

UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL

**PREVISÃO DE DESMATAMENTOS ILEGAIS
BASEADO EM DADOS TEMPORAIS DE
SENSORIAMENTO REMOTO**

ROSANGELA MARIA GUARIENTI VENTURA

PROF. DR. CARLO RALPH DE MUSIS
Orientador

Cuiabá-MT
Março de 2021

UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL

**PREVISÃO DE DESMATAMENTOS ILEGAIS BASEADO EM
DADOS TEMPORAIS DE SENSORIAMENTO REMOTO**

ROSANGELA MARIA GUARIENTI VENTURA

Tese apresentada ao Programa de Pós-Graduação em Física Ambiental da Universidade Federal de Mato Grosso, como parte dos requisitos para obtenção do título de Doutora em Física Ambiental.

**PROF. DR. CARLO RALPH DE MUSIS
ORIENTADOR**

Cuiabá-MT
março de 2021

Dados Internacionais de Catalogação na Fonte.

G915p Guarienti Ventura, Rosangela Maria.
PREVISÃO DE DESMATAMENTOS ILEGAIS BASEADO EM DADOS
TEMPORAIS DE SENSORIAMENTO REMOTO / Rosangela Maria Guarienti
Ventura. -- 2021
61 f. : il. color. ; 30 cm.

Orientador: Carlo Ralph De Musis.
Tese (doutorado) - Universidade Federal de Mato Grosso, Instituto de Física,
Programa de Pós-Graduação em Física Ambiental, Cuiabá, 2021.
Inclui bibliografia.

1. Sentinel 2. 2. Crime ambiental. 3. Desflorestamento. I. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Permitida a reprodução parcial ou total, desde que citada a fonte.



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE MATO GROSSO
PRÓ-REITORIA DE ENSINO DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL

FOLHA DE APROVAÇÃO

TÍTULO: PREVISÃO DE DESMATAMENTOS ILEGAIS BASEADO EM DADOS TEMPORAIS DE SENSORIAMENTO REMOTO

AUTORA: DOUTORANDA ROSANGELA MARIA GUARIENTI VENTURA

Tese defendida e aprovada em **03 de fevereiro de 2021**.

COMPOSIÇÃO DA BANCA EXAMINADORA

1. Prof. Dr. Carlo Ralph De Muis - Examinador Externo - Universidade de Cuiabá – UNIC
2. Prof. Dr. Josiel Maimome de Figueiredo- Examinador Interno - Instituto de Computação- UFMT
3. Prof. Dr. Raphael de Souza Rosa Gomes - Examinador Interno - Instituto de Computação- UFMT
4. Dr. João Paulo Novaes Filho - Examinador Externo - Politec - Perícia Oficial e Identificação Técnica/Mato Grosso
5. Prof. Dr. Victor Hugo de Moraes Danelichen - Examinador Externo - Universidade de Cuiabá - UNIC

Cuiabá-MT, 03/02/2021



Documento assinado eletronicamente por **SERGIO ROBERTO DE PAULO, Coordenador(a) de Programas de Pós-Graduação em Física Ambiental - IF/UFMT**, em 05/02/2021, às 17:26, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **CARLO RALPH DE MUSIS, Usuário Externo**, em 05/02/2021, às 17:33, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **RAPHAEL DE SOUZA ROSA**



GOMES, Docente da Universidade Federal de Mato Grosso, em 05/02/2021, às 17:38, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **JOSIEL MAIMONE DE FIGUEIREDO, Docente da Universidade Federal de Mato Grosso**, em 05/02/2021, às 17:39, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Victor Hugo de Moraes Danelichen, Usuário Externo**, em 05/02/2021, às 17:44, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **João Paulo Novaes Filho, Usuário Externo**, em 10/02/2021, às 10:04, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufmt.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3191285** e o código CRC **B172C32D**.

DEDICATÓRIA

À Deus, minha família, amigos e professores.

AGRADECIMENTOS

Muitas vezes a dureza dos desafios parece tirar o brilho das conquistas, então Deus nos lembra que ele dá o frio conforme o cobertor. Quero agradecer à Deus por ter chegado aqui. Agradeço Ele pelo amor da minha família e do meu esposo, por minha paz, minha saúde, meu lar, meus amigos e meu trabalho. Espero contribuir cada vez mais com o meio ambiente.

Agradeço ao meu pai Osmar e à minha mãe Maria Abraão, que em General Carneiro almejavam ter uma filha doutora. Com oração, apoio e carinho me ajudaram conquistar meus sonhos.

Ao meu marido Thiago, por acreditar e apoiar esse trabalho. Todo seu apoio emocional, familiar e, claro o seu conhecimento técnico me permitiu chegar aqui. Eu não conseguiria ser mãe, trabalhar e estudar sem a sua participação. Obrigada por não me fazer escolher.

Ao meu irmão César, com quem divido a história de poder sair de um lugar simples, com grandes sonhos, e conquistá-los. É um prazer ser um ponto fora da curva como você.

Aos meus amigos que entenderam minha ausência e animaram minha presença durante essa etapa.

Ao tio Mário e tia Beth (*in memoriam*) que sempre incentivaram os meus estudos. Obrigada pela oportunidade. Vocês mudaram minha história.

Ao professor orientador Carlo Ralph De Mussis, por contribuir com minha formação e com esse trabalho. Sua compreensão foi calma em meio à tempestade. Por vezes, desistir deixou de ser opção por saber que poderia contar com seu apoio.

Ao Pe. João e Irmã Aurizena pelas orações e apoio amigo em toda a jornada. Este apoio é explicado assim como a minha Fé: não vejo, mas sinto, e, por isso, não há dúvidas.

Aos colegas de profissão e amigos da POLITEC, pelas trocas de plantões, silêncio na sala de trabalho e por serem aqueles com quem eu pude contar e me espelhar.

Agradeço a toda Coordenação e professores do Programa de Pós-Graduação de Física Ambiental, à Universidade Federal de Mato Grosso (UFMT) e a todos que contribuíram de alguma maneira para a conclusão desse trabalho.

EPÍGRAFE

“Façam como queiram...
Podar-me ou desmatar-me
Saibam que crescerei novamente!”

Julio Aukay.

SUMÁRIO

LISTA DE FIGURAS	ix
LISTA DE TABELAS	x
RESUMO	xi
ABSTRACT	xii
1. INTRODUÇÃO.....	1
1.1. PROBLEMÁTICA	1
1.2. JUSTIFICATIVA	4
1.3. OBJETIVO GERAL	5
1.4. OBJETIVOS ESPECÍFICOS.....	6
1.5. ORGANIZAÇÃO DO TRABALHO.....	6
2. FUNDAMENTAÇÃO TEÓRICA	7
2.1. SENSORIAMENTO REMOTO	7
2.2. DETECÇÃO DE DESMATAMENTOS	9
2.3. SENTINEL-2	11
2.4. BASE PÚBLICA DE DADOS	13
3. MATERIAL E MÉTODO.....	16
3.1. BASE DE DADOS	16
3.1.1. Dados de áreas desmatadas	16
3.1.2. Dados de sensoriamento remoto	17
3.1.3. Definição de áreas conservadas e desmatadas	20
3.2. CONCEPÇÃO DO MÉTODO PARA PREVISÃO DE DESMATAMENTOS ...	22
3.3. IMPLEMENTAÇÃO DA REDE NEURAL ARTIFICIAL	24
3.4. AVALIAÇÃO ESTATÍSTICA.....	26
4. RESULTADOS E DISCUSSÃO.....	28
4.1. CRIAÇÃO DA BASE DE DADOS RSDD	28
4.1.1. Organização dos dados tabulados	28
4.1.2. Organização das imagens.....	30
4.1.3. Características gerais sobre a base RSDD	31
4.2. MÉTODO PARA PREVISÃO DE DESMATAMENTO.....	32
4.2.1. Cenário para os testes.....	32
4.2.2. Resultado dos testes de classificação	34

4.2.3. Processamento com todo o Estado de Mato Grosso	36
5. CONSIDERAÇÕES FINAIS	38
5.1. PRINCIPAS CONTRIBUIÇÕES	38
5.2. PERSPECTIVAS FUTURAS.....	39
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	40

LISTA DE FIGURAS

Figura 1: Biomas do Estado de Mato Grosso.	1
Figura 2: Etapas realizadas para o processamento dos dados, desde a obtenção dos dados relacionados às áreas de desmate até a geração dos arquivos que compõe a base a ser processada.	18
Figura 3: Exemplos de <i>downsampling</i> por meio da média aritmética simples dos valores presentes. Em a) é ilustrado um ponto que contém resolução de 10m, em b) uma resolução de 20m e em c) uma resolução de 60m.	19
Figura 4: Estado de Mato Grosso dividido em 25 quadrantes e pelos seus biomas: Amazônia em verde, Cerrado em amarelo e Pantanal em azul.	21
Figura 5: Definição da entrada da rede neural, no qual são utilizadas as características extraídas das imagens dos momentos selecionados. Neste exemplo são usados dois momentos e extraídos quatro características de cada, como os índices calculados por meio das bandas das imagens.	25
Figura 6: Exemplo de imagens de um único ponto da base RSDD, no qual em a), b) e c) a área parece estar conservada, mas em (d) a área já está totalmente desmatada. A diferença de tempo entre c) e d) foi de menos de 30 dias. A área retangular vermelha consiste em um espaço de 60mx60m, que é a menor resolução espacial do Sentinel-2.	31
Figura 7: Mapa dos pontos utilizados na base RSDD - em verde os pontos conservados e em roxo os pontos que sofreram desmatamento.	33

LISTA DE TABELAS

Tabela 1: Relação entre as bandas espectrais dos sensores do Sentinel-2, sensores S2A e S2B, e seus respectivos comprimentos de onda e resoluções espaciais (ESA, 2020).....	12
Tabela 2: Exemplo de registros dos dados relacionados às áreas desmatadas, contendo longitude, latitude e quantos hectares foi desmatado.....	17
Tabela 3: Limiares inferiores e superiores de NDVI para definição de área desmatada em cada quadrante e bioma do Estado de Mato Grosso. O número do quadrante está na ordem da esquerda para direita, de cima para baixo. Os quadrantes 3, 4, 5 e 25 foram ignorados por possuírem pouca ou nenhuma área de Mato Grosso.....	21
Tabela 4: Exemplo dos dados no arquivo points.csv da base RSDD contendo ID do ponto, coordenada geográfica e área desmatada em hectares.....	29
Tabela 5: Exemplo dos dados no arquivo imgs.csv da base RSDD contendo ID do ponto, data, URL da imagem original e identificador único da imagem na plataforma Earth Explorer.....	29
Tabela 6: Exemplo dos dados no arquivo bands.csv da base RSDD contendo o ID do ponto, data, índice do quadrante e valores das bandas.....	30
Tabela 7: Diferença dos tamanhos de arquivos entre os dados originais e a RSDD.....	31
Tabela 8: Resultados de precisão e acurácia com o conjunto de testes para cada combinação de entradas, considerando a quantidade de momentos utilizados, uso dos valores das bandas e uso de índices pré-calculados.....	34
Tabela 9: Valores de true negatives (TN), false positives (FP), false negatives (FN) e true positives (TP) para cada combinação de entradas.....	35
Tabela 10: Medição de precisão e acurácia para cada quadrante do Estado de Mato Grosso usando os dados de agosto, setembro e início de outubro para prever os desmatamentos no final de outubro.....	36

RESUMO

VENTURA, R. M. G. *Previsão de desmatamentos ilegais baseado, em dados temporais de sensoriamento remoto*. 2021. 61p. Tese (Doutorado em Física Ambiental), Instituto de Física, Universidade Federal de Mato Grosso, Cuiabá, 2021.

Existem poucos estudos voltados a compreender e prever a dinâmica dos desmatamentos ilegais. Dentre as técnicas que possibilitam esse estudo o sensoriamento remoto utilizando imagens de satélites é uma das ferramentas mais difundidas. Porém, a obtenção e processamento dessas imagens não é uma tarefa trivial, principalmente devido à sua complexidade de espaço e também pela dificuldade de se obter pontos validados com a ocorrência do fenômeno desejado, além da criação de métodos para automaticamente realizar a classificação dessas áreas. Assim, é útil a disponibilização de uma base de dados que contém informações já processadas e validadas sobre locais que sofreram desmatamento. Desse modo, a base de dados possibilita reduzir esforços em futuras pesquisas, além de proporcionar uma base de dados comum para ser utilizada como benchmark em soluções de aprendizado de máquina. Nessa situação, foi criado o Remote Sensing Deforestation Dataset (RSDD) contendo dados tabulados e imagens de áreas afetadas pelo desmatamento. O RSDD possui dados sobre desmatamento validado pelo órgão estadual, Secretaria do Estado de Meio Ambiente de Mato Grosso, possui um tamanho menor do que 1% do tamanho das imagens originais de sensoriamento remoto e está disponibilizado gratuitamente. Com uma base como essa, é possível a criação de modelos de aprendizado de máquina para conseguir realizar a previsão de áreas desmatadas. Neste trabalho, foi construído um modelo de Redes Neurais Artificiais capaz de detectar áreas que serão desmatadas, obtendo uma precisão de 93,2% e uma acurácia de 94,8% nos melhores testes.

Palavras-chave: sentinel-2; crime ambiental; desflorestamento.

ABSTRACT

VENTURA, R. M. G. *Prediction of illegal deforestation based, on remote sensing temporal data*. 2021. 61p. Thesis (Doctorate in Environmental Physics), Institute of Physics, Federal University of Mato Grosso, Cuiabá, 2021.

In order to study and understand the dynamics of illegal deforestation, it is necessary to measure information about deforested areas. One of the ways to do this is through remote sensing using satellite images. However, obtaining and processing these images is not easy, mainly due to their complexity of space and also the difficulty of obtaining validated points with the occurrence of the desired phenomenon, in addition to the creation of methods to automatically perform the classification of these areas. Therefore, it is useful to make available a database containing information that has already been processed and validated about spots that have suffered deforestation. This type of database makes it possible to reduce efforts in future research, besides to providing a common database to be used as a benchmark in machine learning solutions. In this situation, the Remote Sensing Deforestation Dataset (RSDD) was created containing tabulated data and images of areas affected by deforestation. The RSDD has data validated by a State agency, it has a size less than 1% of the size of the original remote sensing images, and it is publicly available. With such a base, it is possible to create machine learning models to be able to forecast deforested areas. In this work, a model of Artificial Neural Networks was built capable of detecting areas that will be deforested, obtaining an accuracy of 93.2% and an accuracy of 94.8% on the best tests.

Keywords: sentinel-2; environmental crime; deforestation

1. INTRODUÇÃO

1.1. PROBLEMÁTICA

A região da Amazônia Legal ocupa 5.015.067,749 km² e acumulou até o ano de 2019 um total de 802.898,3 km² de área desmatada, considerando um acréscimo de 11.043,9 km² entre 2018 e 2019, segundo dados disponibilizados pela plataforma TerraBrasilis e o projeto PRODES. Os motivos que levam ao desmatamento são diversos, desde fatores históricos aos serviços ecossistêmicos, sejam eles legais ou não (ASSIS et al., 2019); (VALERIANO et al., 2004); (VANNINI & KUMMER, 2018); (CLECH et al., 2016).

De acordo com Fearnside (2008), a Amazônia Legal possui vários padrões de desmatamento associados a diferentes fatores e históricos de uso e cobertura da terra, e seu desmatamento é um grande problema ambiental. Mato Grosso é um dos Estados que compõe a Amazônia Legal e seu território conta com três biomas, Amazônia, Cerrado e Pantanal (Figura 1). Apesar da diferença de cobertura vegetal, níveis de aproveitamento agropastoril e interesses econômicos, o desmatamento ilegal atinge indiscriminadamente todo o território mato-grossense.

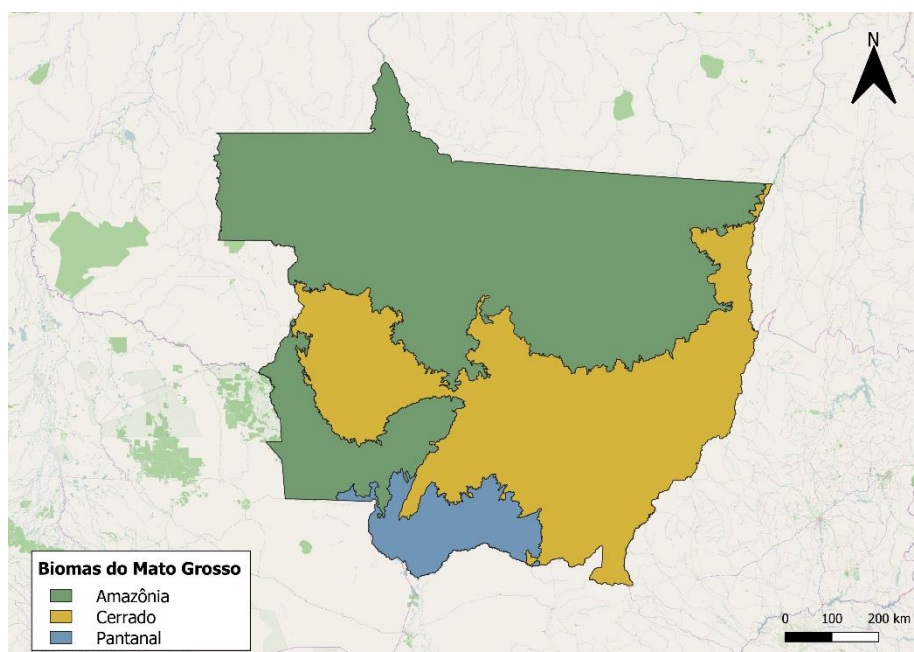


Figura 1: Biomas do Estado de Mato Grosso.

Podem ser citados como exemplos de problemas gerados pelo desmatamento a alterações nos microclimas, assoreamento de canais fluviais, perda de produtividade agrícola, alterações no regime hidrológico, perda de biodiversidade e emissões de gases de efeito estufa. Por motivos como estes, o governo e a sociedade planejam e realizam ações para combater o desmatamento. A lei N° 9.605, de 12 de fevereiro de 1998, Lei dos Crimes Ambientais, dispõe sobre as sanções penais e administrativas derivadas de condutas e atividades lesivas ao meio ambiente. Possui seção de crimes contra a flora, em que prevê pena, detenção, reclusão e multa de acordo com o crime realizado (FONSÊCA et al., 2019) (WATANABE et al., 2018) (SANTOS et al., 2017).

Diversos órgãos públicos do Estado atuam na identificação, combate e responsabilização dos crimes ambientais referentes ao desmatamento ilegal, tais como Secretaria Estadual de Meio Ambiente (SEMA), Delegacia Especializada em Meio Ambiente (DEMA), Ministério Público Estadual de Mato Grosso (MPMT), Corpo de Bombeiros Militar de Mato Grosso (CBMMT), Batalhão de Polícia Militar de Proteção Ambiental de Mato Grosso (BPMPA/PMMT) e Perícia Oficial e Identificação Técnica (POLITEC). Entretanto, a grande extensão territorial e números de ocorrências de desmates ilegais dificultam os órgãos competentes a direcionarem ações de combate e responsabilização criminal. Existe uma diferença temporal entre o início da degradação por desmatamento e finalização total do evento, que varia com o tempo de retirada das árvores de valor comercial da área. Com isso, identificar o início da degradação possibilita ações que direcionem visitas técnicas *in loco*, reduzam o tempo de levantamento de área e identifiquem áreas de maior incidência. Podem propiciar um número maior flagrantos do delito, impedindo a conclusão da degradação ambiental, possibilitando a responsabilização criminal do autor dos fatos.

Este trabalho parte da premissa relatada por profissionais que atua na área que na maioria dos casos de desmate ilegais existem três sequencias de ação. Sendo a 1° o corte seletivo e extração de madeiras de valor comercial, a 2° etapa consiste na limpeza da área com maquinário executando enleiramento da vegetação restante e a 3° etapa a queima da área. Considerando que essas alterações geram respostas sensíveis as bandas e índices analisados o trabalho irá prever o desmatamento a partir do treinamento dessas respostas aos sensores da etapa 1. Vide quadro 01

Quadro 1 - Etapas do desmatamento Ilegal

1º Etapa - Corte Seletivo (exploração). Nesta etapa pode ser encontrada trilhas de arraste das toras, acampamento de beneficiamentos da madeira para otimizar volume de transporte; abertura de estradas clandestinas para entrada de cominhões de transporte e/ou maquinários para derrubada.



2º Etapa - Corte raso, formação de leiras, limpeza. Nessa etapa ocorre com o uso de maquinários por isso é feita de maneira contínua e celere devido a logística envolvida.



3º Etapa - Queima. A queima pode ocorrer com origem nas leiras ou entorno, podendo ou não haver aceiros de contenção da dispersão do fogo.



1.2. JUSTIFICATIVA

Análises de imagens de satélite juntamente com índices possibilitam identificar o início de uma ação de desmate ilegal. Isso porque existem ações antrópicas sutis anteriores ao corte raso, que consistem em abertura de estradas, confecção de guias para o traçado das máquinas e retirada das árvores de valor comercial. Como o Satélite Sentinel-2, possui uma taxa de revisita de até 5 dias, disponibiliza imagens que permitem visualizar a alteração da cobertura de solo, e suas 13 bandas possibilitam cálculos de índices correlacionados a alteração da vegetação, sendo assim, identificam essas alterações.

A análise da mudança da cobertura florestal é necessária para entender a tendência da alteração da cobertura florestal, bem como as alterações do padrão de uso da terra de qualquer região (WOLFF et al., 2018). Esse monitoramento fornece uma visão geral de como o desmatamento está ocorrendo e qual pode ser a possível ameaça à natureza e à comunidade (MBA, 2018). Castelo et al. (2020) relataram que políticas e gestões ambientais são fundamentais para o controle do desmatamento, assim como analisar a dinâmica espaço-temporal do uso e cobertura da terra (LIMA et al., 2020). Mas, para embasar tais discussões, são necessárias informações para entender a realidade atual em relação ao desmatamento, assim como previsões a fim de auxiliar na tomada de decisões.

Em uma área tão ampla como o Estado de Mato Grosso, é fundamental o uso de sensoriamento remoto por imagens de satélite, no qual é possível obter uma grande quantidade de dados de uma extensa área. Além disso, é possível aplicar técnicas de aprendizado de máquina para interpretar os dados obtidos e conseguir classificá-los, tanto identificando áreas desmatadas como realizando previsões do delito.

Dessa forma, uma técnica muito útil é o sensoriamento remoto. As imagens de satélite registram a situação de cobertura vegetal no decorrer do tempo, por isso são ferramentas que auxiliam a prever e, quando não impedido, permitem evidenciar a dinâmica do desmate ilegal realizado com datas e áreas atingidas. Ao fixar a ação do desmate em data definida é possível obter índices correlacionados de cobertura vegetal que, atrelados a outros bancos de dados, possibilitam os estudos das dinâmicas dos desmates ilegais e suas consequências na área atingida.

Existe uma deficiência de pesquisas direcionadas a previsão de desmate ilegal dada a dificuldade de obter dados validados essenciais para uma modelagem do tema. Pesquisas que possam utilizar dados de Mato Grosso possibilitam otimizar previsões com as características locais. É intrínseco à pesquisa os dados de sensoriamento remoto e, para a obtenção destes, há uma exigência computacional por conta da grande extensão da base de dados em número e tamanho de arquivos. Conseqüentemente, para tratar essas informações existem limitações referente a configuração de máquinas para geração dos resultados e trabalho técnico com expertise de interpretação na área, além de demandar tempo na criação do banco de dados. Com isso, esse trabalho visa também facilitar essa etapa para estudos correlacionados, fornecendo uma base validada e com resoluções que necessitam de menor tempo de processamento.

Portanto, é possível utilizar de tais recursos combinados para criar uma ferramenta capaz de fornecer informações importantes sobre desmatamento ilegal para os interessados, sejam eles pesquisadores, agentes do governo ou sociedade civil. As ações de prevenção e combate seriam baseadas em mapas de calor de áreas de possível ação ilegal.

Neste trabalho temos como hipótese que as alterações antrópicas realizadas anterior ao corte raso de uma área que sofrerá desmate ilegal podem ser identificadas a partir de uma sequência de imagens de satélite de períodos diferentes. Assim, pode-se prever se a área será desmatada ou permanecerá conservada.

1.3. OBJETIVO GERAL

Este trabalho tem o objetivo de propor um modelo, baseado em dados de sensoriamento remoto e aprendizado de máquina, para realizar previsões de ações de desmatamentos ilegais no Estado de Mato Grosso.

1.4. OBJETIVOS ESPECÍFICOS

Para atingir o objetivo geral deste trabalho, foi necessário definir e concretizar os seguintes objetivos específicos:

- Obter dados validados de desmates ilegais no Estado de Mato Grosso;
- Construir uma base de dados provenientes de imagens de satélite que possui todas as informações necessárias para conseguir a modelagem de desmatamentos;
- Desenvolver um modelo com redes neurais artificiais capaz de classificar desmatamentos ilegais, tendo como entrada dados de sensoriamento remoto;
- Desenvolver e implementar modelos preditivos para desmatamentos ilegais com os dados de testes e para os meses seguintes.

1.5. ORGANIZAÇÃO DO TRABALHO

Este trabalho está organizado da seguinte forma:

- Capítulo 2: Apresenta os conceitos principais sobre sensoriamento remoto e aprendizado de máquina. Também são descritas informações referentes ao satélite utilizado para uso dos dados de sensoriamento remoto.
- Capítulo 3: Desenvolve os procedimentos utilizados para a geração da base de dados para este trabalho, assim como a proposta do desenvolvimento do método de previsão de desmatamento. Além disso, descreve como a avaliação do modelo foi realizada.
- Capítulo 4: São apresentados em detalhes a base de dados criada para possibilitar o treinamento de modelos no contexto de desmatamento, além das especificações do próprio método criado neste trabalho para realizar as previsões de desmatamentos ilegais.
- Capítulo 5: No capítulo final, são feitas as sínteses deste trabalho, além de listar as principais contribuições e indicar os possíveis trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os tópicos gerais sobre sensoriamento remoto, aprendizado de máquina e do satélite utilizado neste trabalho para obtenção dos dados. É importante o entendimento dos conceitos envolvidos no sensoriamento remoto, no qual as bandas permitem o cálculo de índices que servirão para entender o que está presente no solo. Para conseguir realizar as classificações e previsões, é necessário o uso de técnicas de aprendizado de máquina, cujo fundamento são discutidas neste capítulo também. O mesmo vale para o entendimento das características do satélite escolhido para obtenção dos dados de sensoriamento remoto.

2.1. SENSORIAMENTO REMOTO

O sensoriamento remoto é a técnica de obtenção de imagens dos objetos da superfície terrestre sem que haja um contato físico de qualquer espécie entre o sensor e o objeto (MENESES et al., 2012), inicialmente utilizando fotos aéreas e posteriormente também com os satélites artificiais. Este tipo de abordagem fornece vantagens, proporcionando um conjunto de dados que cobre grandes áreas com alto detalhe espacial e alta frequência temporal (HEROLD et al., 2001). Possibilitando, assim, o registro em imagens do histórico de mudança do uso do solo.

O potencial dos dados de sensoriamento remoto para o monitoramento da dinâmica espaço-temporal da cobertura da vegetação pode ser visto em Novo et al. (2005), como o mapeamento de ecossistemas alagáveis, avaliação da dinâmica sazonal de bacias, monitoramento da dinâmica espaço-temporal das fisionomias de biomas e padronização do mapeamento de uso e cobertura da terra. Há outros exemplos de aplicações, como análise de padrões espaciais de temperaturas em áreas urbanas (GALLO & XIAN, 2014), determinação de tipo de floresta (CARREIRAS et al., 2017), tipologia de paisagem (OLIVEIRA & VENTURIERI, 2018) e degradação de florestas (DUPUIS et al., 2020). Um leque de possibilidades que atendem a diversas demandas ambientais.

Para que essas aplicações sejam possíveis, são utilizados índices gerados a partir dos valores das bandas dos satélites. Em Xue & Su (2017), são descritos vários índices

que podem ser usados para estudos de vegetação, os quais podem ser aplicados para contribuir com a detecção de desmatamentos. Foram avaliados mais de 100 índices de vegetações, discutindo sua aplicabilidade e representatividade de acordo com a vegetação de interesse e precisão. É mencionado que um dos mais utilizados é o Normalized Difference Vegetation Index (NDVI), apesar de ser sensível aos efeitos do brilho do solo, cor do solo, atmosfera, nuvem e sombras. Assim como o NDVI, os outros índices possuem características que trazem vantagens e desvantagens à sua precisão. Ressalta que índices de vegetação combinando bandas visíveis e próximos ao infravermelho melhoraram significativamente a sensibilidade da detecção de vegetação. Xue & Su (2017) concluem que a escolha de um índice de vegetação precisa ser feita considerando suas características e, em seguida, combiná-lo com outros índices para serem aplicados em um ambiente específico.

Além do NDVI, que foi proposto por Rouse et al. (1974), e aplicado em vários estudos como em Soares et al. (2020), Li & Fox (2012) e Jong et al. (2011), podem ser citados também: Normalised Difference Built-up Index (NDBI) que foi utilizada para detectar áreas construídas, útil para estudos comparando áreas urbanas (ZHA et al., 2003; ZHANG et al., 2009); Normalised Difference Water Index (NDWI) utilizado para cobertura vegetal, mas desta vez estimando a quantidade de água líquida presente na vegetação (GAO, 1996) e para monitorar o risco de incêndio por causa de folhas secas (MAKI et al, 2004); Modified Normalised Water Index (MNDWI) que é uma modificação do índice NDWI, diminuindo possíveis ruídos (XU, 2006); Normalised Difference Bareness Index (NDBaI) que foi proposto por Zhao & Chen (2005) para identificar áreas com solo exposto; Soil Adjusted Vegetation Index (SAVI) desenvolvido por Huete (1988) para minimizar os efeitos da interferência da presença de solo na caracterização do dossel das plantas; e Normalized Burn Ratio (NBR) criado por Key et al. (2002) para utilizar do espectro que capta teor de água na vegetação ou solo, detectando vegetação seca ou exposição de solo seco após um incêndio. É possível utilizar índices como esses para auxiliar na detecção de desmatamentos, haja vista que o desmate altera significativamente estas respostas espectrais.

O sensoriamento remoto é amplamente utilizado pelos órgãos responsáveis pela fiscalização e de responsabilização de crime ambiental. O Ministério Público Federal (2020), por exemplo, instaurou somente no mês de maio de 2020, 1.023 ações civis

públicas contra 2.262 réus em razão de desmatamentos ilegais na Amazônia com base em laudos periciais produzidos a partir da análise de imagens de satélite e do cruzamento das informações com bancos de dados públicos. Este trabalho pode ser facilitado com a melhoria das técnicas de detecção de desmatamentos.

2.2. DETECÇÃO DE DESMATAMENTOS

Como mencionado, os dados de sensoriamento remoto podem ser utilizados para monitoramento de áreas desmatadas (LIMA et al., 2020; MAURANO et al., 2019; SILVA et al., 2019; ACHARD et al., 2007; HEROLD & JOHNS, 2007; ALVES, 2002), incluindo sistema de notificação diária (DINIZ et al., 2015). Mapear a probabilidade de desmatamento é crucial, pois pode se tornar uma ferramenta muito importante para evitar futuros desmatamentos (LAMBIN, 1994), que é o caso de alguns trabalhos na área.

Mayfield et al. (2017) estimaram a probabilidade de desmatamentos baseando-se nas características da área, como desmatamentos próximos, áreas protegidas, elevação do solo, estradas próximas, população da região e presença de rios. Em Nayar (2009) foi relatado um método que combinou diferentes modelos para prever desmatamentos de grandes áreas analisando dados de crescimento populacional, produto interno bruto, produção de biocombustíveis, madeiras e culturas agrícolas. Dados semelhantes foram usados em Dlamini (2016).

Uma forma de possibilitar essas detecções e previsões são por meio de técnicas de Inteligência Artificial (IA). Russell & Norvig (2013) diz que a IA tenta compreender e construir entidades inteligentes, abrangendo uma enorme variedade de subcampos de estudo. Ou seja, as técnicas da área de IA podem ser empregadas em diversos tipos de aplicações. Em aplicações no contexto de desmatamento, conjunto de dados são preparadas para mostrar aos métodos de IA as características de áreas que sofrem desmatamento e de áreas que se mantêm conservadas, possibilitando a classificação automática de áreas que ainda não tinham sido analisadas pelo método. Um ramo da área de IA que utiliza dos dados para extrair os padrões existentes é o aprendizado de máquina (ou *Machine Learning* – ML). Segundo Ertel (2017), aprendizado de máquina é um ramo da área de IA que possui métodos para que o

computador entenda as regras de um domínio baseando-se nos dados disponíveis. Diversos trabalhos utilizam de aprendizado de máquina para analisar desmatamentos, como os citados abaixo, e isso é possível porque existe um padrão predominante na sequencias de ações e danos resultantes.

Em Saha et al. (2020) foram testados três métodos de aprendizado de máquina: binary logistic regression, random forest e ensemble rotational forest. O objetivo foi estimar probabilidades de desmatamentos, tendo como dados de entrada índices extraídos das bandas do satélite Landsat. O trabalho de Voight et al. (2019) identificou áreas mais suscetíveis para futuras perdas florestais, utilizando Redes Neurais Artificiais (RNA) e dados relacionados a mudanças recentes de cobertura florestal que foram extraídas de imagens do Landsat.

Há também trabalhos envolvendo processamento de imagens, principalmente com deep learning utilizando redes neurais convolucionais. Esse tipo de rede possui bons resultados com imagens coloridas (LeCun et al., 2015), como em Mhatre et al. (2020) para classificação de florestas ou como em Ortega et al. (2019) para detectar desmatamentos.

Esses exemplos são algumas técnicas de aprendizado de máquina. Descobrir as técnicas mais adequadas para avaliação de desmatamento é um processo contínuo (SAHA et al., 2020). Uma possibilidade, conforme já mostrado em trabalhos anteriores, é o uso de RNAs, que é uma técnica computacional cujo funcionamento baseia-se no contexto da estrutura neural de seres vivos e que realizam tarefas de computação adquirindo conhecimento através de experiência, construída por um processo de aprendizagem, tendo como principais vantagens as características de adaptabilidade, generalização e tolerância a falhas (HAYKIN, 2003).

A técnica de RNA foi escolhida para esse trabalho, assim como o uso de dados obtidos em imagens do satélite da missão Copernicus Sentinel-2. O Sentinel-2 foi escolhido por ter maior resolução espacial e revisita em menor tempo que outros satélites como o Landsat e o CBERS.

Foi construída uma RNA com uma arquitetura comumente utilizada, que é do tipo perceptron de múltiplas camadas, contendo no total três camadas. O número de neurônios na camada de entrada varia de acordo com teste realizado, podendo ter 17, 13 ou 4 neurônios. O número de neurônios na camada intermediária é proporcional ao

tamanho da camada de entrada, no qual é utilizado o Teorema de Kolmogorov-Nielsen (Kovács, 1996) para definir o valor de $2n+1$, onde n é o número de neurônios na camada de entrada. A camada de saída tem apenas dois neurônios, representando as duas possíveis classes: previsão de desmatamento ou continuação da conservação. Por se tratar de um problema de classificação, a função de ativação Softmax foi escolhida para a camada de saída. Para os outros parâmetros foram realizados testes de gridsearch com as principais opções disponíveis na framework utilizada, no qual foi definida a função de ativação Relu para a camada intermediária, o otimizador Adam e Categorical Crossentropy como função de loss. Além disso, 200 épocas foram realizadas para cada treinamento uma vez que foi visto que não havia mais progresso no aprendizado após esse número de épocas.

2.3. SENTINEL-2

A missão Copernicus Sentinel-2 (ESA, 2020) é composta por dois satélites que orbitam a Terra, os quais são providos de um instrumento multiespectral com objetivo de monitorar toda a área continental bem como suas costas marinhas, visando auxiliar aplicações de agricultura, monitoramento de ecossistemas, manejo florestal, qualidade da água e mapeamento de desastres. O primeiro satélite foi lançado em 23 de junho de 2015, e o segundo, em 7 de março de 2017. Até julho de 2020 cerca de 20 milhões de produtos foram gerados e disponibilizados para download, culminando em um total de 10 Petabytes. Há planos para lançar mais dois satélites adicionais, além da substituição dos primeiros. O projeto é mantido pela *Earth Observation Programme* da União Europeia, em parceria com a *European Space Agency* (ESA), *European Organisation for the Exploitation of Meteorological Satellites* (EUMETSAT), *European Centre for Medium-Range Weather Forecasts* (ECMWF) e Mercator Océan. O acesso às imagens é gratuito, mediante cadastro.

Desde o início da missão do Sentinel 2, seus produtos foram utilizados em diversas aplicações nas mais variadas áreas como, por exemplo, na identificação de ferrugem amarela do trigo em plantações (Zheng et al., 2018), no desenvolvimento de um algoritmo para reconhecer áreas queimadas no continente africano (Roteta et al., 2018), na classificação de espécies de árvores na Suécia (Persson et al., 2018), no

mapeamento de algas na costa da Grécia (Traganos & Reinartz, 2018), na classificação da cobertura do solo no Vietnã (Noi & Kappas, 2017), na classificação do tipo de colheita em fazendas na Áustria (Vuolo et al., 2018) e na detecção de mineração artesanal e de pequena escala em Gana (Gallwey et al., 2020), além na detecção de desmatamentos (Candra, 2020).

Os satélites do Sentinel-2 são compostos por instrumentos de captura de informações que resultam em 13 bandas espectrais, possuindo resoluções espaciais variando de 10 a 60 metros, conforme Tabela 1.

No trabalho de Delegido et al. (2011) foi atestado a evolução do satélite, verificando as novas bandas adicionados no Sentinel-2, principalmente nas que auxiliam no cálculo de índices relacionados à vegetação, como o Green Leaf Area Index (LAI).

A resolução espacial do Sentinel-2 de até 10 metros pode trazer uma precisão maior nas classificações. Trabalhos foram realizados utilizando dados desta missão, como em Vargas et al. (2019) que foi feita detecção de desmatamentos alcançando acurácias maiores que 95%. Trabalho com objetivo semelhante foi feito em Bouvet et al. (2018), mas a abordagem nesse caso tenta localizar sombras causadas por árvores na borda do desmatamento. Nesse trabalho também houve acurácias altas nos testes.

Tabela 1: Relação entre as bandas espectrais dos sensores do Sentinel-2, sensores S2A e S2B, e seus respectivos comprimentos de onda e resoluções espaciais (ESA, 2020)

Bandas espectrais dos sensores do Sentinel-2	S2A		S2B		Resolução espacial (m)
	comprimento de onda central (nm)	Largura de Banda (nm)	comprimento de onda central (nm)	Largura de Banda (nm)	
1 – Coastal aerosol	442.7	21	442.3	21	60
2 – Blue	492.4	66	492.1	66	10
3 – Green	559.8	36	559.0	36	10
4 – Red	664.6	31	665.0	31	10
5 – Vegetation red edge	704.1	15	703.8	16	20

6 – Vegetation red edge	740.5	15	739.1	15	20
7 – Vegetation red edge	782.8	20	779.7	20	20
8 – NIR	832.8	106	833.0	106	10
8A – Narrow NIR	864.7	21	864.0	22	20
9 – Water vapour	945.1	20	943.2	21	60
10 – SWIR – Cirrus	1373.5	31	1376.9	30	60
11 – SWIR	1613.7	91	1610.4	94	20
12 – SWIR	2202.4	175	2185.7	185	20

Diferente dos trabalhos citados anteriormente, este trabalho utiliza a classificação para previsão de desmatamento, e não apenas classificação. Neste trabalho bandas do Sentinel-2 são utilizadas para as estimativas de desmatamentos, aproveitando da periodicidade, com capacidade de revisita de 5 dias, e resolução espacial dos dados disponibilizados por este satélite. No desmatamento ilegal os polígonos predominantemente possuem área igual ou superior a 60 hectares, portanto, mesmo a banda com resolução espectral de 60m é eficaz para o objetivo deste trabalho. Na Seção 3.1.2 são descritos os detalhes de como os dados do Sentinel-2 são aproveitados neste trabalho.

2.4. BASE PÚBLICA DE DADOS

É comum os pesquisadores terem o trabalho de aquisição e tratamento de dados, mas eles também podem aproveitar de base de dados compartilhadas por outros pesquisadores. A ImageNet (RUSSAKOVSKY et al., 2015), por exemplo, é uma base de dados que contém milhões de imagens separadas em centenas de categorias. Há informações estruturadas de quais objetos estão presentes em cada imagem, possibilitando que métodos de aprendizado de máquina sejam treinados para conseguir classificar automaticamente as imagens. Essa base de dados foi a responsável por

impulsionar a qualidade de classificação de objetos presente em imagens por computadores, sendo aplicado em vários tipos de aplicações.

Algo semelhante existe para áudios, em que bases com áudios são disponibilizadas para diversos fins, como em Defferrard et al. (2017), com o Free Music Archive, e em Gemmeke et al. (2017), com o Audio Set. O primeiro contém dados de coleções de músicas para que sejam criados sistemas para detecção de gêneros musicais, enquanto o segundo é composto de categorias diversas, como sons humanos (voz, respiração, ações humanas), sons de animais (domésticos e selvagens), sons naturais (vento, água, fogo), sons musicais (instrumentos, gênero), sons de objetos (veículos, alarmes, ferramentas), dentre outros.

Obter, catalogar e extrair informações dos dados primários é uma tarefa que os autores dessas bases fizeram e que outros pesquisadores não necessitam realizar, uma vez que os dados já estão disponibilizados. Ideias semelhantes estão concentradas em repositórios como o UC Irvine Machine Learning Repository (DUA & GRAFF, 2020), contendo diversas bases de dados para auxiliar no treinamento de aprendizado de máquina. Desta forma, há uma economia de tempo significativo nas pesquisas, que podem iniciar as atividades para completar os seus objetivos, evitando retrabalho na comunidade científica. Ganha-se também que uma mesma base de dados subsidiará diversos estudos e suas correlações podem produzir trabalhos mais amplos ao integrar as informações com os demais.

Para a área ambiental, é possível encontrar alguns exemplos de base de dados públicas, como o Forest Fires Data Set (CORTEZ & MORAIS, 2007) que é destinado a prever áreas de queimadas e o Global Forest Change (HANSEN et al., 2013) no qual contém índices e informações previamente calculados para visualização de áreas desmatadas. Entretanto, essas bases de dados já estão com os dados processados, limitando as abordagens dos pesquisadores, ou disponibilizam gigabytes de dados para que os pesquisadores realizem o download dos dados brutos, o que gera as dificuldades de obtenção, armazenamento e processamento de uma grande quantidade de dados.

Por não ser encontrada uma base de dados com características desejáveis, foi necessário a criação de uma nova base. Entretanto, pensando nas dificuldades que os pesquisadores encontram quando realizam estudos nesta área, foi planejado a disponibilização da base criada para auxiliar tais pesquisas, fornecendo dados

validados de áreas desmatadas ilegalmente e em um formato mais fácil de se obter, armazenar e processar do que se comparado às imagens originais de satélite.

3. MATERIAL E MÉTODO

3.1. BASE DE DADOS

Os trabalhos de modelagem exigem uma ou mais bases de dados para que as hipóteses levantadas possam ser testadas e discutidas. Para este trabalho, duas fontes de dados foram necessárias para a criação da própria base de dados: i) banco de coordenadas onde aconteceram desmatamentos; ii) dados de sensoriamento remoto com as características de áreas conservadas e desmatadas.

3.1.1. Dados de áreas desmatadas

Os dados utilizados para este trabalho são referentes ao Estado de Mato Grosso (MT), que é um dos nove estados da Amazônia Legal e apresenta alto índice de desmatamento. Em Mato Grosso há a Secretaria de Estado do Meio Ambiente (SEMA), que tem a missão de implementar políticas públicas visando a conservação ambiental para a sustentabilidade no Estado.

Dias et al. (2018), reuniram análises dos desmatamentos ilegais ocorrido no ano de 2017 em MT, validados por analistas ambientais da SEMA. Nesse relatório, os dados de desmates de 2016 foram utilizados para verificar a diferença de cobertura vegetal em 2017, possibilitando a seleção de apenas novos desmatamentos, ou seja, que ocorreram em 2017.

As áreas de desmates detectadas foram auditadas por diferentes analistas e, depois, cruzadas com outras bases da SEMA, como autos de infração e termos de embargos, para garantir a qualidade da informação. Desta forma, tornou-se possível obter dados de confiança relacionados ao desmatamento da região selecionada no ano de 2017.

Os dados fornecidos contêm informações sobre os polígonos de cada área desmatada. Com base nesses dados, foram extraídas as coordenadas centrais e a quantidade em hectares da área desmatada. A Tabela 2 apresenta exemplos de registros desta base.

Tabela 2: Exemplo de registros dos dados relacionados às áreas desmatadas, contendo longitude, latitude e quantos hectares foi desmatado.

Longitude	Latitude	Hectares
-56,000072	-15,904708	4,5787
-57,484136	-15,556182	69,4194
-57,479366	-15,567944	57,2877
-57,462894	-15,550835	2,6352
...

Além das informações referentes às áreas desmatadas em Mato Grosso, foram utilizados dados de satélite para verificar os valores por sensoriamento remoto de cada área.

3.1.2. Dados de sensoriamento remoto

Para o processamento dos dados foram realizados os passos descritos na Figura 2. Após obter as coordenadas das áreas desmatadas, foi realizado o download das imagens do Sentinel-2 que continham os pontos selecionados. O download foi feito por meio da API disponibilizada pelo portal Earth Explorer da United States Geological Survey (USGS), onde é possível realizar pesquisas com o intuito de encontrar imagens de diversos satélites que satisfazem critérios pré-estabelecidos.

As imagens baixadas eram analisadas para verificar sua utilidade para o trabalho. Primeiramente, pontos que tivessem valores nulos em suas bandas eram descartados. O mesmo foi feito se houvesse a presença de nuvens ou corpos d'água. Para tanto foi utilizado o algoritmo Fmask (Zhu & Woodcock, 2015).

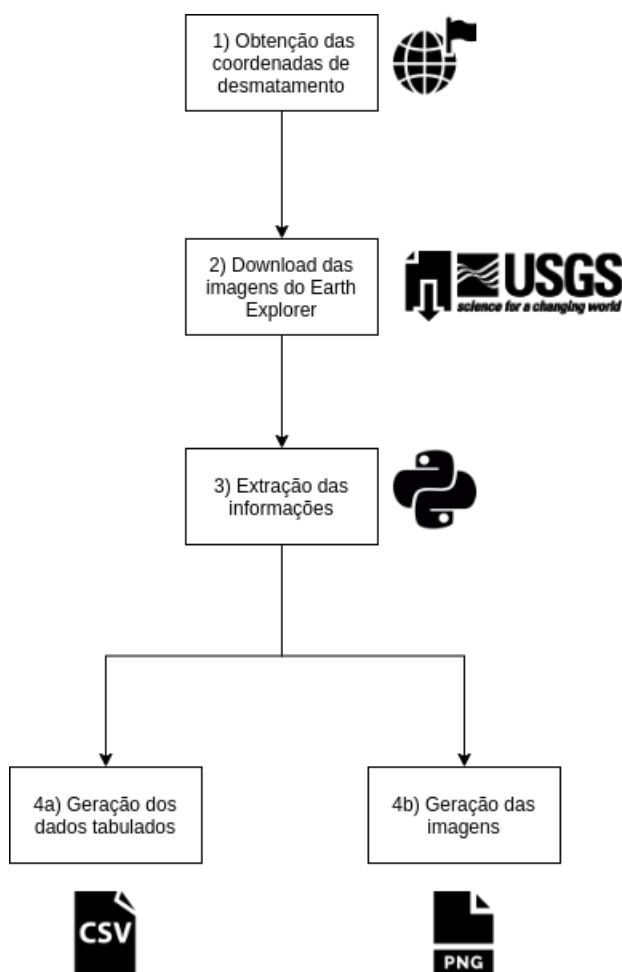


Figura 2: Etapas realizadas para o processamento dos dados, desde a obtenção dos dados relacionados às áreas de desmate até a geração dos arquivos que compõe a base a ser processada.

Posteriormente, foi definido que o tamanho mínimo da área de interesse (AOI) seria de 3.600 m². Devido a diferença da resolução espacial de cada banda, se torna interessante utilizar uma área que não desperdice os dados das bandas com menor resolução espacial. Para as bandas de resolução espacial de 10m, por exemplo, foram consideradas 36 leituras (6x6). Na base de dados foi adicionado o atributo index, representando essas 36 leituras.

Contudo, para processamentos que necessitam a mesma quantidade de informação entre as bandas, foi necessário realizar um procedimento de *downsampling* para padronizá-las. Para as resoluções inferiores à 60m, ou seja, as resoluções de 10m e 20m, a média aritmética simples dos valores existentes no ponto analisado foi obtida para representar à região, conforme ilustrado na Figura 3.

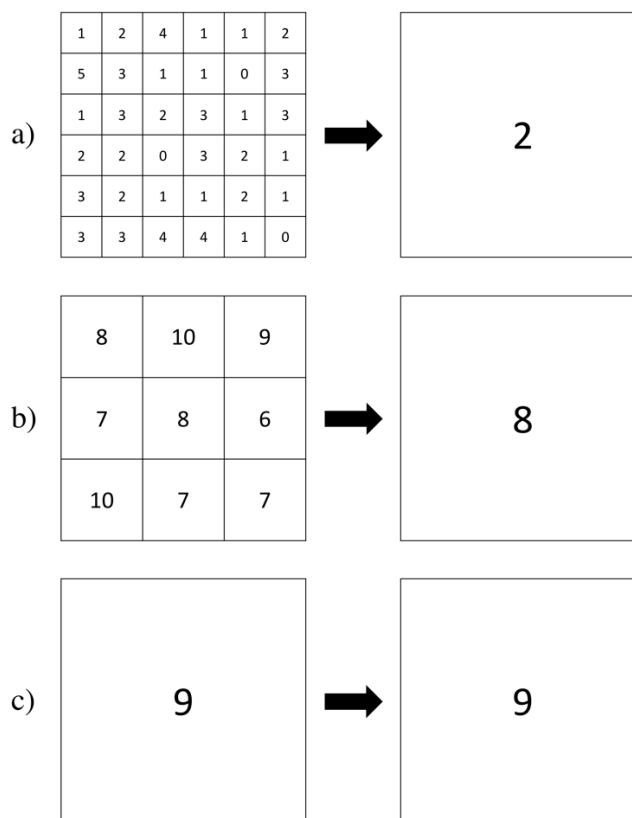


Figura 3: Exemplos de *downsampling* por meio da média aritmética simples dos valores presentes. Em a) é ilustrado um ponto que contém resolução de 10m, em b) uma resolução de 20m e em c) uma resolução de 60m.

Para a manipulação das imagens de satélite foi utilizada a linguagem de programação Python em conjunto com as bibliotecas GDAL¹ e Rasterio², que fornecem um conjunto de ferramentas mais apropriadas para o ambiente Python. Em seguida, foram gerados arquivos com os dados tabulados no formato CSV, assim como as imagens que representam as AOI's em cada imagem de satélite. As imagens foram extraídas da True Color Image (TCI) que acompanha as bandas da imagem, sendo feita uma normalização dos valores das bandas (RGB), possibilitando uma melhor visualização.

Os dados tabulados em arquivos texto no formato CSV (*comma separated values*) são úteis para aplicar em métodos de aprendizagem de máquina, no qual os dados já estão em formato numéricos e estruturados, sendo facilmente importados em diferentes sistemas e linguagens de programação. Já as imagens, estão no formato

¹ <https://gdal.org/>

² <https://rasterio.readthedocs.io/>

PNG (Portable Network Graphics), sendo facilmente carregadas e visualizadas. Elas foram centralizadas no ponto da coordenada que houve o desmate e podem servir para verificação dos resultados e análises mais detalhadas. Formando a base de dados criada neste trabalho, denominada *Remote Sensing Deforestation Dataset* (RSDD).

3.1.3. Definição de áreas conservadas e desmatadas

Foi mencionado a obtenção de áreas desmatadas no ano de 2017, contudo, não há a informação da data exata do desmatamento. Além do mais, não há informação de quais áreas estão conservadas no ano de 2017. Portanto, foi necessário definir uma abordagem para classificar uma determinada imagem como uma área conservada ou desmatada.

Bayma & Sano (2015) compararam índices de vegetação para definir se um local estava desmatado ou não no bioma do Cerrado, determinando limiares dependendo da formação vegetal, no qual o NDVI obteve os melhores resultados. Para este trabalho foi feita uma análise com o mesmo índice para determinar limiares de desmatamento de acordo com as áreas do Estado de Mato Grosso.

A área de Mato Grosso foi dividida em 25 quadrantes, de acordo com a Figura 4. Para cada quadrante foi realizada uma análise visual das imagens existentes a fim de anotar com quais valores de NDVI aconteciam os desmatamentos. Como os valores variavam, foram definidos limiares inferiores e superiores manualmente para cada quadrante, assim como em Bayma & Sano (2015), sendo que valores de NDVI menores que o limiar inferior significam que a área está desmatada e valores maiores que o limiar superior significam que a área estava conservada. Foram evitadas áreas com NDVI entre o limiar inferior e superior por ter um menor grau de confiança na classificação.

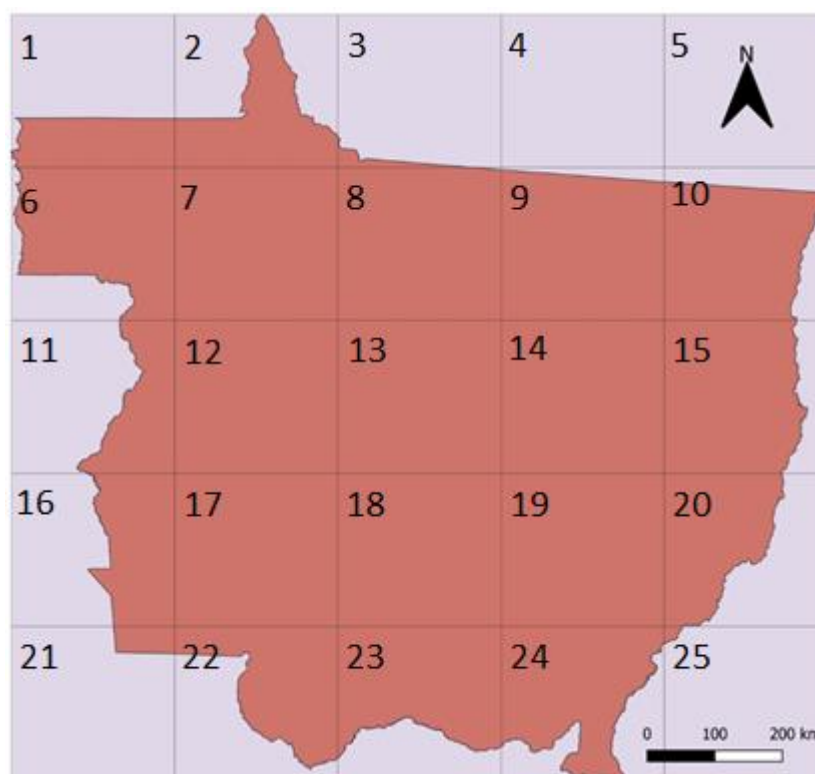


Figura 4: Estado de Mato Grosso dividido em 25 quadrantes.

Além da definição dos limiares, cada quadrante foi também classificado de acordo com o seu bioma predominante. A Tabela 3 apresenta os limiares encontrados para cada quadrante, assim como o respectivo bioma.

Tabela 3: Limiares inferiores e superiores de NDVI para definição de área desmatada em cada quadrante e bioma do Estado de Mato Grosso. O número do quadrante está na ordem da esquerda para direita, de cima para baixo. Os quadrantes 3, 4, 5 e 25 foram ignorados por possuírem pouca ou nenhuma área de Mato Grosso.

Quadrante	Limiar inferior	Limiar superior	Bioma predominante
1	0,51	0,61	Amazônia
2	0,51	0,65	Amazônia
3	-	-	-
4	-	-	-
5	-	-	-
6	0,41	0,65	Amazônia

7	0,48	0,68	Amazônia
8	0,49	0,60	Amazônia
9	0,52	0,58	Amazônia
10	0,49	0,55	Amazônia
11	0,48	0,57	Amazônia
12	0,44	0,57	Cerrado
13	0,50	0,58	Amazônia
14	0,46	0,56	Amazônia
15	0,48	0,55	Cerrado
16	0,48	0,58	Amazônia
17	0,50	0,64	Amazônia
18	0,50	0,59	Cerrado
19	0,46	0,54	Cerrado
20	0,52	0,62	Cerrado
21	0,50	0,63	Amazônia
22	0,49	0,55	Pantanal
23	0,48	0,61	Pantanal
24	0,47	0,55	Cerrado
25	-	-	-

3.2. CONCEPÇÃO DO MÉTODO PARA PREVISÃO DE DESMATAMENTOS

O desmatamento ilegal ocorre normalmente seguindo uma sequência de ações. Primeiramente, é feita a retirada das madeiras mais nobres, depois as madeiras para a construção civil e, por fim, são colhidas as árvores de madeiras leves remanescentes, para a produção de compensados e placas. Depois, as árvores de menor porte são derrubadas e toda a vegetação rasteira é destruída (INPE, 2008).

O método proposto se baseia nessa sequência de ações, tentando detectar modificações em uma determinada área que reflita um desmate seletivo (retirada de árvores selecionadas). Essas ações sendo detectadas, indicaria que posteriormente haverá a ação de desmatamento com corte raso, sendo realizada a previsão do desmate ilegal. O ano das imagens foram o ano de 2017, meses de julho, agosto e setembro.

Para realizar a detecção das primeiras ações foram utilizados dados temporais obtidos nas imagens do Sentinel-2. Para cada área a ser analisada, são considerados momentos diferentes com intervalos de até três meses. Os três meses escolhidos foram os meses de julho, agosto e setembro, que costumam ser os meses de maior incidência de desmate. Desta forma, é possível estudar a evolução do desmatamento, verificando se há um comportamento de desmate seletivo e, posteriormente, o desmatamento de corte raso.

Para cada momento são obtidos os dados das 13 bandas do Sentinel-2. Com os valores dessas bandas são calculados índices para representar as características da área. Foram selecionados quatro índices para servir de dados de entrada para o modelo: NDVI (Equação 1), MNDWI (Equação 2), NBR (Equação 3) e SAVI (Equação 4).

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

$$MNDWI = \frac{green-MIR}{green+MIR} \quad (2)$$

$$NBR = \frac{NIR-SWIR}{NIR+SWIR} \quad (3)$$

$$SAVI = \frac{NIR-red}{NIR+red+L} (1 + L) \quad (4)$$

Nas equações citadas são utilizadas as bandas que representam as refletâncias de infravermelho próximo (NIR), infravermelho médio (MIR), infravermelho de ondas curtas (SWIR), vermelho (red) e verde (green), além de um fator de ajuste para o substrato do dossel (L).

A seleção desses índices foi devido ao fato do NDVI, conforme mostrado, ser amplamente utilizado em aplicação que envolvem vegetação; do MNDWI, pela sua capacidade de obtenção de informação relacionada à água na vegetação, fornecendo uma perspectiva diferente; do NBR, por trazer uma visão relacionada à queimada, que pode estar relacionado ao desmatamento ilegal; e, do SAVI, por conseguir verificar diferentes estágios de cultivos, o que pode auxiliar na diferença de estágios de desmatamentos.

Assim, temos os dados das bandas e dos índices calculados para cada imagem (momento) obtida. Ao utilizar, por exemplo, dois momentos para análise, 34 valores são usados como características para classificar aquela área, sendo 13 valores referente as bandas, mais 4 valores referente aos índices, para cada um dos 2 momentos (um para quando a área estava conservada e outro momento quando há a possibilidade de haver corte seletivo). A mesma lógica pode ser utilizada para três ou mais momentos diferentes ao invés de dois. As características obtidas são utilizadas em uma técnica de aprendizado de máquina, como Rede Neural Artificial, para conseguir realizar a classificação.

3.3. IMPLEMENTAÇÃO DA REDE NEURAL ARTIFICIAL

Conforme mencionado, para cada momento considerado são obtidas 17 características (13 valores de bandas e 4 índices pré-calculados). Pelo menos dois momentos são necessários para se ter a análise de temporalidade, mas dependendo da disponibilidade de imagens mais momentos podem ser utilizados. A Figura 5 ilustra a rede neural criada tendo como entrada as características extraídas de dois momentos diferentes, assegurando que a rede possa abstrair as informações de temporalidade verificando as diferenças entre os momentos.

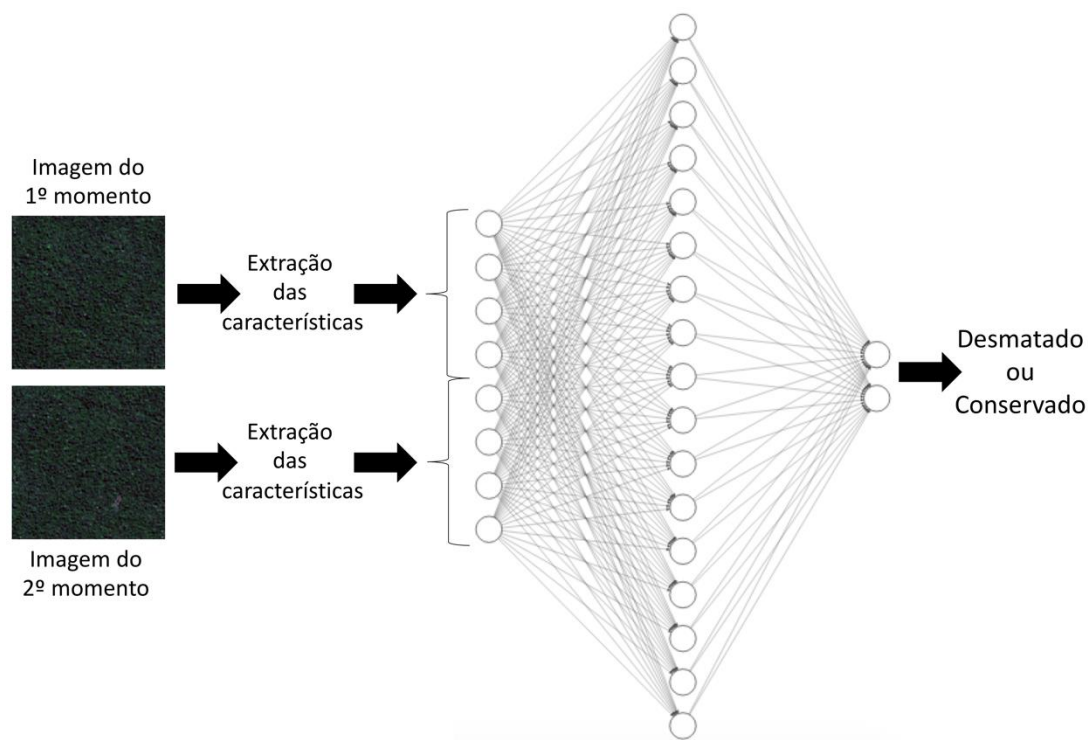


Figura 5: Definição da entrada da rede neural, no qual são utilizadas as características extraídas das imagens dos momentos selecionados. Neste exemplo são usados dois momentos e extraídos quatro características de cada, como os índices calculados por meio das bandas das imagens.

Para este trabalho foram feitos testes com dois e três momentos. E, para cada um deles, foi analisado o uso de todas as 17 características, mas também só com o uso das bandas ou só com o uso dos índices. Assim, a quantidade de características utilizadas como entrada variava para cada teste, gerando um total de seis combinações:

- 2 momentos, utilizando todas as 17 características
- 2 momentos, utilizando apenas os 13 valores das bandas
- 2 momentos, utilizando apenas os 4 valores dos índices pré-calculados
- 3 momentos, utilizando todas as 17 características
- 3 momentos, utilizando apenas os 13 valores das bandas
- 3 momentos, utilizando apenas os 4 valores dos índices pré-calculados

Foi construída uma RNA do tipo perceptron de múltiplas camadas, contendo no total três camadas. O número de neurônios na camada de entrada varia de acordo com teste realizado, podendo ter 17, 13 ou 4 neurônios. O número de neurônios na

camada intermediária é proporcional ao tamanho da camada de entrada, no qual é utilizado o Teorema de Kolmogorov-Nielsen (Kovács, 1996) para definir o valor de $2n+1$, onde n é o número de neurônios na camada de entrada. A camada de saída tem apenas dois neurônios, representando as duas possíveis classes: previsão de desmatamento ou continuação da conservação. Foi utilizada a função de ativação *Relu* na camada intermediária e *Softmax* na camada de saída, otimizador *Adam* e *Categorical Crossentropy* como função de *loss*, 200 épocas foram realizadas para cada treinamento.

Para a implementação da RNA foi utilizada a linguagem de programação Python no ambiente do Google Colaboratory com processador Intel Xeon 2.00GHz e 12GB de memória RAM.

Os códigos desenvolvidos neste trabalho estão disponibilizados como *open source* no repositório <https://github.com/demusis/srdesmatamentos>.

3.4. AVALIAÇÃO ESTATÍSTICA

O método proposto deve ser capaz de classificar as regiões que serão desmatadas e as que continuarão conservadas. Portanto, há quatro situações que podem ocorrer nos testes de acordo com a classificação do método:

- O método previu o desmatamento e a área foi realmente desmatada (*true positive* – TP);
- O método previu o desmatamento, mas na verdade a área manteve-se conservada (*false positive* – FP);
- O método não previu o desmatamento, mas a área foi desmatada (*false negative* – FN);
- O método não previu o desmatamento e a área continuou conservada (*true negative* – TN).

Foram calculados as precisões e acurácias dos testes de acordo com as Equações 5 e 6.

$$\textit{precis\~{a}o} = \frac{TP}{TP+FP} \quad (5)$$

$$\textit{acur\~{a}cia} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

4. RESULTADOS E DISCUSSÃO

Os resultados deste trabalho podem ser divididos em dois. Primeiramente, a criação de uma base de dados para modelagem de desmatamentos, que pode ser útil para outros trabalhos sobre o mesmo tema. Em segundo lugar, a criação de um novo método para previsão de desmatamentos. Esses resultados são apresentados nas seções a seguir.

4.1. CRIAÇÃO DA BASE DE DADOS RSDD

Conforme mencionado, os dados processados foram armazenados tanto no formato CSV (para os dados tabulados) quanto no formato PNG (para as imagens). A junção de todos esses dados compõe a base de dados criada neste trabalho, denominada *Remote Sensing Deforestation Dataset* (RSDD).

4.1.1. Organização dos dados tabulados

Os dados tabulados estão divididos em três arquivos CSV. O primeiro deles, chamado “points.csv” (Tabela 4) contém informações sobre os pontos existentes na base de dados, com seu ID, latitude, longitude e os hectares desmatados. O segundo arquivo é o “bands.csv” (Tabela 5), no qual possui dados extraídos das bandas das imagens de satélite para cada ponto da base de dados, com ID do ponto, data, índice do ponto e os valores das 13 bandas. Por fim, o arquivo “imgs.csv” (Tabela 6) tem informações sobre as imagens geradas, contendo o ID do ponto, data e a URL da imagem original em que foi extraída os dados dos pontos.

Tabela 4: Exemplo dos dados no arquivo points.csv da base RSDD contendo ID do ponto, coordenada geográfica e área desmatada em hectares.

PointId	Longitude	Latitude	Deforested Hectares
0	-56.000072	-15.904708	4,5787
1	-57.484136	-15.556182	69,4194
2	-57.479366	-15.567944	57,2877
		...	
1600	-52.101446	-14.387942	62,7050

Tabela 5: Exemplo dos dados no arquivo imgs.csv da base RSDD contendo ID do ponto, data, URL da imagem original e identificador único da imagem na plataforma Earth Explorer.

PointId	Year	Month	Day	Image URL (Earth Explorer)	ImageId (Earth Explorer)
0	2017	07	01	https://earthexplorer.usgs.gov/download/external/options/SENTINEL_2A/2191125/INVSVC/	2191125
0	2017	07	26	https://earthexplorer.usgs.gov/download/external/options/SENTINEL_2A/3022286/INVSVC/	3022286
0	2017	09	24	https://earthexplorer.usgs.gov/download/external/options/SENTINEL_2A/3022286/INVSVC/	1901701
1	2017	07	01	https://earthexplorer.usgs.gov/download/external/options/SENTINEL_2A/3022616/INVSVC/	3022616
				...	
12588	2017	09	18	https://earthexplorer.usgs.gov/download/external/options/SENTINEL_2A/2158901/INVSVC/	2158901

Tabela 6: Exemplo dos dados no arquivo bands.csv da base RSDD contendo o ID do ponto, data, índice do quadrante e valores das bandas.

PointId	Year	Month	Day	Index	Band 1	...	Band 12
0	2017	07	01	1	1202		1341
0	2017	07	01	2	1202	...	1341
			...				
1599	2017	08	04	36	1142		826

Com esses dados é possível treinar métodos de aprendizado de máquina para realizar as classificações de desmatamento.

4.1.2. Organização das imagens

As imagens geradas estão armazenadas em uma única pasta. A organização acontece no nome do arquivo, o qual contém informações sobre o ID do ponto e data da imagem. Em conjunto com as informações armazenadas em “point.csv” e “imgs.csv”, o pesquisador terá todos os dados necessários para processar tais imagens e, assim, auxiliar em suas aplicações.

Há mais de uma imagem para cada ponto, representando a evolução do desmatamento. Na Figura 6 são apresentados quatro momentos representando áreas que parecem estar conservadas e uma área totalmente desmatada. Ao todo, 6.400 imagens estão disponibilizadas nesta base de dados, elas cobrem uma área de 60mx60m, portanto 3.600m², de locais onde houve desmate ilegal no ano de 2017. As imagens estão no formato PNG com resolução 20x20 pixels, o que abrange a área representada pelo ponto central do desmate e o seu entorno.

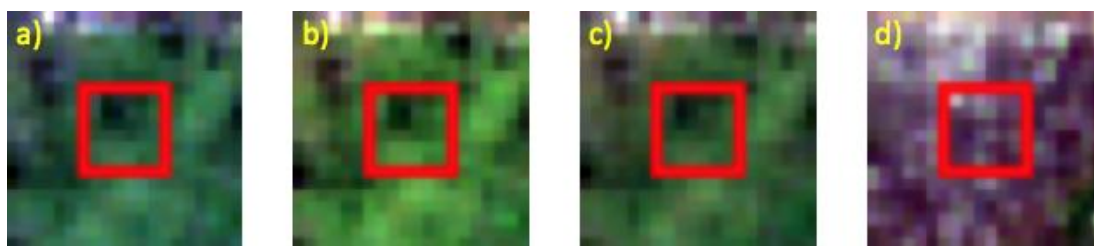


Figura 6: Exemplo de imagens de um único ponto da base RSDD, no qual em a), b) e c) a área parece estar conservada, mas em (d) a área já está totalmente desmatada. A diferença de tempo entre c) e d) foi de menos de 30 dias. A área retangular vermelha consiste em um espaço de 60mx60m, que é a menor resolução espacial do Sentinel-2.

4.1.3. Características gerais sobre a base RSDD

Há vantagens em utilizar a base de dados RSDD. A principal delas é a facilidade em obter dados essenciais para uma modelagem de desmatamento. Caso contrário, há dificuldades na obtenção de dados de sensoriamento remoto pela sua exigência computacional. A Tabela 7 apresenta as diferenças dos tamanhos de arquivos entre esta base de dados e os dados originais.

Tabela 7: Diferença dos tamanhos de arquivos entre os dados originais e a RSDD.

Tipo de arquivo	Dados originais	RSDD
Imagens	267 GBs	26,2 MBs
CSVs	-	53,1 MBs
Total	267 GBs	79,3 MBs

Portanto, são 266,2 GBs de dados a menos a serem baixados, armazenados e processados. A base RSDD tem menos de 1% do tamanho dos dados originais. Nesta base de dados, apenas 79,3 MBs são necessários caso o pesquisador tenha a necessidade de trabalhar com esse tipo de informação, sendo viável praticamente para qualquer pesquisador.

Essa diferença se dá principalmente pelo pré-processamento realizado, deixando somente informações válidas e pertinentes às áreas de desmatamento. As imagens do satélite são disponibilizadas em grânulos de tamanho de 100km por 100km, ocupando cerca de 700MB por imagem. Mas, neste tipo de pesquisa não é

necessário utilizar toda a área que uma imagem pode abranger, mas sim somente as áreas de interesse. Assim, pesquisadores que desejam estudar esse tipo de ação podem baixar o RSDD ao invés das imagens originais, que possuem em sua maioria áreas desnecessárias para o estudo.

Além da diferença de espaço para download, armazenamento e processamento, há também a vantagem desta base de dados disponibilizar o mapeamento e validação de várias áreas de desmate. Isso é algo necessário para o treinamento de métodos estatísticos ou de aprendizado de máquina. Nesta base de dados está agrupado 800 pontos de área de desmate ilegal validadas. Para cada ponto existe a imagem do momento do desmate e três imagens de momentos anteriores ao fato, dentro de um espaço temporal de três meses.

4.2. MÉTODO PARA PREVISÃO DE DESMATAMENTO

Conforme descrito nas Seções 2.2 e 3.3, o método proposto neste trabalho utiliza Redes Neurais Artificiais para prever desmatamentos ilegais. Para tanto, a entrada de dados da RNA recebe os valores das bandas existentes no Sentinel-2 e índices que refletem a vegetação da área que está sendo analisada. A saída do modelo tem apenas duas opções: desmatado ou conservado. Depois do desenvolvimento do modelo foram realizados testes aproveitando da base de dados RSDD. Os testes criados estão descritos na seção a seguir.

4.2.1. Cenário para os testes

O objetivo do teste foi conseguir prever se em um determinado mês uma área será desmatada ou se continuará conservada. Para realizar essa previsão, o modelo tinha a disposição dados dos dois meses anteriores da respectiva área.

O processo da seleção de pontos de desmatamento levou em consideração áreas que tinham no mínimo 4 imagens válidas entre os períodos de julho a setembro do ano de 2017, sendo que a imagem mais antiga tinha que haver vegetação e a imagem

mais recente tinha que estar desmatada. Uma imagem era considerada válida se tivesse os seguintes critérios: a área de interesse na imagem deveria conter dados não nulos em suas bandas e também não devia conter nuvens.

Depois desse processo, 807 pontos de desmatamento foram considerados válidos. Para facilitar a separação dos conjuntos de dados, dentre esses pontos foram selecionados 800. Em seguida, para cada ponto selecionado como de desmatamento, buscou-se um ponto próximo que possui a característica de conservação, ou seja, que tanto a imagem mais antiga quanto a mais recente possuíam vegetação. Assim foi obtido 800 pontos de conservação.

Portanto, foram analisados no total 1.600 pontos diferentes distribuídos pelo Estado de Mato Grosso. O mapa da Figura 7 apresenta os pontos selecionados, no qual em verde estão os pontos que se mantêm conservados durante o período analisado, e em roxo os pontos que estavam conservados no início e ao final do período estavam desmatados.

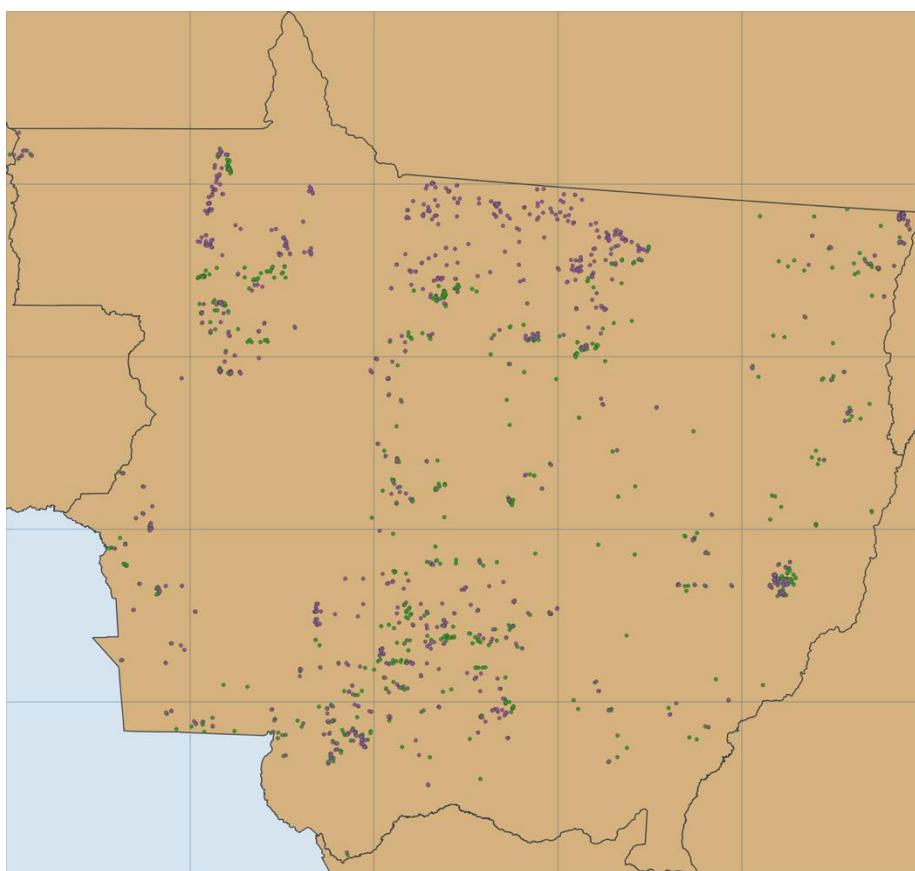


Figura 7: Mapa dos pontos utilizados na base RSDD - em verde os pontos conservados e em roxo os pontos que sofreram desmatamento.

Destes pontos, 1.120 (70%) foram selecionados para treinar o modelo e 480 (30%) para testar o modelo. A separação desses dois conjuntos foi feita de maneira aleatória.

Metade dos pontos representam áreas que do primeiro ao terceiro mês continuaram conservadas, ou seja, não sofreram ações de desmatamento. A outra metade contém pontos que no primeiro mês a área estava conservada, mas no terceiro encontrava-se desmatada.

4.2.2. Resultado dos testes de classificação

Com o cenário de teste estabelecido, o modelo foi calibrado com os pontos do conjunto de treinamento. O treinamento do modelo foi rápido, já que pode ser considerado que há poucos registros na base de dados após todos os processamentos. Em média levou-se 13 segundos para adaptar o modelo aos dados disponíveis.

Após o treino, o modelo tentou prever todos os pontos do conjunto de teste. Vale lembrar que foram testadas diferentes combinações de entradas, considerando dois ou três momentos, e variando o uso de bandas e ou índices pré-calculados. A Tabela 8 mostra os valores de precisão e acurácia para cada teste.

Tabela 8: Resultados de precisão e acurácia com o conjunto de testes para cada combinação de entradas, considerando a quantidade de momentos utilizados, uso dos valores das bandas e uso de índices pré-calculados.

Momentos	Bandas	Índices	Precisão (%)	Acurácia (%)
3	Sim	Sim	93,2	94,8
3	Sim	Não	95,4	93,5
3	Não	Sim	89,0	89,4

2	Sim	Sim	90,6	91,9
2	Sim	Não	93,2	92,1
2	Não	Sim	83,8	85,4

Os melhores resultados foram aqueles que utilizaram 3 momentos como entrada. Isso era esperado, pois o terceiro momento está mais perto do considerado momento de desmatamento e, assim, mais fácil de ser detectado. De qualquer forma, mesmo com apenas 2 momentos, os quais costumam ser mais de 30 dias antes do momento de desmatamento, foi possível obter precisões e acurácias próximos a 90%.

Para uma análise mais detalhada, a Tabela 9 apresenta as classificações positivas e negativas para cada caso. É possível analisar que o uso em conjunto dos valores brutos das bandas e dos índices pré-calculados resultaram em resultados melhores.

Tabela 9: Valores de true negatives (TN), false positives (FP), false negatives (FN) e true positives (TP) para cada combinação de entradas.

Momentos	Bandas	Índices	TN	FP	FN	TP
3	Sim	Sim	235 (48,96%)	16 (6,53%)	9 (3,93%)	220 (45,83%)
3	Sim	Não	241 (50,21%)	10 (4,18%)	21 (9,17%)	208 (43,33%)
3	Não	Sim	226 (47,08%)	25 (9,84%)	26 (11,35%)	203 (42,29%)
2	Sim	Sim	229 (47,71%)	22 (8,76%)	17 (7,42%)	212 (44,17%)
2	Sim	Não	236 (49,17%)	15 (6,15%)	23 (10,04%)	206 (42,92%)
2	Não	Sim	213 (44,38%)	38 (14,23%)	32 (13,97%)	197 (41,04%)

4.2.3. Processamento com todo o Estado de Mato Grosso

Outro teste efetuado para avaliar o modelo foi realizar o processamento com dados de um ano diferente do utilizado durante o treinamento do modelo. Imagens dos meses de agosto, setembro e outubro de 2020 foram baixadas para realizar novas estimativas. O modelo teve como entrada os dados até o início de outubro e tentou estimar as ocorrências de desmatamento do final de outubro. Por serem de um período recente e, assim, não possuírem dados validados de desmatamento pelo órgão competente, a definição de uma área desmatada seguiu a metodologia detalhada na Seção 3.1.3.

Um ponto a cada 600m foi processado para cada imagem existente no Sentinel-2 que possui território mato-grossense. Foram ignorados os pontos que continham dados nulos, presença de nuvens, corpos d'água ou que não eram de uma área conservada. Ao todo, 545.989 pontos foram processados.

Assim, após baixar as imagens entre agosto e outubro de 2020, houve o processo de extração de características e o uso do modelo que obteve maior acurácia nos testes anteriores para realizar a previsão de desmatamento. A Tabela 10 mostra a precisão e acurácia deste teste para cada quadrante do Estado de Mato Grosso.

Tabela 10: Medição de precisão e acurácia para cada quadrante do Estado de Mato Grosso usando os dados de agosto, setembro e início de outubro para prever os desmatamentos no final de outubro.

Quadrante	Precisão (%)	Acurácia (%)
1	88,70%	75,44%
2	90,34%	86,26%
3	-	-
4	-	-
5	-	-
6	37,65%	37,75%
7	99,21%	98,52%
8	91,69%	91,16%
9	69,61%	65,30%

10	87,76%	82,59%
11	45,41%	45,58%
12	93,73%	73,91%
13	93,23%	91,04%
14	79,28%	76,78%
15	34,42%	36,47%
16	59,59%	57,42%
17	60,05%	58,03%
18	94,03%	81,15%
19	78,08%	72,49%
20	41,05%	42,53%
21	67,04%	67,87%
22	91,32%	91,57%
23	98,76%	98,55%
24	80,90%	78,11%
25	-	-

A média do desempenho do modelo entre os quadrantes foi de 75,33% para precisão e de 71,84% para acurácia, tendo como pior desempenho o quadrante 15 e melhor desempenho o quadrante 23. Uma vez que cada quadrante possui um bioma predominante, é possível analisar o desempenho do modelo por bioma, conforme apresentado na Tabela 11. O bioma do Pantanal é o que teve os melhores resultados, seguido pelo bioma da Amazônia e, por fim, do Cerrado.

Bioma	Precisão (%)	Acurácia (%)
Amazônia	74,58%	71,83%
Cerrado	70,37%	64,11%
Pantanal	95,04%	95,06%

5. CONSIDERAÇÕES FINAIS

5.1. PRINCIPAS CONTRIBUIÇÕES

Este trabalho apresentou uma nova base de dados chamada Remote Sensing Deforestation Dataset (RSDD), que contém dados de áreas desmatadas de Mato Grosso obtidas por sensoriamento remoto. Os dados estão em dois formatos: tabulados em arquivos CSV com os valores extraídos das bandas; imagens PNG com resolução 20x20 e aproximadamente 1 KB para cada imagem.

O RSDD visa facilitar as pesquisas que envolvem desmatamento, fornecendo de maneira compacta os dados necessários para modelagem deste tipo de estudo. Assim, os pesquisadores não precisam validar dados de desmate ou executar procedimentos de obtenção e extração em imagens de satélite. Os dados fornecidos são suficientes para processá-los com métodos estatísticos ou de aprendizado de máquina, incluindo *deep learning* com o uso das imagens.

Além da criação da base RSDD, este trabalho também apresentou um método para detecção de desmates ilegais, avaliando a evolução da área analisada em busca de ações de cortes seletivos, indicando a futura ação de desmate de corte raso. O método proposto conseguiu prever desmatamentos ilegais com uma precisão de 93,2% e uma acurácia de 94,8% usando dados do mesmo ano de treinamento e com classificações validadas pelo órgão competente. No teste com ano diferente do treinado e seguindo uma metodologia de classificação de desmate automatizado, o modelo teve uma precisão média de 75,33% e acurácia de 71,84%.

O trabalho possibilita utilizar seus resultados para direcionar ações de fiscalização para áreas com alta probabilidade de sofrerem desmatamentos ilegais. Os códigos desenvolvidos neste trabalho estão disponibilizados como *open source* no repositório <https://github.com/demusis/srdesmatamentos>.

5.2. PERSPECTIVAS FUTURAS

Como perspectivas futuras, é possível realizar melhorias na base de dados RSDD, gerando novas versões. As melhorias podem ser sobre a quantidade de pontos desmatados, incluindo outros anos de dados, e adição de dados de outros Estados, principalmente com biomas diferentes.

Como mencionado, um tipo de atividade que pode ser realizada é a aplicação de *deep learning* na base RSDD, uma vez que esta base disponibiliza também imagens de áreas conservadas e desmatadas. O trabalho de Adarme et al. (2020) mostrou uma performance superior com o uso de *deep learning* quando comparado com outras técnicas para detectar desmatamentos.

Outro estudo a ser considerado é a adição de outros tipos de atributos nos métodos de aprendizado de máquina. Ao invés de utilizar apenas os dados de sensoriamento remoto, pode ser adicionado também dados econômicos, sociais e geográficos, como os usados nos trabalhos de Mayfield et al. (2017), Dlamini (2016) e Nayar (2009). Assim, é esperado um aumento na precisão dos resultados.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Achard, F.; DeFries, R.; Eva, H.; Hansen, M.; Mayaux, P.; Stibig, H-J. Pan-tropical monitoring of deforestation. *Environmental Research Letters*. 2, 1-11, 2007.
- Adarme, M. O.; Feitosa, R. Q.; Happ, P. N.; Almeida, C. A.; Gomes, A. R. Evaluation of deep learning techniques for deforestation detection in the Brazilian Amazon and Cerrado Biomes from remote sensing imagery. *Remote Sensing*, 12, 1-28, 2020.
- Alves, D. S. Space-time dynamics of deforestation in Brazilian Amazônia. *International Journal of Remote Sensing*. 23(14), 2903-2908, 2002.
- Assis, L. F. F. G.; Ferreira, K. R.; Vinhas, L.; Maurano, L.; Almeida, C.; Carvalho, A.; Rodrigues, J.; Maciel, A.; Camargo, C. TerraBrasilis: A Spatial Data Analytics Infrastructure for Large-Scale Thematic Mapping. *ISPRS International Journal of Geo-Information*. 8, 513, 2019.
- Bayma, A. P.; Sano, E. E. Séries temporais de índices de vegetação (NDVI e EVI) do sensor Modis para detecção de desmatamentos no bioma Cerrado. *Bol. Ciênc. Geod.*, 21(4), 797-813, 2015.
- Bouvet, A.; Mermoz, S.; Ballère, M.; Koleck, T.; Toan, T. L. Use of the SAR Shadowing Effect for Deforestation Detection with Sentinel-1 Time Series. *Remote Sens*, 10, 1-19, 2018.
- Candra, D. S. Deforestation detection using multitemporal satellite images. *IOP Conf. Series: Earth and Environmental Science*, 500, 1-13, 2020.
- Carreiras, J. M. B.; Jones, J.; Lucas, R. M.; Shimabukuro, Y. E. Mapping major land cover types and retrieving the age of secondary forests in the Brazilian Amazon by combining single-date optical and radar remote sensing data. *Remote Sensing of Environment*. 194, 16-32, 2017.
- Castelo, T. B.; Adami, M.; Santos, R. B. N. Fronteira Agrícola e a política de priorização dos municípios no combate ao desmatamento no estado do Pará, Amazônia. *Estudos Sociedade e Agricultura*. 28(2), 434-457, 2020.
- Clec'h, S. L.; Oszwald, J.; Dufour, S.; Desjardins, T.; Jegou, N.; Grimaldi, M. Espacialização dos serviços ecossistêmicos na escala local em um contexto

- de desmatamento: que abordagens estatísticas e quais dados? *Revista franco-brasileira de geografia*. 26, 1-15, 2016.
- Cortez, P.; Morais, A. A Data Mining Approach to Predict Forest Fires using Meteorological Data. *Proceedings of the 13th Portuguese Conference on Artificial Intelligence, Portugal*. 512-523, 2007.
- Defferrard, M.; Benzi, K.; Vandergheynst, P.; Bresson, X. FMA: A Dataset For Music Analysis. *Proceedings of the 18th ISMIR Conference*. 316-323, 2017.
- Delegido, J.; Verrelst, J.; Alonso, L.; Moreno, J. Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content. *Sensors*, 11, 7063-7081, 2011.
- Dias, A.; Silva, K. L.; Souza, L. G.; Perdigão, R. J.; Bilio, H.; Oliveira, H. H. C.; Kummer, O. P. Análise dos dados de desmatamento do Estado de Mato Grosso: Período 2016/2017. Relatório técnico nº 003/2018/CGMA/SRMA/SAGA/SEMA-MT, 1-20, 2018.
- Diniz, C. G.; Souza, A. A. A.; Santos, D. C.; Dias, M. C.; Luz, N. C.; Moraes, D. R. V.; Maia, J. S.; Gomes, A. R.; Narvaes, I. S.; Valeriano, D. M.; Maurano, L. E. P.; Adami, M. DETER-B: The New Amazon Near Real-Time Deforestation Detection System. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 8(7), 3619-3628, 2015.
- Dlamini, W.M. Analysis of deforestation patterns and drivers in Swaziland using efficient Bayesian multivariate classifiers. *Modeling Earth Systems and Environment*, 2(173), 1–14, 2016.
- Dua, D.; Graff, C. UCI Machine Learning Repository. Irvine, CA: University of California, School of Information and Computer Science <<http://archive.ics.uci.edu/ml>> Acessado em junho de 2019.
- Dupuis, C.; Lejeune, P.; Michez, A.; Fayolle, A. How Can Remote Sensing Help Monitor Tropical Moist Forest Degradation? — A Systematic Review. *Remote Sens*. 12, 1-24, 2020.
- Ertel, W. *Introduction to Artificial Intelligence*. Springer, 2017.
- ESA. Sentinel-2 Mission, 2020. Disponível em: <<https://sentinel.esa.int/web/sentinel/missions/sentinel-2>>. Acesso em: 15/07/2020.

- Fearnside, P. M. The roles and movements of actors in the deforestation of Brazilian Amazonia. *Ecology and Society*. 13(1): 23, 2008.
- Fonsêca, N. C.; Cunha, J. S. A.; Cunha, J. A. S.; Rodrigues, L. S.; Belém, I. S. M. Expansão agropecuária: o efeito sobre a vegetação e o microclima. *Revista Verde de Agroecologia e Desenvolvimento Sustentável*. 14(4), 2019.
- Gallo, K.; Xian, G. Application of spatially gridded temperature and land cover data sets for urban heat island analysis. *Urban Climate*, 8, 1-10, 2014.
- Gallwey J.; Robiati C.; Coggan J.; Vogt D.; Eyre M. A Sentinel-2 based multispectral convolutional neural network for detecting artisanal small-scale mining in Ghana: Applying deep learning to shallow mining. *Remote Sensing of Environment*, 248, 1-15, 2020.
- Gao, B-C. NDWI — a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58, 257–266, 1996.
- Gemmeke, J. F.; Ellis, D. P. W.; Freedman, D.; Jansen, A.; Lawrence, W.; Moore, R. C.; Plakal, M.; Ritter, M. Audio Set: An ontology and human-labeled dataset for audio events. *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 776-780, 2017.
- Hansen, M. C.; Potapov, P. V.; Moore, R.; Hancher, M.; Turubanova, S. A.; Tyukavina, A.; Thau, D.; Stehman, S. V.; Goetz, S. J.; Loveland, T. R.; Kommareddy, A.; Egorov, A.; Chini, L.; Justice, C. O.; Townshend, J. R. G. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*. 342(6160), 850–853, 2013.
- Haykin, S. *Redes Neurais – Princípios e Prática*. Porto Alegre, RS: Bookman, 2003.
- Herold, M.; Menz, G.; Clarke K. C. Remote Sensing and urban growth models. demands and perspectives. *Proceedings of the Symposium on Remote Sensing of Urban Areas*. Regensburg, Germany, 22-23, 2001.
- Herold, M.; Johns, T. Linking requirements with capabilities for deforestation monitoring in the context of the UNFCCC-REDD process. *Environmental Research Letters*, 2(4), 1-7, 2007.
- Huete, A. R. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295-309, 1988.

- INPE. Monitoramento da cobertura florestal da Amazônia por satélites - Sistemas Prodes, Deter, Degrad e Queimadas 2007-2008. Relatório técnico, São José dos Campos, 2008.
- Jong, R.; Bruin, S.; Wit, A.; Schaepman, M.; Dent, D. L. Analysis of monotonic greening and browning trends from global NDVI time-series. *Remote Sensing of Environment*, 115, 692–702, 2011.
- Key, C. H.; Benson, N.; Ohlen, D.; Howard, S.; McKinley, R.; Zhu, Z. The normalized burn ratio and relationships to burn severity: ecology, remote sensing and implementation. *Proceedings of the 9th Forest Service Remote Sensing Applications Conference*, American Society for Photogrammetry and Remote Sensing, Bethesda, MD. 2002.
- Kovács, Z. L. *Redes Neurais Artificiais: Fundamentos e Aplicações*. Collegium Cognition, 1996.
- LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature*, 521, 436-444, 2015.
- Lambin, E. F. *Modelling Deforestation Processes, A Review*. EUR 15744 EN, TREES Series B: Research Report No. 1. Joint Research Centre. Institute for Remote Sensing Applications, European Space Agency, Luxembourg, Office for Official Publications of the European Community, 1994.
- Li, Z.; Fox, J. M. Mapping rubber tree growth in mainland Southeast Asia using time-series MODIS 250 m NDVI and statistical data. *Applied Geography*, 32, 420–432, 2012.
- Lima, D. L.; Sobrinho, T. A.; Oliveira, A. P. G.; Catalani, T. G.; Dalmas, F. B.; Paranhos Filho, A. C. Identificação e quantificação semiautomática de desmatamento por Sensoriamento Remoto. *Research, Society and Development*. 9(4), 2020.
- Lima, T. P.; França, L. C. J.; Ferraz, F. T.; Aguiar Júnior, A. L.; Silva, D. P.; Acerbi Junior, F. W. Dinâmica Espaço Temporal da Cobertura da Terra em uma Bacia Hidrográfica da Região do MATOPIBA, Brasil. *Anuário do Instituto de Geociências*. 43(1), 162-170, 2020.
- Maki, M.; Ishihara, M.; Tamura, M. Estimation of leaf water status to monitor the risk of forest fires by using remotely sensed data. *Remote Sensing of Environment*, 90, 441–450, 2004.

- Maurano, L. E. P.; Escada, M. I. S.; Renno, C. D. Padrões espaciais de desmatamento e a estimativa da exatidão dos mapas do PRODES para Amazônia Legal Brasileira. *Ci. Fl.* 29(4), 1763-1775, 2019.
- Mayfield, H.; Smith, C.; Gallagher, M.; Hockings, M. Use of freely available datasets and machine learning methods in predicting deforestation. *Environmental Modelling & Software*, 87, 17–28, 2017.
- Mba, E. H. Assessment of environmental impact of deforestation in Enugu, Nigeria. *Resources and Environment*, 8(4), 207–215, 2018.
- Meneses, P. R.; Almeida, T.; Rosa, A. N. C. S.; Sano, E. E.; Souza, E. B.; Baptista, G. M.; Brites, R. S. *Introdução ao processamento de imagens de sensoriamento remoto*. 1ª edição, Brasília: UnB, 2012.
- Mhatre A.; Mudaliar N.; Narayanan M.; Gurav A.; Nair A.; Nair A. Using Deep Learning on Satellite Images to Identify Deforestation/Afforestation. In: Smys S., Tavares J., Balas V., Iliyasa A. (eds) *Computational Vision and Bio-Inspired Computing. ICCVBIC 2019. Advances in Intelligent Systems and Computing*, vol 1108. Springer, 2020.
- MPF instaura mais de mil ações por desmatamento ilegal na 3ª fase do projeto Amazônia Protege. MPF, 2020. Disponível em: <<http://www.mpf.mp.br/pgr/noticias-pgr/mpf-instaura-mais-de-mil-acoes-por-desmatamento-ilegal-na-3a-fase-do-projeto-amazonia-protege>>. Acesso em: 10 de outubro. de 2020.
- Nayar, A. Model predicts future deforestation. *Nature*, 2009.
- Noi, P. T.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors*, 18, 1-20, 2017.
- Novo, E. M. L. M.; Ferreira, L. G.; Barbosa, C.; Carvalho, C.; Sano, E. E.; Shimabukuro, Y.; Huete, A.; Potter, C.; Roberts, D. A.; Hess, L. L.; Melack, J. J.; Yoshioka, H.; Klooster, S.; Kumar, V.; Myneni, R.; Ratana, P.; Didan, K.; Miura, T. Técnicas avançadas de sensoriamento remoto aplicadas ao estudo de mudanças climáticas e ao funcionamento dos ecossistemas amazônicos. *Acta Amazonica*. 35(2), 259–272, 2005.

- Oliveira, R. R. S.; Venturieri, A. Mapeamento Automatizado de Tipologias de Paisagem: Mineração de Dados Utilizando o Plugin GeoDMA. *Revista Brasileira de Cartografia*. 70(1), 258-289, 2018.
- Ortega, M.; Castro, J.; Nigri Happ, P.; Gomes, A.; Feitosa, R. Evaluation of deep learning techniques for deforestation detection in the Amazon forest. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*. 121-128, 2019.
- Persson M.; Lindberg E.; Reese H. Tree Species Classification with Multi-Temporal Sentinel-2 Data. *Remote Sensing*, 10 (11), 1-17, 2018.
- Roteta E.; Bastarrika A.; Padilla M.; Storm T.; Chuvieco E. Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa. *Remote Sensing of Environment*, 222, 1–17. 2019.
- Rouse, J. W.; Hass, R. H.; Schell, J. A.; Deering, D. W. Monitoring vegetation systems in the Great Plains with ERTS. In: *Earth Resources Technology Satellite Symposium*, Washington, 309-317, 1973.
- Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; Berg, A. C.; Fei-Fei, L. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*. 115, 211–252, 2015.
- Russell, S.; Norvig, P. *Inteligência Artificial*. GEN LTC, 2013.
- Saha, S.; Saha, M.; Mukherjee, K.; Arabameri, A.; Ngo, P. T. T.; Paul, G. C. Predicting the deforestation probability using the binary logistic regression, random forest, ensemble rotational forest, REPTree: A case study at the Gumani River Basin, India. *Science of the Total Environment*, 730, 1-20, 2020.
- Santos, T. O.; Andrade Filho, V. S.; Rocha, V. M.; Menezes, J. S. Os impactos do desmatamento de queimadas de origem antrópica sobre o clima da Amazônia brasileira: um estudo de revisão. *Rev. Geogr. Acadêmica*. 11(2), 157-181, 2017.
- Silva, C. A.; Santilli, G.; Sano, E. E.; Rodrigues, S. W. P. Análise Qualitativa do Desmatamento na Floresta Amazônica a partir de Sensores SAR, Óptico e Termal. *Anuário do Instituto de Geociências*. 42(4), 18-29, 2019.

- Soares, J.; Danelichen, V. H. M.; Pereira, O. A.; Martins, A. L. Estudo da dinâmica espaço-temporal do NDVI no Município de Sorriso-MT. *Revista Brasileira de Geografia Física*, 13(2), 834-841, 2020.
- Traganos D.; Reinartz P. Mapping Mediterranean seagrasses with Sentinel-2 imagery. *Marine Pollution Bulletin*, 134, 197-209, 2018.
- Valeriano, D. M.; Mello, E. M. K.; Moreira, J. C.; Shimabukuro, Y. E.; Duarte, V.; Souza, I. M.; dos Santos, J. R.; Barbosa, C. C. F.; de Souza, R. C. M. Monitoring tropical forest from space: The PRODES digital project. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* 35, 272–274, 2004.
- Vannini, I. A.; Kummer, R. Sudoeste Paranaense: Desmatamento como Estratégia de Posse da Terra (1940-1960). *Historia Ambiental Latinoamericana y Caribeña*. 8(1), 92-113, 2018.
- Vargas, C.; Itoh, T.; Tsuji, S.; Koide, T.; Hirose, K.; Okonogi, H. Automatic Deforestation Detection Methodology Using Sentinel-1. In *IEEE International Geoscience and Remote Sensing Symposium*, 2019.
- Voight, C.; Hernandez-Aguilar, K.; Garcia, C.; Gutierrez, S. Predictive modeling of future forest cover change patterns in southern Belize. *Remote Sens.* 11(823), 1-17, 2019.
- Vuolo F.; Neuwirth M.; Immitzer M.; Atzberger C.; Ng W.-T. How much does multi-temporal Sentinel-2 data improve crop type classification? *International Journal of Applied Earth Observation and Geoinformation*, 72, 122–130, 2018.
- Watanabe, M.; Oliveira, F. A.; Nunes, D. D.; Nunes, A. C. S.; Cavalcante, M. M. A.; Aguiar, T. C. Análise do impacto do desmatamento no aporte sedimentar de bacias pareadas na Amazônia ocidental: bacias do rio Mutum-Paraná, Rondônia (Brasil). *Rev. Bras. Geomorf.* 19(3), 601-615, 2018.
- Wolff, N. H.; Masuda, Y. J.; Meijaard, E.; Wells, J. A.; Game, E. T. Impacts of tropical deforestation on local temperature and human well-being perceptions. *Glob. Environ. Chang.* 52, 181–189, 2018.
- Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing.* 27(14), 3025-3033, 2006.

- Xue, J.; Su, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors*, 1-17, 2017.
- Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, 24, 583–594, 2003.
- Zhang, Y.; Odeh, I. O. A.; Han, C. Bi-temporal characterization of land surface temperature in relation to impervious surface area, NDVI and NDBI, using a subpixel image analysis. *International Journal of Applied Earth Observation and Geoinformation*, 11, 256–264, 2009.
- Zheng Q.; Huang W.; Cui X.; Shi Y.; Liu L. New Spectral Index for Detecting Wheat Yellow Rust Using Sentinel-2 Multispectral Imagery. *Sensors*, 18(3), 1-19, 2018.
- Zhao, H. M.; Chen, X. L. Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+. *Geoscience and Remote Sensing Symposium*, 3, 1666–1668, 2005.
- Zhu Z.; Wang S.; Woodcock C. E. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159, 269-277, 2015.