



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA- PPGEE

LUENA OSSANA CANAVIEIRA

**CLASSIFICAÇÃO DE REGIÕES DE DESMATAMENTO VIA IMAGENS DO
SATÉLITE LANDSAT NO NORDESTE DO PARÁ**

DM 38/2023

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém – Pará – Brasil
2023

LUENA OSSANA CANAVIEIRA

**CLASSIFICAÇÃO DE REGIÕES DE DESMATAMENTO VIA IMAGENS DO
SATÉLITE LANDSAT NO NORDESTE DO PARÁ**

AUTORA: LUENA OSSANA CANAVIEIRA

DM 38/2023

Dissertação apresentada à Universidade Federal do Pará, como parte das exigências do Programa de Pós-Graduação em Engenharia Elétrica, para obtenção do título de Mestre.

Área de Concentração: Computação Aplicada

Orientador: Dr. João Crsóstomo Weyl Albuquerque Costa

Coorientador: Dr. Reginaldo Cordeiro dos Santos Filho

UFPA/ITEC/PPGEE

Campus Universitário do Guamá

Belém – Pará – Brasil

2023

**CLASSIFICAÇÃO DE REGIÕES DE DESMATAMENTO VIA IMAGENS DO SATÉLITE
LANDSAT NO NORDESTE DO PARÁ**

AUTORA: LUENA OSSANA CANAVIEIRA

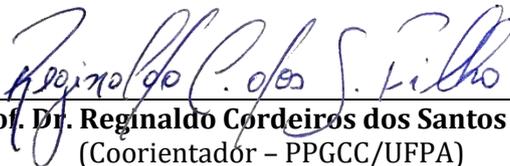
DISSERTAÇÃO DE MESTRADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE MESTRA EM ENGENHARIA ELÉTRICA NA ÁREA DE COMPUTAÇÃO APLICADA.

APROVADA EM: 18/12/2023

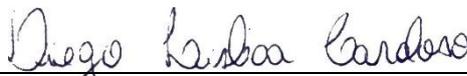
BANCA EXAMINADORA:



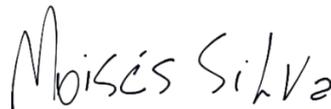
Prof. Dr. João Crisóstomo Weyl Albuquerque Costa
(Orientador – PPGEE/UFPA)



Prof. Dr. Reginaldo Cordeiros dos Santos Filho
(Coorientador – PPGCC/UFPA)

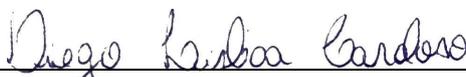


Prof. Dr. Diego Lisboa Cardoso
(Avaliador Interno – PPGEE/UFPA)



Prof. Dr. Moisés Felipe Mello da Silva
(Avaliador Externo – LOS ALAMOS NATIONAL LAB/EUA)

VISTO:



Prof. Dr. Diego Lisboa Cardoso
(Coordenador do PPGEE/ITEC/UFPA)

*Dedico essa dissertação ao meu pai José Ovidio Oliveira Canavieira,
minha mãe Ednilsa Alves Ossana, meu filho Thor e minha amada Raíssa Coutinho.*

AGRADECIMENTOS

Primeiramente, agradeço aos meus pais Ednilsa Alves Ossana e José Ovidio Oliveira Canavieira. Obrigada por sempre estarem ao meu lado, principalmente ao meu pai que não está presente fisicamente, mas sinto seu amor todos os dias, essa vitória não é somente minha, é nossa. Os amo demais.

Não poderia deixar de mencionar o meu filho de quatro patas, Thor. Muitas vezes, pensei que não conseguiria encontrar forças e você me motivou. Obrigada por me fazer ser uma pessoa melhor, trazendo paz aos meus dias.

À minha avó Raimunda Coelho, Tia Célia e Tio Carlos, por serem pessoas especiais na minha vida.

Agradeço à minha companheira de vida, Raíssa Coutinho por cuidar de mim, ser meu abrigo e apoio. Desejo compartilhar todos os dias da minha ao teu lado e tenho a certeza que a nossa Melzinha está ao nosso lado.

Aos meus irmãos de coração Emily Chagas e Moisés Felipe. Crescer ao lado de vocês é um privilégio e sou imensamente grata poder compartilhar cada vitória juntos. Sinto saudades todos os dias.

Aos meus amigos Laisa Epifanio, Luisa Epifanio, Lucienne Amaral, Fernanda Carvalho, Adler Henrique, Layze Amoras, Taianne Melo, Brenda Pinheiro, Romulo Carvalho, Cristina Rêgo e Yuri Rêgo, por cada um fazer parte da minha jornada. Obrigada pela amizade.

À Selma Solange por me apoiar desde a minha entrada na SEMAS, além de ser uma grande amiga, és um exemplo de profissional e pessoa. Obrigada por tudo.

Aos companheiros de trabalho da GEAR Adriane da Rocha, Jaqueline Barbosa, Matheus Almeida, Marcus Vinicius e Yuri Aleph Dias. Obrigada por trazerem leveza na minha rotina, cresci como ser humano e profissional, principalmente, esse ano ao lado vocês.

RESUMO

O desmatamento é um dos principais problemas ambientais enfrentados no mundo, e na Amazônia a perda da cobertura vegetal tem graves consequências para o clima, biodiversidade e sociedade. Nesse contexto, a detecção e classificação de áreas desmatadas é fundamental para o monitoramento e controle do desmatamento. O objetivo deste trabalho foi identificar áreas degradadas e não degradadas na região nordeste paraense por meio de imagens do satélite Landsat 5, 7 e 8, utilizando-se banco de dados de 210 imagens, através de técnicas de aprendizado profundo. Após os resultados foi possível gerar a matriz de confusão, para realizar a classificação digital para avaliação dos resultados das imagens de satélite. Por meio da matriz confusão estimou-se com 69,7% de acurácia os resultados. Os valores determinaram um índice eficaz de classificação das áreas degradadas e não degradadas. O modelo obtido para o treinamento apresentou algumas interferências na imagens sendo possível identificar áreas desmatadas e não desmatadas. Para estudos futuros recomenda-se novas imagens de satélite da região, com destaque para os índices de resolução espectral correspondentes com os sensores determinados pelos satélites do estudo.

Palavras-Chave: Redes neurais; aprendizado profundo; desmatamento ;sensoriamento remoto

ABSTRACT

Logging is one of the major environmental problems in the world, and primarily, in the Amazon forest, the consequences for the climate, biodiversity and society are severe. In this context, logging detection and classification is preemptory for control. The goal of this work is to identify degraded and non-degraded areas with artificial neural networks from satellite images collected from the Landsat 5, 7 and 8 in the northeast region of the Para state. The dataset is composed of imagery from May 05, 2000 through May 17, 2022, matching 210 images in total. The results were evaluated in terms of the confusion matrix using the aforementioned satellite dataset. The results were created following the knowledge discovery in databases process, implemented on the Python language in the Google Collaboratory environment. From the confusion matrix, the resulting sensitivity, specificity, accuracy, efficiency, positive/negative predictive values are 78.9%, 51.1%, 69.7%, 68%, 71.4% and 66.7%, respectively. The developed model presented a effective classification index for degraded and non-degraded areas. The model introduced with effective indexes for training showed interference in the images that reduced few the precision of the approach. In further studies is recommended a larger dataset of the area, highlighting the spectral resolution to the indexes corresponding to the sensors determined by the satellites used in this evaluation.

Key-Words: neural networks; deep learning; deforestation; remote sensing.

LISTA DE FIGURAS

Figura 1 - Apresentação da rede neural	13
Figura 2 - Algoritmo de Canny aprimorado	20
Figura 3 - Rede neural de amostra	23
Figura 4 - Mapa de localização da área de estudo referente a Mesoregião Nordeste paraense	29
Figura 5 - Áreas desmatadas segundo o PRODES 2023 na região do Nordeste paraense	32
Figura 6 - Áreas não desmatadas segundo o PRODES 2023 na região do Nordeste paraense	33
Figura 7 - Conhecimento em mineração de dados	35

LISTA DE GRÁFICOS

Gráfico 1 – Quantitativo do número de imóveis rurais regularizados nas zonas de integração no Estado do Pará	30
Gráfico 2 – Zonas de integração no Estado do Pará referente aos imóveis em regularização ambiental	31
Gráfico 3 – Evolução do rebanho bovino no estado do Pará entre 2000 e 2014	34
Gráfico 4 - Parâmetros de função de perda obtidos após desempenho do modelo em tentativas estimadas durante a execução do algoritmo	39
Gráfico 5 - Parâmetros de alteração nos valores de acurácia da rede neural convolucional, após exxeções destacando o desempenho do modelo.	40

LISTA DE TABELAS

Tabela 1 - Unidade de Medidas.....	15
Tabela 2 - Matriz de confusão da classificação da rede neural convolucional desenvolvida no estudo.....	40

SUMÁRIO

1. INTRODUÇÃO	1
1.1 Contextualização do trabalho	2
1.1.1 Desmatamento na Amazônia	4
1.2 Motivação	5
1.3 Justificativa	6
1.4 Objetivos	7
1.4.1 Objetivo geral	7
1.4.2 Objetivos específicos	7
2. EMBASAMENTO TEÓRICO	10
2.1 Inteligência Artificial	10
1. INTRODUÇÃO	1
1.1 Contextualização do trabalho	2
1.1.1 Desmatamento na Amazônia	4
1.2 Motivação	5
1.3 Justificativa	6
1.4 Objetivos	7
1.4.1 Objetivo geral	7
1.4.2 Objetivos específicos	7
2. EMBASAMENTO TEÓRICO	10
2.1 Inteligência Artificial	10
2.2 Aprendizado de máquinas (<i>machine learning</i>)	11
2.3 Aprendizagem profunda e rede neural convolucional	13
2.4 Sensoriamento remoto e processamento de imagens	14

2.5 Técnicas de segmentação de imagens	17
2.6 Comportamento espectral da vegetação	24
2.6.1 O Índice de vegetação	25
2.7 Descoberta de conhecimento em banco de dados	27
3 MATERIAL E MÉTODOS	29
3.1 Área de estudo	29
3.2 Banco de dados	34
3.3 Processamento dos dados	35
3.4 Métricas avaliativas	37
4 RESULTADOS E DISCUSSÃO	39
5 CONCLUSÃO	43
REFERÊNCIAS	45
2.3 Aprendizagem profunda e rede neural convolucional	13
2.4 Sensoriamento remoto e processamento de imagens	14
2.5 Técnicas de segmentação de imagens	17
1. INTRODUÇÃO	1
1.1 Contextualização do trabalho	2
1.1.1 Desmatamento na Amazônia	4
1.2 Motivação	5
1.3 Justificativa	6
1.4 Objetivos	7
1.4.1 Objetivo geral	7
1.4.2 Objetivos específicos	7
2. EMBASAMENTO TEÓRICO	10
2.1 Inteligência Artificial	10
2.2 Aprendizado de máquinas (<i>machine learning</i>)	11
2.3 Aprendizagem profunda e rede neural convolucional	13

2.4 Sensoriamento remoto e processamento de imagens	14
2.5 Técnicas de segmentação de imagens	17
2.6 Comportamento espectral da vegetação	24
2.6.1 Índice de vegetação por diferença normalizada	25
2.6.2 Índice de vegetação ajustado ao solo	26
2.6.3 índice de área foliar	27
2.7 Descoberta de conhecimento em banco de dados	27
3 MATERIAL E MÉTODOS	29
3.1 Área de estudo	29
3.2 Banco de dados	34
3.3 Processamento dos dados	35
3.4 Métricas avaliativas	37
4 RESULTADOS E DISCUSSÃO	39
5 CONCLUSÃO	43
REFERÊNCIAS	45
2.7 Descoberta de conhecimento em banco de dados	27
3 MATERIAL E MÉTODOS	29
3.1 Área de estudo	29
3.2 Banco de dados	34
3.3 Processamento dos dados	35
3.4 Métricas avaliativas	37
4 RESULTADOS E DISCUSSÃO	39
5 CONCLUSÃO	43
REFERÊNCIAS	45

1. INTRODUÇÃO

O desmatamento é um dos principais problemas ambientais enfrentados em todo o mundo. Na Amazônia, a perda da cobertura florestal tem graves consequências para o clima, a biodiversidade e as populações locais. A detecção e classificação das áreas desmatadas é fundamental para o monitoramento e controle do desmatamento. Nesse contexto, as técnicas de aprendizado profundo têm sido amplamente utilizadas para a classificação de imagens de satélite em regiões de desmatamentos (GUEDES *et al.*, 2019).

O aprendizado profundo é uma subárea da inteligência artificial que utiliza redes neurais artificiais para aprender a partir de grandes quantidades de dados. Essas técnicas têm se mostrado eficazes para a classificação de imagens de satélite, permitindo a identificação de áreas desmatadas com alta precisão e rapidez (SANTOS *et al.*, 2021).

A classificação de imagens de satélite Landsat é uma das principais aplicações das técnicas de aprendizado profundo para a detecção de desmatamentos na Amazônia. As imagens de satélite Landsat permitem uma cobertura temporal e espacial ampla, o que é fundamental para o monitoramento do desmatamento em larga escala. As técnicas de aprendizado profundo são capazes de detectar áreas desmatadas em diferentes estágios de regeneração, o que contribui para o entendimento dos processos de mudança da cobertura florestal na região (GUPTA *et al.*, 2021).

Além disso, a aplicação de técnicas de aprendizado profundo na classificação de imagens de satélite Landsat tem a vantagem de reduzir a necessidade de intervenção humana no processo de detecção e classificação de áreas desmatadas. Isso torna o processo mais rápido e eficiente, permitindo uma resposta mais ágil às mudanças na cobertura florestal (SILVA *et al.*, 2021).

Nesse sentido, a aplicação de técnicas de aprendizado profundo na classificação de imagens de satélite Landsat é uma importante ferramenta para o monitoramento e controle do desmatamento na Amazônia. Essas técnicas permitem a identificação de áreas desmatadas, bem como a detecção de alterações no uso da terra ao longo do tempo.

O aprendizado profundo envolve a utilização de redes neurais artificiais com múltiplas camadas, que são capazes de aprender a partir de grandes quantidades de dados e identificar padrões complexos em imagens de satélite. Com isso, é possível realizar a análise de grandes áreas de forma mais rápida e precisa do que os métodos tradicionais de análise de imagens de satélite.

De acordo com um estudo publicado por Almeida *et al.* (2008), a aplicação de técnicas de aprendizado profundo na classificação de imagens de satélite Landsat permitiu uma precisão média de 96,26% na detecção de áreas desmatadas. Além disso, essas técnicas também foram capazes de identificar áreas de degradação florestal, que são áreas onde a floresta ainda está presente, mas foi severamente danificada por atividades humanas. A identificação dessas áreas é importante para o monitoramento da saúde da floresta e para a implementação de políticas de conservação (LI *et al.*, 2020).

Outra vantagem do uso de técnicas de aprendizado profundo na análise de imagens de satélite é a capacidade de identificar mudanças no uso da terra ao longo do tempo. Isso permite a detecção de atividades como a expansão da agricultura e a construção de estradas, que são fatores que contribuem para o desmatamento. Com essa informação, é possível implementar medidas de controle e fiscalização mais efetivas, visando a redução das taxas de desmatamento na região amazônica (ZHU *et al.*, 2017).

1.1 Contextualização do trabalho

O desmatamento é causado por uma série de fatores, entre eles o avanço da agropecuária, a expansão das estradas e a exploração de madeira. O desmatamento para fins agropecuários é um dos principais motivos do desmatamento na Amazônia (FREITAS *et al.*, 2020). Os autores afirmam ainda que a soja é um dos principais cultivos que impulsionam na região.

Além da perda de biodiversidade, o desmatamento pode ter consequências graves para o clima global. Estudos mostram que a Amazônia é um importante sumidouro de carbono, e que sua degradação pode levar à emissão de grandes quantidades de gases de efeito estufa (DAVIDSON *et al.*, 2012). Segundo os autores, a degradação da floresta pode levar a um aumento de até 50% nas emissões de gases de efeito estufa em relação à taxa atual.

A falta de políticas públicas efetivas para o combate a essa prática é apontada como um dos principais obstáculos para a redução da taxa de desmatamento no Brasil (BUTLER, 2019). Butler (2019) afirma que a ausência de medidas concretas do governo brasileiro para combater o desmatamento, aliada à falta de recursos e pessoal para fiscalização, tem contribuído para o seu aumento na região amazônica.

Por isso, essa prática no Brasil é um problema complexo e multifacetado, que exige a implementação de políticas públicas efetivas para sua redução. É necessário um esforço conjunto de governos, da sociedade civil e do setor privado para combater o desmatamento e garantir a conservação das florestas e da biodiversidade, bem como a mitigação das mudanças climáticas.

A implementação de medidas como a regularização fundiária, o monitoramento por satélite, a intensificação da fiscalização e a promoção de alternativas econômicas sustentáveis para as comunidades locais são algumas das estratégias que podem contribuir para a redução do desmatamento (BARRETO *et al.*, 2018).

Além disso, a educação ambiental e a conscientização da população sobre a importância das florestas e dos serviços ecossistêmicos que elas fornecem são fundamentais para a construção de uma sociedade mais sustentável e consciente. É importante que a sociedade como um todo se mobilize para cobrar ações concretas do governo e do setor privado em relação ao combate ao desmatamento no Brasil.

Em resumo, o desmatamento é um problema ambiental complexo e de grande relevância no Brasil. A implementação de políticas públicas efetivas, o monitoramento constante, a conscientização da população e a busca por alternativas econômicas sustentáveis são algumas das estratégias para enfrentar esse desafio. É necessário um esforço conjunto de todos os setores da sociedade para garantir a conservação das florestas e da biodiversidade, bem como a mitigação das mudanças climáticas.

1.1.1 Desmatamento na Amazônia

O desmatamento na Amazônia é um dos principais problemas ambientais enfrentados pelo Brasil e tem sido objeto de estudo e discussão por diversos pesquisadores e desde a década de 1970, a Amazônia tem sido alvo de forte pressão devido à expansão da agricultura, pecuária, mineração e extração de madeira, o que tem levado a uma perda significativa da cobertura florestal na região (FEARNSIDE, 2021).

Segundo o Instituto Nacional de Pesquisas Espaciais (INPE), a taxa de desmatamento na Amazônia aumentou em 9,5% em 2020 em relação ao ano anterior, totalizando uma área desmatada de 11.088 km² (INPE, 2021). As principais causas dessa atividade na região são a expansão da agricultura e pecuária, a exploração de madeira e a mineração (NEPSTAD *et al.*, 2014).

Além dos impactos imediatos na perda da biodiversidade e alteração do clima local, essa prática também tem implicações globais. A floresta amazônica é considerada um importante regulador do clima mundial devido ao seu papel na absorção de dióxido de carbono da atmosfera. A perda da floresta amazônica pode levar a um aumento significativo da concentração de gases de efeito estufa na atmosfera, contribuindo para o aquecimento global (FEARNSIDE, 2021).

A implementação de políticas públicas efetivas é fundamental para a redução do desmatamento na Amazônia. Entre as medidas propostas, está a regularização fundiária, que visa combater a grilagem de terras e a ocupação desordenada da região. A fiscalização e o monitoramento das atividades econômicas na região também são essenciais para coibir atividades ilegais como a extração de madeira e a mineração clandestina. Além disso, a promoção de atividades econômicas sustentáveis, como o manejo florestal e a agroecologia, pode contribuir para a conservação da floresta e o desenvolvimento socioeconômico da região (PINTO *et al.*, 2020).

Outra medida importante é o fortalecimento das unidades de conservação e terras indígenas, que desempenham um papel fundamental na proteção da biodiversidade e da cultura local. A demarcação e titulação de terras indígenas também é uma demanda histórica dessas comunidades, que enfrentam ameaças constantes de invasão e exploração de seus territórios (FEARNSIDE, 2017).

Em síntese, a redução do desmatamento na Amazônia requer ações integradas e coordenadas, que envolvam desde a regularização fundiária até a promoção de atividades econômicas sustentáveis e a proteção das unidades de conservação e terras indígenas.

Para isso, é necessário um esforço conjunto de governos, organizações não governamentais, empresas e sociedade civil para enfrentar o desafio do desmatamento na Amazônia. A construção de parcerias e diálogos entre esses atores pode contribuir para a implementação de políticas públicas mais efetivas e para o engajamento da sociedade na conservação da floresta (NEPSTAD *et al.*, 2014).

1.2 Motivação

A motivação para este estudo é multifacetada e busca abordar diversos desafios relacionados ao desmatamento na região amazônica. Um dos principais objetivos deste estudo é contribuir para a redução dos índices de desmatamento na região amazônica.

Através da aplicação de técnicas de aprendizado profundo, busca-se desenvolver um sistema eficiente e preciso para a detecção e classificação de áreas desmatadas em imagens de satélite Landsat. Identificar essas áreas de desmatamento de forma automatizada e precisa é um passo fundamental para melhorar as estratégias de conservação e promover ações imediatas de preservação ambiental.

Ao automatizar o processo de detecção de áreas desmatadas, a pesquisa também busca acelerar o processo de descoberta e monitoramento dessas áreas. Através da utilização de técnicas de aprendizado profundo, é possível processar grandes volumes de imagens de satélite, permitindo uma detecção mais rápida e eficiente de áreas desmatadas. Isso possibilita uma resposta mais ágil por parte das autoridades e instituições responsáveis, permitindo ações de fiscalização, combate ao desmatamento ilegal e implementação de políticas de preservação de forma mais eficaz.

Nesse contexto, o presente estudo busca não apenas identificar áreas desmatadas como apresentar ferramentas de aprendizado de máquina como uma alternativa mais rápida e eficaz na hora da classificação das imagens.

A tecnologia de sensoriamento remoto, em particular a utilização de imagens de satélite, tem se mostrado uma ferramenta poderosa para o monitoramento de áreas florestais e detecção de desmatamento. No entanto, a análise manual das imagens pode ser demorada e sujeita a erros. A aplicação de técnicas de aprendizado profundo pode automatizar esse processo, tornando-o mais rápido, preciso e escalável (GOMEZ *et al.*, 2016; LAMBIN *et al.*, 2018).

Por fim, a aplicação de técnicas de aprendizado profundo para a classificação de imagens de satélite tem sido objeto de intensa pesquisa nos últimos anos, com resultados promissores em diversas áreas, incluindo a detecção de desmatamento. A realização de estudos como esse pode contribuir para o desenvolvimento de metodologias mais eficazes e precisas para a detecção de áreas desmatadas, possibilitando uma gestão mais eficiente e sustentável dos recursos naturais (COCCOLO *et al.*, 2018; LI *et al.*, 2019).

1.3 Justificativa

A aplicação de técnicas de aprendizado profundo para a classificação de imagens de satélite Landsat em regiões de desmatamento no nordeste do Pará é um estudo de extrema importância devido à necessidade de monitorar e combater o desmatamento, um problema ambiental crítico e de grande relevância global.

Embora existam inúmeros trabalhos na literatura que abordam a detecção de desmatamento utilizando técnicas de aprendizado de máquina, o diferencial deste trabalho reside na aplicação específica de técnicas de aprendizado profundo.

Além disso, o nordeste do Pará é uma região conhecida por sua alta taxa de desmatamento, tornando-se um local de interesse significativo para a comunidade científica e para os esforços de conservação ambiental.

Outro diferencial deste estudo pode estar relacionado com a utilização de dados obtidos por satélites Landsat. Esses dados são amplamente utilizados em estudos ambientais devido à sua resolução espacial e temporal, além de sua disponibilidade pública. No entanto, a classificação de imagens Landsat para detecção de desmatamento ainda é um desafio, devido à complexidade das informações contidas nas imagens e à necessidade de lidar com grandes volumes de dados. O uso de técnicas de aprendizado profundo pode ajudar a superar esses

desafios, permitindo uma análise mais precisa e eficiente dos dados de satélite Landsat para a detecção de áreas desmatadas.

É importante ressaltar que no ano de 2015, através do Decreto Estadual 1.379, obteve-se a implementação do Programa de Regularização Ambiental na Secretaria de Estado de Meio Ambiente e Sustentabilidade - SEMAS, onde os imóveis rurais que possuem passivos ambientais, devem recuperar as áreas desmatadas. Logo, este trabalho fez o uso do dados da Secretaria, sendo identificados os imóveis que se adequaram ambientalmente.

A partir dos dados apresentados, foi identificado que os municípios do Nordeste Paraense possuem elevado número de rebanho bovino (IBGE,2023). Porém, os municípios do Nordeste são os que menos possuem adesão ao Programa de Regularização, ou seja, não buscam regularização ambiental, impulsionando a importância do monitoramento da região.

Sendo assim, as alterações da cobertura vegetal do Nordeste Paraense causadas pelas atividades agropecuárias são problemas que precisam ser monitorados e são objetos do presente estudo. Diante disto, o problema da pesquisa consiste em classificar áreas desmatadas e não desmatadas por meio da aprendizagem profunda e redes neurais convolucionais. A principal hipótese do trabalho é que a aprendizagem profunda pode gerar produtos classificados com alta acurácia.

1.4 Objetivos

1.4.1 Objetivo geral

Identificar áreas degradadas e não degradadas no nordeste Paraense, através de imagens de satélite Landsat, utilizando a metodologia de aprendizado de máquina por meio de redes neurais artificiais de aprendizado profundo.

1.4.2 Objetivos específicos

- Aplicar Sensoriamento Remoto para o tratamento de imagens de satélite e Inteligência Artificial para a classificação das imagens;

- Utilizar imagens de satélite, especificamente sensor LandSat 7 e LandSat 8 para a análise do comportamento espectral da vegetação;
- Treinar a rede neural para a identificação de áreas desmatadas e não desmatadas.

1.5 Trabalhos relacionados

Um estudo relevante nesse contexto de monitoramento de áreas através de aprendizado profundo é o de Vale, Delgado e Silva (2022), que exploraram a classificação de imagens hiperespectrais utilizando redes neurais convolucionais para caracterizar a ocupação desordenada do solo em um recorte do Parque Nacional da Tijuca, RJ. Os autores demonstraram que as redes neurais convolucionais são eficazes na identificação e mapeamento de áreas desmatadas, fornecendo resultados precisos e detalhados sobre a degradação florestal.

Outro estudo importante é o de Kuck et al. (2021), que utilizaram dados SAR multitemporais em banda X e técnicas de Aprendizado de Máquina para mapear a degradação florestal causada pela exploração seletiva de madeira na Amazônia Legal. Os resultados mostraram que a combinação dessas técnicas permite identificar áreas de desmatamento com alta precisão, contribuindo para a detecção e o monitoramento eficaz da degradação florestal.

Para otimizar o processo de classificação de imagens e melhorar os resultados obtidos, técnicas como a otimização por enxame de partículas (Particle Swarm Optimization - PSO) têm sido aplicadas. Estudos como os de Shi e Eberhart (1999), Santos et al. (2018) e Santos et al. (2019) exploraram o PSO em diferentes contextos de otimização, demonstrando sua eficácia na busca por soluções ótimas.

Além das técnicas de aprendizado profundo e otimização, é fundamental considerar o contexto socioeconômico e as políticas públicas relacionadas ao desmatamento. Estudos como os de Souza et al. (2014), Villela e Bueno (2017) e Rafaela Soares (2017) analisaram as políticas públicas e as práticas discursivas relacionadas ao desenvolvimento territorial e ao desmatamento na Amazônia, fornecendo uma compreensão mais ampla dos fatores sociais, econômicos e políticos que influenciam as mudanças no uso da terra.

A aplicação de técnicas de aprendizado profundo para a classificação de imagens de satélite Landsat em regiões de desmatamento tem se mostrado uma

abordagem promissora para o monitoramento e a compreensão dos impactos ambientais. Estudos recentes, como os citados, têm explorado essa abordagem, combinando diferentes técnicas e abordagens para obter resultados mais precisos e abrangentes. A utilização de redes neurais convolucionais em imagens hiperespectrais, como mostrado por Vale, Delgado e Silva (2022), permitiu uma caracterização detalhada da ocupação desordenada do solo, identificando áreas desmatadas de forma eficiente.

Além disso, a combinação de dados SAR multitemporais em banda X com técnicas de Aprendizado de Máquina, conforme investigado por Kuck et al. (2021), revelou-se uma abordagem eficaz para mapear a degradação florestal causada pela exploração seletiva de madeira na Amazônia Legal. Esses estudos demonstram a importância de utilizar diferentes fontes de dados e algoritmos avançados para obter uma compreensão mais completa das mudanças no uso da terra.

Nesse sentido, a aplicação de técnicas de aprendizado profundo para a classificação de imagens de satélite Landsat em regiões de desmatamento é uma abordagem promissora e multidisciplinar. A integração de diferentes estudos e abordagens permite obter uma compreensão mais completa e precisa das mudanças no uso da terra, contribuindo para a conservação dos ecossistemas florestais e o desenvolvimento de políticas públicas mais eficazes para a mitigação do desmatamento.

Por fim, estes estudos mostram que as técnicas de aprendizado profundo têm grande potencial para a classificação de imagens de satélite Landsat em regiões de desmatamento, permitindo identificar áreas de desmatamento com alta precisão e monitorar a situação da cobertura florestal ao longo do tempo.

1.6 Organização do trabalho

Este estudo está organizado em cinco tópicos, sendo o primeiro referente a introdução do trabalho. Nesse tópico são abordadas informações que apresentam o tema, as motivações desta pesquisa, a justificativa deste estudo, os objetivos e outros estudos relacionados a este tema. O segundo é referente ao embasamento teórico que contribuiu como base para as sustentações técnicas e científicas deste trabalho abordando a inteligência artificial, aprendizado de máquinas, sensoriamento remoto e as técnicas de segmentação de imagens. O terceiro tópico é o material e métodos explicando o objeto do estudo e a metodologia utilizada para obtenção dos

resultados desta pesquisa. O quarto tópico apresenta os resultados e discussão dos dados obtidos. O quinto e último é a conclusão do trabalho

2. EMBASAMENTO TEÓRICO

2.1 Inteligência Artificial

A Inteligência Artificial (IA) é um campo de pesquisa interdisciplinar que envolve a criação de sistemas capazes de realizar tarefas que, até então, eram consideradas exclusivas de seres humanos, como reconhecimento de fala, visão computacional, processamento de linguagem natural, tomada de decisão e aprendizado de máquina (RUSSELL, 2019).

Essa tecnologia utiliza técnicas de Machine Learning (aprendizado de máquina), *Deep Learning* (aprendizado profundo), Redes Neurais Artificiais, Processamento de Linguagem Natural, entre outras, para possibilitar que os sistemas aprendam com dados e aprimorem seus resultados ao longo do tempo (LEITA *et al.*, 2018).

As técnicas de IA estão sendo aplicadas em diversas áreas, como saúde, finanças, transporte, marketing e segurança. Na área financeira, os algoritmos de inteligência artificial são capazes de realizar análises de risco, prever tendências e automatizar processos (MACHADO *et al.*, 2018).

Apesar dos desafios, essa tecnologia apresenta um grande potencial para transformar a sociedade e melhorar a qualidade de vida das pessoas, desde que seja utilizada de forma responsável e ética. A evolução contínua das técnicas de IA e sua aplicação em novas áreas prometem revolucionar a forma como vivemos, trabalhamos e nos relacionamos (RODRIGUES *et al.*, 2018).

A IA também tem sido utilizada para resolver problemas complexos que antes não tinham solução, como a detecção de fraudes em sistemas financeiros, previsão de eventos climáticos, análise de dados de tráfego e otimização de processos industriais. Com o uso de algoritmos de Machine Learning, é possível extrair insights e padrões dos dados que não seriam detectáveis por seres humanos, possibilitando a tomada de decisão baseada em informações mais precisas e confiáveis (FERREIRA *et al.*, 2018).

Outra área é a robótica, com a criação de robôs capazes de realizar tarefas complexas e trabalhar em conjunto com seres humanos. Esses robôs podem ser utilizados em diversas áreas, como manufatura, serviços domésticos, saúde e segurança (SOUSA *et al.*, 2018).

Em resumo, a IA representa uma grande oportunidade para avançar em diversas áreas, mas também apresenta desafios e riscos que devem ser considerados. É necessário um esforço conjunto de pesquisadores, empresas, governos e sociedade em geral para garantir que essa tecnologia seja utilizada de forma ética, responsável e segura, visando sempre o bem-estar humano e o progresso da sociedade (SILVA *et al.*, 2018).

2.2 Aprendizado de máquinas (*machine learning*)

Machine Learning é um subcampo da IA que visa resolver problemas mais complexos, é o primeiro pilar da IA tornando possível que os computadores possam aprender e evoluir. Para essa finalidade, os sistemas são estimulados a solucionar problemas de forma autônoma, aprendendo sobre eles e propondo uma saída.

Nesse contexto, os algoritmos analisam dados de entrada, processam e preveem saídas possíveis dentro de um intervalo definido, utilizando de diferentes abordagens para otimizar sua capacidade de chegar ao resultado. Os sistemas artificiais de *Machine Learning* possuem duas fases: treinamento e execução. No primeiro caso, eles são alimentados por um conjunto de entradas de informações, que os contextualiza com a situação específica em que o problema se encontra. Na segunda fase, eles começam a traçar caminhos do que aprenderam na etapa anterior. Essas duas fases são diferentes, dependendo do tipo de algoritmo (LAVECCHIA, 2015).

Os algoritmos de *Machine Learning* podem ser divididos em três grandes categorias - aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço. O aprendizado supervisionado é útil nos casos em que uma propriedade (conjunto de informações inseridas) encontra-se disponível para um determinado conjunto de dados (conjunto de treinamento). O aprendizado não supervisionado é útil nos casos em que o desafio é descobrir relacionamentos implícitos de dados não agrupados. O aprendizado por reforço fica entre esses dois

extremos - existe alguma forma de feedback disponível para cada etapa ou ação preditiva, mas não há agrupamento de ou mensagem de erro precisa (KUMARI, 2017).

Aprendizado supervisionado é uma técnica de aprendizagem através da qual todo o processo é regulado e supervisionado. Esse tipo de aprendizado tem por finalidade antever os resultados de um conjunto de amostras de treinamento (rótulo) junto com seus dados de treinamento. Pode ser conhecido como classificação de pontos de dados. Então, esses dados são passados para o sistema de aprendizagem, que tem como função descobrir caminhos e ajustar seu próprio modelo para chegar aos resultados esperados. Baseia-se em inserir durante o treinamento o algoritmo daquilo que deve ser previsto (LECUN *et al.*, 2016).

Aprendizado não supervisionado que, ao contrário da aprendizagem supervisionado, não usa rótulos de treinamentos para os exercícios. Os algoritmos são formulados de forma a encontrar estruturas e padrões apropriados nos dados. Assim que esses padrões consistentes estiverem visíveis, dados semelhantes podem ser resumidos em grupos (clusters), agrupando grupos semelhantes em determinadas classes. É como um aprendizado sem assistência humana, com o sistema ditando os caminhos a seguir. Nesse modelo de algoritmo, as estratégias comuns são clustering (agrupamento) e redução de dimensão (LAVECCHIA, 2015).

Aprendizado por Reforço é um tipo de *Machine Learning* que possui um agente, por exemplo, que aprende como deve se comportar em determinado ambiente, determinando ações e calculando resultados. Se o agente obtém uma resposta correta, adquire um ponto de recompensa, desta forma há o aumento da confiança do agente e, assim, este executará mais ações de maneira correta. O Aprendizado por Reforço está baseado na existência de críticas externas ao ambiente, onde ocorre uma avaliação da ação tomada, mas sem indicar explicitamente a ação correta, utilizando-se de estrutura composta de estados, ações e recompensas (CASTAÑO, 2018).

Embora isso não seja novo, o *Machine Learning* é um campo de conhecimento em desenvolvimento, em que há novas perspectivas e muitas vezes novos modelos. Este desenvolvimento está tentando resolver problemas ainda mais precisos e oferecer para cada ação uma ferramenta adequada ao problema. A *Machine Learning* estuda métodos computacionais para adquirir novos

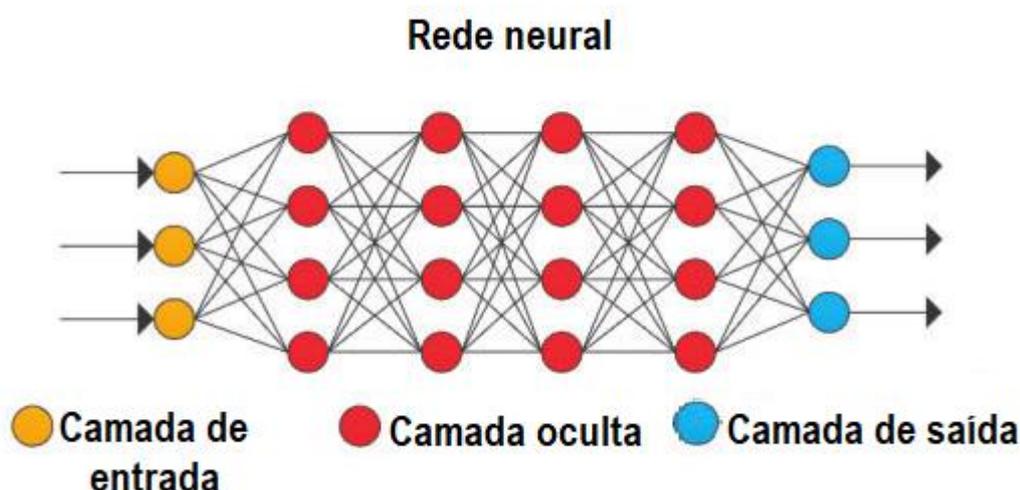
conhecimentos, novas habilidades e novos meios de organizar o conhecimento já existente (ZHANG, 2016).

2.3 Aprendizagem profunda e rede neural convolucional

Deep Learning (DL) está integrado ao *Machine Learning*, sendo uma subcategoria deste. Assim como o *Machine Learning*, também contém aprendizado supervisionado, não supervisionado e de reforço. Foi inspirado em redes neurais artificiais, também conhecidas como RNA, em que há o funcionamento análogo às redes neurais biológicas humanas. Embora seja uma maneira, dentre outras, de executar o *Machine Learning*, o *Deep Learning* se dá em um nível mais profundo, com algoritmos de alto nível para o treinamento de neurais artificiais. Por isso, para entender as aplicações de DL funcionam é necessário conhecer como as redes neurais artificiais (RNA) funcionam (LECCHI; DE ALMEIDA; VIVACQUA, 2022).

De acordo com os estudos de Falcão *et al.* (2019), assim como o cérebro humano, as RNA atua transmitindo e conectando informações de um ponto a outro, as informações são processadas, após isso, a resposta pretendida é enviada para a camada de saída (Figura 1).

Figura 1 - Apresentação da rede neural



Fonte: Adaptado de Falcão *et al.* 2019

Um dos benefícios dos algoritmos de *Deep Learning* é a capacidade de aprender através de grandes quantidades de dados de maneira autônoma,

tornando-se uma ferramenta valiosa para analisar grandes conjuntos de dados; portanto, de extrema importância ao *Big Data*, onde a maioria dos dados é do tipo não-estruturado (LECUN *et al.*, 2016).

O *Deep Learning* é amplamente usado para criar aplicativos de IA. A tecnologia terá grandes evoluções nos próximos anos, destacando-se muito no seu paradigma de “capacitar o computador a aprender a partir da observação dos dados”. Redes Neurais Artificiais (RNA) são algoritmos que reproduzem a estrutura biológica do cérebro humano. Nas RNAs existem neurônios com múltiplas camadas e conexões, em que cada nível seleciona uma função específica para aprender, tais como curvas e bordas para o reconhecimento de uma imagem. Há neurônios de entrada que são ativados através de sensores, permitindo perceber o ambiente, outros neurônios são ativados através de conexões ponderadas de neurônios previamente ativos (CASTAÑO, 2018).

Existem diferentes maneiras de aplicar o *Deep Learning*. Qualquer método possui uma aplicação específica, por exemplo: tipo de dados que se utiliza; se há aprendizado supervisionado ou não; em quais tarefas deseja-se aplicar e em que tipos de tarefas deseja-se resolver com os dados já agrupados etc. Alguns métodos de aprendizado aprofundados são: Rede Neural Convolutiva; Rede Neural Recorrente; Memória de longo prazo, entre outros. Ao usar métodos baseados em *Deep Learning*, cientistas da computação, engenheiros e desenvolvedores de software obtiveram bons resultados em várias aplicações. Usando várias camadas de processamento de dados não lineares, pode-se obter uma representação complexa e abstrata dos dados de maneira hierárquica. Nos últimos anos, o *Deep Learning* fomentou muitos avanços em áreas diversas, como a percepção do objeto, a tradução automática e o reconhecimento de todos os tópicos de pesquisa de voz, que tem sido por muito tempo, difícil para os pesquisadores de IA (PEREIRA, 2016).

2.4 Sensoriamento remoto e processamento de imagens

O sensoriamento remoto é uma importante técnica de observação da Terra que é capaz de adquirir imagens e obter informações sobre objetos sem fazer

contato físico, através do uso de sensores em satélites ou aeronaves (ZANG *et al.*, 2018).

A segmentação de imagens no sensoriamento remoto, em particular, é um ramo importante, cujos produtos têm sido amplamente utilizados em muitas aplicações do mundo real, como mapeamento global da cobertura da terra, monitoramento da vegetação, monitoramento da qualidade da água, clima urbano e estudos ambientais, detecção de incêndios florestais, exploração mineral, detecção de derramamento de óleo e agricultura de precisão (ZANOTTA, 2019).

A radiação eletromagnética do sol é refletida em formas de ondas eletromagnéticas junto a velocidade da luz, com a frequência mensurada em hertz e múltiplos em quilohertz e megahertz, assim como o comprimento de onda medido em metro e submúltiplos em micromêtro e nanômetro (RACHEL, 2013).

Tabela 1 - Unidades de medidas

Velocidade da Luz	300.000 Km/s
Hertz	Hz
Quilohertz	1kHz = 10^3
Megahertz	1mHz = 10^6
Comprimento de Onda	λ
Micrômetro	1 μ m = 10^{-6}
Nanômetro	1nm = 10^{-9}

As imagens de sensoriamento remoto podem ser caracterizadas por três resoluções, a saber: 1) a resolução espacial (por exemplo, QuickBird, IKONOS e o sistema chinês Gaofen-1; 2) a resolução espectral (por exemplo, Hyperion, AVIRIS, HYDICE e ROSIS); e 3) a resolução temporal (por exemplo, MODIS); (AMANI *et al.*, 2020).

A disponibilidade dessas imagens com altas resoluções espaço-espectrais-temporais oferece grande potencial para o desenvolvimento do processamento de dados, incluindo: 1) o pré-processamento de recursos (representação de recursos e seleção de banda hiperespectral); e 2) as aplicações específicas (classificação

supervisionada e não supervisionada e detecção de alterações); (HOSSAIN *et al.*, 2019). Isso significa que as resoluções espaço-espectro-temporais das imagens de sensoriamento remoto se tornaram maiores. Tanto os volumes de dados quanto a dimensionalidade aumentam significativamente, e a distribuição de dados no espaço de recursos tornam-se mais complexa e esparsa.

A distribuição gaussiana comumente usada não pode modelar esses dados de sensoriamento remoto muito bem. Assim, as abordagens tradicionais podem não funcionar bem em certas tarefas no processamento de imagens de sensoriamento remoto, por exemplo, classificação e detecção de mudanças, visto que algumas das abordagens tradicionais (por exemplo, o classificador de máxima verossimilhança (MLC) e a máquina de vetores de suporte (SVM) ocasionam problemas de classificação ou regressão e os resolvem usando amostras de treinamento conhecidas para prever os atributos correspondentes (por exemplo, o rótulo de classe ou mudar de rótulo), eles só podem alcançar um desempenho satisfatório sob certas suposições (uma distribuição normal ou outra) ou condições (um pequeno número de amostras de treinamento); (TOURE *et al.*, 2019). No entanto, a complexidade dos dados dificulta a satisfação dessas suposições ou condições.

Para lidar com a complexidade dos dados, diferentes modelos foram projetados com uma poderosa capacidade de otimização para lidar com problemas multiobjetivos (LI *et al.*, 2021). Por um lado, certas tarefas de processamento de imagem de sensoriamento remoto podem ser transformadas em um problema de otimização contínua (por exemplo, agrupamento) ou um problema de mochila (por exemplo, seleção de banda hiperespectral, extração de *endmember* e detecção de mudança), que é um problema NP-difícil representativo (SILVA *et al.*, 2021).

Embora os métodos tradicionais de otimização sejam eficientes em termos de tempo no processamento de imagens de sensoriamento remoto, eles são sensíveis aos valores iniciais e é fácil para eles ficarem estagnados em determinados locais (BIZON *et al.*, 2020). Por outro lado, em certas tarefas de processamento de imagem de sensoriamento remoto (por exemplo, agrupamento), como resultado da incerteza da estrutura de dados, funções objetivas projetadas de forma diferente são muitas vezes conflitantes e atingem diferentes desempenhos em diferentes imagens de sensoriamento remoto, o que pode influenciar drasticamente a capacidade de generalização das abordagens tradicionais (ZHOU, 2020).

Portanto, modelos recém-projetados, que precisam levar em consideração múltiplas funções objetivo, torna-se complexo. Assim, diferentes funções objetivo são frequentemente combinadas em uma única função com a ajuda de um parâmetro de regularização. Entretanto, a determinação do parâmetro de regularização não é uma tarefa fácil.

2.5 Técnicas de segmentação de imagens

A segmentação de imagens é uma tarefa básica na análise de imagens em que a imagem é dividida em regiões significativas cujos pontos têm quase as mesmas propriedades, por exemplo, níveis de cinza, valores médios ou propriedades texturais. subsequente (por exemplo, precisão da classificação do uso/cobertura da terra); (RAMESH *et al.*, 2021).

Por isso, o processo de segmentação é um dos primeiros passos na análise de imagens de sensoriamento remoto, visto que a imagem é dividida em regiões que melhor representam os objetos relevantes na cena. Os atributos da região, como área, forma, parâmetros estatísticos e textura, podem ser extraídos e usados para análise posterior dos dados (WANG *et al.*, 2020).

A tarefa de segmentação pode ser realizada de duas maneiras: 1) dividindo as imagens em várias regiões homogêneas, cada uma com um rótulo único; 2) determinando limites entre regiões homogêneas de diferentes propriedades. Essas técnicas de segmentação são conhecidas como segmentação baseada em região e detecção de borda, respectivamente (SARAVANA; KARTHIGAIVEL; MAGUDEESWARAN, 2021).

Cada abordagem é afetada de forma diferente por vários fatores. Para alguns aplicativos, a abordagem de detecção de borda não foi bem-sucedida. A causa principal é a presença de pequenas lacunas nos limites das bordas que permitem a fusão de regiões diferentes. Outras desvantagens são que essas técnicas também costumam ser muito sensíveis à intensidade das variações locais e os contornos obtidos geralmente não são fechados. Portanto, para produzir limites fechados, as arestas devem estar ligadas (PARE *et al.*, 2020).

Por outro lado, a segmentação baseada em região sempre fornece regiões de contorno fechadas, o que é um requisito em muitas aplicações. Além disso, é

muito simples e eficaz em muitas aplicações. Erros nos limites das regiões são a principal desvantagem dessa abordagem: os pixels da borda podem ser unidos a qualquer uma das regiões vizinhas. Entre as abordagens de segmentação com base na região estão os métodos de crescimento de região que serão discutidos na próxima seção com mais detalhes (BAFFA; COELHO; CONCIN, 2021).

Outras técnicas de segmentação combinam as informações baseadas em bordas e baseadas em regiões, levando em consideração sua natureza complementar (MOIGNE; TILTON, 1995) ou usam informações auxiliares para orientar o procedimento de segmentação (MASON *et al*, 2018). A seleção de qualquer uma dessas abordagens de segmentação depende muito do tipo de dados que estão sendo analisados e da área de aplicação.

É válido destacar que segmentação é tipicamente associada ao problema de reconhecimento de padrões. Este é o primeiro passo no processo de reconhecimento de padrões e é chamado de isolamento de objetos. É uma tarefa desafiadora para imagens pobres ou de baixo contraste que podem resultar na difusão dos limites do tecido (KAUR; KAUR, 2021). Requer conhecimento prévio. Vários métodos de detecção de bordas são propostos na literatura, mas a seleção de um método particular depende do tipo de imagem e do domínio do problema.

Ye *et al.* (2018), propôs um novo algoritmo para detecção de borda de imagem de sensoriamento remoto usando filtros guiados rápidos para suavizar a imagem, então um operador Sobel aprimorado de máscara 3X3 e modelo de 8 direções é usado para encontrar gradientes e direções dos gradientes. Limites altos e baixos são selecionados usando um novo método Ostu bidimensional, pois o método tradicional é muito sensível ao ruído e ao tamanho do alvo. O conjunto de dados usado para o experimento é a imagem PNG da Universidade da Academia Chinesa de Ciências e o tamanho do pixel é 1280X659. O s/w é implementado usando MATLAB 2016. Embora o algoritmo proposto forneça mais detalhes de borda, contornos claros e contínuos, o algoritmo sofre de duas limitações – alta complexidade de tempo, ineficiente para ruído de alta intensidade.

Fawwaz *et al* (2018), compara o desempenho do Filtro Gaussiano e do Filtro Bilateral. Os filtros são aplicados a várias imagens de satélite usando Canny, Robert e Frie Chen. A vantagem de um filtro bilateral sobre o filtro gaussiano é que os filtros bilaterais usam dois filtros de kernel – kernel espacial e kernel de intervalo. O kernel espacial é a distância entre os pixels de uma imagem. O kernel de alcance é a

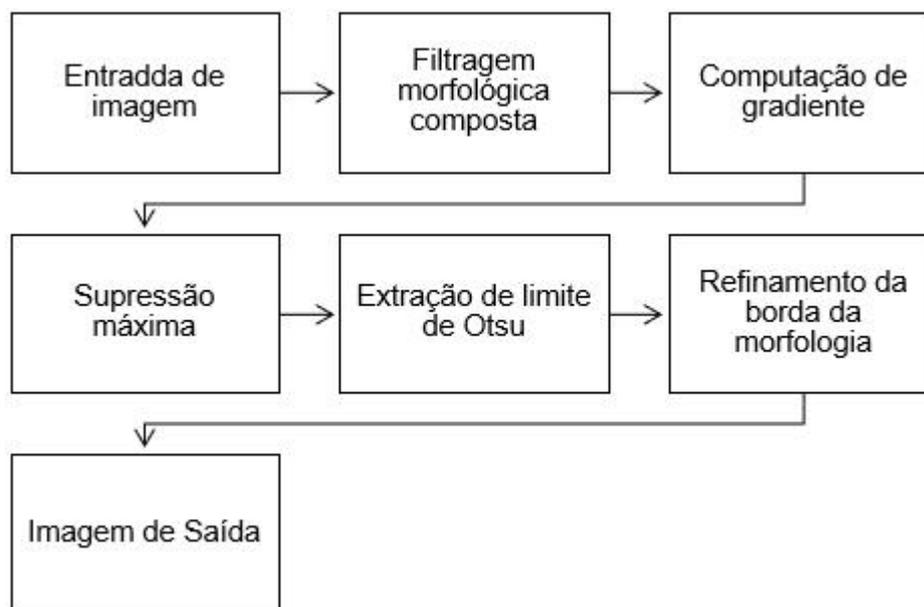
semelhança de intensidade entre os dois pixels na imagem. Dois parâmetros de comparação MSE e valor PSNR são usados. Verificou-se que quanto maior o valor de MSE, pior a qualidade da imagem enquanto for maior o valor de PSNR melhor será a qualidade da imagem. Eles usaram as imagens de satélite de mares e lagos, aplicaram os métodos de Canny, Robert e Frei Chen para diferentes valores MSE e PSNR nessas imagens usando filtros bilaterais e filtros gaussianos. Experimentos mostram que filtros bilaterais com o operador Canny com o menor valor de MSE e os maiores valores de PSNR fornecem o melhor resultado.

Zaaj *et al* (2018), propuseram uma abordagem híbrida para a extração de edifícios em imagens de satélite THR usando detecção de recursos. O método Canny é usado para encontrar bordas de um objeto. Se a imagem for ruidosa, a detecção de contornos é muito difícil. Portanto, o operador de Harris junto com Canny é usado para criar um método para extração de edifícios. A extração de edifícios a partir de imagens de satélite THR requer a detecção de arestas e cantos. Às vezes, o método astuto não consegue detectar todos os pontos de borda. O operador de Harris é usado para a detecção de cantos. O operador de Harris é usado como um determinante de limiar. O novo método combinado é usado em imagens diferentes com três limiares diferentes. Os limites escolhidos são combinados com a saída de Harris. Os resultados foram considerados mais eficazes para o valor de limite mais alto.

Kumar *et al* (2017), usaram um processo de três estágios para extração da rede rodoviária a partir de imagens de satélite multiespectrais de alta resolução. Imagens multiespectrais são imagens com três ou mais bandas espectrais. A extração de feições viárias é necessária para planejamento urbano, gerenciamento de tráfego e atualização de mapas. O método proposto utiliza três etapas para a extração de feições viárias. As arestas são detectadas usando diferentes operadores de detecção de arestas (Canny, Sobel, Prewitt). O ruído presente na imagem resultante é detectado usando a operação morfológica. Filtros de mediana são usados para reduzir o ruído presente em uma imagem.

Guiming *et al* (2016), introduziram um algoritmo de Canny melhorado, baseado nas desvantagens do algoritmo de Canny convencional. O processo é representado na Figura 2.

Figura 2 - Algoritmo de Canny aprimorado



Fonte: Guiming, 2016

O método convencional usa Filtros Gaussianos que às vezes detectam o ruído como uma borda. Limites altos e baixos são selecionados manualmente, portanto, fornecem uma borda incompleta (DHRUV; MITTAL; MODI, 2019).

O algoritmo aprimorado de Canny usa filtros morfológicos para suavizar a imagem e fornece contornos de imagem precisos. Usando o método Otsu, os dois limites são escolhidos automaticamente, o que fornece uma borda clara e contínua. Ambos os métodos foram aplicados em três imagens diferentes para detectar as bordas do Lena e duas imagens de sensoriamento remoto de estrada e navio (DAR; PADHA, 2019).

Lavanya *et al* (2014), implementou o método Sobel no conjunto de blocos Simulink e testou em três imagens diferentes. A técnica de Sobel é usada principalmente para implementação em hardware e detecção de borda em tempo real. A vantagem do método Sobel é menos complexidade e fácil cálculo. O resultado do operador Sobel não é preciso, pois usa apenas duas máscaras. Resultados precisos com operadores de Sobel podem ser obtidos usando um conjunto maior de máscaras.

Wang *et al* (2014), propuseram um novo método para detectar bordas de imagens de sensoriamento remoto de alta resolução espacial. A imagem é transformada por Fourier para o espectro de magnitude obtido, a análise do espectro

é feita usando amostras de raio e ângulo para identificar bordas. Dois filtros log Gabor são multiplicados com espectro de frequência e IFFT é aplicado no resultado para detectar as bordas. A imagem Quickbird da região de Nanjing, na China, é usada para um experimento que dá bons resultados.

Vijayarani *et al* (2013), descreveram dois métodos diferentes para detectar bordas em imagens faciais: Canny Edge Detector e Sobel Edge Detector. A vantagem de detectar a borda de uma imagem é reduzir a quantidade de dados a serem armazenados. Neste trabalho de pesquisa, ambas as técnicas são aplicadas a um conjunto de imagens e, em seguida, o desempenho das duas técnicas em termos de precisão e velocidade é comparado. Após comparação, verifica-se que o algoritmo de Canny leva 34,7 segundos, enquanto Sobel leva 34,9 segundos para detectar a borda. A precisão de Canny é de 87,5%, enquanto a precisão de Sobel é de 75%. Pode-se concluir que o algoritmo de Canny supera bem quando comparado ao algoritmo de Sobel para detecção de bordas.

Kaur *et al.* (2013), discutiu métodos recentemente propostos para segmentação automática de imagens. Aplicações de segmentação automática de imagens estão em tarefas de visão de máquina como reconhecimento de objetos, compressão de imagens, edição de imagens, busca de imagens etc.

Xi *et al.* (2012), usaram o operador Canny e a transformada de Hough para detectar bordas de imagens de sensoriamento remoto. As imagens de sensoriamento remoto contêm uma grande quantidade de dados corrompidos pelo ruído, portanto, os métodos tradicionais não são eficazes para a detecção de bordas de imagens de sensoriamento remoto porque os métodos tradicionais às vezes detectam o ruído como uma borda e podem perder as bordas reais. Experimentos mostram que a abordagem combinada detecta bordas mais precisas de imagens de sensoriamento remoto.

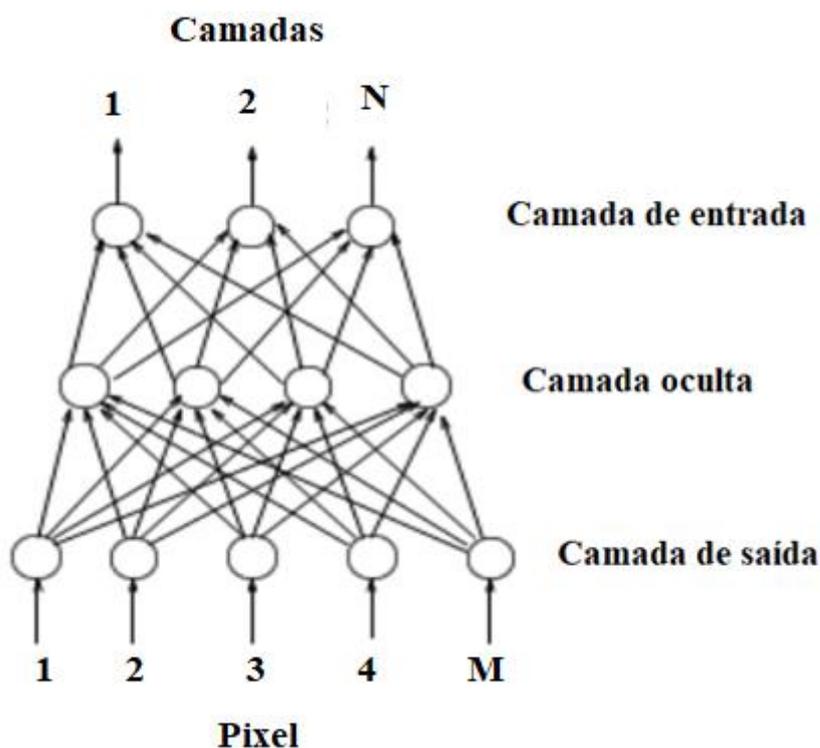
Kaur *.et al.* (2021), usou Morfologia Matemática para detecção de borda da imagem de 'Saturno'. Os resultados são comparados com Sobel Edge Detection, Canny Edge Detection, LOG, detecção de borda Prewitt. Observou-se que a detecção de arestas usando morfologia matemática dá um resultado melhor do que o método tradicional. O algoritmo proposto aplicou elementos de estrutura de vários tamanhos em diferentes direções com operadores de morfologia. Ele seleciona e calcula a média das arestas tomando a diferença entre imagens erodidas e dilatadas em todas as direções.

Senthilkumaran *et al.* (2019), explicou vários métodos de detecção de borda para segmentação de imagem, como Robert, Prewitt e Sobel. Ele também explica algumas abordagens de computação flexível. Os métodos foram aplicados a uma imagem real de uma universidade e mostra o resultado dos diferentes métodos.

Yi *et al* (2014), propuseram um novo algoritmo para detecção de borda da imagem de sensoriamento remoto multi-fonte usando lógica difusa. A detecção de bordas de imagens de sensoriamento remoto por métodos convencionais requer cálculos complexos. No método proposto, uma nova função de pertinência é projetada, modificados os métodos de realce fuzzy, e usando critérios de avaliação de arestas trata o procedimento iterativo automaticamente. Imagens do Landsat-7, Ikonos, Quickbird e SPOT-5 são usadas para o experimento. Conclui-se que o novo método é mais adequado para aplicação em tempo real e detecção de bordas de imagens de satélite multifonte e multitemporais.

Zhang (1997), demonstrou o uso de rede neural para segmentar uma imagem. Os recursos da imagem são usados como vetores de entrada. Nesta técnica, as características do objeto que devem ser segmentadas são identificadas pela análise manual de amostras de objetos e, em seguida, os programas são criados. Existem vários tipos de redes neurais: rede ART (Adaptive Resonance Theory), SIM (mapas auto-organizados) e rede recorrente. Um exemplo de rede neural para segmentar uma imagem é mostrado na Figura 3. Cada pixel em uma imagem corresponde a neurônios. Os conjuntos de amostras de treinamento são usados para treinar a rede neural para determinar a conexão e o peso entre os nós. Usando essa rede neural recém-treinada, novas imagens são segmentadas.

Figura 3 - Rede neural de amostra



Fonte: adaptado de Zhang (1997).

Maini *et al.* (2009); Jayakumar *et al.* (2014); Dharampal *et al.* (2015); Bala Krishnan *et al.* (2017); classificaram a detecção de borda em duas categorias: baseada em gradiente e baseada em laplaciano.

Os métodos baseados em gradiente detectam as bordas encontrando os valores máximo e mínimo nas derivadas de primeira ordem da imagem. Métodos baseados em laplacianos detectam bordas encontrando cruzamentos zero na derivada de segunda ordem de uma imagem. Esses métodos geralmente são aplicados às imagens que foram suavizadas usando filtros de suavização Gaussiana (MAINI *et al.*, 2009; JAYAKUMAR *et al.*, 2014; DHARAMPAL *et al.*, 2015).

Os métodos baseados em gradiente incluem principalmente um operador Canny, operador Sobel, operador Prewitt, operador Cross de Robert. Entre esses métodos, Canny é um algoritmo mais eficiente para extração de objetos, pois retorna menos arestas falsas, portanto, também é chamado de 'Optimal Edge Detector' (KRISHNAN *et al.*, 2017).

Sendo assim, a seleção de uma técnica particular depende da aplicação utilizada. Entre os métodos, o de canny dá melhores resultados em comparação com os outros métodos de detecção de borda, mas às vezes o canny dá falsa

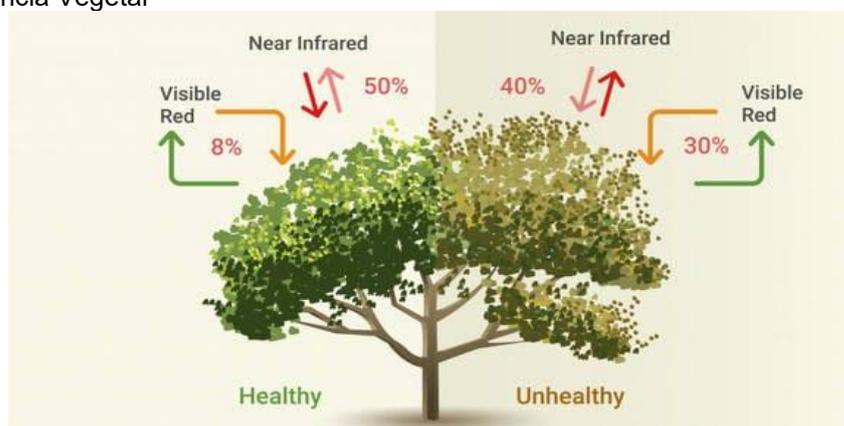
detecção de canto, portanto, o operador de Harris pode ser usado para detecção de canto eficiente. Além disso, para suavização de imagem, filtros bilaterais ou filtros morfológicos funcionam melhor do que filtros gaussianos. Por isso, a escolha do parâmetro para segmentar uma imagem também desempenha um papel importante na detecção de características (BEDOYA *et al.*, 2019).

2.6 Comportamento espectral da vegetação

A análise da vegetação na cobertura do solo por meio do sensoriamento remoto permite identificar se ocorreu desmatamento e qual possível alteração na cobertura do solo. Para o estudo da vegetação desmatada são utilizadas três faixas espectrais: azul, vermelho e o infravermelho próximo, com a faixa azul apresentando baixa absorção, vermelha baixa reflectância e a banda do infravermelho próximo a com a maior reflectância vegetal (ALMEIDA, 2009), representada pela figura 4.

Fatores naturais como a radiação solar, iluminação do ambiente e nuvens contribuem para interferências nos padrões de radiância espectral (NOGUEIRA, 1996). Nos estudos de classificação do solo para identificar áreas desmatadas é fundamental que as imagens possuam as mínimas interferências. A diferenciação pelo classificador entre as áreas desmatadas e não desmatadas pode apresentar equívocos, uma vez que ambas as classes possuem semelhança espectral, sobreposição de classes e baixa resolução espectral dos sensores (AMARAL *et al.*, 2009).

Figura 4 - Reflectância Vegetal



Fonte: EOS Data Analytics

No estudo de Rodrigues, Ferreira e Silva (2023) foi utilizada a imagem do satélite CBERS WPM para a classificação supervisionada identificando de forma mais precisa a conversão da cobertura florestal para uso alternativo do solo. O uso de imagem com altas resolução espacial ao mesmo tempo que permite maior precisão na identificação das alterações na cobertura do solo exigem um maior processamento e acurácia na classificação uma vez que os valores radiométricos possuem maior variação (RODRIGUES; FERREIRA, SILVA, 2023).

A importância da espectrometria das bandas para identificação de vegetação está no fato de permitirem o cálculo de índices que possibilitam identificar o grau de vegetação em determinada área.

2.6.1 Índice de vegetação por diferença normalizada

O Índice de vegetação por diferença normalizada (NDVI) é um exemplo que auxilia na identificação de áreas com maiores coberturas vegetais (OLIVEIRA; AQUINO, 2020). O cálculo do NDVI consiste em:

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)} \quad (1.)$$

Onde:

NIR = reflectância no comprimento de onda correspondente ao infravermelho próximo (0,76 a 0,9 μm para o Landsat 5 e 0,85 a 0,88 para o landsat 8);

Red = reflectância no comprimento de onda correspondente ao vermelho (0,63 a 0,69 μm para Landsat 5 e 0,64 a 0,67 μm para o Landsat 8).

Oliveira e Aquino (2020) por meio da aplicação do NDVI na Bacia hidrográfica do Rio Gurguéia identificaram diminuição da atividade fotossintética e o aumento do desmatamento na região. O mesmo estudo permitiu demonstrar que além de classificações supervisionadas e não-supervisionadas, é possível por meio de índices de vegetação obter marcas de alterações da cobertura do solo e impactos ambientais na vegetação local.

Braga *et al.* (2022) identificaram padrões de corte seletivo na Floresta Nacional do Jamari e também fora, o que poderia ser indício de degradação ambiental como abertura de estradas, pátios de estocagem de madeira e clareiras

geradas pelo corte das árvores. Importante ressaltar que os autores utilizaram imagens com alta resolução espacial da missão *Planet* (3 a 5 metros) e alta revisita (1 a 3 dias), contudo seu modelo *deep learning* teve exatidão de 67%. Isso permite inferir que não basta o banco de dados ser composto por imagens com boa resolução espacial, é fundamental que o modelo seja bem construído.

2.6.2 Índice de vegetação ajustado ao solo

Segundo GILBERT (*et al* 2002; BERNARDI, *et al.*, 2014), o índice de vegetação ajustado ao solo (SAVI) tem a finalidade de reduzir os efeitos causados pela reflectância do solo, desta forma, minimizando a densidade da vegetação para normalizar a densidade média para que obtenha pouca influência do solo destacando as respostas espectrais do vegetação.

O SAVI utiliza a constante “L” que tem a finalidade de minimizar o efeito do solo como resultado final do índice (CANAVIEIRA, 2018), onde a constante “L” varia conforme a densidade do dossel referente a área de estudo.

Os valores para os seguintes graus de densidade do dossel: L = 1 para densidades baixas de vegetação L = 0,5 para densidades médias L = 0,25 para densidades baixas.

$$SAVI = \frac{(NIR - R)}{NIR + R + L} \times (1 + L) \quad (2.)$$

Em que:

SAVI - índice de vegetação ajustado ao solo.

NIR - reflectância de superfície no comprimento de onda do infravermelho, banda 5.

RED - reflectância de superfície no comprimento de onda do vermelho banda 4.

L – Constante para o ajustamento do efeito do solo sobre a reflectância do dossel. O valor da constante L é em função do índice de área foliar.

2.6.3 índice de área foliar

O índice de área foliar (IAF) ele reflete a biomassa de cada pixel de uma imagem, determinado pela razão entre a área foliar de toda a vegetação pela unidade de área do campo em uso, por essa vegetação (CANAVIEIRA.2018).

Segundo BRAZ (2015), Tem-se na equação os valores numéricos 0,69/ 0,59 e 0,91 adotados como constantes para o cálculo e IVAS representando os dados calculados no SAVI. Logo, para realizar o índice de área foliar é necessário gerar antes o SAVI.

$$IAF = - \frac{\ln \left(\frac{0,69 - IVAS}{0,59} \right)}{0,91} \quad (3.)$$

Em que:

IAF - índice de área foliar.

0,69/ 0,59 e 0,91 – é respectivamente a área foliar total da vegetação dividido pela área da vegetação e dividido pela vegetação (BRAZ, 2015 apud ALLEN *et al* 2002).

IVA- são os dados resultantes do índice de vegetação do solo – SAVI.

2.7 Descoberta de conhecimento em banco de dados

KDD pode ser compreendido como um processo composto por três processos denominados pré-processamento, mineração de dados e pós-processamento. A etapa de pré-processamento consiste no conjunto de funções que relacionam-se com a organização e tratamento dos dados permitindo identificar a melhor forma de trabalhar com os dados; na etapa de mineração do dados é realizada a busca de conhecimentos úteis no contexto da pesquisa, ou seja, no que o pesquisador pretende identificar nos seus dados e por fim o pós-processamento abrange o tratamento do conhecimento obtido na mineração dos dados (BOENTE; ROSA, 2007).

O KDD possui três grupos de atividades sendo elas desenvolvimento tecnológico, execução de KDD e aplicação de resultados (BOENTE; ROSA, 2007). A metodologia deste trabalho está no grupo execução de KDD, que integra as atividades voltadas à busca efetiva de conhecimentos em bases de dados.

A utilização do KDD consistiu nos seguintes procedimentos: seleção dos dados por meio do dataset; transformação dos dados, consistindo no pré-processamento das imagens com sua segmentação e redimensionamento de acordo com suas assinaturas espectrais; mineração dos dados, identificando padrões que possivelmente sejam válidos para serem utilizados no modelo da rede neural convolucional; conhecimento, etapa de interpretação das informações obtidas (MAMATHA; RASHMI, 2022).

O KDD empregado no modelo utilizado neste trabalho identificado pelas imagens de satélite especificado para o desenvolvimento de um aprendizado profundo, permitiu o emprego da modelagem para as classes desmatado e não desmatado. Dentro dessa metodologia é importante abordar o aprendizado profundo, compreendido como uma subárea do aprendizado de máquina (CANAVIEIRA, 2022).

O aprendizado profundo é uma modelagem que fundamenta-se em hipóteses constituídas por regras e dados disponíveis para o controle de informações a partir de um banco de dados obtidos, pela arquitetura intitulada *Deep Learning* (DL). Esta utiliza camadas de neurônios artificiais para processar dados. A rede neural de aprendizado profundo possui camadas ocultas, além das de saída e entrada o que permite que um conjunto de neurônios realizem operações aritméticas entre os estímulos de entrada e seus pesos sinápticos (CANAVIEIRA, 2022).

A escolha do KDD para este trabalho consiste na capacidade que o KDD possui em cooperar amplamente na busca do conhecimento embutido nos dados, por meio da combinação de métodos e ferramentas de estatística, inteligência artificial, visualização e banco de dados para identificar padrões nos dados (MANILLA *et al.*, 1997). O KDD possui um conjunto de atividades contínuas que geram um conhecimento a respeito da base de dados que somada a necessidade da mineração de dados justificam o emprego desta ferramenta no trabalho. O processo de descoberta de padrões e tendências do KDD em grandes conjuntos de dados, tendo como principal etapa a mineração dos dados, consistindo na execução da análise e de algoritmos específicos, que mesmo com limitações computacionais,

produz relações particulares de padrões (FAYYAD *et al.*, 1996; DANTAS *et al.*, 2008).

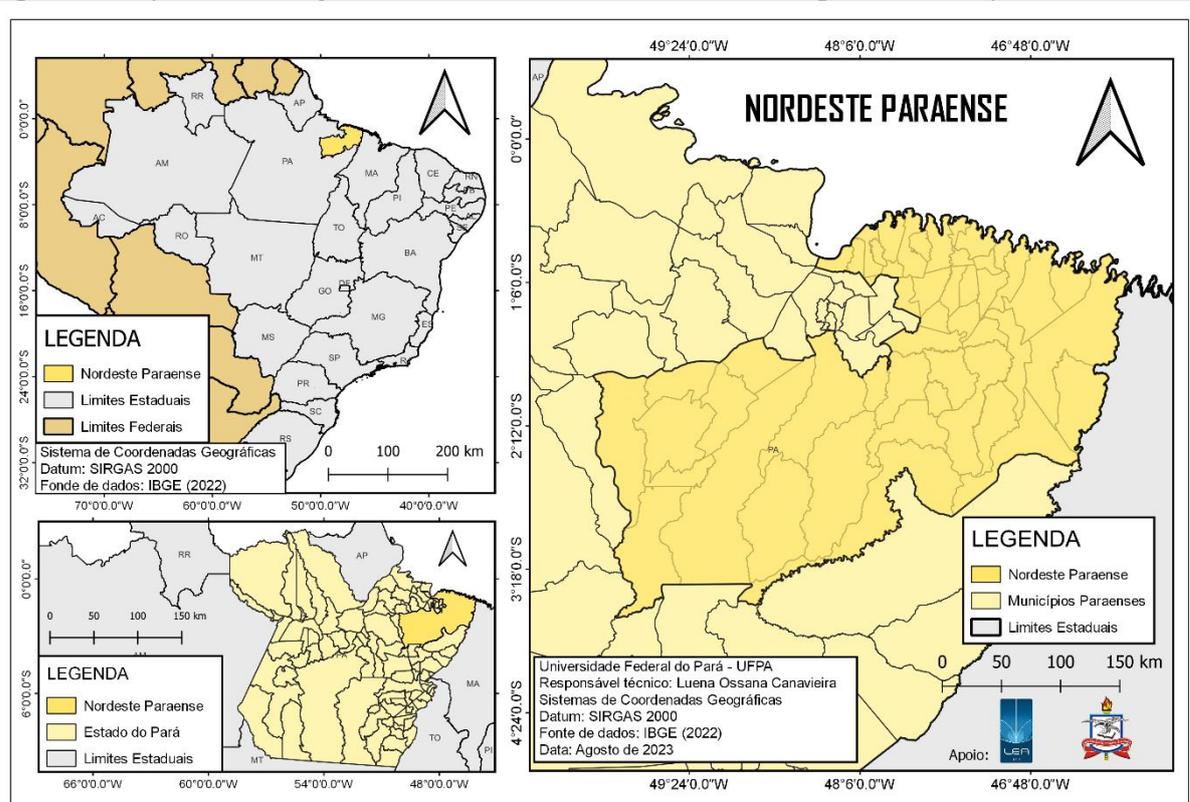
3 MATERIAL E MÉTODOS

Este capítulo mostra a descrição da área de estudo e das etapas para a realização da classificação de desmatamento nas imagens do Landsat 5,7 e 8.

3.1 Área de estudo

A área objeto da pesquisa é a mesoregião nordeste paraense, composta por 49 municípios localizados entre as latitudes 0° 00' e 3° 40" S e longitudes 50° 42' O e 45° 30' O (Figura 4).

Figura 4 - Mapa de localização da área de estudo referente a Mesoregião Nordeste paraense



Fonte: autoria própria

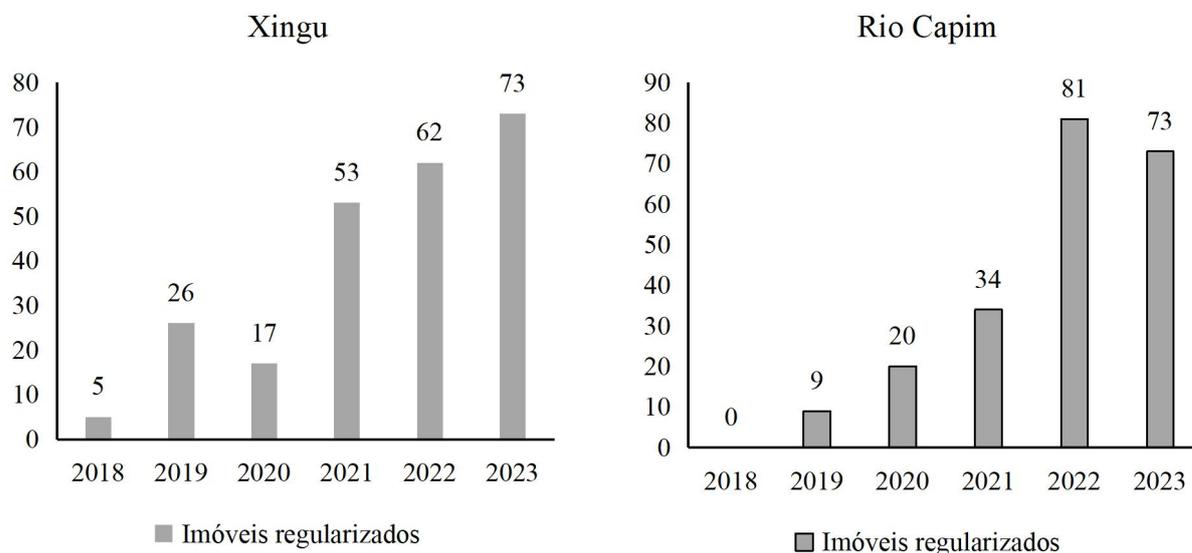
A região, segundo classificação de Koppen, apresenta clima tropical chuvoso de floresta e clima tropical de monção com precipitação variando entre 1.750 mm até 3.000 mm (ALVARES, 2013; IBGE, 2016).

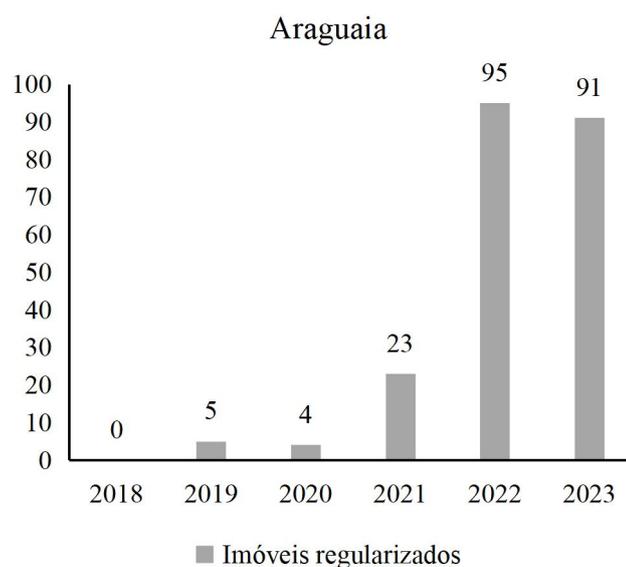
O nordeste paraense é formado por cinco microrregiões que são: Bragantina, Cametá, Guamá, Salgado e Tomé-açu que juntas abrangem 49 municípios, que juntos somam 49 municípios. A população aproximadamente é de 1.789.387 habitantes com uma taxa de crescimento populacional de 8,5% (IBGE, 2016).

O clima da região varia de mesotérmico a megatérmico úmido, caracterizado pela ocorrência de temperaturas elevadas variando de 23 °c a 34 °c com alta umidade relativa do ar entre 85% a 91%. As chuvas não são uniformes durante o ano, com maior concentração no período de janeiro a junho (MARTORANO *et al.*, 1993).

A região possui poucos dados referente a regularização ambiental dos imóveis rurais, o que dificulta inferir se os mesmos estão contribuindo para o desmatamento na região. O gráfico 1 a seguir apresenta o número de imóveis regularizados nas principais zonas de integração do estado do Pará.

Gráfico 1 – Quantitativo do número de imóveis rurais regularizados nas zonas de integração no Estado do Pará





Fonte: autoria própria

A partir de 2021 o número de imóveis regularizados nessas regiões aumentou, principalmente as regiões Xingu e Rio Capim. Araguaia apenas em 2022 teve um crescimento significativo com 95 imóveis regularizados. O gráfico 2 apresenta a porcentagem do número de imóveis rurais regularizados no Pará por zona de integração.

Gráfico 2 – Zonas de integração no Estado do Pará referente aos imóveis em regularização ambiental

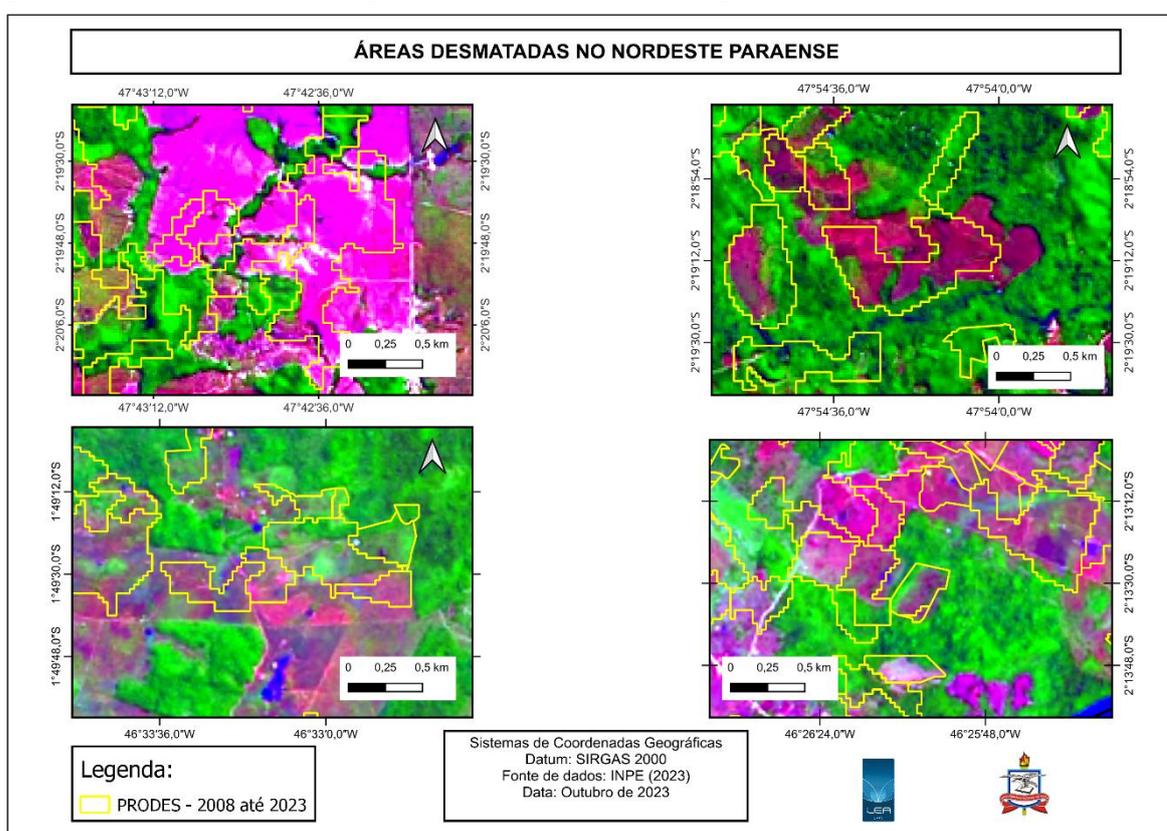


Fonte: autoria própria

Xingu, Araguaia, Rio Capim e Tapajós representam 76% dos imóveis do Pará que estão em processo de regularização. O nordeste paraense abrange as zonas do Caeté, Guamá e Tocantins com 7%.

A figura 5 apresenta exemplos de áreas desmatadas segundo o Instituto Brasileiro de Pesquisas Espaciais (INPE) por meio do banco de dados vetoriais do Projeto de Estimativa do desmatamento da Amazônia (PRODES). As áreas de cor roxa sobrepostas aos polígonos em amarelo representam as áreas que foram identificadas como desmatadas, ou seja, ocorreu a conversão da vegetação nativa para algum uso do solo de forma ilegal.

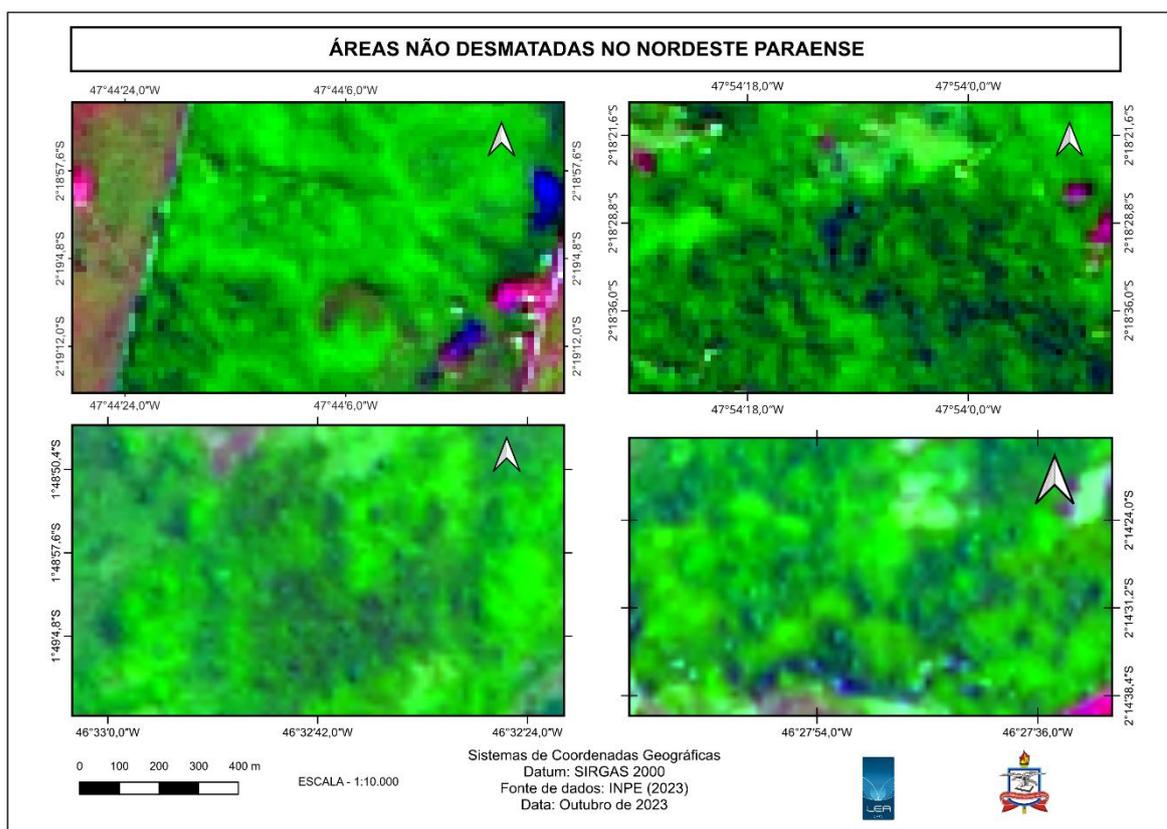
Figura 5 - Áreas desmatadas segundo o PRODES 2023 na região do Nordeste paraense



Fonte: Imagens do INPE (2023).

A figura 6 apresenta imagens de áreas não desmatadas. É possível identificar a coloração verde densa, representando a vegetação de floresta densas ombrófilas que predominam na região.

Figura 6 - Áreas não desmatadas segundo o PRODES 2023 na região do Nordeste paraense

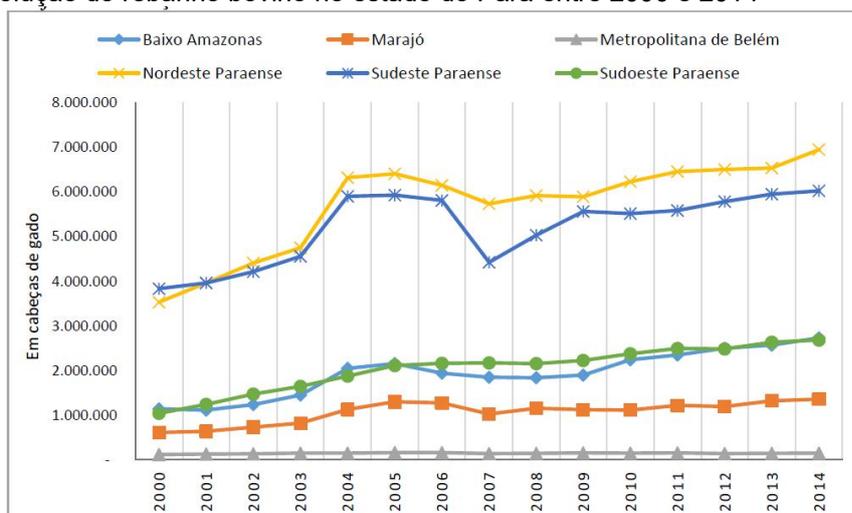


Fonte: Imagens do INPE (2023).

A composição das imagens do satélite Landsat 8 utilizadas para a identificação de áreas desmatadas e não desmatadas, seguem a composição de bandas Infravermelho-média (B6), Infravermelho-próximo (B5) e Vermelha (B4) o que realça a cobertura do solo sendo o verde denso usualmente áreas de florestas densas e a cor roxa referente a solo exposto.

O gráfico 3 apresenta a evolução do rebanho entre 2000 e 2014 no Pará, com a região do Nordeste paraense sendo a com maior aumento de cabeças de gado. A alteração da vegetação para a pastagem do gado seguiu um crescimento quase linear. Isso ressalta a importância do monitoramento regular do uso do solo por meio de imagens de satélite, pois algumas atividades no solo como a bovinocultura são realizadas em imóveis regularizados, o que não infere em desmatamento ilegal mas em um uso do solo ambientalmente correto.

Gráfico 3 – Evolução do rebanho bovino no estado do Pará entre 2000 e 2014



Fonte: IBGE (2014)

3.2 Banco de dados

A segmentação do banco de dados das imagens Landsat 7/8 ocorreu no período de 17/05/2000 a 17/05/2022, totalizando em torno de 210 imagens. Importante frisar que inicialmente tinham sido adquiridas 120 imagens, número reduzido e com intuito de melhorar a classificação e o modelo computacional, foi inserida a técnica de aumento de dados baseado nas imagens adquiridas, aumentando o número de imagens já processadas e redimensionadas em 75%. Isso permitiu a obtenção de 210 imagens. A figura 7 apresenta as etapas para extração de conhecimento na mineração de dados.

Figura 7 - Conhecimento em mineração de dados



Fonte: autoria própria

Por meio dos resultados do modelo foi possível gerar a matriz de confusão, com o intuito de realizar classificações digitais para avaliação de resultados provenientes das imagens satélites. Desta forma, a análise não linear dos valores pós processamento foram determinantes para avaliação do modelo, sendo utilizado o método hold-out com divisão da base de dados em 75% para treinamento do algoritmo de aprendizado de máquina e 25% para a etapa de testes com a finalidade de extrair métricas estatísticas para posterior análise.

3.3 Processamento dos dados

Inicialmente para o desenvolvimento do modelo desse estudo as imagens utilizadas foram classificadas como “desmatado” ou “não desmatado”. Os dados foram extraídos utilizando o processo de descoberta de conhecimento o que auxiliou na presente metodologia.

A linguagem de programação utilizada para montar o código referente ao processamento dos dados foi a Python e foi rodado na plataforma *Collaboratory*, que permite escrever e executar códigos em Python.

Para o treinamento do modelo, foi utilizada a função documentada pelo Python identificada como ReLU (Unidade Linear Retificada). Através desta função foi implementado o algoritmo desenvolvido para a classificação das áreas desmatadas e não desmatadas, apresentando a probabilidade dos dados serem nas classes 0 e 1. Por meio da DL várias camadas são definidas para o modelo bem treinado, pois pode apresentar recursos e métricas estatísticas correspondente ao grande acerto na classificação das imagens pós processamento (LALKHEN, 2022).

No conjunto de treinamento, a validação representa o ajuste de parâmetros na arquitetura geral do modelo proposto, para evitar *overfitting* (sobre-ajuste dos dados pós processados), utilizando a técnica “*EarlyStopping*” da plataforma de aprendizagem profunda *Keras*.

O método abrandou o número de parâmetros gerados pelo modelo, chamados de número de épocas de um treinamento com objetivo de parada antecipada para analisar a necessidade da manipulação de um modelo de bom desempenho. Para este estudo, foi utilizada a média entre as paradas antecipadas do modelo, no desenvolvimento do algoritmo foi identificado um parâmetro de critério de parada na 34^a época. Isto representa que o modelo para nesta época por conter informações necessárias para o desempenho da rede neural convolucional.

Estima-se nesse processo, que o modelo com técnicas de predição e paradas realizou a classificação real para identificar as áreas desmatadas e não desmatadas baseadas no banco de dados adquirido para, enfim, gerar a matriz de confusão para a verificação de dados reais do modelo proposto para a classificação das imagens e interpretação da matriz de confusão.

Medeiros e Tsen (2021) utilizaram a biblioteca Tensor Flow, um código aberto que permitiu treinar e executar modelos de aprendizado de máquina e redes neurais fazendo uso da ferramenta *Google Colaboratory*. No mesmo estudo o modelo para identificação de áreas de manguezais, rios e margens foi de de 92,5% de acurácia, F1-score de 89% e um Precision e Recall de 88% indicando que uma *Deep Learning* se mostrou eficiente para a classificação da imagens quanto ao quer era manguezal, rio e margem.

Estudos como esse são exemplos do potencial da *Deep Learning* aplicados na classificação de imagens, uma vez que permitem de maneira eficiente e rápida o processamento de grande volume de imagens para a identificação com alta precisão

de um objeto. Informações precisas como essas podem fornecer subsídios para planejamento e políticas públicas eficientes.

A obtenção da acurácia, *recall*, precisão e f-score se deu por meio da matriz de confusão. Esta apresenta as frequências de classificação de cada classe do modelo utilizado (SOUZA, 2019). A matriz é composta por verdadeiro positivo: ocorre quando no conjunto real, a classe que se busca a precisão está correta; falso positivo: a classe para qual está sendo buscada a precisão está errada; falso verdadeiro: ocorrendo quando no conjunto real, a classe para qual não estava sendo buscada a previsão foi prevista corretamente; e falso negativo: ocorrendo quando no conjunto real, a classe para a qual não estava sendo buscada a previsão foi prevista corretamente (NICOLAU, 2020).

3.4 Métricas avaliativas

Como forma de avaliar as métricas de precisão e a conclusão do modelo, foram utilizadas a matriz de confusão, acurácia, sensibilidade, precisão e erro.

A Matriz de confusão registra os valores corretos e incorretos para cada classe sendo esses valores classificados como verdadeiros positivos (VP) que representa a classificação correta do modelo, verdadeiros negativos (VN) que representam a classificação correta da classe negativa, falso positivo (FP) ocorrendo quando o modelo classifica como classe positiva quando o valor real é classe negativa e falso negativo (FN) quando o modelo classifica como classe negativa o que é classe positiva (PAUL, 2021).

A acurácia pode ser definida como a taxa de acerto global do experimento, isto é, a proporção de predições corretas em relação ao conjunto de dados, e pode ser calculada por meio da seguinte equação:

$$Acurácia = \frac{\text{total de acertos}}{\text{total de dados no conjunto}} = \frac{(VP + VN)}{(P + N)} \quad (1)$$

Onde VP representa verdadeiros positivos, VN verdadeiros negativos, P os positivos e N os negativos.

Precisão é a taxa de acertos dos exemplos classificados como positivos e que são positivos quanto que erro pode ser definido como a medida que compara a

classe verdadeira de cada exemplo com o rótulo atribuído pelo classificador induzido (MONARD; BARANAUSKAS, 2003). A precisão é dada pela equação:

$$Precisão = \frac{VP}{VP + FP} \quad (2)$$

onde FP representa os falsos positivos. A sensibilidade, ou *Recall*, é a métrica que considera quantas classificações estão corretas considerando as situações de classe positiva com valor esperado. A mesma é utilizada em situações onde os falsos negativos são considerados mais ofensores do que falsos positivos. A sensibilidade é dada pela equação:

$$Sensibilidade = \frac{VP}{VP + FN} \quad (3)$$

onde FN representa falso negativo.

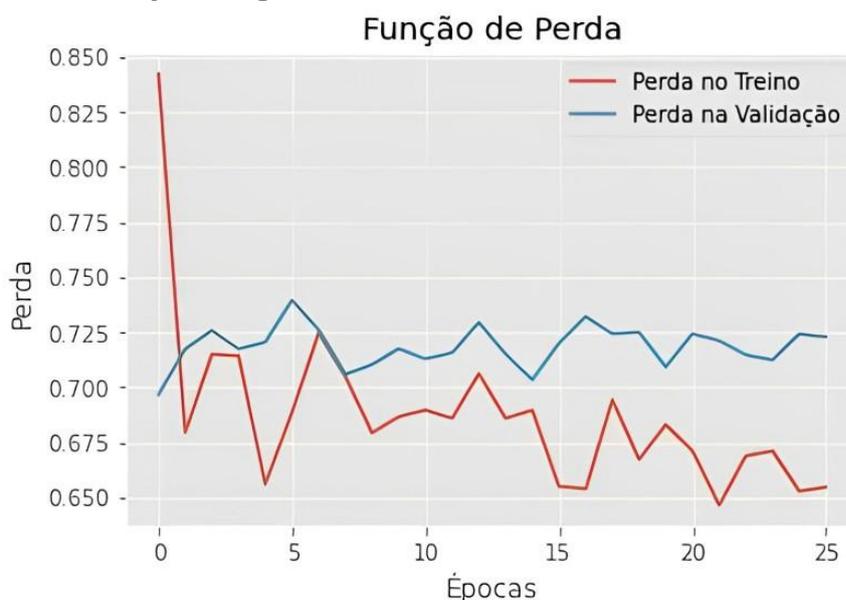
As equações regidas são determinadas para avaliar as métricas de precisão e a conclusão do modelo, através da matriz de confusão determinada na predição da técnica de treinamento. A predição correta é uma medida para determinar os desbalanceamentos do conjunto de dados, possibilitando conclusões a respeito do desempenho do modelo.

Para a proporção de verdadeiros negativos, convém a capacidade de prever o modelo e a ausência da condição para os casos que realmente há acertos positivos. Determina-se a proporção de verdadeiros positivos em relação a todas as predições positivas, evitando o desbalanceamento dos outros acertos já gerados. A medida em análise também faz parte da predição gerada pelo desbalanceamento, identificando a proporção dos verdadeiros negativos para predições negativas

4 RESULTADOS E DISCUSSÃO

A matriz de confusão apresenta os resultado referente ao valores de desempenho necessário para a extração de métricas avaliativas de um modelo computacional treinamento, como: acurácia, sensibilidade, especificidade, precisão e erro. O modelo desempenhou as duas classes, desmatado e não desmatado, descritivas por 0 e 1 respectivamente. O gráfico 4 apresenta a função de perda relacionando parâmetros da rede em desempenho. No mesmo gráfico é possível verificar que a Perda de Validação manteve-se variando entre 0.7 e 0.75, quanto que a Perda no Treinamento diminuiu de 0.84 para próximo a 0.65.

Gráfico 4 - Parâmetros de função de perda obtidos após desempenho do modelo em tentativas estimadas durante a execução do algoritmo

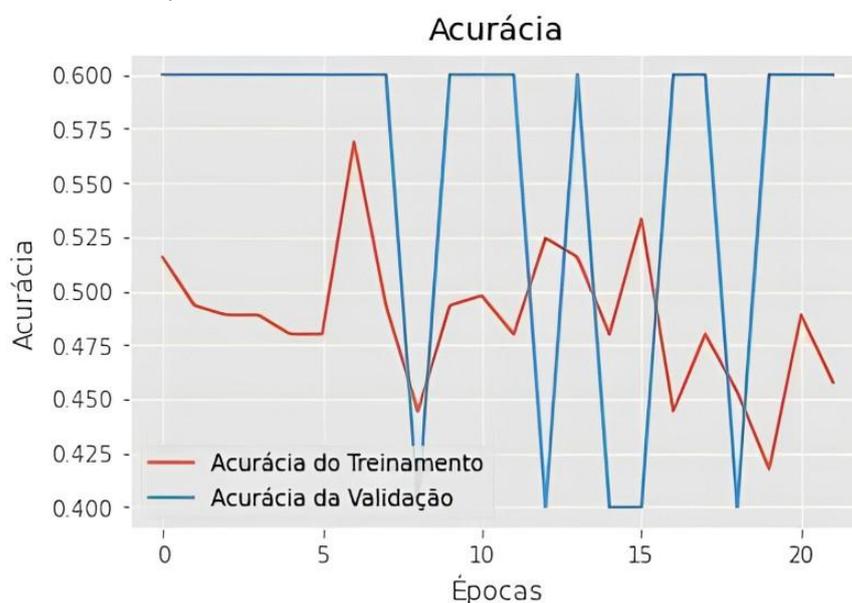


Fonte: autoria própria

A avaliação da acurácia apresenta picos variados originados após o desenvolvimento de testes de o modelo preditivo, ilustrados no gráfico 5. A classificação da rede neural combina sensibilidade em TP (probabilidade da área desflorestada), especificidade (probabilidade negativa da área não desflorestada). Em FP, valor preditivo positivo, em que a probabilidade da área desflorestada é real e o valor preditivo negativo (TN e FN), probabilidade de valores preditivos

classificando as áreas desflorestadas a partir das imagens do Landsat como áreas reais.

Gráfico 5 - Parâmetros de alteração nos valores de acurácia da rede neural convolucional, após execuções destacando o desempenho do modelo.



Fonte: autoria própria

Com a avaliação da matriz de confusão, obteve-se acurácia de 69,7% para acertos reais do modelo. Estes valores foram obtidos após a execução do modelo equivalente à oito vezes para avaliação em valores satisfatórios. No entanto, os valores determinaram um índice baixo de classificação das áreas degradadas e não degradadas da região paraense e isto é justificável a partir dos dados quantificados relacionados ao banco de dados fornecido pelo órgão público SEMAS para o desenvolvimento da pesquisa.

A tabela 1 apresenta a matriz de confusão desenvolvida no estudo com o resultado das classificações.

Tabela 2 - Matriz de confusão da classificação da rede neural convolucional desenvolvida no estudo

		VALOR PREDITO	
		Sim	Não
REAL	Não	15	6
	Sim	4	8

Fonte: autoria própria

Para um modelo com acurácia superior a 70%, seria necessário o acesso a um maior número de imagens do satélite Landsat, pois a concisão da rede neural estimada apresentaria maiores acertos. Contudo, durante a metodologia da classificação do modelo testes foram desenvolvidos e relatados, por meio de gráficos estabelecidos, a importância da técnica de aprendizado profundo para utilização na região nordeste do Pará. Isso demonstra a importância de um banco de dados maior e com imagens de qualidade sem possíveis interferências como nuvens.

Vale, Delegado e Silva (2022) obtiveram 88% de acurácia no modelo utilizado para reconhecimento de padrões para mapeamento da cobertura e uso da terra, apontado quatro elementos fundamentais para avaliação do desempenho da classificação como matriz de confusão, acurácia geral, precisão, revocação e o *F-score*.

A limitação do número de imagens para o banco de dados é um fator crucial para a classificação, pois interfere diretamente no processo de diferenciação das classes.

Além das classificação por meio de alterações na cobertura vegetais em grandes áreas, as estradas podem servir de subsídio para compreender a dinâmica de expansão em uma região. Costa *et al* (2023) identificaram expansão de rodovias em áreas da amazônia legal por meio da rede neural convolucional, enfatizando que a expansão das estradas antecede desmatamento e são principais indicadoras de ocupação de terras públicas.

Braga *et al.* (2022) em seu estudo com o modelo neural convolucional indicaram não apenas a necessidade do uso de várias imagens como a importância dos critérios a serem adotados no modelo, pois mesmo com imagens com boa resolução espacial ainda assim é possível o modelo apresentar alta taxa de erros.

Silva *et al.* (2017) identificaram que a metodologia KDD apresentou bom desempenho para as classificações supervisionadas e não supervisionadas de imagens Landsat 8 para os cultivos de soja e milho, com altos valores de índice Kappa sendo 99,75% para o algoritmo *Random Forest* (RF), 99,8% para *Classification and Regression Trees* e 95,05% para *Extreme Learning Machine*. As técnicas de aprendizagem de KDD se apresentaram eficazes para a classificação da

cultura, mostrando que pode ser uma ferramenta com grande potencial de mapeamento de culturas agrícolas (SILVA *et al.*, 2017).

Dependendo do banco de dados este pode requerer grandes etapas de pré-processamento antes de ser analisado, além de processadores eficazes que possam identificar padrões de forma ágil. Almeida *et al.* (2023) utilizou a metodologia KDD no processamento de imagens para índices de vegetação, cujos resultados apresentaram elevados índices Kappa com 96% para o método RF. O uso de métodos avançados para análises de dados ambientais permitiu avaliar as melhores áreas para cultivo e desenvolvimento de culturas anuais além de identificar áreas sensíveis as alterações antrópicas (ALMEIDA *et al.*, 2023).

Técnica de inteligência artificial somadas ao uso e aprendizagem de máquina auxiliam no acompanhamento, qualificação e quantificação dos elementos da cobertura do solo (ALBA *et al.*, 2022). Ainda que um classificador possua alta confiabilidade na sua classificação, é importante frisar a possibilidade de ocorrer confusão entre classes em virtude destas serem similares. Ainda em Alba *et al.* (2022), o RF apresentou 96% de precisão na classificação, contudo entre as classes solo exposto e áreas agrícolas ocorreram erros pois sua espectrometria são semelhantes.

Diante disto é importante frisar que mesmo um modelo apresentando grande precisão na sua classificação, um grande número de classes e com espectrometria semelhantes podem equivocar o pesquisador em relação aos resultados.

5 CONCLUSÃO

O trabalho utilizou a rede neural convolucional para a classificação de imagens dos satélites Landsat 5, 7 e 8 para a identificação de áreas desflorestadas com base no sensoriamento remoto, para oferecer aos produtores agrícolas maior produtividade nas regiões do nordeste do Pará.

O modelo apresentado obteve algumas interferências nas imagens, o que dificultou a identificação de áreas desmatadas e não desmatadas com alta precisão. Um fator que contribuiu para isso foi o banco de imagens limitado e a interferência de fatores físicos como nuvens nas imagens. A escolha das imagens para o banco de dados acarretou em um número reduzido do que o esperado, porém bastante eficaz, onde mesmo sendo um empicilho não impediu o modelo de obter acurácia de aproximadamente 70% de acertos reais após a validação, treinamento e teste.

Nesse estudo para a identificação das classes desmatadas e não desmatadas as bandas do vermelho e vermelho próximo foram utilizadas, com isso os possíveis erros no modelo podem ser explicados por conta que as classes possuem espectrometria próxima dificultando uma diferenciação. A não desmatada seria referente a cobertura do solo por floresta e a desmatada é referente a outros tipos como uso do solo, pasto e vegetação secundária, esta última por possuir espectrometria semelhante a vegetação primária também contribui para falhas no modelo.

Para estudo futuros, recomenda-se novas imagens satélites da região proposta pelo estudo, com destaque para os índices de resolução espectral correspondentes com os sensores determinados pelos satélites do estudo. No caso do Landsat 7/8, em muitas imagens foram identificadas grandes penetrações de água e mapeamento de águas costeiras, contudo foi escolhida pela descrição de sua banda 3, sendo adequada para os tipos de solo e agricultura e perturbações ocorridas pela chuva.

Além do aumento de amostras para o banco de dados, recomenda-se utilizar imagens de outros satélites que possuam tempo de revisita menor que do satélite Landsat 7 e 8 (16 dias) como imagens do Sentinel-2a e 2b com 5 dias de revisita.

Recomenda-se considerar primordialmente a resolução radiométrica do Landsat 7/8 com intuito de escolher um dataset adequado para resolução do modelo. Estima-se em um banco de dados aproximadamente com 500 a 1.000 imagens para validação do modelo buscando até 95% de acurácia.

REFERÊNCIAS

- ALBA, Elisiane *et al.* Comparação entre Algoritmos de Aprendizado de Máquina para a Identificação de Floresta Tropical Sazonalmente Seca. *Anuário do Instituto de Geociências*, v. 45, p. 1-10, 2022.
- ALMEIDA, G. A.; *et al.* Utilização de Redes Neurais Artificiais na Previsão de Preços de Ações. In: *Anais do XII Encontro Nacional de Inteligência Artificial*. 2019.
- ALMEIDA, T. S. Sensibilidade da reflectância de uma floresta tropical em 460 nm, 650 nm e 850 nm aos parâmetros ópticos e arquitetônicos do dossel, *Dissertação*, Universidade Federal de Viçosa, Minas Gerais, Brasil, 2009.
- ALMEIDA, Luiz *et al.* Engenharia e extração de recursos em séries temporais de índice de vegetação para mapeamento de área com culturas anuais. 2023.
- ALVES, G. M. *et al.* Detecção de desmatamento na Amazônia utilizando redes neurais convolucionais. In: *Simpósio Brasileiro de Sensoriamento Remoto*, 19., 2019, Santos. *Anais...* Santos: INPE, p. 258-265, 2019.
- ALVARES, C. A. *et al.* Koppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, Stuttgart, v. 22, n. 6, p. 711–728, 2013.
- ALMEIDA, Diego Araújo de *et al.* Utilização de imagens Landsat-5/TM e SRTM para reconhecimento e análise das mudanças na paisagem da Serra de Pacaraima, Roraima, Brasil. 2008.
- AMORIM, João Victor Alves; VALLADARES, Gustavo Souza; PORTELA, Mirya Grazielle Torres. Classificação não-supervisionada de imagens rapideye no mapeamento da cobertura das terras do delta do Parnaíba, Piauí. *Geosaberes: Revista de Estudos Geoeducacionais*, v. 12, p. 88-106, 2021. Acesso em: 29/03/2022.
- AMARAL, M. V. F. *et al.* Avaliação e comparação de métodos de classificação de imagens de satélites para o mapeamento de estádios de sucessão florestal. *Revista Árvore*, v.33, n.3, p.575-582, 2009.
- AMANI, Meisam *et al.* Plataforma de computação em nuvem Google Earth Engine para aplicativos de big data de sensoriamento remoto: uma revisão abrangente. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 13, p. 5326-5350, 2020. See More
- ANDRADE, R. B. *et al.* Evaluation of semantic segmentation methods for deforestation detection in the amazon. *ISPRS Archives*; 43, B3, v. 43, n. B3, p. 1497-1505, 2020.
- BAFFA, Matheus de Freitas Oliveira; COELHO, Alessandra Martins; CONCI, Aura. Segmentação de imagens infravermelhas para detecção do câncer de mama utilizando u-net cnn. In: *Anais do XXI Simpósio Brasileiro de Computação Aplicada à Saúde*. SBC, 2021. p. 119-128.
- BEM, Pablo Pozzobon de *et al.* Change detection of deforestation in the Brazilian Amazon using landsat data and convolutional neural networks. *Remote Sensing*, v. 12, n. 6, p. 901, 2020.
- BIZON, Artur Ricardo *et al.* Estudo da classificação de imagens a partir de sensoriamento remoto de Santa Catarina. *Anais do Computer on the Beach*, v. 11, n. 1, p. 074-076, 2020.
- BOENTE, Alfredo Nazareno Pereira; ROSA, José Luiz Dos Anjos. Utilização de ferramentas de KDD para Integração de aprendizagem e tecnologia em busca da gestão estratégica do conhecimento na empresa. *Anais do Simpósio de Excelência em Gestão e Tecnologia*, v. 1, p. 123-132, 2007.

BRAGA, Daniel et al. Mapeamento de corte seletivo na Amazônia Brasileira usando imagens Planet de alta resolução espacial e inteligência artificial: estudo de caso na Floresta Nacional do Jamari. In: GEOINFO. 2022. p. 204-210.

CANAVIEIRA; L. O. Análise espectral da cultura da palma de óleo no município do Tailândia-PA, Acesso em 10/03/2020

CANAVIEIRA, Luena Ossana. Modelo cnn final transfer~encia de aprendizado aumento de dados. Disponível em: https://colab.research.google.com/drive/1bphXpGb3xw-_EeoEpR8KsTAYExedvRiJ#scrollTo=rvboD3OSQrcr

Chollet, F. & others, 2021. Keras. Acesso em: 29/03/2022

DANTAS, Eric Rommel G. et al. O uso da descoberta de conhecimento em base de dados para apoiar a tomada de decisões. V Simpósio de Excelencia em Gestão e Tecnologia, p. 1-10, 2008.

DAR, Aarish Shafi; PADHA, Devanand. Segmentação de imagens médicas: Uma revisão de técnicas recentes, avanços e uma comparação abrangente. Int. J. Comput. ciência Eng , v. 7, n. 7, pág. 114-124, 2019.

DHRUV, Bhawna; MITTAL, Neetu; MODI, Megha. Técnicas de segmentação de imagens médicas e sua relevância na imagiologia contemporânea. In: 2020 8ª Conferência Internacional sobre Confiabilidade, Tecnologias Infocom e Otimização (Tendências e Direções Futuras)(ICRITO) . IEEE, 2020. pág. 1306-1310.

FALCÃO, João Vitor Regis et al. Redes neurais deep learning com tensorflow. RE3C-Revista Eletrônica Científica de Ciência da Computação, v. 14, n. 1, 2019.

FAYYAD, U. M., Piatetsky Shapiro, G., Smyth, P. & Uthurusamy, R. "Advances in Knowledge Discovery and Data Mining", 1996.

FEARNSIDE, Philip M. Brazil's Amazonian forest carbon: the key to Southern Amazonia's significance for global climate. Regional Environmental Change, v. 18, p. 47-61, 2018.

FEARNSIDE, Philip M. Desmatamento na Amazônia brasileira: história, índices e consequências. Destruição e Conservação da Floresta Amazônica, v. 1, p. 7-19, 2005.

FERREIRA, A. S.; et al. Uma Revisão Sistemática sobre o uso de Algoritmos Genéticos em Problemas de Otimização. In: Anais do XXXVIII Congresso da Sociedade Brasileira de Computação. 2018.

FERREIRA, R. S. et al. Aplicação de redes neurais artificiais na previsão de séries temporais de demanda de energia elétrica. Revista Brasileira de Energia, Rio de Janeiro, v. 24, n. 1, p. 1-13, 2018.

From Data Mining to Knowledge Discovery in Databases. Disponível em: <https://www.kdnuggets.com/gpspubs/aimag-kdd-overview-1996-Fayyad.pdf>

GOODFELLOW, Ian. Machine Learning, p.80-96. Acesso em: 12/05/2022.

GOODFELLOW, Ian. Convolutional Networks. 330-371. Acesso em 10/04/2022.

GUEDES, R. L. et al. Detecção de áreas de desmatamento na Amazônia brasileira utilizando redes neurais convolucionais e dados Landsat. Boletim de Ciências Geodésicas, v. 25, n. 2, p. 259-274, 2019.

HOSSAIN, Mohammad D.; CHEN, Dongmei. Segmentação para Análise de Imagem Baseada em Objeto (OBIA): Uma revisão de algoritmos e desafios da perspectiva de sensoriamento remoto. Jornal ISPRS de Fotogrametria e Sensoriamento Remoto , v. 150, p. 115-134, 2019.

Instituto Brasileiro de Geografia e Estatística. Cidades. 2016 b. Disponível em: <http://www.ibge.gov.br/home/geociencias/cartograia/default_territ_area.shtm>. Acesso em: 29 ago. 2023

Instituto Brasileiro de Geografia e Estatística. Produção Agrícola Municipal. 2014. Disponível em: <<[Introdução ao Sensoriamento Remoto. Instituto de Pesquisas Espaciais. Disponível em: \[http://www.dpi.inpe.br/AlunosPG/Jarvis/SR_DPI7.pdf\]\(http://www.dpi.inpe.br/AlunosPG/Jarvis/SR_DPI7.pdf\)](http://www.sidra.ibge.gov.br/bda/acervo/acervo9.asp?e=c&p=PA&z=t&o=11/>> Acesso em: 01/10/2023.</p>
</div>
<div data-bbox=)

Informações sobre o Cadastro Ambiental Rural. Disponível em: <https://www.meioambiente.go.gov.br/component/content/article/136-meio-ambiente/cadastro-ambiental-rural-car/1374-car.html?Itemid=101>

Instituto Nacional de Pesquisas Espaciais (INPE). (2021). Monitoramento do desmatamento na Amazônia Legal por satélite. Disponível em <http://terrabrasilis.dpi.inpe.br/app/dashboard/deforestation/bdqueimadas>

Instituto Nacional de Pesquisas Espaciais (INPE). (2021). Monitoramento do desmatamento na Amazônia Legal por satélite. https://www.inpe.br/noticias/noticia.php?Cod_Noticia=5707.

INSTITUTO NACIONAL DE METEOROLOGIA. Glossário. 2016. Acesso em: 5 jul.2023.

JACKSON, Colbert M.; ADAM, Elhadi. A machine learning approach to mapping canopy gaps in an indigenous tropical submontane forest using WorldView-3 multispectral satellite imagery. *Environmental Conservation*, v. 49, n. 4, p. 255-262, 2022.

JOHNSON, Brian Alan; MA, Lei. Image segmentation and object-based image analysis for environmental monitoring: Recent areas of interest, researchers' views on the future priorities. *Remote Sensing*, v. 12, n. 11, p. 1772, 2020.

Keras documentation. Disponível em: <http://docset/Keras.docset/Contents/Resources/Documents/index.html>

KUHN, A. L. et al. Mapeamento de áreas de desmatamento na Amazônia Legal com o uso de redes neurais convolucionais e imagens Landsat-8. In: Simpósio Brasileiro de Sensoriamento Remoto, 19., 2019, Santos. Anais... Santos: INPE, p. 1142-1149, 2019.

LALKHEN, Abdul Ghaaliq. Statistics I: data and correlations. *Continuing Education in Anaesthesia, Critical Care & Pain*, v. 7, n. 3, 2007. Acesso em 10/05/2022.

LALKHEN, Abdul Ghaaliq. Statistics I: data and correlations. *Continuing Education in Anaesthesia, Critical Care & Pain*, v. 7, n. 3, 2007. Acesso em 10/05/2022.

LAMBIN, Eric F. et al. Effectiveness and synergies of policy instruments for land use governance in tropical regions. *Global environmental change*, v. 28, p. 129-140, 2018.

LE MOIGNE, Jacqueline; TILTON, James C. Refining image segmentation by integration of edge and region data. *IEEE transactions on geoscience and remote sensing*, v. 33, n. 3, p. 605-615, 1995.

LECCHI, Thiago; DE ALMEIDA, Gustavo Maia; VIVACQUA, Rafael Peixoto Derenzi. Algoritmo de Deep Learning baseado na rede neural YOLOv5 para classificação dermatoscópica e detecção de câncer (melanoma). *Concilium*, v. 22, n. 7, p. 813-827, 2022.

LEITE, J. C. S. *et al.* Uma revisão sistemática sobre a utilização de inteligência artificial em sistemas de recomendação. *Revista de Informática Teórica e Aplicada, Porto Alegre*, v. 25, n. 1, p. 67-81, 2018.

LI, Wei et al. Deep Learning for Mapping Forest Cover and Change in Forests Using Multispectral Landsat Imagery. *Remote Sensing*, v. 9, n. 1, p. 1-22, 2017.

LI, Yansheng; MA, Jiayi; ZHANG, Yongjun. Recuperação de imagens de big data de sensoriamento remoto: uma pesquisa. *Fusão de informações*, v. 67, p. 94-115, 2021. See More

MACHADO, F. C. S. *et al.* A aplicação da inteligência artificial na previsão de demanda de energia elétrica: uma revisão sistemática. *Revista de Gestão da Tecnologia e Sistemas de Informação*, São Paulo, v. 15, n. 1, p. 1-22, 2018.

MARTORANO, L. *et al.* Estudos climáticos do Estado do Pará, classificação climática (Köppen) e deiciência hídrica (Thornthwhite Mather). Belém: SUDAM; Rio de Janeiro: EMBRAPA-SNLCS, 1993. 53 p.

Mamatha Bai, B. G., and N. S. Rashmi. "Data Mining Techniques in the Agricultural Sector." *Emerging Research in Computing, Information, Communication and Applications*. Springer, Singapore, 2022. 87-107. Acesso em 10/05/2022

MANILLA, H. *Data mining: Machine learning, statistic and databases*. Department of computer science, University of Helsinki, 1997.

MEDEIROS, Iago Nogueira; TSEN, Jonathan Martinez. Deep learning para identificação e classificação em imagens de manguezais. 2021.

MOHAMED, Soha; FAYED, Mahmoud S. Modelling of Received Signals in Molecular Communication Systems based machine learning: Comparison of azure machine learning and Python tools. arXiv preprint arXiv:2112.10214, 2021. Acesso em 20/05/2022.

MONARD, Maria Carolina; BARANAUSKAS, José Augusto. Conceitos sobre aprendizado de máquina. *Sistemas inteligentes-Fundamentos e aplicações*, v. 1, n. 1, p. 32, 2003.

NADAL, Carlos Aurélio. Tratamento estatístico de observações geodésicas, *Sistemas de Referência e Tempo em Geodésia*, Universidade Federal do Paraná. Acesso em: 12/07/2023

NICOLAU, Vinicius Nakalski. Mapeamento de cobertura e uso de solo a partir de dados de sensoriamento remoto utilizando redes neurais. 2020. Trabalho de Conclusão de Curso. Universidade Tecnológica Federal do Paraná.

NOGUEIRA, J. L. M., *et al.* Reflectância de uma superfície vegetada em função da geometria de aquisição. In: VIII Simpósio Brasileiro de Sensoriamento Remoto, Salvador, Brasil, 14-19 abril 1996, INPE, p. 941-949.

OLIVEIRA, Livanía Norberta; AQUINO, Cláudia Maria Sabóia. Índice da Vegetação da Diferença Normalizada (NDVI) na sub-bacia hidrográfica do rio Gurguéia, Piauí-Brasil: análise do efeito da expansão agrícola. *Revista Geoaraguaia*, v. 10, n. 2, p. 126-143, 2020.

OLIVEIRA, Willer Carlos de; BERTOLINI, Geysler Rogis Flor. Uma revisão sistemática sobre a contribuição das cooperativas para a sustentabilidade da agricultura familiar. *Research, Society and Development*, v. 11, n. 2, 2022.

PAN, Xia *et al.* Detailed and automated classification of land use/land cover using machine learning algorithms in Google Earth Engine. *Geocarto International*, v. 37, n. 18, p. 5415-5432, 2022.

PARE, S. *et al.* Segmentação de imagens usando limiarização multinível: uma revisão de pesquisa. *Jornal Iraniano de Ciência e Tecnologia, Transações de Engenharia Elétrica*, v. 44, p. 1 a 29 de 2020.

PAUL, Sérgio. Detecção automática de mudança de azimute em antenas para redes 5G-IoT utilizando deep learning. 2021. 55. Dissertação, Programa de Pós-Graduação em Engenharia Elétrica (Computação aplicada) - Instituto de Tecnologia da UFPA, UFPA, Belém, 2021.

Piatetsky-Shapiro, G.; Brachman, R.; Khabaza, T.; Kloesgen, W.; and Simoudis, E., 1996. An Overview of Issues in Developing Industrial Data Mining and Knowledge Discovery Applications. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), eds. J. Han and E. Simoudis, 89–95. Menlo Park, Calif.: American Association for Artificial Intelligence.

Projeto Peabiru. Disponível em: <https://peabiru.org.br/areas-de-atuacao/>. Acesso em: 10/05/2022.

RAMESH, KKD et al. Uma revisão dos algoritmos de segmentação de imagens médicas. Transações endossadas pela EAI sobre saúde e tecnologia generalizadas , v. 7, n. 27, pág. e6-e6, 2021.

Revista Humanidades e Inovação v.7, n.9 – 2020. Disponível em: <https://revista.unitins.br/index.php/humanidadeseinovacao/issue/view/7>. Acesso em: 10/05/2022.

RODRIGUES, L. M. et al. Inteligência artificial e aprendizado de máquina na detecção de fraudes em sistemas financeiros. Revista de Administração de Empresas, São Paulo, v. 58, n. 6, p. 604-617, 2018

RODRIGUES, Ciro da Cunha; FERREIRA, Cícero Paulo; SILVA, Izael Costa. Dinâmica do desmatamento e uso do solo em assentamento rural da reforma agrária brasileira na amazônia legal. Revista Brasileira de Agroecologia, v. 18, n. 2, p. 05-25, 2023.

RUSSELL, S.; NORVIG, P. Inteligência artificial: uma abordagem moderna. 3. ed. São Paulo: Pearson, 2019.

SANTOS, Pedro Thiago Cutrim dos. Aprimoramento da detecção de áreas de garimpo na Região dos Tapajós através de Redes adversárias de super-resolução. In: Simpósio Brasileiro de Sensoriamento Remoto, 20., 2021, Natal. Anais... Natal: INPE, p. 1764-1771, 2021.

SARAVANAN, S.; KARTHIGAIVEL, R.; MAGUDEESWARAN, V. Uma técnica de segmentação de imagem de tumor cerebral em processamento de imagem usando algoritmo ICA-LDA com modelo ARHE. Journal of Ambient Intelligence and Humanized Computing , v. 12, p. 4727-4735, 2021. See More

Secretaria Especial do Desenvolvimento Agricultura e Desenvolvimento Agrário. Disponível em: https://www.cfn.org.br/wp-content/uploads/2017/10/Agricultura_Familiar.pdf

SILVA, Francisca Cassia et al. Técnicas de sensoriamento remoto para delimitação de áreas queimadas no Tocantins através do processamento digital de imagens Landsat 8-TM. Journal of Biotechnology and Biodiversity, v. 8, n. 1, p. 006-011, 2020.

SILVA, M. C.; et al. Avaliação do Impacto da Inteligência Artificial na Análise de Crédito. In: Anais do XV Congresso Brasileiro de Sistemas. 2018.

SILVA, R. H.; SOUSA, G. R. Proposta de um Sistema de Recomendação baseado em Redes Neurais Artificiais. In: Anais do XI Simpósio Brasileiro de Sistemas de Informação. 2019.

SILVA, Laíza Cavalcante De Albuquerque et al. Algoritmos de Aprendizado de Máquina para classificação de soja em imagens Landsat-8. Acta Iguazu, v. 6, n. 3, p. 65-78, 2017.

SOUSA, G. R.; SILVA, R. H. Sistemas de Recomendação baseados em Inteligência Artificial: Uma Revisão Sistemática. In: Anais do X Simpósio Brasileiro de Sistemas de Informação. 2018.

SOUZA, E. G. de. Entendendo o que é Matriz de Confusão com Python. 2019. Disponível em: Entendendo o que é Matriz de Confusão com Python | by Emanuel G de Souza | Data Hackers | Medium. Acesso em: 18/08/2023.

Tratamento estatístico de observações geodésicas, Sistemas de Referência e Tempo em Geodésia, Universidade Federal do Paraná. Acesso em: 26/05/2022

TOURE, Seynabou et al. Detecção costeira usando sensoriamento remoto óptico: uma revisão. *ISPRS International Journal of Geo-Information*, v. 8, n. 2, pág. 75, 2019.

VALE, Italo Guimarães do; DELGADO, Angel Ramon Sanchez; SILVA, Robson Mariano. Classificação de Imagens hiperespectrais utilizando redes neurais convolucionais para caracterização da ocupação desordenada do solo sobre um recorte do Parque Nacional da Tijuca, RJ. *Revista Brasileira de Geografia Física*, v. 15, n. 1, p. 561-571, 2022.

WANG, Eric Ke e cols. Uma técnica de segmentação de imagens médicas baseada em aprendizagem profunda no domínio Internet-of-Medical-Things. *Future Generation Computer Systems*, v. 108, p. 135-144, 2020.

ZANOTTA, Daniel Capella; FERREIRA, Matheus Pinheiro; ZORTEA, Maciel. *Processamento de imagens de satélite. Oficina de Textos*, 2019.

ZHOU, Guoqing. *Sensoriamento Remoto Urbano de Alta Resolução: Algoritmos e Modelagem*. CRC Press, 2020.

ZHU, Xiao Xiang et al. Machine learning for big earth observation data: Recent advances and future prospects. *Remote Sensing of Environment*, v. 240, p. 111626, 2020

KUCK, Tahisa Neitz et al. Mapeamento da degradação florestal causada pela exploração seletiva de madeira na Amazônia Legal por meio de dados SAR multitemporais em banda X e técnicas de Aprendizado de Máquina. 2021.

SHI, Yuhui; EBERHART, Russell C. Empirical study of particle swarm optimization. In: *Proceedings of the 1999 congress on evolutionary computation-CEC99* (Cat. No. 99TH8406). IEEE, 1999. p. 1945-1950.

SANTOS, Reginaldo et al. A semi-autonomous particle swarm optimizer based on gradient information and diversity control for global optimization. *Applied Soft Computing*, v. 69, p. 330-343, 2018.

SANTOS, Reginaldo et al. Empirical study on rotation and information exchange in particle swarm optimization. *Swarm and Evolutionary Computation*, v. 48, p. 312-328, 2019.

SOUZA, Edilson Almeida de et al. *Políticas públicas e práticas discursivas do desenvolvimento territorial na Amazônia: uma análise do Programa Territórios da Cidadania no Nordeste Paraense*. 2014.

VILLELA, Raphael; BUENO, Rafaela Soares. A expansão do desmatamento no estado do Pará: população, dinâmicas territoriais e escalas de análise. *Anais*, p. 1-15, 2017.

KENNEDY, James; EBERHART, Russell. Particle swarm optimization. In: *Proceedings of ICNN'95-international conference on neural networks*. IEEE, 1995. p. 1942-1948.

EBERHART, Russell; KENNEDY, James. A new optimizer using particle swarm theory. In: *MHS'95. Proceedings of the sixth international symposium on micro machine and human science*. IEEE, 1995. p. 39-43.