution Remote Sensing Imagery Using Deep Learning. **Remote Sensing**, v. 10, n. 9, p. 1461, 13 set. 2018b.

YAKUBOVSKIY, P. **Welcome to Segmentation Models's documentation! — Segmentation Models documentation**. Disponível em: <<https://smp.readthedocs.io/en/latest/>>. Acesso em: 18 dez. 2022.

YAMASHITA, R. et al. **Convolutional neural networks: an overview and application in radiology**. Springer Nature, , 2018a.

YAMASHITA, R. et al. Convolutional neural networks: an overview and application in radiology. **Insights into Imaging**, v. 9, n. 4, p. 611–629, ago. 2018b.

YANG, L.; SHAMI, A. On hyperparameter optimization of machine learning algorithms: Theory and practice. **Neurocomputing**, v. 415, p. 295–316, 20 nov. 2020.

YANG, M.-D. et al. Semantic Segmentation Using Deep Learning with Vegetation Indices for Rice Lodging Identification in Multi-date UAV Visible Images. **Remote Sensing**, v. 12, n. 4, p. 633, jan. 2020.

YANI, M.; BUDHI IRAWAN, S., Si. ., M. T.; CASI SETININGSIH, S. T., M. T. Application of Transfer Learning Using Convolutional Neural Network Method for Early Detection of Terry's Nail. **Journal of Physics: Conference Series**, v. 1201, p. 012052, maio 2019.

ZAJMI, L.; AHMED, F. Y. H.; JAHARADAK, A. A. Concepts, Methods, and Performances of Particle Swarm Optimization, Backpropagation, and Neural Networks. **Applied Computational Intelligence and Soft Computing**, v. 2018, p. e9547212, 3 set. 2018.

ZHANG, C. et al. 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, p. 189, 11 abr. 2019.

ZHANG, H. et al. Feature-Level Change Detection Using Deep Representation and Feature Change Analysis for Multispectral Imagery. **IEEE Geoscience and Remote Sensing Letters**, v. 13, n. 11, p. 1666–1670, nov. 2016.

ZHANG, Z.; LIU, Q.; WANG, Y. Road Extraction by Deep Residual U-Net. **IEEE Geoscience and Remote Sensing Letters**, v. 15, n. 5, p. 749–753, maio 2018.

ZHAO, Z.-Q. et al. Object Detection With Deep Learning: A Review. **IEEE Transactions on Neural Networks and Learning Systems**, v. 30, n. 11, p. 3212–3232, nov. 2019.

ZHONG, L.; HU, L.; ZHOU, H. Deep learning based multi-temporal crop classification. **Remote Sensing of Environment**, v. 221, p. 430–443, 1 fev. 2019.

ZHOU, Z. et al. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. arXiv, , 18 jul. 2018. Disponível em: <<http://arxiv.org/abs/1807.10165>>. Acesso em: 13 dez. 2022

ZHU, X. X. et al. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. **IEEE Geoscience and Remote Sensing Magazine**, v. 5, n. 4, p. 8–36, dez. 2017.