

**UNIVERSIDADE FEDERAL DE UBERLÂNDIA - *CAMPUS* MONTE
CARMELO
INSTITUTO DE GEOGRAFIA
CURSO DE ENGENHARIA DE AGRIMENSURA E CARTOGRÁFICA**

RAFAEL CORDEIRO PEREIRA MENDONÇA

**DETECÇÃO DE PRAGAS E PATÓGENOS DO CAFEIEIRO ATRAVÉS DE
IMAGENS MULTIESPECTRAIS DE BAIXO CUSTO E ALGORITMOS
BASEADOS EM APRENDIZADO DE MÁQUINA**

MONTE CARMELO

2022-1

RAFAEL CORDEIRO PEREIRA MENDONÇA

**DETECÇÃO DE PRAGAS E PATÓGENOS DO CAFEIEIRO ATRAVÉS DE
IMAGENS MULTIESPECTRAIS DE BAIXO CUSTO E ALGORITMOS
BASEADOS EM APRENDIZADO DE MÁQUINA**

Trabalho de Conclusão de Curso apresentado na Universidade Federal de Uberlândia *campus* Monte Carmelo como requisito para aprovação na disciplina de Trabalho de Conclusão de Curso 2 no curso de Engenharia de Agrimensura e Cartográfica.

Orientador: Prof. Dr. George Deroco Martins.

Coorientador: M.e Vinicius Silva Werneck Orlando

MONTE CARMELO

2022-1

RESUMO

A agricultura é uma das principais atividades econômicas nacionais, devido a sua grande participação no Produto Interno Bruto (PIB). Proporcionando ao Brasil o segundo lugar como maior consumidor de café e o primeiro lugar como maior produtor e exportador a nível mundial do produto cultivado. Contudo, diversas variáveis podem prejudicar o potencial produtivo do cafeeiro nas lavouras, principalmente doenças, como a Cercosporiose e o inseto praga conhecido como Bicho-Mineiro. Medidas de controle por meio de pesticidas agrícolas onerosos e ambientalmente invasivos, diagnóstico *in loco* e a investigação contínua da ocorrência de plantas afetadas por esses patógenos, são ações caras e nem sempre tecnicamente viáveis. Assim, tendo como base a Agricultura de Precisão (AP), este estudo teve como objetivo identificar um conjunto de atributos que facilitem a discriminação desses patógenos nos cafeeiros, a partir de análises de imagens multiespectrais de baixo custo, além de informações obtidas *in situ*, e através da utilização de algoritmos baseados em aprendizado de máquina. Por fim, resultou que os algoritmos *Support Vector Machine* (SVM) e Redes Neurais Artificiais (RNA), são capazes de discriminar as plantas saudáveis das infectadas, e que SVM apresentou melhor desempenho que as RNA. Ademais identificou-se um conjunto de atributos que contribuiu para a discriminação das plantas.

Palavras chaves: Cafeicultura, Doenças, Imagens Multiespectrais, Monitoramento Remoto de Pragas e Patógenos.

LISTA DE FIGURAS

- Figura 1** Folha de cafeeiro apresentando lesões causadas pela Cercosporiose
- Figura 2** Fruto do cafeeiro apresentando lesões causadas pela Cercosporiose
- Figura 3** Folha de cafeeiro apresentando lesões causadas pelo bicho-mineiro
- Figura 4** Diferentes níveis de coleta de dados do sensoriamento remoto
- Figura 5** (a) Imagem digital
(b) Imagem digital com área em zoom de grupo de pixels em valores de cinza
(c) Imagem digital com área em zoom de grupo de pixels em valores digitais.
- Figura 6** Modelo artificial de neurônio biológico
- Figura 7** (A) Minas Gerais
(B) Mancha Urbana de Monte Carmelo e Universidade Federal de Uberlândia
(C) Área de Estudo
- Figura 8** (A) ARP *Phantom 4 Advanced*
(B) Câmara Mapir acoplada a ARP *Phantom 4 Advanced*
- Figura 9** Receptor GNSS (*Rover*) utilizados no georreferenciamento das parcelas experimentais.
- Figura 10** Fluxograma da pesquisa
- Figura 11** (A) Ortomosaico da câmera RGB
(B) Ortomosaico da câmera Mapir
- Figura 12** Comportamento espectral das bandas RGB
- Figura 13** Comportamento espectral das bandas RGN

LISTA DE QUADROS

- Quadro 1** Índices Vegetativos
- Quadro 2** Matriz de Confusão e resultados da classificação por meio do *Support Vector Machine* das bandas RGB e seus índices
- Quadro 3** Matriz de Confusão e resultados da classificação por meio do *Support Vector Machine* das bandas RBN e seus índices
- Quadro 4** Matriz de Confusão e resultados da classificação por meio das Redes Neurais Artificiais das bandas RGB e seus índices
- Quadro 5** Matriz de Confusão e resultados da classificação por meio das Redes Neurais Artificiais das bandas RGN e seus índices

LISTA DE SIGLAS E ABREVIATURAS

AP	Agricultura de Precisão
ARP	Aeronave Remotamente Pilotada
BMC	Bicho-Mineiro do Cafeeiro
GLI	<i>Green Leaf Index</i>
GNDVI	<i>Green Normalized Difference Vegetation Index</i>
GNSS	<i>Global Navigation Satellite System</i>
GPS	<i>Global Positioning System</i>
GSD	<i>Ground Sample Distance</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
LAMFO	Laboratório de Aprendizado de Máquina em Finanças e Organizações
MCC	<i>Mapir Camera Control</i>
MG	Minas Gerais
NDVI	<i>Normalized Difference Vegetation Index</i>
NGRDI	<i>Normalized Green Red Difference Index</i>
PIB	Produto Interno Bruto
REM	Radiação Eletromagnética
RGB	<i>Red, Green, Blue</i>
RGN	<i>Red, Green, Near-infrared</i>
RMSE	<i>Root Mean Square Error</i>
RNA	Redes Neurais Artificiais
ROI'S	Regiões de Interesse
SR	<i>Simple Ratio</i>
SVM	<i>Support Vector Machine</i>
TGI	<i>Triangular Greenness Index</i>
VARI	<i>Visible Atmospherically Resistant Index</i>

SUMÁRIO

1 INTRODUÇÃO.....	09
2 OBJETIVOS.....	11
3 JUSTIFICATIVA.....	12
4 FUNDAMENTAÇÃO TEÓRICA.....	14
4.1 DOENÇAS.....	14
4.1.1 Cercosporiose.....	14
4.1.2 Bicho-Mineiro do Cafeeiro.....	15
4.2 SENSORIAMENTO REMOTO APLICADO NA DETECÇÃO DE DOENÇAS.....	16
4.3 PROCESSAMENTO DIGITAL DE IMAGENS.....	18
4.4 MÉTRICA DE QUALIDADE.....	20
4.4.1 Classificação de Imagens.....	20
4.4.1.1 Classificação Não-Supervisionada.....	20
4.4.1.2 Classificação Supervisionada.....	20
4.4.2 Classificação de Variáveis.....	20
4.4.2.1 Variáveis Quantitativas.....	21
4.4.2.2 Variáveis Qualitativas.....	21
4.4.3 Algoritmos.....	21
4.4.3.1 Redes Neurais Artificiais (RNA).....	21
4.4.3.2 Support Vector Machine (SVM).....	23
5 MATERIAL E MÉTODOS.....	24
5.1 ÁREA DE ESTUDO.....	24
5.2 MATERIAL.....	25

5.3 MÉTODOS.....	26
5.3.1 Levantamento Geodésico.....	27
5.3.2 Avaliação Epidemiológica.....	28
5.3.3 Planejamento e Execução do Aerolevanteamento.....	28
5.3.4 Elaboração e Georreferenciamento dos Ortomosaicos.....	29
5.3.5 Calibração Radiométrica do Ortomosaico da Mapir.....	29
5.3.6 Extração dos Valores Digitais de cada planta.....	30
5.3.7 Cálculo dos Índices de Vegetação.....	30
5.3.8 Classificação dos dados e Validação dos Resultados.....	31
6 RESULTADOS E DISCUSSÕES.....	33
7 CONCLUSÃO.....	38
REFERÊNCIAS.....	39

1 INTRODUÇÃO

No início da história brasileira, o Pau-brasil, a Cana-de-açúcar, o Fumo e o Ouro deram ao Brasil sua característica econômica, depois iniciou-se a marcha triunfal do café, que foi trazido em 1727 para o Pará por Vigilenzo Francisco de Mello Palheta, e de lá espalhou-se para todo território brasileiro, habituando-se melhor no Estado de São Paulo (MURTA, 1978). Atualmente, os cafés brasileiros são produzidos nas cinco regiões geográficas do Brasil, com a receita bruta estimada, no ano de 2021, em 33,51 bilhões de reais, montante que teve como base de cálculo e referência os preços médios recebidos pelos produtores de janeiro a junho deste ano (CONSÓRCIO PESQUISA CAFÉ, 2021).

Dentre as principais doenças que afetam o cafeeiro, destacam-se a Cercosporiose (*Cercospora coffeicola* Berk & Cooke) e o Bicho-Mineiro do Cafeeiro (BMC) (*Leucoptera coffeella*) (*Lepidoptera: Lyonetiidae*), cujo controle dessas doenças eleva os custos de produção do café. De acordo com Kimati et al. (2005), a Cercosporiose do cafeeiro, também conhecida como mancha de olho-pardo, é uma das doenças mais antigas das Américas, tendo sido exposta no Brasil em 1887. Já o Bicho-Mineiro do cafeeiro é uma marcante praga da cultura, em virtude da sua ocorrência nos cafezais e devido os prejuízos quantitativos e econômicos provocados por esse inseto na produção de café (SOUZA et al., 1998).

O principal método de controle das doenças é por meio de aplicações de defensivos químicos, causando impactos ambientais, elevando os custos de produção e de reduzindo a sustentabilidade do agroecossistema cafeeiro (GARCIA JÚNIOR, 2002). Além destes danos, essas aplicações também violam a meta dos agricultores, que buscam obter alta produtividade com o menor custo possível, sempre com o objetivo de reter o melhor retorno econômico da atividade agrícola (MOLIN et al., 2015).

Diante deste problema, nota-se a necessidade da introdução de tecnologias modernas, que visam aumentar a produtividade das lavouras e fornecer informações para tomar melhores e rápidas decisões de manejo, bem como reduzir o custo de defensivos químicos e fertilizantes, visando aumentar as margens de lucro e limitar a poluição que agricultura pode causar (FORMAGGIO; SANCHES, 2017).

Dentre as várias tecnologias, o Sensoriamento Remoto, técnica que utiliza sensores orbitais, aéreos e terrestre, tem contribuído com grandes avanços e melhorias na agricultura, desde a fase inicial da germinação até a colheita, proporcionando maior

controle de produção e prevenção dos danos causados por pragas (RODRIGUES; BARROS, 2019).

Estes avanços e melhorias ocorrem devido as imagens demonstrarem grande potencial no reconhecimento da variabilidade espacial do desenvolvimento das plantas, tornando-se possível estudar os responsáveis dessa variedade. Assim como, após essa variação for outorgada a certo fator, como ataque de pragas ou problemas de fertilidade, sugerir intervenções (MOLIN et al., 2015).

Assim sendo, a partir de análises de imagens multiespectrais e da utilização de algoritmos baseados em aprendizado de máquina, além de informações obtidas *in situ*, amparadas com a caracterização ambiente favorável ao desenvolvimento do patógeno, espera-se identificar um conjunto de atributos que facilitem a discriminação dos patógenos nos cafeeiros, sendo esses atributos submetidos a um processo de classificação supervisionada. Dentre os algoritmos de classificação serão utilizados aqueles baseados em aprendizado de máquina, tais como as Redes Neurais Artificiais (RNA) e Máquinas de Vetores Suporte (SVM, do inglês *Support Vector Machine*).

Os algoritmos de classificação supervisionada baseados em aprendizado de máquina, tais como as Máquinas de Vetores Suporte (SVM), as Redes Neurais Artificiais (RNA) têm sido muito aplicadas na agricultura (WEISS; JACOB; DUVEILLER, 2020), pois além de serem capazes de indicar o grau de importância das variáveis ou atributos, também caracterizam as classes de informações a serem mapeadas.

Portanto, sabendo que os algoritmos são aptos em diferenciar o nível de importância das variáveis e de caracterizar os tipos de informações a serem representadas, e que o sensoriamento remoto é capaz de identificar a variabilidade espacial do desenvolvimento das plantas, será possível indicar um conjunto de atributos que seja capaz de distinguir os patógenos nos cafeeiros?

2 OBJETIVOS

O objetivo geral deste trabalho é identificar um conjunto de atributos que facilitem a discriminação da Cercosporiose e do Bicho-Mineiro nos cafeeiros, a partir de imagens multiespectrais obtidas durante um aerolevante executado com uma Aeronave Remotamente Pilotada (ARP). Bem como, mediante a utilização de algoritmos baseados em aprendizado de máquina, além de informações obtidas *in situ* através de avaliações epidemiológicas, as quais estão amparadas com a caracterização ambiente favorável ao desenvolvimento do patógeno. Para tal, são apresentados os seguintes objetivos específicos:

- Investigar os atributos extraídos de imagens multiespectrais, na discriminação do cafeeiro em diferentes condições de sanidade.
- Avaliar o desempenho dos algoritmos baseados em aprendizado de máquina na discriminação das infecções, utilizando como base os dados obtidos das imagens multiespectrais e das avaliações epidemiológica.

3 JUSTIFICATIVA

A origem do termo “Agricultura de Precisão” está respaldada no fato de que as lavouras não são iguais no espaço nem no tempo, assim, foi preciso o desenvolvimento de estratégias para gerir os problemas decorrentes da desuniformidade das lavouras ou de outros problemas, com variados níveis de complicação (MOLIN et al., 2015).

Dentre as várias culturas produzidas no país, a cafeicultura foi e é essencial para o desenvolvimento e modernização do Brasil (NAGAY, 1999), no entanto, a cafeicultura pode ter sua produtividade diminuída por vários fatores, tais como condições climáticas adversas, carência nutricional e doenças (BOLDINI, 2001).

Com relação às principais doenças do cafeeiro responsáveis por prejuízos quantitativos, com redução no rendimento na produção e na qualidade do produto final, destacam-se a Cercosporiose do cafeeiro que está presente em quase todas as regiões que possuem condições favoráveis, a qual apresenta como sintomas manchas nas folhas com centro claro e desfolhamento (OLIVEIRA et al., 2001). Além desta, o bicho-mineiro é uma das principais pragas da cultura do café, em virtude atingir alto nível de controle todo ano nas regiões produtoras. O BMC ocorre principalmente no período seco do ano, ocasionando a redução do potencial fotossintético pela morte e queda das folhas (ASSIS et al., 2009).

Com a alta dos custos de produção, além do crescimento na competição do mercado interno e externo, torna-se essencial eliminar quaisquer fontes de perdas nas lavouras cafeeiras (GARCIA JÚNIOR, 2002), e isso pode ser possível com as aplicações de procedimentos produzidos pela agricultura de precisão.

Na atualidade, a área com maior potencial de crescimento na AP é a de sensores, tanto quanto para equipamentos como em aplicações, pois é possível o reconhecimento e o mapeamento de vários parâmetros do solo e da planta. Além do que, os dados adquiridos, os quais podem ser obtidos por diferentes sistemas sensores e em diferentes períodos, podem ser informações valiosas no controle das unidades de gestão (MOLIN et al., 2015). Trabalhos que estudam a relação entre patógeno, hospedeiro e ambiente durante o ciclo da cultura, são fundamentais para serem tomadas medidas adequadas de manejo, de maneira a evitar grandes prejuízos na produção, ou gastos desnecessários para o controle (BOLDINI, 2001).

Portanto, considerando os objetivos da pesquisa e que estudos sobre a relação entre patógeno, hospedeiro e ambiente durante o ciclo da cultura são relevantes, conforme Boldini (2001), esse trabalho justifica-se pela possibilidade de caracterizar espectralmente a ocorrência de duas doenças no cafeeiro, de maneira remota, não destrutiva, e conseqüentemente mais sustentável. Atualmente pesquisas sobre este assunto tem apresentado uma boa resposta, pois apontam maneiras de utilizar os defensivos agrícolas de forma mais racional, aumentando assim a sustentabilidade do campo.

4 FUNDAMENTAÇÃO TEÓRICA

4.1 DOENÇAS

Fungos e infecções são notáveis por arruinar partes das lavouras e trazerem consigo grandes prejuízos aos agricultores, plantas geneticamente modificadas são criadas, no entanto resistem por pouco tempo a esses “predadores” (GALVÃO et al., 2018). Entre algumas doenças, destacam-se a Cercosporiose do Cafeeiro e Bicho-Mineiro do cafeeiro.

4.1.1 Cercosporiose

A Cercosporiose, também conhecida como mancha de olho-pardo, é causada pelo fungo *Cercospora coffeicola*, da família *Dematiaceae*, ordem *Moniliales*, classe *Deuteromycetes*, após entrar em contato com o cafeeiro, os conídios germinam e o tubo germinativo penetra diretamente por meio da cutícula ou aberturas naturais (KIMATI et al., 2005).

Vale ressaltar que, plantas com escassez nutricionais, principalmente de nitrogênio, são mais atacadas pelo patógeno, e quando ocorre infecção de frutos, o patógeno pode alcançar as sementes e ser transmitido através das mesmas. Os sintomas nas folhas apresentam-se através de manchas de formato circular, com coloração pardo-clara ou marrom-escuro, com centro branco-acinzentado, contornada por um anel roxeado, dando a impressão de um olho (KIMATI et al., 2005), como ilustrado na Figura 1.

Figura 1 – Folha de cafeeiro apresentando lesões causadas pela Cercosporiose



Fonte: O Autor (2022).

Nos frutos, os sintomas se apresentam na forma de pequenas manchas castanhas, que se estende para as extremidades. As lesões surgem nos frutos ainda pequenos e ficam até o amadurecimento, à proporção que as manchas envelhecem, elas sofrem ressecamento e escurecem, fazendo com que a casca, fique grudada à semente, como ilustrado na Figura 2 (KIMATI et al., 2005).

Figura 2 – Fruto do cafeeiro apresentando lesões causadas pela Cercosporiose



Fonte: Rehagro (2020).

O manejo da Cercosporiose inicia por meio de cuidados na formação das mudas, buscando-se evitar condições propícias a doenças e por meio de aplicações preventivas de fungicidas quinzenalmente. A época de maior incidência de Cercosporiose no campo é de janeiro a abril, necessitando começar o controle em dezembro, no começo da granação dos frutos (KIMATI et al., 2005).

4.1.2 Bicho-Mineiro do Cafeeiro

O BMC tem como origem o continente Africano, e recebeu este nome pelo fato de a lagarta minar as folhas do cafeeiro. É uma praga exótica e monófaga, por isso só ataca o cafeeiro, surgiu no Brasil em 1851, provavelmente introduzido por meio de mudas de café trazidas das Antilhas e das ilhas Bourbon (SOUZA et al., 1998). Em níveis populacionais elevados, pode causar a queda de até 70% das folhas, provocando a redução da fotossíntese e queda na produção de até 50% (SCALON; MATEUS; ZACARIAS, 2013 *apud* SOUZA; REIS; RIGINATO, 1998).

Os insetos sofrem metamorfose completa, passando pelos estágios de ovo, lagarta, crisálida e adulta, mas só são prejudiciais aos cafeeiros na fase de lagarta, na idade adulta causa danos colaterais, pois é nesse período que as fêmeas são fecundadas e serão

produzidos seus ovos. O estágio da lagarta pode durar de 9 a 40 dias, dependendo das condições climáticas, e à proporção que a lagarta se alimenta, a doença se desenvolve. O ciclo evolutivo completo pode durar de 19 a 87 dias, dependendo das condições climáticas, especialmente temperatura, umidade relativa e precipitação (SOUZA et al., 1998).

Devido ao ataque dos bicho-mineiro, formam-se minas ou lesões nas folhas, que se transformam em áreas mortas de necrose, como ilustrado na Figura 3, por consequência, numa menor área fotossintética. Além disso, durante o período de seca do ano, dependendo da intensidade da infestação da praga, as folhas minadas caem, acontecendo desfolhas drásticas dos cafeeiros na sua época de floradas, resultando na diminuição da produção de café durante a safra (SOUZA et al., 1998).

Figura 3 – Folha de cafeeiro apresentando lesões causadas pelo bicho-mineiro



Fonte: Souza et al (1998).

4.2 SENSORIAMENTO REMOTO APLICADO NA DETECÇÃO DE DOENÇAS

De acordo com Jensen (2009), Sensoriamento Remoto é o registro de informação das bandas do ultravioleta, visível, infravermelho e micro-ondas do espectro eletromagnético, sem contato, através dispositivos como câmeras, escâneres, lasers, encontrados em plataformas como aeronaves ou satélites, e a análise da informação adquirida por meio visual ou processamento digital de imagens.

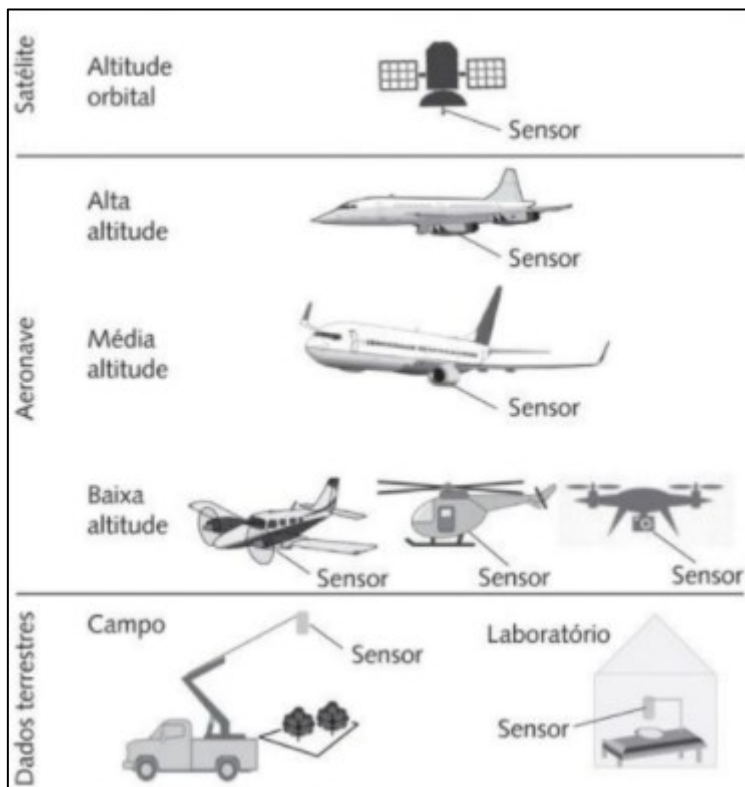
O sensoriamento remoto é praticado usando dispositivos que são nomeados de sensores, os quais na maior parte das vezes registram a REM (Radiação Eletromagnética) por reflexão ou radiância (JENSEN, 2009). As ondas eletromagnéticas, podem ser vistas como “termômetros-mensageiros” do Sensoriamento Remoto, pois elas não apenas

recebem as informações das principais características das feições terrestres, como também as transportam até os satélites (FIGUEIREDO, 2005). A principal fonte de radiação eletromagnética, da superfície da terrestre, é o Sol (NOVO, 2010).

Sensores que medem como a radiação interage com os alvos, especialmente a interação com a energia refletida, funcionam em poucas bandas espectrais, que se caracterizam por diferentes intervalos de faixas do espectro eletromagnético, por exemplo, as bandas nas regiões do verde (490 nm a 565 nm) ou do infravermelho próximo (760 nm a 1500 nm), levando o nome de sensores multiespectrais (MOLIN et al., 2015).

Em relação à plataforma que operam os sensores, existem basicamente três níveis de coleta de dados: orbital, aéreo e o terrestre. A escolha da plataforma depende do objetivo do usuário, pois a partir dos seus objetivos, ele deve estabelecer as especificações mínimas requeridas, referente as características dos sensores, da disponibilidade dos dados e do custo (MOLIN et al., 2015). A seguir, a Figura 4 ilustra os diferentes níveis de coleta de dados do sensoriamento remoto.

Figura 4 – Diferentes níveis de coleta de dados do sensoriamento remoto



Fonte: Formaggio e Sanches (2017).

O Sensoriamento Remoto é uma ferramenta que pode identificar de forma rápida e eficiente plantas doentes em grandes e pequenas áreas, além disso, também pode detectar plantas infectadas sem coletar e processar de amostras em laboratório, alcançando investigações precisas e confiáveis em um curto espaço de tempo, independentemente da área em questão (NAUE et al., 2011). Como exemplo, a pesquisa de Pithan et al. (2019) que aplicou o Sensoriamento Remoto Hiperespectral na discriminação de doenças fúngicas da videira, com intuito de identificar os comprimentos de onda que discriminam os tipos de doenças fúngicas em plantas de videiras.

O estudo de Lemes (2017), demonstra também a possibilidade de utilizar o Sensoriamento Remoto na detecção de seringais infestados por *Meloidogyne exigua*. Para tal estudo, utilizou-se imagens da constelação de satélites *RapidEye* (resolução de 5 metros), além do uso dos índices espectrais de vegetação - Índice de Vegetação por Diferença Normalizada (NDVI) e Índice de Razão Simples (SR) - gerados a partir de imagens.

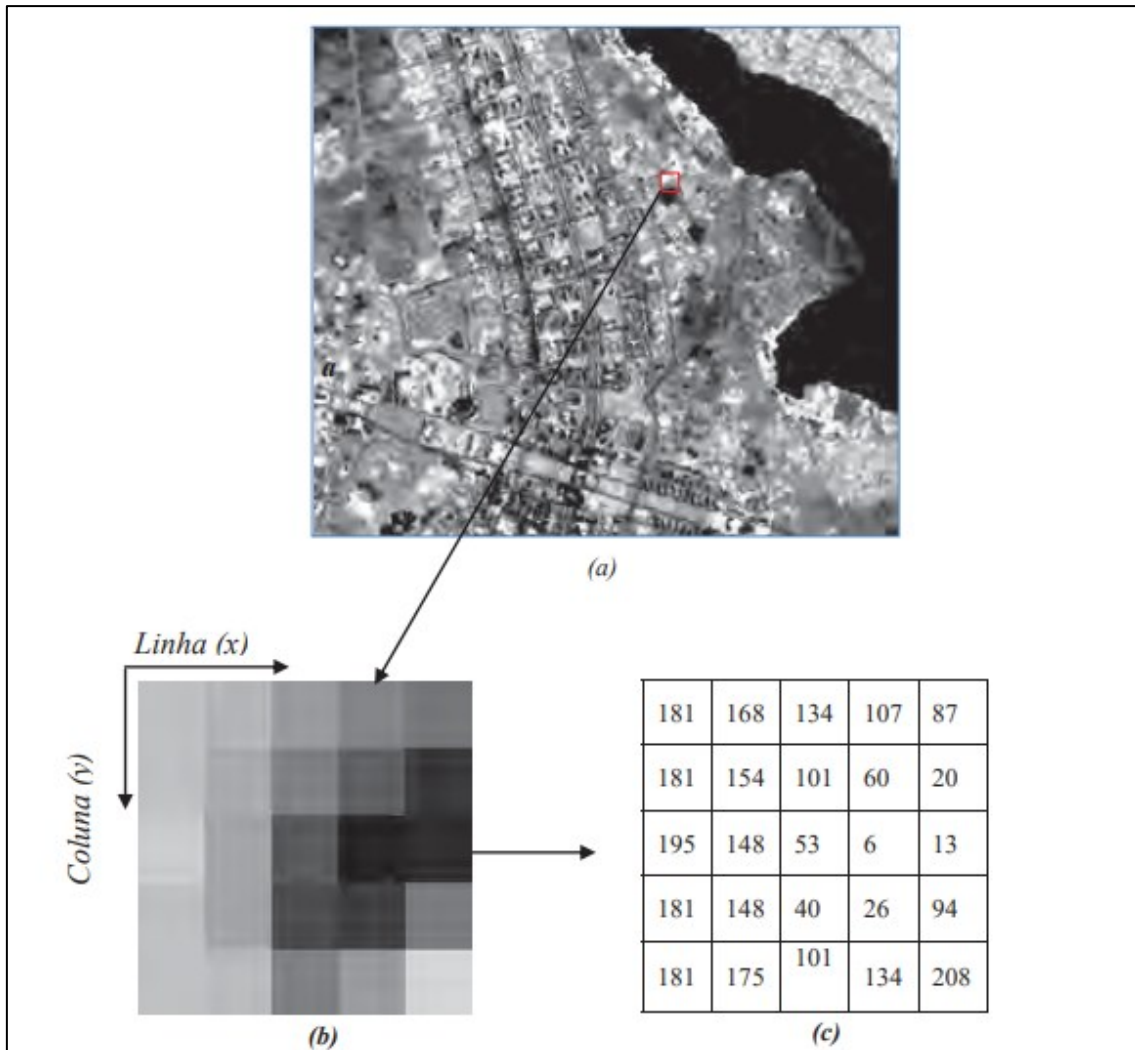
O Sensoriamento Remoto pode ser aplicado em outros tipos de estudos como no estudo de Cicerelli e Galo (2015), o qual utilizaram Sensoriamento Remoto Multifonte aplicado na detecção do fitoplâncton em águas interiores, visto que florações de fitoplâncton podem constituir em riscos à saúde humana e biota aquática, sendo necessários o monitoramento da comunidade fitoplanctônica e a adoção de mecanismos visando à prevenção de sua ocorrência. Para esta análise foram adquiridos dados hiperespectrais em campo e imagens multiespectrais *Modis* e *RapidEye*, os quais foram relacionados a variáveis limnológicas indicadoras do comportamento fitoplanctônico; clorofila a e ficocianina.

4.3 PROCESSAMENTO DIGITAL DE IMAGENS

Uma imagem digital pode ser definida como um conjunto de pontos, cujo cada ponto (pixel) representa a uma informação do terreno, produzida através de uma função bidimensional $f(x,y)$, onde x e y são coordenadas espaciais e o valor de f no ponto (x,y) corresponde o brilho da área equivalente ao pixel (FIGUEIREDO, 2015). A forma digital dos dados é o que torna possível usar computadores para processar imagens, com o objetivo de transformá-los em uma imagem com melhor qualidade espectral e espacial, tornando-a mais adequada para uma determinada aplicação, usando-se de processamentos matemáticos, estatísticos e probabilísticos dos dados (MENESES; ALMEIDA, 2012).

Adiante, a Figura 5 ilustra uma imagem digital com zoom em um grupo de pixels e seus correspondentes valores digitais.

Figura 5 - Imagem digital (a) com área em zoom de grupo de pixels em valores de cinza (b) e correspondentes valores digitais (c).



Fonte: Meneses; Almeida (2012).

Vale destacar que não existe nenhum passo a passo de processamento pré-definido que o usuário é obrigado a fazer. A função básica das técnicas de transformação no domínio espectral e espacial, é preparar uma imagem de alto contraste visual para o analista, com intuito de facilitar a interpretação da cena, explorar todas as informações contidas na imagem e extrair informações de interesse e/ou combiná-las com parâmetros estatísticos, num processo que tem como objetivo melhorar ao máximo a análise dos dados, em termos de eficiência, tempo e custo (MENESES; ALMEIDA, 2012).

4.4 MÉTRICA DE QUALIDADE

4.4.1 Classificação de Imagens

Classificação de imagens multiespectrais é o processo de relacionar os pixels da imagem a um número limitado de classes, as quais que reproduzem os objetos do mundo real, a partir dos seus valores digitais que são geralmente analisados por medidas estatísticas, o pixel é atribuído à classe que melhor corresponde com os critérios estabelecidos (MENESES; ALMEIDA, 2012). Dentre as técnicas de classificação de imagens, tem a classificação supervisionada e não supervisionada.

4.4.1.1 Classificação Não-supervisionada

Classificação Não-supervisionada que não requer nenhuma informação sobre classes de interesse, apenas examina os agrupamentos espectrais do alvo e os dividem em classes, sem a participação do investigador. Esse método de classificação utiliza uma lógica denominada de *clustering*, útil quando não há informações sobre a região imageada, sendo as classes definidas no algoritmo de classificação (TANGERINO; LOURENÇO, 2013 *apud* ROSA, 2009).

4.4.1.2 Classificação Supervisionada

A classificação supervisionada é o processo de identificação de regiões espectralmente semelhantes em uma imagem, onde o usuário identifica inicialmente regiões conhecidas de interesse na superfície terrestre, e posteriormente, um algoritmo selecionado extrapola essas características espectrais selecionadas para outras regiões da imagem, realizando a classificação (CATTANI et al., 2013 *apud* CASTILLEJO-GONZÁLEZ et al., 2009). Portanto, a classificação supervisionada trata-se de designar os tipos de cobertura, amostrar e criar assinaturas dos pixels considerados pertencentes à mesma classe, deste modo, realizar o método de classificação de probabilidade máxima, e finalmente filtrar e refinar os agrupamentos (OLIVEIRA et al., 2014)

4.4.2 Classificação de Variáveis

Quando fazemos uma amostragem, coletamos não apenas a informação sobre a característica de interesse, mas diversas outras informações que auxiliarão no entendimento desta característica. Essas características da população amostrada são denominadas de uma variável. As variáveis podem assumir diferentes valores, que basicamente podem ser separados em Quantitativos e Qualitativos.

4.4.2.1 Variáveis Quantitativas

Os dados de natureza quantitativa são quaisquer atributos sujeitos de alteração numérica e de alguma classificação. Quando um valor é atribuído a uma característica ou propriedade de um objeto sob investigação, supõe-se que a característica ou propriedade tenha aspecto de uma variável quantitativa. Portanto, refere-se de algumas características ou propriedade distinguíveis de um determinado objeto de investigação que podem ser medidos ou manipulados (RAUEN, 2012).

4.4.2.2 Variáveis Qualitativas

Os dados de natureza qualitativos representam uma relação de similaridade ou diversidade, as variáveis qualitativas ou categóricas podem ser nominais, quando as categorias não possuem uma ordem natural, ou podem ser ordinal, onde os dados apresentam uma relação de hierárquica entre as classes (RAMOS et al., 2016).

4.4.3 Algoritmos

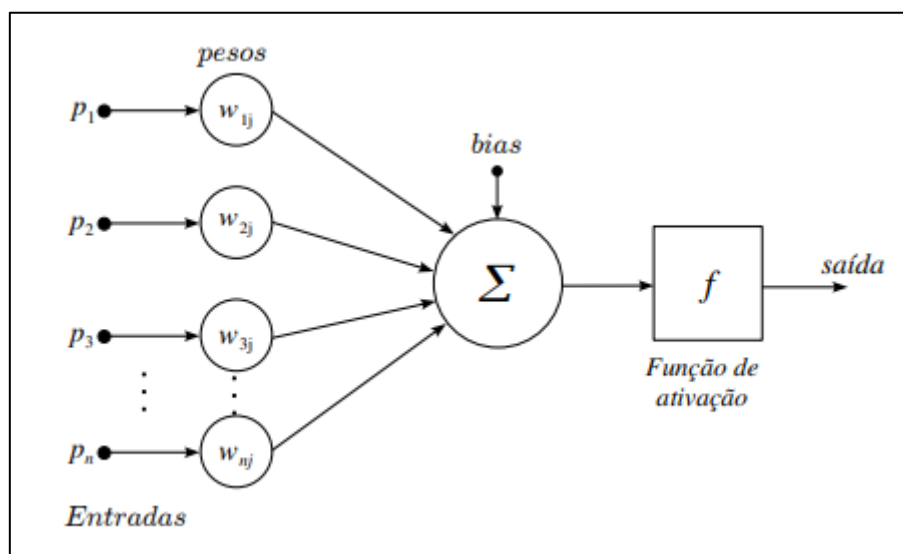
A computação é muito importante para a indústria, onde técnicas de inteligência artificial são cada vez mais usadas em várias áreas existentes, ocasionalmente, algoritmos de aprendizado de máquina também são usados para que os robôs possam reconhecer novos padrões que não foram especificados anteriormente (Galvão et al., 2018). Entre os algoritmos, destacam-se as Redes Neurais Artificiais e o *Support Vector Machine*.

4.4.3.1 Redes Neurais Artificiais (RNA)

A Redes Neurais Artificiais (RNA) é um mecanismo de processamento de dados e recebe esse nome porque compartilha algumas características de funcionamento com as

redes neurais biológicas (Galvão et al., 2018). As redes neurais possuem neurônios artificiais que imitam o comportamento dos neurônios biológicos, esse neurônio artificial possui várias entradas, que representam as conexões sinápticas com outras unidades semelhantes a ele, e uma saída, onde o valor depende de modo direto da soma ponderada de todas as saídas de outros neurônios ligados a este neurônio (MATSUNAGA, 2012). A Figura 6 adiante ilustra aparência de neurônio artificial.

Figura 6 - Modelo artificial de neurônio biológico



Fonte: Matsunaga (2012).

A arquitetura de uma rede neural pode possuir diferentes formas, pois variam de acordo com as necessidades da aplicação, sendo diretamente relacionadas com o algoritmo de aprendizagem utilizado para treinar a rede, mas, basicamente, os itens que constituem a estrutura da rede neural são as Camadas intermediárias, Quantidade de neurônios, Função de transferência e Algoritmo de aprendizado (MATSUNAGA, 2012).

O aprendizado das redes neurais pode ser supervisionado ou não supervisionado. No aprendizado supervisionado, os analistas verificam o quanto a rede está perto de uma solução aceitável e ajustam os pesos entre os neurônios no projeto de treinamento para fornecer diferenças menores entre as saídas desejadas. Na aprendizagem não supervisionada, as saídas da rede são desconhecidas e diferentes categorias de padrões são distinguidas dos dados apresentados à rede, por meio de algoritmos de aprendizagem que geralmente são apoiados no conceito de vizinhança e agrupamento (MATSUNAGA, 2012).

4.4.3.2 *Support Vector Machine* (SVM)

O algoritmo *Support Vector Machine* (SVM), foi criado pelo pesquisador russo Vladimir Vapnik e seus colaboradores, em 1995, com o intuito de solucionar problemas de classificação de padrões, o algoritmo tem uma arquitetura de classificação de imagem baseado na teoria de aprendizagem estatística e apoiada no conceito de Minimização do Risco Estrutural, sendo que este método calcula o erro com relação ao conjunto de treinamento e faz a diminuição do erro global (MARTINES et al., 2019).

Segundo o Laboratório de Aprendizado de Máquina em Finanças e Organizações (LAMFO) *apud* Vapnik et al. (1997), *Support Vector Machine* é um algoritmo de aprendizado supervisionado, cujo objetivo é classificar determinado conjunto de pontos de dados que são mapeados para um espaço de características multidimensionais usando uma função kernel, a qual é responsável pela suavização mais adequada da estrutura dos conjuntos de dados.

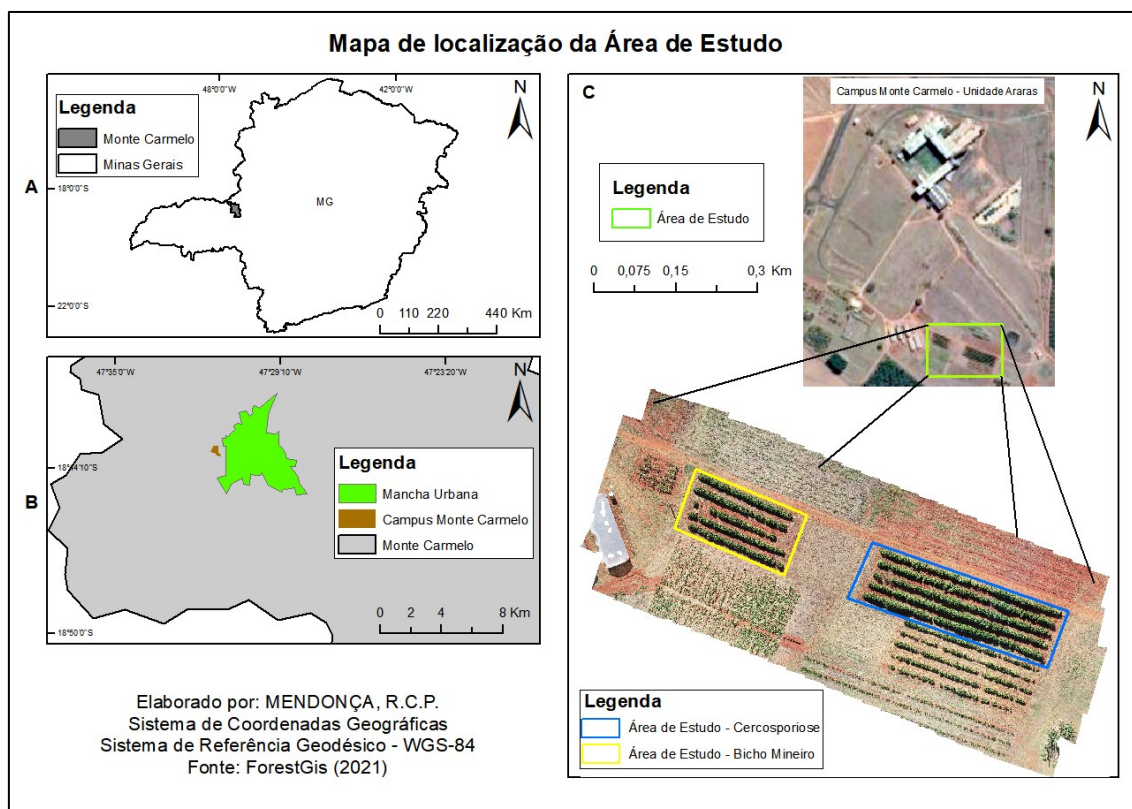
De acordo com Lorena e Carvalho (2003) *apud* Smola et al. (1999), as características do SVM é a boa capacidade de generalização, medida calculada por sua competência na classificação de dados que não pertence ao conjunto usado em seu treinamento, além de mostrar robustez em grandes dimensões e convexidade da função objetivo, sendo isso uma vantagem sobre as Redes Neurais Artificiais, em que existe a presença de mínimos locais na função objetivo a ser reduzida. O SVM destaca-se também diante de classificadores convencionais, principalmente quando o número de amostras de treinamento é pequeno e a dimensionalidade dos dados é grande (ANDREOLA, 2009).

5 MATERIAL E MÉTODOS

5.1 ÁREA DE ESTUDO

A área de estudo para realização dessa pesquisa (Figura 7) é localizada na Universidade Federal de Uberlândia, campus Monte Carmelo – Unidade Araras, o qual se localiza na rodovia estadual LMG-746, Km 1. Esta área pertence ao município de Monte Carmelo-MG que está situado na Mesorregião do Triângulo Mineiro e Alto Paranaíba.

Figura 7 - Mapa de localização da Área de Estudo. (A) Minas Gerais. (B) Mancha Urbana de Monte Carmelo e Universidade Federal de Uberlândia. (C) Área de Estudo



Fonte: O Autor (2022).

O talhão de estudo da Cercosporiose do cafeeiro é formado por 8 cultivares diferentes, entre elas estão, Acaiá Cerrado MG-1474, Mundo Novo IAC 379-19, Bourbon Amarelo IAC J10, Catuaí Vermelho IAC 99, Topázio MG-1190, Acauã Novo, IAC 125 RN e Paraíso MG H 419-1. O talhão de estudo do Bicho-Mineiro do cafeeiro é formado

pela cultivar Topázio MG-1190, e como na área da Cercosporiose, ambos talhões apresentam irrigação por gotejamento.

5.2 MATERIAL

Esta pesquisa foi dividida em duas etapas, sendo a primeira etapa de obtenção de dados em campo e a segunda etapa de processamento computacional das informações. Os materiais necessários para a primeira etapa estão descritos a seguir:

- Câmera Mapiir, GPS integrado e operando nas regiões do vermelho, verde e infravermelho próximo do espectro eletromagnético (Figura 8B).
- *ARP Phantom 4 Advanced*, equipado com câmera RGB de 20 megapixels e GPS integrado (Figura 8A).

Figura 8 – (A) *ARP Phantom 4 Advanced*. (B) Câmera Mapiir acoplada a ARP



Fonte: O Autor (2022).

- Receptores GNSS da marca *Topcon*, modelo *Hiper V*, para georreferenciamento das parcelas experimentais (Figura 9).

Figura 9 – Receptor GNSS (*Rover*) utilizados no georreferenciamento das parcelas experimentais.



Fonte: O Autor (2022).

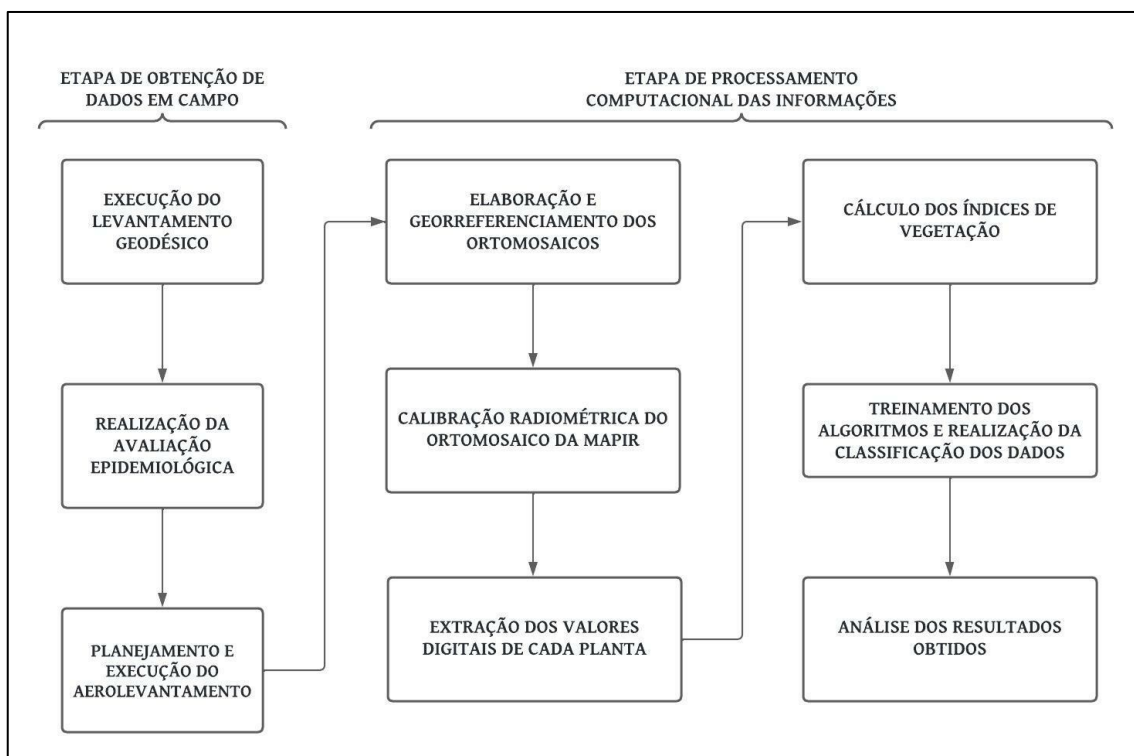
Posteriormente, para realização da segunda etapa foram necessários os seguintes hardwares e softwares para o processamento das informações:

- Notebook *Samsung Expert*, Intel Core i5, 8GB RAM com sistema operacional de 64 bits para processos e análises computacionais;
- *Smartphone* modelo M30 da Samsung para armazenar e executar o plano de voo;
- Software *DroneDeploy* para elaborar o plano de voo;
- Software *AgiSoft PhotoScan Professional 1.3.3*, versão gratuita de 30 dias, para os processamentos fotogramétricos (Criação do ortomosaicos da área);
- Software *Qgis 3.8.2* para georreferenciamento dos ortomosaicos, através dos pontos levantados em campo;
- Software *Envi Classic 5.0* para medir os valores digitais de cada planta;
- Software *Weka 3.9.4* para realizar a classificação dos dados através de algoritmos baseados em aprendizado de máquina.

5.3 MÉTODOS

A metodologia deste estudo está descrita no Fluxograma abaixo (Figura 10), cujo as atividades são: (I) Efetuar o levantamento geodésico; (II) Realizar a avaliação epidemiológica da Cercosporiose e do Bicho Mineiro; (III) Planejar e executar o aerolevanteamento; (IV) Gerar e Georreferenciar os ortomosaicos; (V) Calibrar radiometricamente do ortomosaico da MapiR; (VI) Extrair os valores digitais de cada planta; (VII) Calcular os índices de vegetação; (VIII) Treinar os algoritmos e realizar a classificação dos dados e (IX) Analisar os resultados obtidos e a classificação realizada.

Figura 10 – Fluxograma da pesquisa



Fonte: O Autor (2022).

5.3.1 Levantamento Geodésico

O levantamento geodésico foi a primeira atividade desenvolvida na pesquisa e teve como propósito levantar as coordenadas das plantas avaliadas, além de alguns pontos no centro e nas extremidades da área de estudo. Posteriormente, com o ortomosaico gerado, estes pontos foram utilizados para georreferenciar a imagem, realizando assim a correção geométrica.

Para realização do levantamento geodésico foi utilizado receptores GNSS da marca *Topcon*, modelo *Hiper V*, cujo a “Base” foi instalada no marco de centro forçado,

pertencente a Rede Planialtimétrica do Instituto Brasileiro de Geografia e Estatística (IBGE), localizado no estacionamento da Universidade Federal de Uberlândia, campus Monte Carmelo – Unidade Araras, e com “*Rover*” realizou-se o levantamento dos pontos de interesse.

5.3.2 Avaliação Epidemiológica

A avaliação epidemiológica foi a segunda atividade desenvolvida na pesquisa, a qual ocorreu no mês de julho de 2021, e teve como intuito identificar as plantas saudáveis e as doentes dos dois talhões de cafés. Para identificar as plantas infestadas pela Cercosporiose, foram analisadas na própria planta e de maneira não destrutiva, as folhas a partir do terceiro ou quarto par, de 4 ramos escolhidos aleatoriamente no terço médio/superior da planta, sendo 2 ramos ao Norte e 2 ao Sul.

Para identificar as plantas infectadas pelo Bicho Mineiro, na própria planta e de maneira não destrutiva, foram analisadas 10 folhas localizadas no terceiro ou quarto par, de ramos escolhidos aleatoriamente no terço médio/superior da planta, sendo 5 folhas ao Norte e 5 ao Sul.

5.3.3 Planejamento e execução do aerolevanteamento

Simultaneamente ao levantamento geodésico e a avaliação epidemiológica, ocorreu o planejamento e a execução do aerolevanteamento por meio de uma ARP e duas câmaras multiespectrais. Para realização de um aerolevanteamento, primeiramente, foi necessário realizar um planejamento, pois horário e condições atmosféricas poderiam afetar o nível de informações registrada pela fotografia aérea, além de que alguns parâmetros são necessários previamente serem definidos para o êxito do aerolevanteamento.

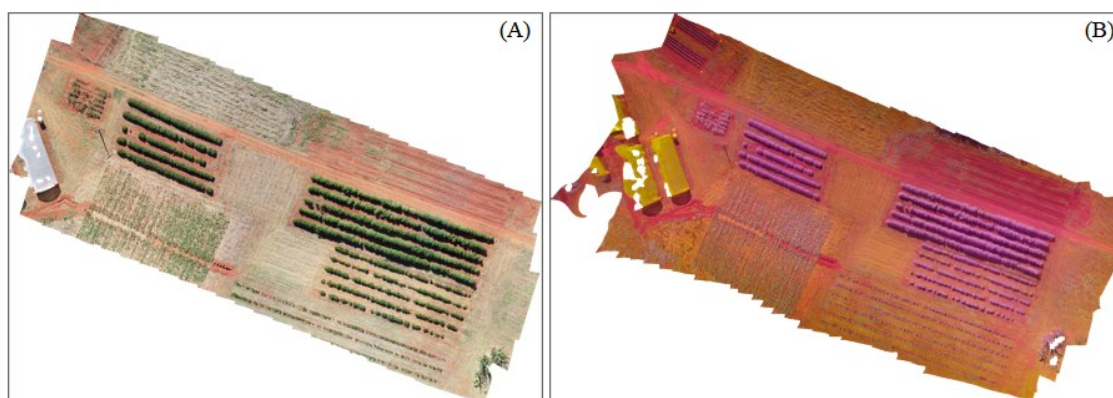
Para a inserção e cálculo dos parâmetros necessários para o planejamento de voo, foi utilizado o software *DroneDeploy*, no qual foi determinado que as fotografias fossem tiradas com uma sobreposição lateral de 70% e uma sobreposição longitudinal de 75%, na altitude de 30 metros, com velocidade de aproximadamente de 10 m/s, totalizando 62 fotografias, tiradas em 3 faixas no decorrer do voo de duração de 3 minutos e 34 segundos.

Assim, após o planejamento do voo, foi efetuado o aerolevanteamento, o qual foi realizado em um dia com condições atmosféricas adequadas, ou seja, céu limpo, sem previsão de chuva, realizado entre onze horas da manhã e meio-dia, visando imagens com maior incidência da REM e menor sombreamento no cafeeiro.

5.3.4 Elaboração e georreferenciamento dos ortomosaicos

Posteriormente, a partir das imagens obtidas no aerolevanteamento, deu-se início a construção dos ortomosaicos por meio do software *AgiSoft PhotoScan Professional 1.3.3*, versão gratuita de 30 dias, sendo um relativo as imagens da câmera do RGB (Figura 11A) e outro referente as imagens da câmera Mapir (Figura 11B).

Figura 11 – (A) Ortomosaico da câmera RGB (B) Ortomosaico da câmera Mapir



Fonte: O Autor (2022).

Em seguida, utilizando o software Qgis 3.8.2 inicializou-se o georreferenciamento dos ortomosaicos, com intuito de realizar a correção das distorções geométricas presente nas imagens, para tal ação foi necessário adicionar os pontos de controle, obtidos no levantamento geodésico, juntamente com os ortomosaicos no programa. O georreferenciamento dos ortomosaicos ocorreu de forma manual, sendo realizado um de cada vez, e refere-se a uma ação que trata da identificação dos pontos de controle no ortomosaico, e após a sua identificação, atribuir suas coordenadas.

5.3.5 Calibração Radiométrica do ortomosaico da Mapir

A calibração radiométrica foi realizada apenas no mosaico gerado a partir das imagens da Mapir, usando-se o software *Mapir Camera Control* (MCC). A correção radiométrica é utilizada para eliminação dos efeitos de interferências da atmosfera e da topografia. No software, foi selecionado o modelo, a lente e o filtro da câmara Mapir utilizados, com base nessas informações o software calculou os valores de calibração necessários. Após o processo de calibração criou-se, automaticamente, uma pasta com mosaico calibrado.

5.3.6 Extração dos valores digitais de cada planta

Após os ortomosaicos serem produzidos, iniciou-se o processo de extração dos valores digitais de cada planta, por meio dos ROI'S (Regiões de interesse), tal operação foi realizada através do software *Envi Classic 5.0* e de maneira manual, em um ortomosaico de cada vez. Para auxiliar na identificação das plantas avaliadas e conseqüentemente na extração de seus valores digitais, juntamente com o ortomosaico, abriu-se *shapefile* com os pontos das plantas examinadas, pontos os quais foram obtidos no levantamento geodésico descrito anteriormente.

Os ROI'S expressam em valores numéricos cada pixel do recorte do *raster* de interesse, possibilitando, por exemplo, a análise de média, desvio padrão, máximo e mínimo. Para esta pesquisa os ROI'S foram gerados com intuito de coletar os valores digitais das bandas espectrais de cada planta, posteriormente, associar esses valores ao estado de saúde de cada planta.

5.3.7 Cálculo dos Índices de Vegetação

Com os valores digitais de cada planta, sendo saudável ou infectada, foram calculados os índices de vegetação com intuito de auxiliar os algoritmos na classificação dos dados. Estes índices são relevantes em virtude de serem sensíveis a variações biofísicas e bioquímicas do vegetal.

Os índices de vegetação calculados a partir dos comprimentos de onda dos sensores de estudo estão ilustrados na Quadro 1 a seguir:

Quadro 1 – Índices Vegetativos

Nome	Abreviatura	Equação
Normalized Green Red Difference Index	NGRDI	$(G - R) / (G + R)$
Green Leaf Index	GLI	$(2 \times G - R - B) / (2 \times G + R + B)$
Triangular Greenness Index	TGI	$G - (0.39 * R) - (0.61 * B)$
Visible Atmospherically Resistant Index	VARI	$(G - R) / (G + R - B)$
Simple Ratio	SR	IR / R
Green Normalized Difference Vegetation Index	GNDVI	$(IR - G) / (IR + G)$
Normalized Difference Vegetation Index	NDVI	$(IR - R) / (IR + R)$

Fonte: O Autor (2022).

5.3.8 Classificação dos dados e Validação dos Resultados

Após a extração dos valores digitais das bandas espectrais e calculado os índices vegetativos, subseqüentemente, utilizou-se esses dados para formar arquivos no formato arff, sendo estes utilizados na validação e treinamento dos algoritmos. Os arquivos são compostos com as classes das plantas (saudável ou infectada), com os valores digitais das bandas RGB e RGN e com índices de vegetação calculados.

Para treinamento no software *Weka* 3.9.4, as plantas saudáveis foram adotadas como classe 0, enquanto as plantas infectadas por Cercosporiose pela classe 1, e as infectadas por BMC pela classe 2. Destaca-se, que as plantas avaliadas apresentam diferentes graus das doenças, no entanto, não foram discriminadas nesse estudo.

Ademais, para não comprometer o treinamento dos classificadores as classes das plantas foram balanceadas, sendo assim, o arquivo de treinamento foi dividido em: 27 plantas saudáveis, 27 infectadas por BMC e 27 por Cercosporiose. Seguindo a mesma metodologia, o arquivo de classificação foi composto por 7 plantas saudáveis, 7

infectadas por BMC e 7 por Cercosporiose, totalizando 34 plantas por classe e consequentemente, utilizados 80% dos dados para treinamento e 20% para classificação.

Por fim, com os arquivos desenvolvidos realizou-se a discriminação das plantas por meio de algoritmos baseados em aprendizado de máquina, como o *Support Vector Machine* e as Redes Neurais Artificiais, através do software *Weka* 3.9.4 e de maneira supervisionada.

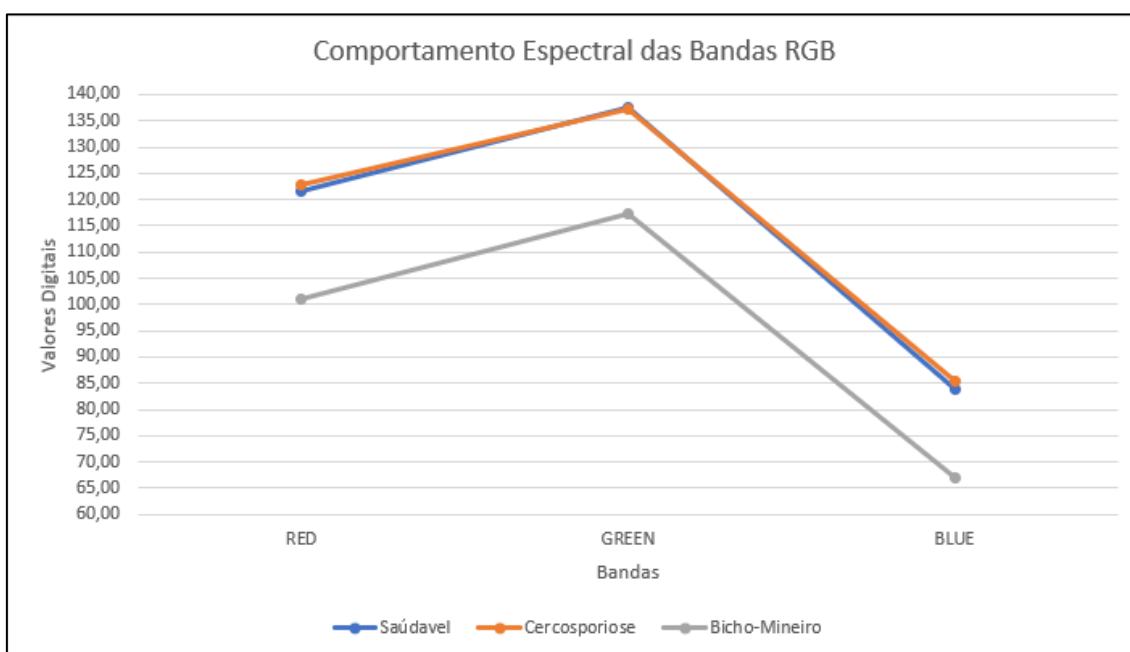
Para validação dos resultados utilizou-se a matriz de confusão, visto que a matriz de confusão representa a qualidade obtida de uma classificação, sendo expressa mediante a correlação de informações dos dados de referência, compreendido como verdadeiro, com os dados classificados, logo, permitindo uma análise mais detalhada do que a mera proporção de classificações corretas.

Além das matrizes de confusão, usou-se o Coeficiente Kappa para validação dos resultados. O coeficiente Kappa é definido como uma medida de associação utilizada para descrever e aferir o grau de concordância, isto é, a confiabilidade e a precisão da classificação (PERROCA; GAIDZINSK, 2001 *apud* KOTZ S, 1983). O valor do coeficiente Kappa pode variar de 0 a 1, visto que quanto mais próximo de 1, maior o grau de confiabilidade e a precisão da classificação.

6 RESULTADOS E DISCUSSÕES

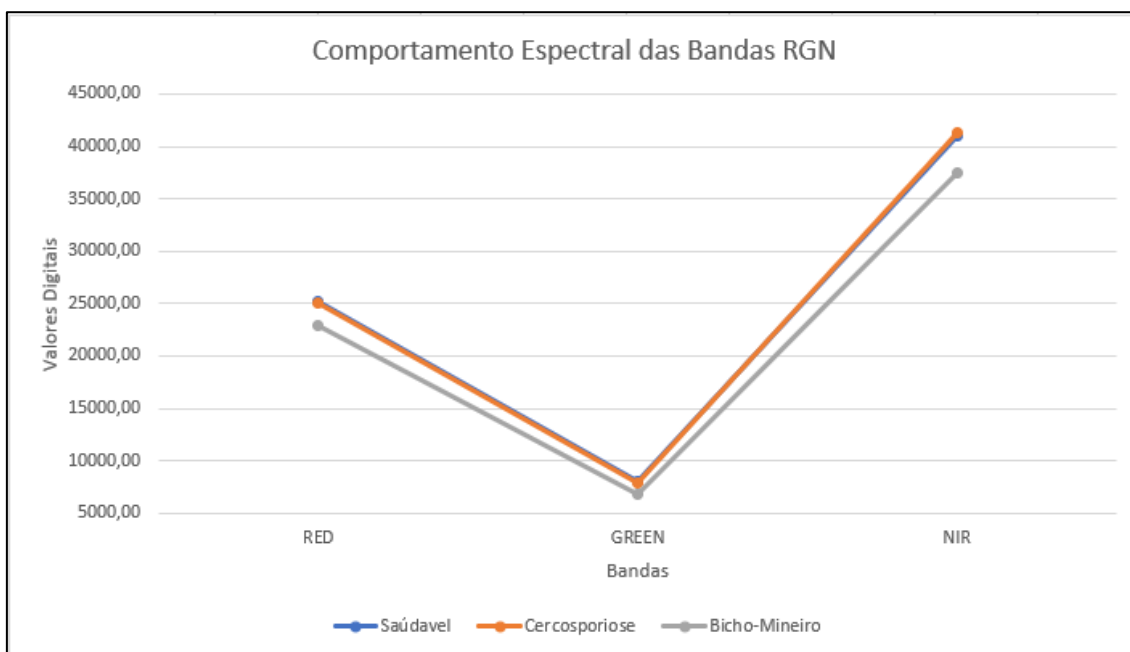
Após a realização de todas as etapas descritas na metodologia, observa-se ao final do treinamento e classificação, que os dois algoritmos foram capazes de discriminar partes das plantas saudáveis, infectadas por BMC e infectadas por Cercosporiose. Afim de auxiliar a análise dos resultados, assim como compreender a estrutura de como algoritmo trabalha, foram feitos dois gráficos, um representando o comportamento espectral das bandas RGB e outro das bandas RGN, das amostras utilizadas na classificação. (Figura 12 e 13).

Figura 12 – Comportamento espectral das bandas RGB



Fonte: O Autor (2022).

Figura 13 – Comportamento espectral das bandas RGN



Fonte: O Autor (2022).

Ao analisar o comportamento das bandas nota-se que as plantas saudáveis e as infectadas por Cercosporiose apresentam valores digitais similares, constatando, conseqüentemente, que as plantas com Cercosporiose avaliadas neste estudo não se encontram em um estado avançado. Com isto, as plantas infestadas apresentaram valores digitais semelhantes às plantas saudáveis na faixa de luz verde e infravermelho, visto que a luz verde é empregada para medir a refletância da vegetação verde saudável e a infravermelho a biomassa da vegetação. Logo, essa semelhança ocasiona certa confusão nos algoritmos e contribuindo assim para uma fraca discriminação.

Observa-se também, em contrapartida, que as plantas infestadas por BMC manifestaram valores digitais menores nas bandas RGB e RGN, tornando-se mais fácil o treinamento do algoritmo e conseqüentemente a classificação. Vale destacar, que o BMC se encontrava em um estado avançado, e tem como característica a redução da absorção de clorofila, e por conseqüência da fotossíntese, devido aos ataques dos bichos mineiros formarem lesões nas folhas, que se transformam em áreas mortas de necrose, por conseqüência, numa menor área fotossintética, justificando assim os menores valores digitais.

A seguir, os quadros 2, 3, 4 e 5, contêm os resultados encontrados, em que se nota que os algoritmos apresentaram desempenhos semelhantes, no entanto, o *Support Vector Machine* apresentou melhor desempenho que as Redes Neurais Artificiais na

discriminação das ocorrências nas plantas. Vale ressaltar, antes da apresentação dos resultados, que as plantas sadias foram nomeadas como “a”, as infectadas por Cercosporioses como “b” e por BMC como “c”. Além disto, que na cor verde estão representados os números de plantas classificadas corretamente, em laranja/pastel, a baixo da diagonal, os falsos positivos (FP) e em cinza, a cima da diagonal, os falsos negativos (FN).

Quadro 2 – Matriz de Confusão e resultados da classificação por meio do *Support Vector Machine* das bandas RBG e seus índices

RBG - Support Vector Machine		Dados Classificados			Total	Erro de Comissão (%)	Erro de Omissão (%)	Kappa	Acurácia Global (%)
		a	b	c					
Dados de Referência	a	5	0	2	7	54,55	28,57	0,21	47,62
	b	5	0	2	7	100,00	100,00		
	c	1	1	5	7	44,44	28,57		
Total		11	1	9	21				

Fonte: O Autor (2022).

Quadro 3 – Matriz de Confusão e resultados da classificação por meio do *Support Vector Machine* das bandas RGN e seus índices

MAPIR - Support Vector Machine		Dados Classificados			Total	Erro de Comissão (%)	Erro de Omissão (%)	Kappa	Acurácia Global (%)
		a	b	c					
Dados de Referência	a	3	2	2	7	57,14	57,14	0,21	47,62
	b	4	1	2	7	75,00	85,71		
	c	0	1	6	7	40,00	14,29		
Total		7	4	10	21				

Fonte: O Autor (2022).

Quadro 4 – Matriz de Confusão e resultados da classificação por meio das Redes Neurais Artificiais das bandas RGB e seus índices

RBG - Redes Neurais Artificiais		Dados Classificados			Total	Erro de Comissão (%)	Erro de Omissão (%)	Kappa	Acurácia Global (%)
		a	b	c					
Dados de Referência	a	3	3	1	7	40,00	57,14	0,21	47,62
	b	2	2	3	7	71,43	71,43		
	c	0	2	5	7	44,44	28,57		
Total		5	7	9	21				

Fonte: O Autor (2022).

Quadro 5 – Matriz de Confusão e resultados da classificação por meio das Redes Neurais Artificiais das bandas RGN e seus índices

MAPIR - Redes Neurais Artificiais		Dados Classificados			Total	Erro de Comissão (%)	Erro de Omissão (%)	Kappa	Acurácia Global (%)
		a	b	c					
Dados de Referência	a	0	5	2	7	100,00	100,00	0,07	38,10
	b	4	1	2	7	83,33	85,71		
	c	0	0	7	7	36,36	0,00		
Total		4	6	11	21				

Fonte: O Autor (2022).

Observando os resultados, nota-se que o SVM nas duas classificações, classificou corretamente 47,62% das plantas, obtendo coeficiente Kappa igual 0,21, enquanto RNA classificou corretamente 47,62% das plantas na classificação das bandas do visível e 38,10% na classificação das bandas RGN, resultando em coeficiente Kappa igual 0,21 e 0,07 respectivamente. Portanto, o *Support Vector Machine* apresentou melhor desempenho.

Percebe-se ao examinar a matriz de confusão da classificação por SVM das bandas do visível, que o algoritmo discriminou das 7 plantas sadias, 5 de forma correta e 2 de forma incorreta, portanto associou a outra classe. Ademais, nota que as plantas infectadas por Cercosporiose não foram classificadas corretamente, todas as 7 plantas estão correlacionadas a outra condição. Repare-se também, que o algoritmo discriminou 5 plantas infectadas por BMC corretamente, apenas 2 de forma errada.

Ao avaliar a matriz de confusão da classificação por SVM das bandas do RGN, verifica-se que o algoritmo discriminou das 7 plantas sadias, 3 de forma correta e 4 de forma incorreta. Além de que, apenas 1 planta infectada por Cercosporiose foi classificada corretamente, todas as outras 6 plantas estão correlacionadas a outra condição. Percebe-se também, que o algoritmo discriminou 6 plantas infectadas por BMC corretamente, apenas 1 de maneira errada.

Por fim, observa-se que as melhores classificações do SVM ocorreram na discriminação para as plantas infectadas por BMC, visto que apresentaram menor Erro de Comissão e Omissão. Em virtude que o Erro de Comissão informa a quantidade de classificações associadas a falsos positivos, enquanto o Erro de Omissão, informa a quantidade percentual de erros de omissões, ou seja, classes que deveriam ter sido detectadas, mas que por alguma razão foram omitidas da classificação.

Ao analisar a matriz de confusão da classificação feita com a RNA das bandas do visível, nota-se que das 7 plantas saudáveis, o RNA discriminou 3 de forma correta e 2 como infestada por Cercosporiose e 2 como infestada por BMC. Das plantas infectadas por Cercosporiose 2 foram discriminadas de maneira corretas e 5 estão associadas a outra classe, contudo o maior resultado dentre as outras classificações. Por último, das 7 plantas infectadas por BMC, 5 foram discriminadas corretamente, 2 como infestada por Cercosporiose.

Quanto a classificação feita com a RNA das bandas do RGN, ao examinar a matriz de confusão verifica-se que o algoritmo discriminou das 7 plantas saudáveis, nenhum de forma correta. Ademais, que as plantas infectadas por Cercosporiose apenas 1 foi classificada corretamente, todas as outras 6 plantas estão correlacionadas a outra condição. Por fim, percebe-se que o algoritmo discriminou todas as plantas infectadas por BMC corretamente e semelhante as classificações por SVM, observa-se que as melhores classificações ocorreram na discriminação das plantas infectadas por BMC, em virtude de apresentar menor Erro de Comissão e Omissão.

7 CONCLUSÃO

Ao final deste trabalho concluiu-se que existe a possibilidade de discriminar as plantas infectadas das saudáveis, por meio dos algoritmos baseados em aprendizado de máquina, e através do conjunto de atributos desenvolvidos, além de que o algoritmo SVM apresentou melhor desempenho que as Redes Neurais Artificiais na classificação.

Além disto, o BMC foi mais discriminável nas imagens que a Cercosporiose. Para trabalhos futuros, indica-se um maior aprimoramento dos atributos e mais avaliações de campo, com o objetivo de elaborar um banco de dados mais robusto e capaz alcançar melhores desempenho na discriminação da Cercosporiose e superiores resultados para coeficiente Kappa.

REFERÊNCIAS

- ANDREOLA, R. **Support Vector Machines na classificação de imagens hiperespectrais**. 2009. Dissertação (Mestrado em Sensoriamento Remoto) – Centro Estadual de Pesquisas em Sensoriamento Remoto e Meteorologia, Porto Alegre, 2009.
- ASSIS, G.A. et al. Incidência de Bicho-Mineiro (*Leucoptera coffeella*) (Lepidoptera: Lyonetiidae) em cafeeiro submetido a diferentes critérios de irrigação e densidades de plantio. 2009. **In: Simpósio de Pesquisa dos Cafés do Brasil (6.: 2009: Vitória, ES). Anais Brasília**. Brasília, DF: Embrapa - Café, 2011.
- CATTANI, C.E.V. et al. Desempenho de algoritmos de classificação supervisionada para imagens dos satélites RapidEye. **IN: Anais XVI Simpósio Brasileiro de Sensoriamento Remoto-SBSR**, Foz do Iguaçu, PR, Brasil, v. 13, 2013.
- BOLDINI, J.M. **Epidemiologia de ferrugem e da Cercosporiose em cafeeiro irrigado e fertirrigado**. 2001. Dissertação (Mestrado em Agronomia, área de concentração em Fitopatologia) – Universidade Federal de Lavras, Lavras. 2001.
- CICERELLI, R.E.; GALO, M. de L.B.T. Sensoriamento remoto multifonte aplicado na detecção do fitoplâncton em águas interiores. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 19, p. 259-265, 2015.
- Consórcio Pesquisa Café**. Disponível em:
<http://www.consorcioquesquisacafe.com.br/index.php/imprensa/noticias/1071-2021-07-19-20-43-30/> Acesso em: 10 out. 2021.
- DE MORAES NOVO, E.M.L. **Sensoriamento Remoto: princípios e aplicações**. Editora Blucher, 2010.
- FIGUEIREDO, D. **Conceitos básicos de sensoriamento remoto**. São Paulo, 2005.
- FORMAGGIO, A.R.; SANCHES, I.D.A. **Sensoriamento remoto em agricultura**. Oficina de Textos, 2017.
- GALVÃO, G.F.P. et al. **Visão computacional para detecção de doenças fúngicas na agricultura**. Única Cadernos Acadêmicos, v. 2, n. 1, 2018.
- GARCIA, D. **Incidência e severidade da Cercosporiose do cafeeiro (*Coffea arabica* L.) em função de doses de potássio e cálcio em solução nutritiva**. 2002. Dissertação (Mestrado em Agronomia, área de concentração em Fitopatologia) – Universidade Federal de Lavras, Lavras. 2002.
- HALLAK, R.; PEREIRA FILHO, A.J. Metodologia para análise de desempenho de simulações de sistemas convectivos na região metropolitana de São Paulo com o modelo ARPS: sensibilidade a variações com os esquemas de advecção e assimilação de dados. **Revista Brasileira de Meteorologia**, v. 26, n. 4, p. 591-608, 2011.

JENSEN, J.R. **Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres**. São José dos Campos: Parêntese, 2009.

KIMATI, H. et al. **Manual de fitopatologia**. 2005. 3. ed. São Paulo: Agronômica Ceres, 1997.

LEMES, E.M. **Sensoriamento remoto para detecção de seringais**. 2017. Dissertação (Doutorado em Agronomia, área de concentração em Fitopatologia) – Universidade Federal de Uberlândia, Uberlândia, 2017.

LORENA, A.C.; CARVALHO, A.C.P.L.F. **Introdução às máquinas de vetores suporte (Support Vector Machines)**. Laboratório de Inteligência Computacional, ICMC/USP, São Carlos, 2003.

MARTINES, M.R. et al. Comparação dos classificadores Maximum Likelihood Classification e Support Vector Machine para o mapeamento de fragmentos florestais na Mata Atlântica. **In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), XIX, 2019**, Santos, SP, Brasil, p. 78-81.

MATSUNAGA, V.Y. **Curso de Redes Neurais utilizando o MATLAB**. Belém, 2012.

MENESES, P.R.; ALMEIDA, T. de. **Introdução ao processamento de imagens de sensoriamento remoto**. Universidade de Brasília, Brasília, 2012.

MOLIN, J.P.; DO AMARAL, L.R.; COLAÇO, A. **Agricultura de precisão**. 1. ed. São Paulo: Oficina de textos, 2015.

MURTA, D. F. **Breve história do café no Brasil**. Seiva (Brasil) v. 38 (85) p. 18-41, 1978.

NAGAY, J. H. C. **Café no Brasil: dois séculos de história**. Formação Econômica, Campinas, n. 17-23, jun. 1999

NAUE, C.R. et al. Sensoriamento remoto como ferramenta aos estudos de doenças de plantas agrícolas: uma revisão. **Revista Brasileira de Geografia Física**, v. 3, n. 3, p. 190-195, 2011.

OLIVEIRA, C. A. et al. Escala diagramática para avaliação da severidade de Cercosporiose em folhas de cafeeiro. 2001. **In: Simpósio de Pesquisa dos Cafés do Brasil (2.: 2001: Vitória, ES). Resumos**. Brasília, DF.: Embrapa Café, 2001. 181p.: il.

OLIVEIRA, G. et al. Classificação supervisionada como ferramenta para avaliação da dinâmica do uso e cobertura do solo. **Enciclopedia biosfera**, v. 10, n. 18, 2014.

PANTOJA-GOMEZ, L. M.; CORRÊA, A. S.; DE OLIVEIRA, L. O.; GUEDES, R. N. **C. Common origin of Brazilian and Colombian populations of the neotropical coffee leaf miner, Leucoptera coffeella (Lepidoptera: Lyonetiidae)**. Journal of economic entomology, 112(2), 924-931. 2019

PERROCA, M.G.; GAIDZINSKI, R.R. Avaliando a confiabilidade interavaliadores de um instrumento para classificação de pacientes: coeficiente Kappa. **Revista da Escola de Enfermagem da USP**, v. 37, n. 1, p. 72-80, 2003.

PITHAN, P.A. et al. **Sensoriamento remoto hiperespectral aplicado na discriminação de doenças fúngicas da videira**. Simpósio Brasileiro de Sensoriamento Remoto, 2017.

Laboratório de Aprendizado de Máquina em Finanças e Organizações (LAMFO). Disponível em: <https://lamfo-unb.github.io/2017/07/13/svm/>. Acesso em: 28/05/2021

RAMOS, A.P.M. et al. Avaliação qualitativa e quantitativa de métodos de classificação de dados para o mapeamento coroplético. **Revista Brasileira de Cartografia**, v. 68, n. 3, 2016.

RAUEN, F.J. Pesquisa científica: discutindo a questão das variáveis. **Publicado nos Anais do IV Simpósio sobre Formação de Professores–SIMFOP da Universidade do Sul de Santa Catarina, Campus de Tubarão, de**, v. 7, p. 1-14, 2012.

Rehagro. Disponível em: <https://rehagro.com.br/blog/Cercosporiose-cercospora-coffeicola-do-cafeeiro/> Acesso em: 10 out. 2021.

RODRIGUES, S.S; BARROS, M.M de. O Uso do VANT e do Sensoriamento Remoto como ferramenta de melhoria das Técnicas de Agricultura de Precisão. **In: Revista Digital Simonsen**, nº10, maio, 2019. Disponível em www.simonsenbr/revistasimonsen. ISSN: 2446-5941.

SCALON, J.D.; MATEUS, A.L.S.S.; ZACARIAS, M.S. **Análise espaço-temporal do nível de infestação do bicho-mineiro 347 *Leucoptera Coffeella*, (Guérin-Menèveille & Perrottet, 1842)(Lepidoptera: Lyonetiidae) em cafezal orgânico (*Coffea arabica* L.)**. Embrapa Café-Artigo em periódico indexado (ALICE), 2013.

SOUZA, J.C de; REIS, P. R.; RIGITANO, R.L de O. **Bicho-mineiro do cafeeiro: biologia, danos e manejo integrado**. Belo Horizonte: EPAMIG, 1998.

TANGERINO, D. F.; LOURENÇO, R. T. Comparação da exatidão de métodos de classificação supervisionada e não supervisionada a partir do índice kappa na microbacia do Ribeirão Duas Águas em Botucatu/SP. **XVI Simpósio Brasileiro de Sensoriamento Remoto-SBSR**, v. 4093, 2013.

WEISS, M.; JACOB, F.; DUVEILLER, G. **Remote sensing for agricultural applications: A metareview**. Remote Sensing of Environment, v. 236, p.111402. 2020.