

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MÉTODO PARA CONTAGEM DE PLANTAS DE MILHO
BASEADO NO PROCESSAMENTO DIGITAL DE
IMAGENS MULTIESPECTRAIS UTILIZANDO DRONES
EM AMBIENTE DE CAMPO**

WILBUR NAIKE CHIUYARI VERAMENDI

ORIENTADOR: PROF. DR. PAULO ESTEVÃO CRUVINEL

São Carlos – SP

Setembro, 2022

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MÉTODO PARA CONTAGEM DE PLANTAS DE MILHO
BASEADO NO PROCESSAMENTO DIGITAL DE
IMAGENS MULTIESPECTRAIS UTILIZANDO DRONES
EM AMBIENTE DE CAMPO**

WILBUR NAIKE CHIUYARI VERAMENDI

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação, área de concentração: Processamento de Imagens e Sinais.

Orientador: Prof. Dr. Paulo Estevão Cruvinel

São Carlos – SP

Setembro, 2022



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Wilbur Naike Chiuyari Veramendi, realizada em 15/03/2022.

Comissão Julgadora:

Prof. Dr. Paulo Estevão Cruvinel (EMBRAPA)

Prof. Dr. Diego Furtado Silva (UFSCar)

Prof. Dr. Elmer Alexis Gamboa Penaloza (UFPel)

AGRADECIMENTOS

Aos meus pais, Oscar Chiuyari e Alejandrina Veramendi, pelo apoio e amor em minha trajetória de vida.

Ao Prof. Dr. Paulo Estevão Cruvinel, pela dedicação na orientação deste trabalho. Por apoiar e acreditar na realização deste projeto de pesquisa, sempre com palavras de apoio e compreensão nos momentos mais difíceis

À Embrapa Instrumentação, pela oportunidade de desenvolvimento da pesquisa.

À Universidade Federal de São Carlos (UFSCar) e ao Programa de Pós-Graduação em Ciência da Computação (PPGCC), pela oportunidade de desenvolvimento profissional e pessoal. Aos professores do programa, sempre dedicados ao ensino. Aos colaboradores do setor administrativo.

Aos meus colegas da Pós-Graduação com os quais pude compartilhar momentos de vitórias e dificuldades, sobre tudo, os colegas Matheus, Alex, Ricardo e Bruno.

À CNPq (Processo 133517/2019-5) pelo auxílio e apoio durante todo o curso.

Finalmente, agradeço a todos que, diretamente ou indiretamente, contribuíram para a elaboração deste trabalho.

RESUMO

O processamento de imagens multiespectrais obtidas com câmera embarcada em veículos aéreos não tripulados (Drones) têm trazido novas oportunidades para a agricultura de precisão. Inserido neste contexto, este trabalho apresenta um método para a avaliação da emergência de plantas de milho (*Zea mays L.*) em área de cultura. Foram utilizadas imagens RGB adquiridas com um sensor multiespectral embarcado em Drone, técnicas de processamento digital de imagens, reconhecimento de padrões e inteligência computacional para auxílio de decisão baseado no uso de aprendizado de máquina. Para as etapas de aquisição das imagens aéreas, foi utilizada técnica para o processamento de ortomosaicos nos canais espectrais do vermelho, do verde e do azul (RGB), sendo assim possível registrar e organizar um banco de imagens da área de cultura. Para a etapa de pré-processamento, foram avaliadas técnicas para transformações geométricas, ajustes de brilho e contraste em forma global, enquanto ajustes locais foram avaliados com base no uso de técnicas de equalização adaptativa, os quais foram explorados em espaços de cores (YCbYcr), (HSV) e (CIELAB). Para etapa de pós-processamento, uma segmentação baseada em técnica de limiares de cores foi avaliada no espaço HSV, além do uso de filtragem Gaussiana e operações morfológicas. Para viabilizar a etapa de reconhecimento de padrões, técnicas que utilizam mapas de distâncias foram avaliadas considerando o uso de Distância Euclidiana (DE). Assim, a localização de padrões do dossel das plantas de milho foi estudada com algoritmo de *template matching* e máscara padrão de Chamfer. Para as etapas de extração de características foram avaliadas as técnicas de código de cadeia e mapa de padrões circulares. Essa análise possibilitou estabelecer vetores de características baseados em padrões do número de ocorrências das emergências das plantas de milho. Adicionalmente, foram consideradas duas etapas de calibração, uma sobre a altura das plantas e o raio de abertura do dossel e a segunda sobre a relação estabelecida entre o número de sementes plantadas e os raios identificados com as técnicas de reconhecimentos de padrões. Finalmente, uma etapa de classificação dos padrões da abertura do dossel da planta de milho foi estabelecida, tendo sido utilizado um conjunto de classificadores baseados em Máquina de Vetores de Suporte (SVM). O método desenvolvido mostrou-se adequado para contagem e classificação dos padrões em função da abertura do dossel das plantas de milho em estágio pós-emergência.

Palavras-chave: Processamento de imagens digitais, Drone, contagem de plantas, transformada de distância, *template matching*, imagens multiespectrais, SVM.

ABSTRACT

The processing of multispectral images acquired with an embedded camera in unmanned aerial vehicles (Drones) has brought new opportunities for precision agriculture. In such a context, this work presents a method for evaluating the emergence of corn seeds (*Zea mays L.*) in a crop area. For validation a database of spectral images has been organized from flights over a real agricultural area, and digital image processing techniques have been applied, i.e., taking into account the concept of intelligent processing. Therefore, the image processing techniques based on pattern recognition and models to aid decision making by using machine learning were also used. In addition, after images acquisition it has been used the processing of the orthomosaics in the spectral channels, i.e., red (R), green (G) and blue (B), being possible to register and organize all the images. Likewise, for the pre-processing stage, techniques for geometric transformations, brightness and contrast adjustments were also evaluated in a global way, while local adjustments were evaluated based on the use of adaptive equalization techniques, which were explored in the color spaces (YCbYcr), (HSV) and (CIELAB). For the post-processing step, it has been considered a segmentation based on the best observed color threshold technique together with Gaussian filtering and morphological operations. To enable the pattern recognition step, techniques that use distance maps were evaluated considering the use of Euclidean Distance (DE). Thus, the location of canopy patterns in maize plants was studied with a template matching algorithm and a Chamfer pattern mask. For the feature extraction steps, the chain code and circular pattern map techniques have been considered. The analyses made it possible to establish vectors of features based on the patterns related to the number of emergence occurrences for the maize seeds. Finally, two calibration steps have been considered, one of them related to the plant height versus the canopy opening radius, and other related to the number of seeds planted into soil for each position in the crop area versus the identified radii by the developed model. Last, but not least, the classification step has been established, using a set of classifiers based on Support Vector Machine (SVM), and the developed method, proved to be adequate for counting the seeds of the maize plants in the post-emergence stage.

Keywords: Digital image processing, Drone, plant counting, distance transform, template matching, multispectral images, SVM.

LISTA DE SIGLAS

AIBO	Object-based Image Analysis
AP	Agricultura de Precisão
AUC	Área sob a Curva ROC
BOW	Bag of Visual of Words
CCD	Charged-Coupled Device
CLAHE	Contrast Limited Adaptive Histogram Equalization
CMOS	Complementary Metal-Oxide Semiconductor
CONAB	Companhia Nacional de Abastecimento
DE	Distância Euclidiana
DEM	Digital Elevation Model
DLS	Downwelling Light Sensor
DSM	Digital Surface Model
DTM	Digital Terrain Model
FAO	Food and Agriculture Organization
GIS	Geographic Information Systems
GLCM	Gray Level Co-occurrence Matriz
GPS	Global Positioning Systems
HSI	Hue Saturation Intensity
HSL	Hue Saturation Lightness
HSV	Hue Saturation Value
IR	Infrared

LBP	Local Binary Patterns
ML	Machine Learning
NIR	Near-infrared
NCN	Normalized Cross-Correlation
NDVI	Normalized Difference Vegetation Index
PCA	Principal Component Analysis
PDI	Processamento de Imagens
PID	Processamento de Imagens Digitais
PR	Pattern Recognition
RGB	Red Green Blue
ROC	Característica de Operação do Receptor
SfM	Structured from Motion
SR	Sensoriamento Remoto
SVM	Support Vector Machine
TM	Template Matching
TGI	Triangular Greenness Index
VC	Vapnik-Chervonenkis

LISTA DE FIGURAS

Figura 1 – Estádios fenológicos da cultura do milho.	17
Figura 2 – Detalhe da área experimental LANAPRE.	31
Figura 3 – Exemplo de quadrante para contar e estudar plantas.	31
Figura 4 – Diagrama de blocos do método desenvolvido.	33
Figura 5 – Ciclo da agricultura de precisão.	35
Figura 6 – Principais componentes de um Drone	36
Figura 7 – Curva de reflectância para sensor multiespectral	38
Figura 8 – Sensores e painel de reflectância para imageamento aéreo.	39
Figura 9 – Geometria da imagem aérea com Drone	40
Figura 10 – Sobreposição frontal e lateral de imagens aéreas	42
Figura 11 – Padrão de sobreposição para cobertura do solo	42
Figura 12 – Mapeamento durante a sobreposição das imagens aéreas.	43
Figura 13 – Marcado de pontos de controle em campo.	44
Figura 14 – Produtos derivados a partir do processo SfM.	45
Figura 15 – Processo de produção de uma ortofoto digital.	46
Figura 16 – Exemplo de nuvem de pontos esparsos a partir de canais RGB.	47
Figura 17 – Exemplo de ortomosaico gerado com dois fluxos de trabalho diferentes	47
Figura 18 – Imagens multiespectrais RGB com seus respectivas representações em formato raster e vetor	48
Figura 19 – Ortomosaico final recortado em bloco	49
Figura 20 – Etapas do processamento digital de imagens.	51
Figura 21 – Interpolação de CLAHE utilizando coordenadas relativas (x', y')	56
Figura 22 – Espaço de cor RGB representado como um cubo	57
Figura 23 – Representações geométricas de espaços de cores (a) HSI (b) HSL e (c) HSV	59
Figura 24 – Representação geométrica do espaço de cor CIELa*b*	60
Figura 25 – Processo de filtragem espacial por convolução.	62
Figura 26 – Exemplos de filtragem da média para suavização	63
Figura 27 – Exemplo de um filtro Gaussiano	64
Figura 28 – Diagrama funcional do filtro da mediana	64
Figura 29 – Exemplo de objeto e elemento estruturante para PDI	65
Figura 30 – Exemplo de operações morfológicas de erosão e dilatação	66

Figura 31 – Exemplo de resultado da aplicação da transformada de distância	70
Figura 32 – Geometria do template matching	72
Figura 33 – Máscara de chamfer com distância octogonal	75
Figura 34 – Exemplo de representação de região por código de cadeia	76
Figura 35 – Geração de um cone a partir de um círculo	79
Figura 36 – Elementos de um cone circular	79
Figura 37 – Seção meridiana	80
Figura 38 – Relação entre modelo de seções cónicas e estádios fenológicos de planta de milho	80
Figura 39 – Exemplos de hiperplanos de separação	82
Figura 40 – Vetores de suporte com hiperplano otimizado	84
Figura 41 – Interpretação geométrica das variáveis de folga	85
Figura 42 – Exemplo de mapeamento não linear com função kernel	86
Figura 43 – Diagrama de blocos para o conjunto de classificadores SVM.	88
Figura 44 – Definição da grade de amostragem e dos blocos da área experimental.	91
Figura 45 – Instruções recomendadas para calibração.	92
Figura 46 – Interface principal para configuração da câmera Micasense®.	93
Figura 47 – Processo de calibração com captura manual da imagem de referência.	94
Figura 48 – Modelo de estimativa para crescimento das plantas de milho	96
Figura 49 – Ortomosaicos gerais para bandas vermelho, verde e azul respectivamente.	97
Figura 50 – Ortomosaicos da área experimental das plantas de milho.	97
Figura 51 – Bloco 30 com suas respectivas bandas espectrais vermelho, verde, azul.	98
Figura 52 – Bloco número 30 multiespectral RGB utilizado para validação do método.	98
Figura 53 – Transformação geométrica em torno ao objeto	99
Figura 54 – Exemplo de ajustes para brilho, contraste e CLAHE com espaço de cor YCbCr	101
Figura 55 – Exemplo de ajustes para brilho, contraste e CLAHE com espaço de cor HSV	101
Figura 56 – Exemplo de ajustes para brilho, contraste e CLAHE com espaço de cor $L^*a^*b^*$	102
Figura 57 – Histograma do espaço de cor HSV	103
Figura 58 – Resultado da segmentação pelo espaço de cor HSV	103
Figura 59 – Resultado da transformada de distância	104
Figura 60 – Mapa de distâncias para máscara padrão de chamfer	105
Figura 61 – Picos localizados após o uso do algoritmo de template matching com coeficiente de correlação e máscara de chamfer	105
Figura 62 – Localizações pós-emergência V2 com descritores geométricos	106
Figura 63 – Histograma para bloco número 30	107
Figura 64 – Histograma espectral de raios	108
Figura 65 – Modelo de estimativa para contagem das plantas de milho	109
Figura 66 – Matriz de confusão do Classificador 1 com <i>kernel Sigmoide</i> e relação 50%-50%. 112	
Figura 67 – Curva ROC do Classificador 1 com <i>kernel Sigmoide</i> e relação 50% - 50%.	112

Figura 68 – Matriz de confusão do Classificador 2 com <i>kernel Sigmoide</i> e relação 50% - 50%.	112
Figura 69 – Curva ROC do Classificador 2 com <i>kernel Sigmoide</i> e relação 50% - 50%. . .	113
Figura 70 – Matriz de confusão do Classificador 1 com <i>kernel Sigmoide</i> e relação 70%-30%. 113	
Figura 71 – Curva ROC do Classificador 1 com <i>kernel Sigmoide</i> e relação 70% - 30%. . .	114
Figura 72 – Matriz de confusão do Classificador 2 com <i>kernel Sigmoide</i> e relação 70%-30%. 114	
Figura 73 – Curva ROC do Classificador 2 com <i>kernel Sigmoide</i> e relação 70% - 30%. . .	114
Figura 74 – Matriz de confusão do Classificador 1 com <i>kernel Sigmoide</i> e relação 80%-20%. 115	
Figura 75 – Curva ROC do Classificador 1 com <i>kernel Sigmoide</i> e relação 80% - 20%. . .	115
Figura 76 – Matriz de confusão do Classificador 2 com <i>kernel Sigmoide</i> e relação 80%-20%. 116	
Figura 77 – Curva ROC do Classificador 2 com <i>kernel Sigmoide</i> e relação 80% - 20%. . .	116
Figura 78 – Matriz de confusão do Classificador 1 com <i>kernel Gaussiano</i> e relação 50%- 50%.	117
Figura 79 – Curva ROC do Classificador 1 com <i>kernel Gaussiano</i> e relação 50% - 50%. . .	117
Figura 80 – Matriz de confusão do Classificador 2 com <i>kernel Gaussiano</i> e relação 50%- 50%.	117
Figura 81 – Curva ROC do Classificador 2 com <i>kernel Gaussiano</i> e relação 50% - 50%. . .	118
Figura 82 – Matriz de confusão do Classificador 1 com <i>kernel Gaussiano</i> e relação 70%- 30%.	118
Figura 83 – Curva ROC do Classificador 1 com <i>kernel Gaussiano</i> e relação 70% - 30%. . .	119
Figura 84 – Matriz de confusão do Classificador 2 com <i>kernel Gaussiano</i> e relação 70%- 30%.	119
Figura 85 – Curva ROC do Classificador 2 com <i>kernel Gaussiano</i> e relação 70% - 30%. . .	119
Figura 86 – Matriz de confusão do Classificador 1 com <i>kernel Gaussiano</i> e relação 80%- 20%.	120
Figura 87 – Curva ROC do Classificador 1 com <i>kernel Gaussiano</i> e relação 80% - 20%. . .	120
Figura 88 – Matriz de confusão do Classificador 2 com <i>kernel Gaussiano</i> e relação 80%- 20%.	121
Figura 89 – Curva ROC do Classificador 2 com <i>kernel Gaussiano</i> e relação 80% - 20%. . .	121

LISTA DE TABELAS

Tabela 1 – Estádios da planta de milho	32
Tabela 2 – Especificações do Drone DJI Matrice® 100.	37
Tabela 3 – Bandas espectrais do sensor MicaSense® RedEdge-M	38
Tabela 4 – Parâmetros do sensor MicaSense® RedEdge-M	39
Tabela 5 – Imagens Multiespectrais.	50
Tabela 6 – Tipos de Kernels	87
Tabela 7 – Parâmetros identificados e utilizados para fase pré-voo	91
Tabela 8 – Dados coletados manualmente no campo para Voo 02	94
Tabela 9 – Dados coletados manualmente no campo para Voo 05	95
Tabela 10 – Dados coletados manualmente no campo para Voo 07	95
Tabela 11 – Médias e desvios padrão dos dados coletados	95
Tabela 12 – Principais parâmetros para a transformação geométrica.	100
Tabela 13 – Parâmetros utilizados para ajustes de brilho, contraste e CLAHE	100
Tabela 14 – Parâmetros mínimos e máximos para plantas de milho no espaço de cor HSV	103
Tabela 15 – Quantidade de sementes emergidas para cada um dos sítios específicos (blo- cos) na cultura de milho.	110
Tabela 16 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com <i>kernel Sigmoide</i> utilizando a relação de 50%-50%.	111
Tabela 17 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com <i>kernel Sigmoide</i> utilizando a relação de 70%-30%.	113
Tabela 18 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com <i>kernel Sigmoide</i> utilizando a relação de 80%-20%.	115
Tabela 19 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com <i>kernel Gaussiano</i> utilizando a relação de 50%-50%.	116
Tabela 20 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com <i>kernel Gaussiano</i> utilizando a relação de 70%-30%.	118
Tabela 21 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com <i>kernel Gaussiano</i> utilizando a relação de 80%-20%.	120

SUMÁRIO

CAPÍTULO 1–INTRODUÇÃO	15
1.1 Motivação	15
1.2 Revisão bibliográfica	17
1.2.1 Aquisição de imagens por Drone	18
1.2.2 Método da transformada de distância	19
1.2.3 Segmentação de imagens	21
1.2.4 Método de template matching	22
1.2.5 Extração de características	24
1.2.6 Máquina de vetores de suporte	26
1.3 Objetivo geral e objetivos específicos	29
1.3.1 Objetivos específicos	29
1.4 Contribuições do trabalho	29
1.5 Organização do trabalho	29
CAPÍTULO 2–MATERIAIS E MÉTODOS	30
2.1 Área experimental	30
2.1.1 Densidade de plantio	31
2.1.2 Fenologia e fenotipagem do milho	32
2.2 Abordagem geral do modelo e principais materiais	33
2.3 Agricultura de precisão	35
2.3.1 Sensoriamento Remoto	36
2.4 Aquisição e organização das imagens	37
2.4.1 Sensor multiespectral	37
2.4.2 Geometria de imagens aéreas	39
2.4.3 Sobreposição de imagens aéreas	41
2.4.4 Pontos de controle em campo	43
2.5 Caracterização das imagens multiespectrais	44
2.5.1 Ortorretificação e ortomosaicos	45
2.5.2 Representação de imagens aéreas	48
2.5.3 Formato de armazenamento de imagens aéreas	49
2.6 Processamento Digital de Imagens	50
2.6.1 Pré-processamento das imagens da cultura de milho	50
2.6.2 Ajustes de brilho e contraste	54
2.6.3 Ajuste de histograma	55
2.6.4 Espaços de cores	56

2.6.4.1	Espaço de cor RGB	57
2.6.4.2	Espaço de cor XYZ	57
2.6.4.3	Espaço de cor YCbCr	58
2.6.4.4	Espaço de cor HSV	58
2.6.4.5	Espaço de cor CIEL*a*b*	60
2.6.5	Filtragem no domínio espacial	62
2.6.5.1	Filtro da média	63
2.6.5.2	Filtro de Gaussiano	63
2.6.5.3	Filtro da mediana	63
2.7	Processamento morfológico	64
2.7.1	Dilatação e erosão	66
2.7.2	Abertura e fechamento	67
2.8	Segmentação de imagen	67
2.8.1	Binarização	67
2.8.2	Transformada de distância	69
2.9	Reconhecimento de padrões	70
2.9.1	Template matching	72
2.9.2	Template matching com correlação cruzada normalizada	73
2.9.3	Template matching com coeficiente de correlação	74
2.9.4	Template matching com algoritmo de chamfer	74
2.10	Descritores de objetos e extração de características	75
2.10.1	Código da cadeia	76
2.10.2	Características geométricas	77
2.11	Modelo baseado em secções conicas	79
2.12	Classificação de padrões por aprendizado de máquina	81
2.12.1	Máquina de vetores de suporte	82
2.12.2	SVM linear	82
2.12.3	SVM com margen suaves	84
2.12.4	SVM não linear	86
2.12.5	Conjunto de classificadores SVM	87

CAPÍTULO 3—RESULTADOS E DISCUSSÕES 90

3.1	Resultados em campo experimental	90
3.2	Resultados para aquisição das imagens	90
3.2.1	Parâmetros para fase pré-voo	91
3.2.2	Parâmetros para planejameto de voo	92
3.2.3	Calibração para câmera em cada voo	92
3.3	Calibração de medidas reais em campo	94
3.4	Resultados para caracterização de imagens	96
3.5	Qualificação do banco de imagem RGB V2	98

3.6	Resultados obtidos através da transformação geométrica	99
3.7	Resultados obtidos para ajustes de brilho e contraste	100
3.8	Resultados obtidos para segmentação de imagens	102
3.9	Resultados obtidos para transformada de distância	104
3.10	Resultados obtidos para a etapa de template matching	104
3.11	Resultados para a etapa de extração de características	106
3.12	Resultados obtidos através da etapa de treinamento dos classificadores SVM . . .	110
3.12.1	Avaliação da etapa de treinamento do conjunto de classificadores SVM . . .	121
CAPÍTULO 4–CONCLUSÕES		123
4.1	Principais conclusões	123
4.2	Sugestões para trabalhos futuros	124
REFERÊNCIAS		125

Capítulo 1

INTRODUÇÃO

Este Capítulo apresenta a motivação contextualizada com base em estudos realizados por diversos autores, bem como avanços recentes com uma revisão sistemática da literatura para cada uma das técnicas que foram utilizadas no seu desenvolvimento. Por fim, apresenta-se o objetivo geral, expondo seus objetivos específicos identificados.

1.1 Motivação

O aumento da população mundial traz desafios futuros para alcançar a segurança alimentar. Estudos prospectivos indicam que haverá necessidade de aumentar em 70% a produção agrícola para alimentar os habitantes do planeta até 2050 (POPP; LAKNER, 2013). Desta forma, relatórios recentes da Organização das Nações Unidas para a Alimentação (FAO, do inglês *Food and Agriculture Organization*), mencionam que a segurança alimentar enfrenta muitos desafios devido a fatores como mudanças climáticas, recessões econômicas e a recente pandemia causada pelo COVID19 (LABORDE et al., 2020; FAO et al., 2021).

Inseridos neste contexto, os grãos de milho, junto ao arroz e o trigo têm merecido especial atenção, uma vez que provém da ordem de 30% das calorias em relação à todos os demais alimentos, onde o milho é um dos ingredientes chave para a produção de rações de animais, produção de biocombustíveis e alimentação humana, o que tem levado à necessidade de se buscar maior competitividade para que os preços no mercado se apresentem mais estáveis (SHIFERAW et al., 2011). De acordo com os relatórios da Companhia Nacional de Abastecimento (CONAB), a produção global de milho é mantida firmemente em quatro países: Estados Unidos, China, Brasil e Argentina, dos quais o Brasil responde por dois terços da produção global.

Neste contexto, existe uma necessidade constante de melhorar a produção agrícola nas culturas, onde uma destas informações está relacionada com a quantidade e a arquitetura¹ das plantas, também chamado de densidade de plantios (LIU et al., 2017). Outras destas informações está relacionada com a altura da planta (BENDIG et al., 2015) e a fenologia baseada em campo

¹ Arquitetura de plantas: ramo da morfologia vegetal que estuda as partes das plantas.

que podem ser utilizadas como uma abordagem promissora para fornecer o rendimento necessário em termos de número de plantas e populações para uma descrição precisa das características nos sistemas de cultivo (WINTERHALTER et al., 2013). Desta forma, a determinação do número de plantas por are² e a estimação de altura representam um índice importante para avaliar a densidade de plantas, bem como a emergência em campos agrícolas, já uma análise do rendimento final pode ser melhor determinado incluindo o número exato de plantas (TESTA et al., 2016). Sendo assim que a quantificação dessas informações tem aplicações essenciais na estimativa do crescimento, planejamento e manutenção de campos agrícolas (KOC-SAN et al., 2018) fornecendo a possibilidade de analisar a qualidade do plantio, prevendo a produtividade e tomando decisões importantes durante o manejo dos ciclos de cultura (KITANO et al., 2019).

Métodos manuais de contagem de plantas utilizam sistematicamente uma área mensurável de acordo com sua população (POUND; CLEMENTS, 1898), assim como diferentes variações de contagem (OLIVER; TANSLEY, 1904; WEAVER, 1918) são consideradas como o ponto de partida das pesquisas atuais. Portanto, esses métodos tradicionais, aliados aos avanços tecnológicos na agricultura de precisão, estão utilizando o sensoriamento remoto como meio de aplicações não invasivas, seja por instrumentos ópticos (HASITUYA et al., 2020) ou satélites (AHMAD et al., 2018). Embora em menor escala, essas soluções apresentam limitações devido ao alto custo ou à presença de distúrbios climáticos como as nuvens (SU et al., 2019), porém estão sendo substituídas por sensoriamento remoto de baixa altitude, entre os quais estão incluídos os veículos aéreos não tripulados (UAV, do inglês *Unmanned Aerial Vehicle*) ou comumente conhecido como Drones (KESTUR et al., 2018).

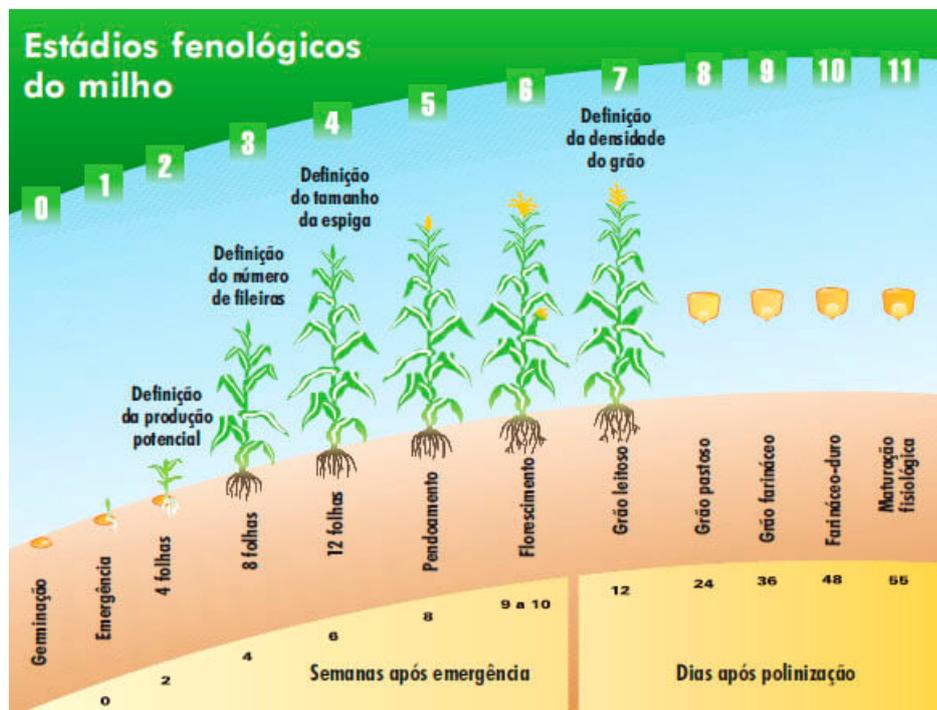
Dentro dos avanços dos Drones, estão relacionados aos sensores das câmeras dos quais apresentam configurações de diferentes bandas, sendo geradas imagens multiespectrais (MARI-NELLO, 2017) ou hiperespectrais (LU et al., 2020). Por sua vez, essas características permitem a coleta de dados em alta definição, embora isso também implica novos desafios quando são aplicadas à avaliação da densidade de plantas, especificamente na contagem de plantas. No entanto, se houver muitas áreas de plantação que podem conter mais de 1 000 plantios, os métodos de contagem manual com auxílio de *software* podem representar um processo demorado, caro e cansativo (SRESTASATHIERN; RAKWATIN, 2014).

Pesquisas atuais na agricultura de precisão utilizam sensores multiespectrais ou hiperespectrais embarcados em Drones. Neste contexto, há pesquisadores que usam uma variedade de modelos para estimar a densidade de plantas, sendo categorizados em três abordagens: identificação, contagem e estimativa dos estádios fenológicos (Figura 1). Esses modelos, por sua vez, são baseados em diferentes técnicas de processamento digital de imagens entre eles se têm: análise de correlação (CRUVINEL; MINATEL, 2002), espaços de cores (GNÄDINGER; SCHMIDHALTER, 2017), regiões de interesse (KOC-SAN et al., 2018), assinaturas espectrais (HALL et al., 2018), número de folhas (KUMAR; DOMNIC, 2019), índices de vegetação (BAH

² Are: unidade de medida para superfícies agrárias correspondendo à 100 m².

et al., 2020) entre outros. Por outro lado, a estimativa da altura de plantas ou árvores em fazendas usando Drones, estão baseados em técnicas como: nuvens de pontos tridimensionais a partir de sensores LIDAR (do inglês *Light Detection And Ranging*) (MALAMBO et al., 2018), modelos digitais de elevação (DEM, do inglês *Digital Elevation Model*) a partir de sensores multiespectrais (JOHANSEN et al., 2018), ângulo de elevação do sol com valores azimute (KOC-SAN et al., 2018), entre outros. Nesse sentido, pretende-se propor uma metodologia para contagem de plantas de milho em pós-emergência, bem como estimar sua altura com base em seções cônicas (VERAMENDI; CRUVINEL, 2021) utilizando imagens multiespectrais de campos agrícolas.

Figura 1 – Estádios fenológicos da cultura do milho.



A Figura 1 ilustra os estádios fenológicos do milho (*Zea Mays L.*), divididas por suas fases vegetativas (V) e reprodutivas (R) seguindo os padrões estabelecidos por Ritchie e Hanway (1989). Desta forma, para o desenvolvimento da metodologia proposta, a planta de milho em análise encontra-se nas semanas pós-emergência, especificamente na numeração 2 da Figura 1 sendo finalmente codificada como V2.

1.2 Revisão bibliográfica

Esta seção apresenta uma descrição dos principais trabalhos relacionados ao tema desta pesquisa, os quais envolvem a aquisição de imagens por Drone, métricas de avaliação da qualidade de imagens, segmentação de imagens, transformada de distância, template matching, extração de características e máquina de vetores de suporte. A revisão bibliográfica foi baseada no uso de palavras-chave nas seguintes bases bibliográficas: *IEEE Xplore*, *ACM Digital Library*,

Springer, ScienceDirect, Google Scholar, ArXiv e periódicos CAPES, resultando na seleção aproximada de 80 artigos pertinentes e de interesse para a pesquisa. Enquanto para o gerenciamento dos dados bibliográficos foi utilizada a ferramenta de código aberto *Zotero*.

1.2.1 Aquisição de imagens por Drone

Jin e colaboradores em 2017, desenvolveram e avaliaram um método para estimar a densidade de plantas de trigo no estágio de emergência com base em imagens de alta resolução obtidas por Drone em altitudes muito baixas com aplicação para fenotipagem de alto desempenho em condições naturais (JIN et al., 2017). Os autores obtiveram as imagens com uma câmera Sony ILCE® α 5100L de 24 megapíxeis embarcado em um hexacóptero Atechsys, sua distância focal foi de 60 mm mantendo a altitude de voo a uma altura razoável de 3 m, 5 m e 7 m. Esta configuração técnica, permitiu uma resolução de solo de 0,20 mm a 0,45 mm, fornecendo 59% - 77% de sobreposição entre as imagens. A câmera foi posicionada em um ângulo zenital de 45° em uma direção da bússola perpendicular à direção da linha para maximizar a visão da seção transversal das plantas e minimizar o efeito do vento criado pelos rotores. Os autores usaram o software Agisoft Photoscan® para derivar a posição das câmeras para cada imagem. Os autores projetaram as imagens na superfície do solo para finalmente extrair subamostras e estimar a densidade das plantas. Finalmente, os experimentos mostraram que, após uma análise de sensibilidade, a resolução do solo das imagens deve ser maior que 0,40 mm para facilitar a classificação dos píxeis verdes, embora o método seja bom quando as imagens são tiradas de geometrias observacionais semelhantes.

Fuentes-Peaolillo e colaboradores em 2018, compararam o Índice de Vegetação por Diferença Normalizada (NDVI, do inglês *Normalized Difference Vegetation Index*) de câmeras multiespectrais com quatro índices obtidos a partir de canais vermelho, verde e azul (RGB, do inglês *Red, Green e Blue*) para a identificação de solo e vegetação em imagens capturadas com um veículo aéreo não tripulado. Esta comparação foi realizada utilizando o NDVI como o *ground-truth*, obtendo-se duas classes de dados que posteriormente foram comparados com os demais índices, contando-se os píxeis correspondentes para cada classe. No caso dos índices RGB, a média foi definida como o centro dos dados e como o ponto de corte de ambas classes. Os resultados dos autores indicaram que é possível identificar os mesmos padrões espaciais utilizando os índices RGB, onde o Índice Triangular Verde (TGI, do inglês *Green Triangular Index*) apresentou melhor desempenho. Porém, apesar de a contagem de píxeis apresentaram resultados semelhantes, a inspeção visual dos resultados indicou que os índices RGB apresentam erros na identificação da vegetação, principalmente na área da linha. Isso indica que para delimitar com precisão as áreas correspondentes à vegetação ao solo é necessário o uso de técnicas de agrupamento mais complexas.

Em 2021, Pallottino e colaboradores propuseram o uso de Drones de baixo custo para aplicações de fenotipagem in situ na área de cereais (PALLOTTINO et al., 2021). Para a aplicação

de fenotipagem de campo, usaram um Drone DJI® SPARK de baixo custo, enquanto para reconstruções de ortoimagem usaram o software 3DF Zephyr®, e para calibração de cores usaram um comparador de cores GretagMacbeth. As plantas foram dispostas em parcelas de 10 m^2 , com densidade de semeadura de 350 sementes/m^2 para trigo duro e 450 sementes/m^2 para trigo mole, obtendo-se 90 parcelas de trigo duro e 108 de trigo mole. Para a aquisição das imagens, a câmera não intercambiável registrou a cada 2 segundos a uma velocidade do obturador a $1/2000\text{ s}$, com 15 m de altura, com 75% de sobreposição e resolução espacial de $0,52\text{ cm/pxel}$. Usando o software 3DF Zephyr, obtiveram o modelo digital de superfície (DSM, do inglês *Digital Surface Model*), o modelo digital de terreno (DTM, do inglês *Digital Terrain Model*) e a ortoimagem, as métricas de altura foram obtidas fazendo a diferença entre o DSM e o DTM. A calibração de cores foi o resultado da aplicação do método TPS (do inglês *Thin-plate Spline*) em um conjunto de pontos de controle que foram obtidos a partir das plotagens capturadas, a ideia principal dos autores era resolver um sistema linear de forma fechada, mapeando assim cada ponto de referência correspondente para os gráficos capturados (MENESATTI et al., 2012). Tendo a função de interpolação TPS, os autores finalmente obtiveram as coordenadas do ColorChecker RGB. Finalmente os autores avaliaram a metodologia com um campo experimental das parcelas de trigo duro e mole obtendo informações sobre a cor real (com precisão inferior de $12/256$), e para a altura de cada parcela apresentaram informações visuais representando variedades de ciclo médio-precoce com plantas de altura média.

1.2.2 Método da transformada de distância

Dorj e colaboradores em 2017, desenvolveram um algoritmo para detecção e contagem de frutas cítricas em ambientes naturais de árvores usando processamento digital de imagens (DORJ et al., 2017). Os autores tiveram como objetivo desenvolver e testar um sistema para estimar a produtividade de citros de uma única árvore antes da colheita. Os autores adquiriram imagens RGB com uma câmera Panasonic® DMC-ZS10, as imagens consistiram em frutas cítricas amarelas maduras de 21 árvores de diferentes tons de verde. Para que as frutas fossem separadas, os autores processaram as imagens no canal HSV, condicionando o plano H em uma faixa de 60 a 260 para separar a fruta cítrica do restante dos objetos. Em seguida os autores utilizaram a escala de cinza juntamente com um filtro mediano para redução de ruído, para posteriormente criar uma imagem binarizada. Os autores calcularam uma transformação de distância euclidiana para forçar os píxeis a pertencerem à fruta cítrica. Adicionalmente calcularam os marcadores para o objeto usando o algoritmo de *Watershed*, criando componentes conectados da fruta cítrica. Por fim, os autores contaram cada componente conectado, comparando-o com dois trabalhos anteriores de estimativa de desempenho. Os resultados encontrados pelos autores indicaram um erro menor que 0% e maior que 29,02%, assim como a análise de regressão foi de $R^2 = 0,93$.

Em 2018 Nguyen e colaboradores desenvolveram um sistema de avaliação automática para o cálculo da taxa de germinação de sementes de arroz (NGUYEN et al., 2018). O sistema

contém uma etapa de segmentação semântica baseada em uma arquitetura *U-Net*, seguida do cálculo da transformada de distância juntamente com um limiar e finalmente uma etapa de classificação baseada em um *ResNet2* para duas classes: sementes germinadas e não germinadas. Para a aquisição das imagens, os autores utilizaram uma câmera Canon® DSLR sobre um fundo de papel utilizado para contar as sementes, coletando assim as imagens em condições naturais e base rugosa não uniforme. O conjunto de dados incluiu 8 variedades de sementes de arroz com intervalo de 6 a 10 dias, gerando assim 1276 imagens. Para os experimentos, os autores geraram um mapa segmentado feito por uma *U-Net*, com o objetivo de calcular as distâncias euclidianas pela transformada de distância juntamente com um limiar para separar regiões das sementes, finalmente os autores compararam o classificador da arquitetura *ResNet* com: (1) *Harr-Adaboost* e (2) *Gist-Random Forest*, obtendo resultados de F1 score com 95,38% e 95,66% respectivamente superando os métodos 1 e 2.

Para a detecção automática de pepinos cultivados em campos agrícolas, Fernandez e colaboradores em 2018, projetaram um algoritmo baseado em processamento de imagens e aprendizado de máquina (FERNANDEZ et al., 2018). Onde uma transformação de distância euclidiana foi usada principalmente para separar manchas correspondentes às folhas e flores jovens do campo agrícola, destacando assim os pepinos que deveriam ser detectados. Para isso, os autores organizaram os algoritmos seguindo a seguinte sequência de etapas: aquisição de imagem RGB, transformações de espaço de cores, classificador de mapa baseado em píxeis, operações morfológicas para gerar máscaras de ponto, cálculo da transformada de distância euclidiana para separar sobreposições, recorte de objeto baseado em *Convex Hull*, classificador baseado em *Bag of Visual Words* (BoW), segmentação pelo algoritmo de *Watershed* e cálculo de centróides orientados. Os autores adquiriram imagens com uma câmera ®Prosilica GC2450 localizada em um robô trator a 700 mm do solo, gerando 105 imagens, onde 15 imagens foram rotuladas manualmente como: pepinos, folhas e fundo. Os autores separaram os píxeis pelos canais azul e verde do RGB, além de quantificar o canal Hue do HSV. Adicionalmente realizaram operações morfológicas para remover píxeis de fundo e pequenos buracos nos pepinos. Em seguida, aplicaram a transformada de distância euclidiana para possível identificação de píxeis pertencentes ao objeto. Seguido por um limiar global de Otsu, gerando assim marcadores do objeto. Os autores descreveram o objeto usando *Convex Hull* preservando assim os marcadores, já para quantificar as características, usaram o descritor *SURF*, com o qual obtiveram 5161 características para cada classe. Para o classificador BoW, os autores usaram 500 palavras no vocabulário do modelo, além de características reduzidas usando K-means, verificaram as sobreposições finais por segmentação *Watershed* e finalmente representaram os centróides dos pepinos. Os autores validaram o algoritmo, com uma alta taxa de detecção de 91,79%, bem como uma taxa de falsos positivos de 2,56%. Por fim, a precisão do algoritmo atingiu 85,65% e um recall de 90,10%.

Halik e colaboradores em 2018, desenvolveram um novo método para detectar tomates maduros plantados em estufas com mudanças complexas de fundo e iluminação, bem como a

separação de tomates sobrepostos (MALIK et al., 2018). Os autores adquiriram imagens com uma câmera digital Canon PowerShot® SD880IS, coletando assim 53 fotos de tomates vermelhos e não vermelhos. Para a segmentação os autores encontraram melhores resultados no espaço de cores HSV, onde os parâmetros matiz (*Hue*) estavam próximos de 0,07, a saturação (*Saturation*) estavam próximos de 0,35 e o valor (*Value*) estava próximos de 0,4, todos esses parâmetros foram usados juntos como limiar no método Otsu. Os autores separaram os tomates sobrepostos com o algoritmo Watershed, embora os autores tenham detectado problemas menores, como a segmentação excessiva e para uma separação correta, aplicaram um filtro Sobel, seguido de operações morfológicas com abertura e fechamento. Por fim, para criar uma diferença mais acentuada, utilizaram a transformada de distância, que foi posteriormente sobreposta a cada píxel da imagem segmentada. Os resultados finais foram 81,6% detectados corretamente e 18,3% detectados incorretamente.

1.2.3 Segmentação de imagens

Hernández e colaboradores em 2017, descreveram a criação de um aplicativo para dispositivos móveis denominado pCAPS (do inglês *portable Classification Application for Plants and Soil*) que integra múltiplas técnicas de visão computacional para a segmentação e análise de plantas em imagens de culturas (HERNÁNDEZ-HERNÁNDEZ et al., 2017). A ferramenta desenvolvida pelos autores permitiu o monitoramento das safras agrícolas em tempo real, fornecendo informações que podem ser utilizadas para automação e otimização do cálculo das necessidades hídricas. Os três módulos principais do pCAPS foram: aquisição e recorte, análise de imagens e por último o registo histórico. Primeiro, propuseram um algoritmo que detecta marcadores retangulares com base em projeções integrais que são colocadas em um campo agrícola, as imagens foram recortadas para obter uma análise uniforme por longos períodos de tempo. Em seguida, aplicaram a segmentação de cores usando uma abordagem probabilística baseada em histogramas sobre um espaço de cores ideal. Finalmente, realizaram um processo de contagem de objetos com base em componentes conectados em imagens binarizadas. Os experimentos avaliaram a eficiência e precisão em um conjunto de 140 imagens de alta resolução de imagens coloridas, onde 24 imagens eram de couve-rábano (*Brassica oleracea L. var. Gongylodes*) e 116 imagens de 4 diferentes parcelas de cultivo eram de alface (*Lactuca sativa L. var. Ferro*). A precisão pelo algoritmo de detecção do retângulo ficou abaixo de 0,7 cm, com erro máximo de 1,64 cm, as áreas de corte obtiveram em média 98,3% de intersecção com as áreas corretas.

Para melhorar a precisão da sementação em sobreposições de folhas, Wang e colaboradores apresentaram um método de segmentação baseado no modelo de Chan-Vese e no operador Sobel (WANG et al., 2018). O método consistiu em três etapas. Primeiro, os autores usaram uma característica que identifica tons com níveis relativamente altos de verde para extrair a região das folhas e eliminar o fundo. Em segundo lugar, implementaram o modelo de Chan-Vese e o opera-

dor Sobel aprimorado para extrair os contornos das folhas e detectar as bordas, respectivamente. Terceiro, extraíram uma folha objetivo com um fundo complexo e sobreposto, combinando os resultados obtidos pelo modelo de Chan-Vese e o operador Sobel. Para verificar a eficácia do algoritmo proposto, os autores realizaram experimentos de segmentação com 30 imagens de folhas de pepino. A taxa média de erro do método proposto foi de 0,0428, o que representa um decréscimo de 6,54% em relação à taxa média de erro do método de nível estabelecido.

Khan e colaboradores propuseram o uso da álgebra geométrica de Clifford para o aprimoramento de imagens segmentadas adquiridas desde Drones em diferentes campos agrícolas (KHAN et al., 2019). Os autores identificaram que a álgebra geométrica pode superar as limitações relacionadas ao espaço RGB, levando a um processamento de imagem mais genuíno em espaços de cores, com base em um espaço de cores CIELab. Especificamente, os autores usaram a subálgebra chamada quatérnios de Clifford para a segmentação de primeiro e segundo plano, na forma de uma máscara de quatérnio 3×3 a partir de uma conversão RGB. Os resultados obtidos foram imagens segmentadas sintonizadas com diferentes fatores de escala.

Riehle e colaboradores desenvolveram um algoritmo que executava segmentação automática robusta do fundo e das plantas sob várias condições de captura, como superexposição ou subexposição, juntamente com diferentes cores de culturas e fundos (RIEHLE et al., 2020). O algoritmo foi baseado em um método de índices de vegetação para pré-segmentação aproximada e usou esta primeira aproximação para calcular o valor limiar para uma segmentação usando um modelo de espaço de cores. O desempenho do algoritmo foi avaliado com 200 imagens de 4 câmeras com diferentes configurações. Além disso, a qualidade do algoritmo foi comparada com os métodos de segmentação baseados em índices de vegetação, onde todos os índices foram combinados usando o limiar de Otsu. Os resultados mostraram um desempenho mais forte e confiável do que outros métodos de segmentação baseados em índices de vegetação. As investigações mostraram excelentes resultados de algoritmo usando espaço de cores CIELab com precisão de segmentação semântica de planta de até 97,4%.

1.2.4 Método de template matching

Kalantar e colaboradores e 2017, propuseram um algoritmo de contagem de palmeiras de óleo usando template matching e correlação cruzada usando imagens Drone. Os autores adicionaram a análise de imagem baseada em objeto (AIBO, do inglês *Object-based Image Analysis*) ao *template matching* (TM) para superar a variabilidade geométrica e escala (KALANTAR et al., 2017). Para realizar o TM, calcularam a correlação cruzada, comparando sua distância euclidiana, onde o template consistia nas copas das árvores. Para AIBO, consideraram informações espectrais de píxeis semelhantes com base em: cor, tamanho, textura, forma e informações textuais, gerando uma área segmentada. Os experimentos para TM foram utilizados coeficiente de correlação selecionando um limiar ótimo de 0,7, para AIBO os parâmetros de segmentação foram: 85, 0,4 e 0,5 para escala, forma e compacidade, respectivamente. Para

avaliação, usaram o classificador máquina de vetores de suporte (SVM, do inglês *Support Vector Machine*) com 10 imagens de árvore e 10 de fundos. Compararam o TM com o TM+AIBO e um ground truth (509 árvores), usando métricas de: precisão, recall e F-measure. Os autores concluíram que obtiveram uma redução do erro de estimativa em torno de 27% onde o número de árvores contadas foi: 790 para TM e 582 para TM+AIBO.

Xie e colaboradores em 2019, propuseram um novo algoritmo de Correlação de Fase (CF) estendida com base na Filtragem Log-Gabor (CFFLG) para imageamento de imagens com diferenças radiométricas não lineares e diferenças geométricas em grande escala entre pares de imagens (XIE et al., 2019). De esta forma os autores construíram um atlas multiescala supercompleto da imagem original com base nas magnitudes filtradas obtidas usando filtros Log-Gabor com diferentes frequências centrais. O algoritmo CFFLG eliminou a interface de diferenças radiométricas não lineares significativas sobrepondo espectros estruturais geométricos multiescala, de modo que a translação pode ser bem determinada. Os experimentos foram testados com imagens sintéticas e sensoriamento remoto multimodal de áreas urbanas. Os autores compararam com outros 4 métodos no estado da arte, cujas métricas foram: escala, rotação, translação, RMSE e erro máximo. Os autores encontraram certas limitações quando os modelos que utilizam transformações complexas. Finalmente, demonstraram a racionalidade e eficácia do CFFLG.

Usando uma estratégia de programação dinâmica, um algoritmo de template matching rápido baseado em correlação cruzada normalizada (NCN do inglês *Normalized Cross Correlation*) foram propostos por Cui et al. (2020), para superar a complexidade computacional e a velocidade de cálculo em algoritmos de *matching*. Os autores dividiram o método em: (1) cálculo da matriz de soma cumulativa e soma quadrática usando programação dinâmica, (2) cálculo do grau de similaridade de correspondência de imagem com base em medidas de correlação, (3) otimização para reduzir o número de janelas deslizantes e o processo de amostragem. Para os experimentos foram utilizadas imagens de caracteres com fontes de 40 pontos (E1), e imagens de cenas de 1150x748 píxeis (E2). Avaliaram a precisão e o tempo de matching, juntamente com um método do estado da arte. Os resultados para a acurácia em E1 foram: 0,84, 0,50, 0,54; do tempo foram: 62 ms, 65 ms, 64 ms em três tipos de fontes, respectivamente. Para a eficiência do casamento E2 os tempos foram: 6,002 s, 6,123 s, 6,330 s na imagem original, transformados e filtrados respectivamente. Os autores concluíram que alcançaram alta precisão e eficiência no tempo.

Um método que aborda a NCN, o TM e espaço de cores CIELab foram desenvolvidos por García-Martínez et al. (2020) para a contagem de plantas de milho em imagens RGB capturadas por Drones. As imagens foram obtidas aos 23, 44 e 65 dias após o plantio, sobre 30 metros de altura, em uma área de 2500 m² e com sensores RGB de 12, 16 e 20 megapíxeis. Os autores descreveram que, para a contagem de plantas e sua distribuição na lavoura, era necessário que as plantas estivessem livres de ervas daninhas. Para a contagem foram utilizados

os ortomosaicos, onde um NCN de 4, 8 e 12 templates de plantas foram selecionados de acordo com seu diâmetro, sendo processados no canal a^* de CIELab que permitiram a segmentação da planta. Para validação automática, foi comparada com a contagem manual, onde a precisão média de 99% foi obtida para um píxel de tamanho 0,49 cm, seu erro médio foi de 2,2%, o coeficiente de determinação foi de 0,90 em 44 dias após o plantio. Os autores encontraram boas precisões nos primeiros fenótipos de crescimento, quando as plantas não se sobrepunham e não havia ervas daninhas.

1.2.5 Extração de características

Lottes e colaboradores propuseram um sistema baseado em visão que realiza detecção de vegetação, extração de características de plantas sob medida e classificação para obter uma estimativa da distribuição de culturas e ervas daninhas no campo agrícola coletadas por um Drone (LOTES et al., 2017). Os autores descreveram as principais etapas do sistema: (1) aplicaram pré-processamento para normalizar os valores de intensidade em uma escala global e detectar a vegetação em cada imagem, (2) extraíram características apenas para as regiões, que correspondem à vegetação, aproveitando a combinação de uma abordagem baseada em objeto e uma abordagem baseada em pontos de interesse e (3) aplicaram a classificação com *Random Forest* para várias classes e obteve uma distribuição de probabilidade para os rótulos de classe previstos. As imagens foram coletadas em diferentes conjuntos de dados: (a) 97 imagens a uma altura de 2-3 m, (b) 31 imagens a uma altura de 15 m e (c) 15 imagens a uma altura de 3m, todas as imagens tinham beterraba-açucareira em colheitas e ervas daninhas. Para a detecção da vegetação, removeram o fundo usando os canais RGB+NIR, explorando as características espectrais das plantas no canal NIR+RED através do NDVI. No caso do RGB, foram baseados no índice ExG para calcular o mascaramento com base a um limiar. A partir de uma máscara de vegetação, buscaram objetos pertencentes ao segmento vegetal, e para os pontos de interesse definiram um vetor de características a partir de uma grade densa com blocos de 20x20 píxeis, juntos essas duas características reuniram informações estatísticas: (i) valores de intensidade, (ii) representações de gradiente, (iii) espaços de cores e (iv) informações de textura. Além disso, consideraram características geométricas para a separação ótima entre as linhas das culturas, usando assim um Transformada de Hough (HT, do inglês *Hough Transform*) linear com a ideia de buscar linhas paralelas máximas com suas respectivas penalidades. Finalmente, essas características juntas foram usadas com o classificador *Random Forest* e os resultados foram avaliados com uma área sob a curva (AUC) e gráficos de precisão-recall, para a beterraba obtiveram um recall de 78% com uma precisão de 90%, um recall de 95% para a erva-sal e 87% para a camomila, e em geral o recall para outras ervas daninhas resultou em 45%.

Lu e colaboradores desenvolveram um algoritmo para detectar e contar frutas cítricas imaturas em imagens de copas de árvores (LU et al., 2018). Para a aquisição das imagens, os autores obtiveram um total de 70 imagens coloridas com uma câmera digital Canon® EOS M,

a uma distância de cerca de 1,5 - 2 m em diferentes condições de luz natural e algumas das aquisições foram feitas à noite. A resolução original era de 5184x3456 píxeis que foi interpolado bilinearmente para 960x540 píxeis. Os autores utilizaram características de forma e textura do componente verde, após aplicação de filtro de media e Wiener para remoção de ruído, assim encontraram os máximos locais da intensidade das imagens, formando os componentes conectados, onde o centro de cada máximo local representou os pontos de interesse. Posteriormente, os autores usaram o descritor visual de padrões binários locais (LBP do inglês *Local Binary Patterns*) normalizado depois que os pontos de interesse foram exibidos em blocos 16x16, essas características foram consideradas como entrada de um classificador *ensemble RUSBoost*. As previsões positivas foram consideradas como candidatas, mas muitos pontos e fundos foram erroneamente classificados como frutas, assim os autores propuseram um novo método de análise chamado análise de contorno hierárquico que, juntamente com a Transformada Hough Circular (THC do inglês *Transform Hough Circle*) previram frutas através de um raio predeterminado. O algoritmo foi avaliado com conjuntos de teste de 25 imagens, alcançando uma taxa de verdadeiro positivo de 80,4%, uma taxa de precisão de 82,3% e uma métrica F1 de 81,3%.

Em 2019 Su e colaboradores apresentaram novos algoritmos de processamento de imagem para o reconhecimento automático e contagem individual de uma única cápsula de algodão em condições de iluminação natural no campo (SUN et al., 2019). A câmera utilizada pelos autores foi um Panasonic Lumix® G6 montado em um robô móvel, a altura da câmera foi de 1,3 m com um ângulo de inclinação para baixo de 45° e a distância das cápsulas de algodão foi de 70-200 cm, três dias antes da colheita. Os autores utilizaram as características espaciais e de cor quantitativamente pelos espaços de cores RGB e HSV, desenvolvendo assim um algoritmo de duplo limiar com crescimento da região. Para estimar o número de cápsulas de algodão, usaram algoritmos baseados em características geométricas: (1) características de linha detectadas pela TH, (2) a distância limite mínima entre duas regiões para mesclar regiões não conectadas divididas por ramos e rebarbas, (3) a razão da área e do alongamento entre os eixos principais e secundários foram usados para separar cápsulas sobrepostas em clusters. Para validação, os autores utilizaram um total de 210 imagens, onde o score F1 ficou em torno de 0,98; enquanto a melhor precisão para a contagem de cápsulas foi em torno de 84,6%. Finalmente, para analisar a relação entre o número de cápsulas e o rendimento no nível das tramas, as imagens de cada trama foram unidas usando a ferramenta *Auto Blend* no Adobe® Photoshop CC 2017, então os autores usaram quinze tramas para construir uma regressão linear entre o número estimado de cápsulas e o rendimento total com um valor de R^2 de 0,53. O rendimento foi avaliado por outras dez tramas com MAPE de 8,92% e RMSE de 99 g.

Owen e colaboradores apresentaram um método para classificar a cobertura do solo exposta, vegetação palha/estepe e vegetação verde, todos baseados em descritores de cores, textura e gradientes orientados do espectro RGB (OWEN et al., 2020). Para a aquisição das imagens, utilizaram câmera digital Canon Nikon® D7100 em campo e casa de vegetação com 5 tipos diferentes de solo com palha e sem palha, essas imagens originalmente tinham o

tamanho de 4000x6000 píxeis, após um processo de correção foi normalizado aproximadamente para 2000x3500 píxeis, onde o número total de imagens foram de 123. Para o processo de segmentação e extração de características, primeiro classificaram manualmente um método automatizado *SamplePoint* padrão para determinar a cobertura do solo com base em um método de grade. O método gerou blocos de 12x12 para finalmente obter 144 blocos por imagem. Para as características de cor verde, usaram índices de contraste da vegetação ExG e excesso de vermelho ExR, já para as características não coloridas usaram Histograma de Gradientes Orientados (HOG do inglês *Histogram of Oriented Gradients*) e elementos de textura como a entropia. Finalmente, para a classificação, 50 imagens foram usadas para treinamento e 73 para teste em um grupo de algoritmos de aprendizado de máquina. As classificações obtidas foram comparadas com a classificação manual baseada no *SamplePoint* cujos valores R^2 foram de 0,86, 0,87 e 0,96 correlacionando bem com a cobertura de classificação automática, sendo o mais notável o algoritmo da árvore de decisão.

1.2.6 Máquina de vetores de suporte

Zhou et al. (2018) propuseram um novo algoritmo para contagem de espigas de trigo em campo. Aquisiram imagens RGB sistematicamente com iluminação variada, usando câmera em 45° acima da horizontal com um veículo terrestre tripulado. Usaram um método de agrupamento baseado em superpixel para gerar um patch a partir de imagens selecionadas, rotulando manualmente cada patch de 20×20 em duas categorias: espigas e fundo, desta forma 700 imagens foram usadas para treinamento e 300 imagens para teste. Para as características, usaram coocorrência de matrizes em níveis de cinza (GLCM) e um descritor de borda de histograma, essas características foram juntadas com kernel de análises de componentes principais (PCA). Um modelo twin-SVM segmentation (TWSVM-Seg) foi treinado para diferenciar os dois tipos de patches, obtendo uma classificação correta das amostras testadas. Para contagem, os autores usaram algoritmo baseado em imagens binárias de píxeis conectados contra contagem manual. Enquanto, a validação para algoritmo de segmentação indicadores Qsef, Sr, SSIM, precisão e revocação. Os resultados se correlacionam fortemente com uma precisão de 0,79 – 0,82 em um tempo médio de 0,1 s.

Hong et al. (2019) desenvolveram um novo framework para SVMs esparsos atingindo problemas a grande escala ou seja treinamento para um grande número de amostras e características em ultra-altas dimensões. O framework está baseado em técnicas de *screening* que procuram estimativas precisas do ótimo primal e dual, conseguindo identificar simultaneamente características e amostras inativas nas primeiras fases de treinamento, levando desta forma a uma redução substancial no custo computacional sem sacrificar a precisão. Além disso, o framework consegue atingir problemas multiclasse. Os autores começaram estudando a técnica emergente chamada de *screening* (GHAOUI et al., 2010) para a aceleração no processo de treinamento, mas sem resolver nenhum problema de otimização. Após estudar diversos autores

que tentam resolver o problema de otimização, identificaram que o estado da arte conseguia resolver o treinamento bem reduzindo o número de amostras ou de características mas não simultaneamente. Com tudo isso os autores estudaram o método de [Shibagaki et al. \(2016\)](#) que resolvia otimização utilizando amostras e características simultaneamente da forma dinâmica (regras *screening* durante o treinamento), mas o método dependia fortemente da lacuna da dualidade. Nesse contexto os autores contribuíram com um novo método *screening* chamado de *Simultaneously identify Inactive Features and Samples* (SIFS) com abordagem estática (regras *screening* antes de treinamento), além disso demonstraram limitações da abordagem dinâmica com uma série de lemas. Enquanto aos experimentos, os autores utilizaram conjuntos de dados sintéticos e reais com métricas: aceleração, proporção de escala e proporção de rejeição. A avaliação foi comparada com os métodos do estado da arte: *Accelerated Proximal Stochastic Dual Coordinate Ascent* (AProx-SDCA) ([SHALEV-SHWARTZ; ZHANG, 2014](#)) para classes binárias e [Shibagaki et al. \(2016\)](#) para multiclasse. O número de pares de parâmetros (α e β) encontrados no problema de dualidade foram de 1000 conseguidos com auxílio da Eigen library do C++. Porém, avaliaram o SIFS em 3 conjuntos de dados sintéticos (syn1, syn2 e syn3) onde a proporção de escala conseguiu aproximadamente o 99,9% e a aceleração atingida foi de até 76,8. No caso para os dados de syn2 precisou de menos de três minutos para ser solucionado. Para experimentos reais cinco conjuntos de dados (real-sim, rcv1-train, rcv1-test, url, e kddb) do projeto LibSVM ([CHANG; LIN, 2011](#)) foram utilizadas, onde o conjunto de dado kddb tem 20 milhões de amostras com 30 milhões de características. Assim a comparação de SIFS ao lado do método [Shibagaki et al. \(2016\)](#), conseguiu identificar aproximadamente o 98% das características inativas, o que levou a ganhar velocidades de até 300 vezes nos dados: real-sim, rcv1-train and rcv1-test. No caso de conjunto do dado kddb o método resolveu problemas de dualidade até 13 horas, já o método de Shibagaki levou 11 dias para realizar a mesma tarefa. Experimentos para SVMs esparsos multiclasse usou tres conjunto de dados sintéticos (syn-multi1, syn-multi2, e syn-multi3) e reais (news20 e rcv1-multiclass), no caso foram identificadas aproximadamente 95% de amostras inativas e 99% das características inativas o que levou velocidades de até 200 vezes no dado news20 e 300 vezes no dado syn-multi3 resolvendo assim os problemas em meia hora. Finalmente concluíram que o framework conseguiu reduzir dados identificando características e amostras em SVMs esparsos, contribuindo na precisão de estimação do ótimo primal e dual baseado na forte convexidade.

[Chen et al. \(2019\)](#) propuseram um método de segmentação baseado em visão monocular e algoritmo SVM para segmentar árvores cítricas sob diferentes luminosidade e condições de plantas daninhas. Para aquisição de imagens RGB utilizaram uma câmara embarcada em Drone com altura de 10 – 15 m, constituindo um conjunto de dados em seis categorias: condições de brilho, cobertura do solo, brilho insuficiente, brilho suficiente, faixa pequena de solo, faixa media do solo e faixa grande do solo. Usaram a equalização de histograma por região para compensar a iluminação, aumentando o contraste do brilho do primeiro plano sem alterar seu matiz e saturação. Para segmentação, combinaram aberração cromática com método Otsu, e

finalmente para extrair características de texturas e cores foi estabelecido uma segmentação por SVM para as árvores cítricas. Desta forma, os autores obtiveram 14 características de cores, 5 texturas estatísticas e padrões binários locais. A avaliação das árvores segmentadas incluíram acurácia com IoU, o qual representa corretamente a área segmentada. Para as regiões foram calculadas pelo SVM, verificadas em conjuntos de dados de diferentes brilhos e ervas daninhas, portanto a acurácia alcançada foi de $85,27\% \pm 9,43\%$.

Pereira et al. (2019), descreveram um algoritmo para analisar e processar imagens, a fim de reconhecer pêssegos, calcular as dimensões como volume e pesos em condições naturais das árvores em campo. As imagens foram obtidas por uma câmera de alta resolução Nikon® D3200, por sua vez essas as árvores tinham muita folhagem, proporcionando altos níveis de oclusão e sombras. Os autores utilizaram, segmentação por cores baseados nas cores quentes do pêssego, avaliando em espaços de cores HSV, YCbCr e $L^*a^*b^*$. Métricas de arredondamento, foram utilizadas como características para treinar o SVM, onde os contornos foram encontradas com auxílio de operador Sobel, além disso, outra característica como a área foi utilizada para treinamento do SVM. Essas duas métricas foram normalizadas para garantir bom desempenho nos treinamento e teste. Em relação às funções de kernels, os autores utilizaram função polinomial de segundo grau. Para predição da produção a característica de tamanho foi calibrada por um fator de peso e diâmetro, onde o pêssego com 400 g e 9 cm , na conversão de píxeis correspondentes a dimensões dos objetos foram de $29,3$ tonelada por hectare, em uma área cultivar de $1\,237,5\text{ m}$. Finalmente, em geral, a precisão alcançada pelo modelo SVM foi de 72% para pomares de 808 árvores/ha.

Wang et al. (2019) propuseram um método para detecção e enumeração de palmeiras de óleo individuais usando Drones. Os autores coletaram imagens RGB com 5 cm para píxeis da distância da amostragem do solo, com árvores maduros e árvores jovens, além de três espécies misturadas nas imagens coletadas. O método dos autores, prosseguiu uma abordagem hierárquica, primeiramente classificando as imagens como vegetação e não vegetação utilizando classificador supervisionado SVM. Em sequência, as árvores de palmeira de óleo são detectadas em áreas de vegetação em níveis de imagem patch. Para a primeira abordagem, autores selecionaram manualmente amostras positivas (500 imagens) e negativas ($1\,000$ imagens). Desta forma o método de HOG extraiu características de todas as amostras com blocos de 16×16 píxeis e deslocamentos de 8 píxeis. Assim, essas características foram constituídas em elementos de vetor com dimensão de 26×1 os quais foram utilizadas no SVM. Além disso, a detecção foi avaliada em vários tamanhos, garantindo mudanças em escala com parâmetro de regularização para margens suaves de SVM em $2,5$ e parâmetro gama de $1,0$. As detecções finais foram sobrescritas com círculos, descrevendo centro e diâmetro da copa dos árvores. A avaliação foi conduzida com resultados manuais, detectando $2\,590$ arvores de palmeiras de óleo, com uma acurácia geral da ordem de $99,21\%$. Além disso, os autores observaram situações de cenários complexos existentes entre as fronteiras e a sobreposição das árvores.

1.3 Objetivo geral e objetivos específicos

O objetivo geral deste trabalho consiste no desenvolvimento de um método para a contagem de plantas de milho em emergência sobre áreas de cultura usando técnicas de processamento de imagens multiespectrais e aprendizado de máquina.

1.3.1 Objetivos específicos

Para atingir o objetivo principal, os seguintes objetivos específicos foram definidos:

- Prover e avaliar técnicas de processamento de imagens digitais que envolvam: transformações geométricas, melhora de contraste e brilho e normalização de intensidade;
- Aplicar segmentação usando espaços de cores, visando separar plantas de milho do solo;
- Detectar as plantas de milho, representando-as com padrões circulares correspondentes ao raio de abertura do dossel;
- Avaliar técnicas de extração de características e o reconhecimento desses padrões;
- Contar e classificar as plantas de milho no estágio fenológico de pós-emergência (V2), utilizando técnicas do processamento de imagens e máquina de vetores de suporte (SVM).

1.4 Contribuições do trabalho

As principais contribuições deste trabalho estão relacionadas ao desenvolvimento de uma metodologia para a contagem de plantas de milho em ambiente de campo, utilizando imagens multiespectrais digitais obtidas por câmera multiespectral embarcada em Drone. Este método possibilita suporte a decisão para estudos relacionados à cultura do milho, sendo de grande interesse para pesquisadores, extensionistas e produtores agrícolas.

1.5 Organização do trabalho

A organização deste trabalho está estruturada em quatro Capítulos, onde o presente Capítulo está dedicado à introdução, o Capítulo 2 apresenta materiais e métodos, incluindo a fundamentação teórica das técnicas e os pseudo códigos dos algoritmos desenvolvidos, O Capítulo 3 apresenta resultados e discussões e finalmente, no Capítulo 4 são apresentados as conclusões.

Capítulo 2

MATERIAIS E MÉTODOS

Este Capítulo apresenta materiais e métodos, com aprofundamento teórico das técnicas utilizadas para o desenvolvimento do trabalho. Da mesma forma, são apresentados os aspectos da modelagem para calibração das medidas realizadas em ambientes de campo, assim como detalhamento dos pseudocódigos para os algoritmos computacionais implementados.

2.1 Área experimental

De acordo com [Lal e Stewart \(2015\)](#), os solos agrícolas são essenciais para os sistema de produção, da mesma forma que os solos desempenham um papel crucial na proteção da qualidade ambiental do ar e da água. Portanto, a agricultura de hoje implica novos desafios para atender à crescente demanda por alimentos, com menos espaço, adaptando-se às mudanças climáticas e sem usar tantos tratamentos químicos para cumprir regulamentações rígidas nacionais e internacionais ([SYLVESTER, 2018](#)).

Diversos autores relatam técnicas de delineamento experimental aplicadas na agricultura, onde os conceitos de campo, bloco e parcela costumam ser diferenciados a fim de fazer comparações de tratamentos, abordando também a variabilidade do campo (atributos relacionados à textura do solo) e conseqüentemente melhorar o manejo do solo ([HOSHMAND, 2006](#)).

Nesse contexto, a área experimental onde foi desenvolvida a presente pesquisa seguiu os padrões de estudo de Embrapa Instrumentação e contou com apoio de projeto financiado pela FAPESP (Processo PITE 2017/19350-2). Enquanto para a análise do solo, os estudos foram realizados pela Embrapa Pecuária Sudeste, e a adubação do solo seguiu a recomendação do Instituto Agrônomo de Campinas.

A Figura 2 ilustra a área experimental do Laboratório Nacional de Referência em Agricultura de Precisão (LANAPRE) localizada a 860 m e com coordenadas geográficas: 21°57'13,9"S e 47°51'10,9"W. Arranjo experimental organizado para a avaliação dos riscos agrícolas e logísticos em área da cultura de milho (*Zea Mays. L*) foi estabelecido em uma área de 4 000 m².

Figura 2 – Detalhe da área experimental LANAPRE.

2.1.1 Densidade de plantio

A densidade de plantio pode ser considerado como uma variável de fato muito importante. Onde a quantidade de plantas individuais são medidos por unidade de comprimento, área ou volume. Esta unidade pode ser claramente identificada como planta individual em uma cultura de plantas de milho (*Zea mays L.*). Enquanto Benjamin (2017), afirma que não existe uma resposta única e clara sobre o que deve ser a unidade da população, mas é importante estar ciente de que essa dificuldade pode surgir e é necessária uma consistência na escolha da unidade da população.

Existe uma diversidade de técnicas para estimar a densidade de plantio, dentre as mais utilizadas em pesquisas in situ, pode-se citar os *transectos*, que utiliza uma determinada área geralmente pré-definida da forma retangular. Enquanto uma segunda técnica usa os *quadrantes*, que são estabelecidos diretamente no solo. Geralmente tem um metro quadrado e pode ter um fio nos pontos médios para dividi-lo em quatro quadrados de 50 cm ou subdivisões menores (SUTHERLAND, 2006). Um exemplo de quadrante é ilustrado na Figura 3 na qual pode-se observar que o quadrante permite uma amostragem de plantas.

Figura 3 – Exemplo de quadrante para contar e estudar plantas.

A densidade de plantas geralmente esta baseada na contagem de plantas enraizadas em um quadrante e, funciona melhor com plantas de caules distintos e, diâmetro razoavelmente pequeno. A medida que o tamanho da área da planta aumenta e, cruza a área da amostra do solo, fica mais complicado decidir se a planta ao longo da fronteira está fora do quadrante (ELZINGA; SALZER, 1998).

2.1.2 Fenologia e fenotipagem do milho

Segundo Schwartz et al. (2003), definem a fenologia como um estudo das fases recorrentes do ciclo de vida de plantas, especialmente as relações que eles têm com o tempo e clima, o que inclui amplamente as mudanças sazonais no desenvolvimento e metabolismo da planta. Assim, as plantas geralmente vêm se adaptando à sazonalidade de seu ambiente e esses ecossistemas estão agora sendo influenciados pelas mudanças climáticas (CLELAND et al., 2007).

Dessa forma, a fenologia do milho (*Zea mays. L*) ou especificamente seu desenvolvimento estão divididos em três grupos: (i) superprecoce que esta em torno de até 110 dias, (ii) precoce que esta entre 110 e 145 dias, e (iii) tardio que é maior a 145 dias. Nesse contexto, as plantas milho plantadas na área experimental de acordo com as condições de clima do solo do Estado de São Paulo apresentou um ciclo variável entre 110 a 160 dias.

Independente do ciclo da planta de milho, escalas fenológicas foram propostas pelos pioneiros Ritchie e Hanway (1989) com o objetivo de tomar decisões ou ações no manejo da cultura. De acordo a esta escala são definidos estádios fenologicos do milho em duas grandes estádios: estágio vegetativo (V) e estágio reprodutivo (R), conforme é mostrado na Tabela 1. As subdivisões dos estádios vegetativos são designados numericamente como V0, V1, V2 até VT os quais podem ser até 18 estádios e representam a quantidade de folhas desenvolvidas. Por outro lado o estágio reprodutivo decorre desde a emissão do embonecamento (cabelos de milho) até o ponto de maturidade fisiológica (grão de milho atingiu a máxima matéria seca).

Tabela 1 – Estádios da planta de milho

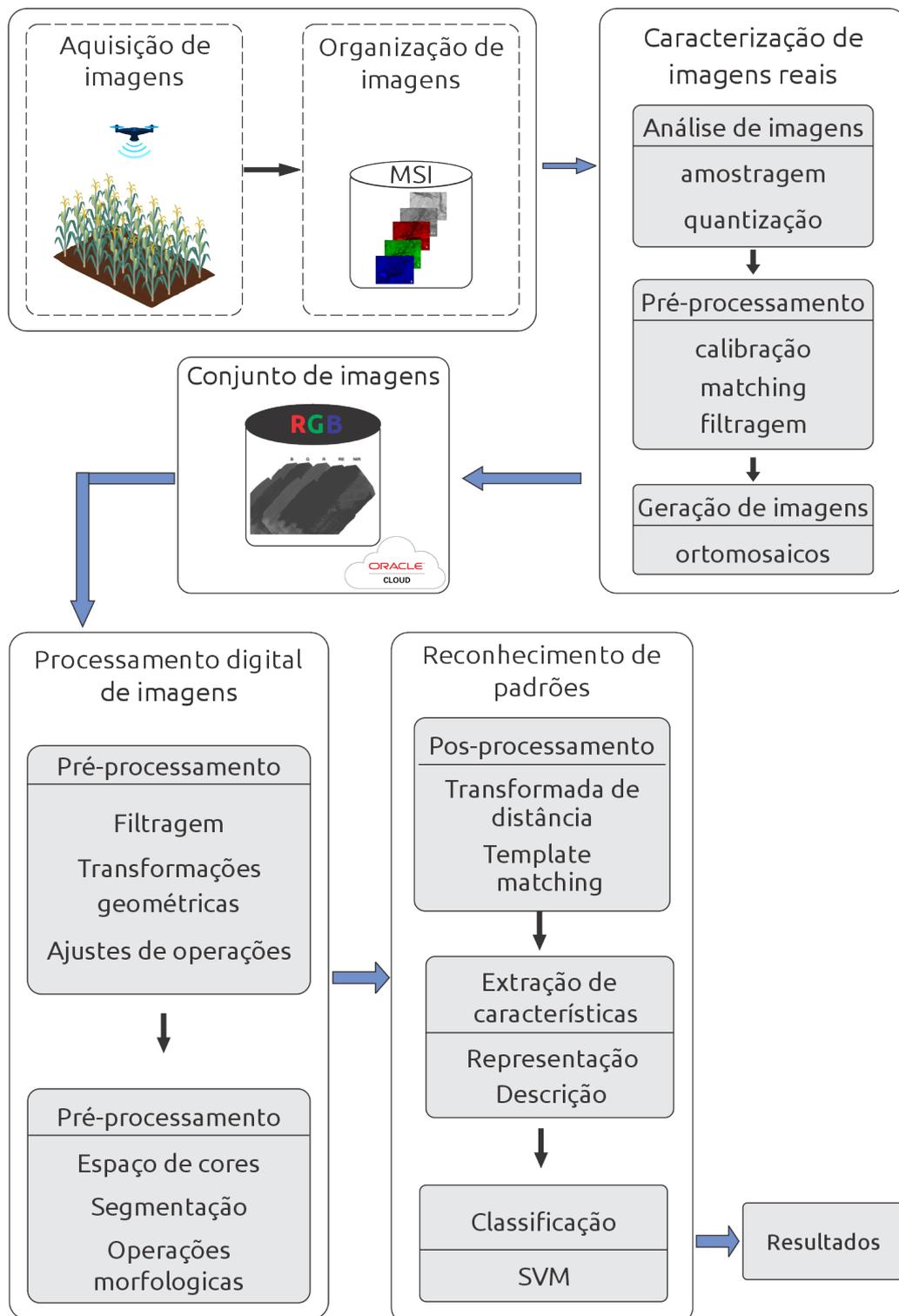
Vegetativo	Reprodutivo
VE, Emergência	R1, Embonecamento
V0, 1 ^a folha desenvolvida	R2, Bolha de água
V1, 2 ^a folha desenvolvida	R3, Leitoso
V2, 3 ^a folha desenvolvida	R4, Pastoso
V(n), n ^a folha desenvolvida	R5, Formação de dente
VT, Pendoamento	R6, Maturidade fisiológica

Portanto, para esta pesquisa foram estudados os estádios vegetativos de pós-emergência (caracterizados pela abetura e presença das folhas de milho). Assim, de acordo aos estudos de Magalhaes e Durães (2006), foram seleccionados imagens do estágio V2, porque é o estadio que define o ponto de crescimento dos estádios iniciais.

2.2 Abordagem geral do modelo e principais materiais

Nesta seção, é apresentado em diagrama de blocos (Figura 4) a organização do método para contagem de plantas e a classificação do estágio fenológico pós-emergência (V2), bem como as principais técnicas utilizadas para seu desenvolvimento.

Figura 4 – Diagrama de blocos do método desenvolvido.



Para a aquisição de imagens foram previamente definidos pontos de controle na área experimental com o fim de coseguir a comunicação entre o GPS do Drone e a área plantada. Desta forma, o sensor multiespectral embarcado no Drone coletou imagens em cinco bandas multiespectrais, porém as imagens foram organizadas com auxílio da interface do Drone. Para a caracterização das imagens coletadas, foi utilizado o software Pix4Dmapper®¹ com objetivo da criação dos ortomosaicos das imagens adquiridas. Durante o processo de sua construção, algoritmos já disponíveis para o referido software realizam a amostragem dos pontos de interesse a partir das imagens capturadas pelo Drone, principalmente o algoritmo que permite a geração do ortomosaico faz uso do matching dos pontos de interesse. Finalmente, o ortomosaico do mapa completo é cortado em 40 blocos multiespectrais com auxílio de software QGIS², sendo seleccionados os blocos dos canais vermelho, verde e azul para a composição do conjunto de imagens RGB.

O módulo de processamento digital de imagens, refere-se às etapas de verificação de contraste, brilho, suavização, aprimoramento de características e redimensionamento das imagens explorando diversos espaços de cor com a finalidade de obter maiores informações das características nos canais RGB. A seguir é realizada uma etapa de segmentação com o objetivo de separar objeto e fundo das culturas de milho, além de operações morfológicas aplicadas no pré-processamento.

O módulo de reconhecimento de padrões considera o uso da transformada de distância junto com o TM para detectar e localizar os objetos segmentados e por sua vez sobrepor círculos nesses objetos e assim fazer a contagem das plantas de milho. Em seguida, a proposta de um modelo de cone estima a altura das plantas de milho, utilizando a informação de raio de círculo encontrado pelos descritores de contornos em relação à abertura da folha do milho. Consequentemente, tanto raio de círculo encontrado quanto a altura da planta encontrado pelo modelo de cone permitem o estabelecimento das características dos objetos em estudo. Por fim, a etapa de classificação faz uso do SVM para classificar o estágio vegetativo pós-emergência (V2) das plantas de milho nas imagens processadas.

Para o desenvolvimento do método proposto, foi necessário organizar um ambiente computacional com serviços em nuvem que suportassem processamento digital de imagens, aprendizado de máquina e armazenamento em nuvem. Então foi considerado principalmente o suporte ao armazenamento de dados através de um sistema de arquivos distribuídos da Oracle Cloud³. Após configuração dos serviços com fim de gerenciar a integração da rede web com o computador desktop, foi possível organizar um ambiente de trabalho. Portanto o desenvolvimento do método foi considerado a utilização de diversas bibliotecas Python com suporte para

¹ <<https://www.pix4d.com/>>

² <<https://qgis.org/>>

³ <<https://www.oracle.com/br/cloud/>>

processamento digital de imagens como: OpenCV⁴, scikit-image⁵, além de suporte ao aprendizado de máquina: scikit-learn⁶, suporte para ferramentas de visualização como: matplotlib⁷ e manipulação de arquivos csv: pandas⁸.

2.3 Agricultura de precisão

A agricultura de precisão (AP) de acordo com Zhang (2016), é entendida como o manejo de cultivos em sítios específicos do solo considerando tanto a variabilidade espacial como a variabilidade temporal. Para a prática de manejo dos cultivos, a AP considera quatro componentes: (i) coleta de dados em escala e frequência apropriada, (ii) processamento e análise desses dados (iii) toma de decisão e (iv) aplicação de uma resposta de manejo e tempo apropriados. Desta forma, a Figura 5 ilustra os componentes envolvidos na AP.

Figura 5 – Ciclo da agricultura de precisão. Fonte: Pusch et al. (2019).



Nesse contexto Srinivasan (2006), lista alguns principais ferramentas tecnológicas utilizadas para a implementação da AP: sistemas de posicionamento global (GPS, do inglês *Global Positioning Systems*), sistemas de informação geográfica (GIS, do inglês *Geographic Information Systems*), sensoriamento remoto e tecnologias de taxa variável (VRT, do inglês *Variable Rate Technology*). Assim mesmo é importante ressaltar que essas tecnologias da AP não substituem o papel do agricultor no processo de decisão (SHANNON et al., 2020).

⁴ <<https://opencv.org/>>

⁵ <<https://scikit-image.org/>>

⁶ <<https://scikit-learn.org/>>

⁷ <<https://matplotlib.org/>>

⁸ <<https://pandas.pydata.org/>>

2.3.1 Sensoriamento Remoto

Sensoriamento remoto (SR) é o termo utilizado para a identificação e coleta de dados sem contato físico com o objeto de estudo, refere-se a dados coletados por dispositivos que detectam radiação eletromagnética, luz visível, luz infravermelha e luz infravermelha próxima (ZACHARIAH, 2018). Da mesma forma, o SR é considerado como um componente da tecnologia de detecção que inclui imagens de satélite e fotografia aérea. Sendo de principal interesse o uso de sensores remotos que podem ser agrupados em: sensores de solo, sensores aéreos e sensores de satélite (SRINIVASAN, 2006; ZACHARIAH, 2018). Por sua vez, o SR é utilizado em conjunto com tecnologias de análise de componentes, como o GIS.

De acordo com o Krishna (2018) os sensores aéreos para o sensoriamento remoto de baixa altitude até agora se tornaram os mais razoáveis com a chegada dos Drones porque oferecem imagens aéreas de pontos de vantagem acima de uma cultura, com uma visão panorâmica, mas com um grande detalhe analítico e precisão obtida usando sensores visuais sofisticados: infravermelho (IR, do inglês *Infrared*), infravermelho próximo (NIR, do inglês *Near-Infrared*) e térmicos.

Uma definição descomplicada de Frazier e Singh (2021) para o Drone, consiste em uma aeronave não pilotada. Enquanto Jr e Ellis (2020) afirma que os Drones são menos caros e menos complexos de voar em comparação com aeronaves tripuladas. Portanto os principais componentes e sistemas que compõem um Drone são os seguintes: enlace de dados, acelerômetro, barômetro, magnetômetro, giroscópio, sensores e atuadores. A Figura 6 ilustra os principais componentes de um Drone.

Figura 6 – Principais componentes de um Drone. Fonte: Frazier e Singh (2021).



Desta forma, se utilizou um Drone multirrotor DJI Matrice® 100 devido à estabilidade e facilidade que fornece no controle nas missões de voo. Por conta disso, principais características do Drone utilizado na coleta das imagens, são apresentadas na Tabela 2. Enfatizando sua capacidade de navegação customizada através da própria interface, além da capacidade de controlar o voo de forma programável.

Tabela 2 – Especificações do Drone DJI Matrice® 100.

	Parâmetros	Especificações
Desempenho	Precisão em voo (modo GPS)	H: 2,5 m V: 0,5 m
	Tempo de voo (2 baterias TB48D)	40 min
	Velocidade máxima (modo GPS)	17 m/s
	Velocidade máx. de ascensão	5 m/s
	Velocidade máx. de descensão	4 m/s
Configuração de sistema	Modelo	N1
	Controle de voo	Programável
	Navegação	Frame customizado

2.4 Aquisição e organização das imagens

A aquisição das imagens é realizada com a ajuda de diferentes tipos de sensores que são usados para transformar a energia eletromagnética incidente sobre os elementos que pertencem a uma cena (GONZALEZ; WOODS, 2018). Assim, a tecnologia de sensores de imagem tiveram mudanças ao longo dos anos desde os tubos de vácuo à dispositivos de estado sólido, notavelmente baseados nos Dispositivos de Carga Acoplada (CCDs, do inglês *Charge-Coupled Devices*) e no Semicondutor Complementar de Metál-Óxido (CMOS, do inglês *Complementary Metal-Oxide Semiconductor*) (MARQUES, 2011).

Sensores utilizados em Drones podem ser embarcados com sensores passivos, ou seja, que captam a energia refletida ou emitida pelo alvo a partir da iluminação de uma fonte de luz natural. Da mesma forma, também podem ser embarcados com sensores laser ou ativos, onde neste caso os próprios sensores emitem sua própria fonte de luz. Por conta disso, nesta pesquisa desenvolvida utilizou-se um sensor de tipo pasivo embarcado em Drone.

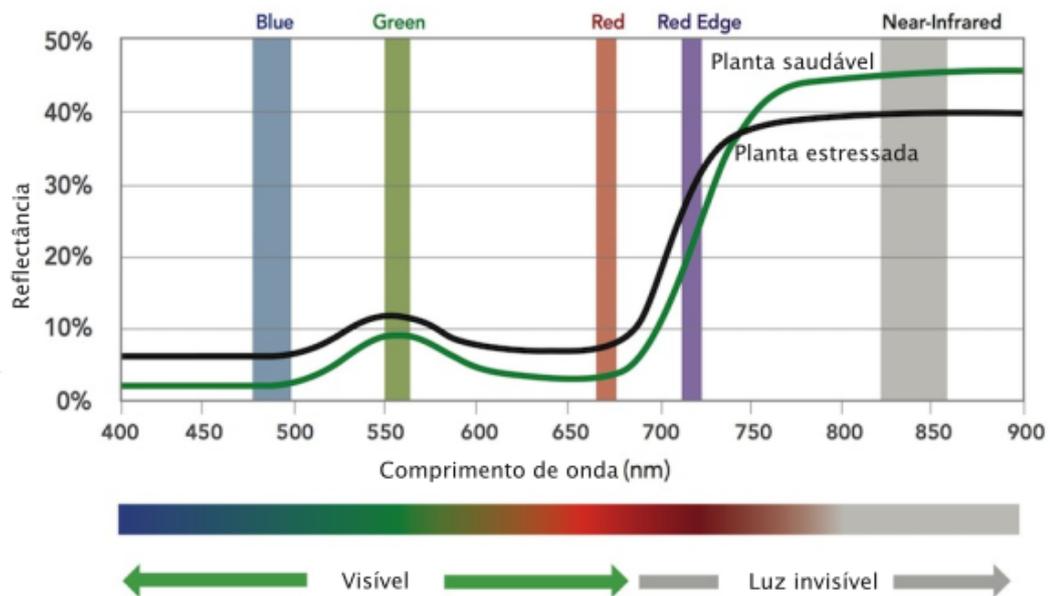
2.4.1 Sensor multiespectral

Muitas vezes no espectro visível as bandas escolhidas correspondem às cores azul, verde e vermelho que nos permitem reconstruir as imagens em suas cores naturais (TUPIN et al., 2014), embora o espectro eletromagnético contenha centenas de bandas estreitamente separadas, e com a ajuda dos sensores multiespectrais captura-se várias dessas bandas com a finalidade de capturar suas refletâncias.

Alguns sensores para Drones incluem entre 3 a 10 comprimentos de onda de luz e, no mercado são projetados especificamente para capturar refletância: infravermelho próximo (NIR, do inglês *Near-infrared*) e borda vermelha (RE, do inglês *Red edge*). Onde, o *Red edge* procura mudanças abruptas na reflectância em regiões de vegetação. Enquanto NIR, reflete os pigmentos da folha, que por sua vez é de estudo valioso para a saúde das plantas com uma diversidade de índices de vegetação.

Nesse contexto, a Figura 7 exibe as curvas de refletâncias de um sensor multiespectral para vegetação saudável (linha verde) e para vegetação estressada (linha preta), juntamente com suas bandas aproximadas: azul (*Blue*), verde (*Green*), vermelho (*Red*), borda vermelha (*Red Edge*) e infravermelho próximo (*Near-infrared*). Assim, a partir da Figura 7 pode-se destacar uma pequena crista nas curvas de reflectância da banda verde. O que nos diz essa crista, é a presença da clorofila na vegetação refletida pela luz mais verde relativa.

Figura 7 – Curva de reflectância para sensor multiespectral.



Desta forma, descreve-se na Tabela 3 as regiões de comprimento de onda que são capturadas por cada banda do sensor multiespectral MicaSense® RedEdge-M, o mesmo que foi utilizado na aquisição das imagens.

Tabela 3 – Bandas espectrais do sensor MicaSense® RedEdge-M

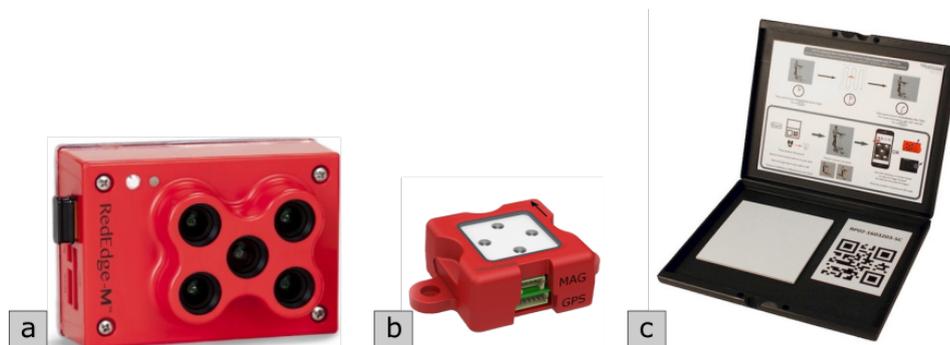
Banda	Comprimento de onda (nm)
Azul	475
Verde	560
Vermelho	668
Red Edge	717
NIR	840

As características do sensor multiespectral MicaSense® RedEdge-M, são ilustrados na Tabela 4. Onde é importante salientar: o tamanho do sensor e, a distância focal, por estar relacionada com a altitude do voo, a escala da imagem e a resolução da imagem. Desta forma, é possível coletar imagens com maior precisão em relação a sua geometria (FRAZIER; SINGH, 2021).

Tabela 4 – Parâmetros do sensor MicaSense® RedEdge-M

Descrição	Valores	Unidades
Imagem raw	12	bit
Distância focal	5,4	mm
Taxa de captura	1	captura/seg
Tamanho do sensor	4,8 × 3,6	mm
Resolução espacial	1280 × 960	pixels

Aquisição de imagens aéreas com sensores, são influenciados fortemente pelas condições naturais da luz do sol, assim como as nuvens do céu. Por conta disso, os valores dos pixels podem ser modificados quando estão sendo imageados com ajuda de um sensor multiespectral. Nesse contexto, é importante utilizar um sensor de luz (DLS, do inglês *Downwelling Light Sensor*), o qual tem por objetivo medir a irradiância da luz ambiente para cada um das bandas de um sensor multiespectral. Além disso, o DLS oferece a capacidade de correções nas mudanças globais da luz, as mesmas que flutuam durante o voo. Para que essas medidas de reflectância sejam corretas, é necessário de uma calibração radiométrica, os quais são disponibilizados por meio de um painel de reflectância. Dito isso, a Figura 8, ilustra o sensor Micasense® RedEdge-M (Figura 8-a), o sensor DLS (Figura 8-b), e seu correspondente painel de reflectância (Figura 8-c).

Figura 8 – Sensores e painel de reflectância para imageamento aéreo.

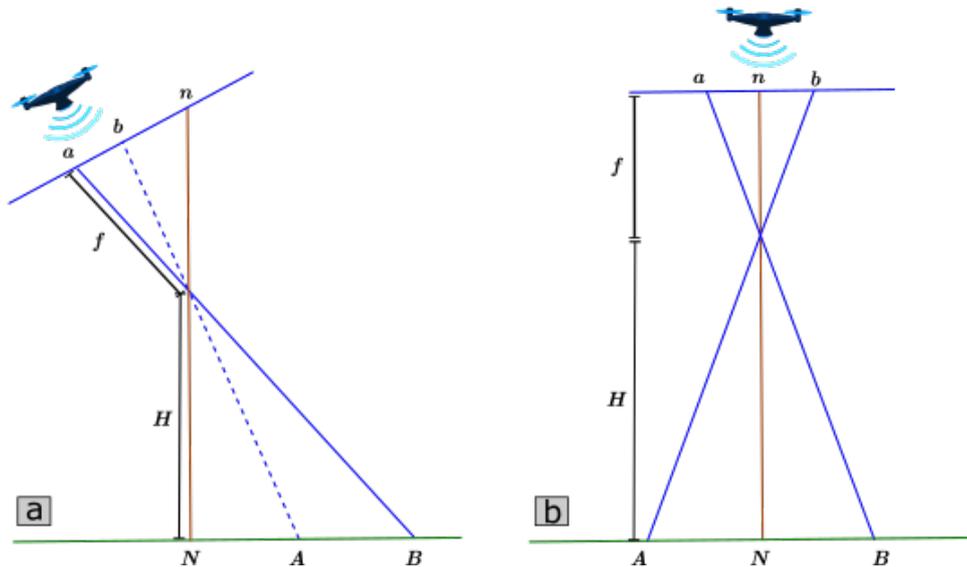
2.4.2 Geometria de imagens aéreas

Entender a geometria de uma imagem aérea é considerado uma regra básica, pois o tamanho do sensor da câmera, a distância focal e a altura do voo podem determinar a escala da imagem ou a resolução espacial da imagem. Assim, à medida que a luz do sol reflete em um objeto terrestre, parte dessa luz entra na lente da câmera e pode-se relacionar o solo à imagem por meio de princípios trigonométricos (FRAZIER; SINGH, 2021).

As imagens aéreas podem ser classificadas de acordo com o ângulo do eixo óptico da câmera em relação ao ângulo vertical (nadir, representado por N na Figura 9). Uma imagem aérea capturada com inclinação de $\pm 0^\circ$ a partir de nadir, é chamada de imagem aérea vertical verdadeira (Figura 9-b). Se a inclinação da imagem aérea capturada for maior que 0° e menor que

$\pm 3^\circ$, é chamada de imagem verticalmente inclinada. Se a inclinação da imagem aérea capturada for superior a $\pm 3^\circ$, é chamada de imagem oblíqua (Figura 9-a).

Figura 9 – Geometria da imagem aérea com Drone.



A partir do fenômeno físico da luz do sol e a câmera embarcado em Drone, pode-se expressar uma relação entre a imagem e o chão. Nesse sentido, a partir da Figura 9, o ponto do chão "A", localizado na borda esquerda da cobertura do solo, é projetado no plano da imagem "a"; e o ponto do chão "B", é projetado no plano da imagem "b". Desta forma, tanto "a" quanto "b" estão localizados na matriz do sensor multiespectral. Essas relações nos permitem calcular a escala da imagem e, a distância da amostra do solo (GSD, do inglês *Ground Sample Distance*).

A escala da imagem aérea, de acordo com [Frazier e Singh \(2021\)](#), é a razão entre a distância da imagem e a distância correspondente no solo. Nesse sentido, a distância em relação do solo (AB) é projetado na distância do plano da imagem (ab). Porém, a escala da imagem pode ser calculada utilizando a seguinte equação:

$$S = \frac{ab}{AB} \quad (2.1)$$

Conforme a Equação 2.1, a escala da imagem aérea também pode ser determinada pela distância focal (f) da câmera, e a altura de voo (H) acima do solo, conforme é definido pela Equação 2.2:

$$S = \frac{f}{H} \quad (2.2)$$

É importante destacar, que de acordo a definição padrão da óptica dos lentes, a métrica para a distância focal (f) é expressado em milímetros (mm). Enquanto, a distância em relação a altura do voo (H) é expressado em metros (m).

Embora, para as imagens digitais a escala no sensor não é de muito interesse, por que a escala da imagem digital pode mudar quando é visualizado em um dispositivo de exibição, tornando-se uma propriedade desse dispositivo. Nesse contexto, a resolução original da imagem não muda com a escala da exibição variável, e o tamanho do objeto visível depende diretamente do tamanho dos pixels. Por tanto, uma propriedade mais adequada para a escala da imagem é o GSD (COMER et al., 1998).

Dito isso, o GSD se refere a distância que o sensor pode registrar os pixels de uma amostra da superfície do chão. A unidade do GSD é expressa em cm/pixel, onde por cada pixel é possível representar \overline{AB} centímetros do chão. Desta forma, queda registrada a área coberta pela imagem capturada com Drone, e para seu cálculo, o GSD é definido pela seguinte Equação 2.3:

$$GSD = \frac{L_s \times H}{L_i \times f} \quad (2.3)$$

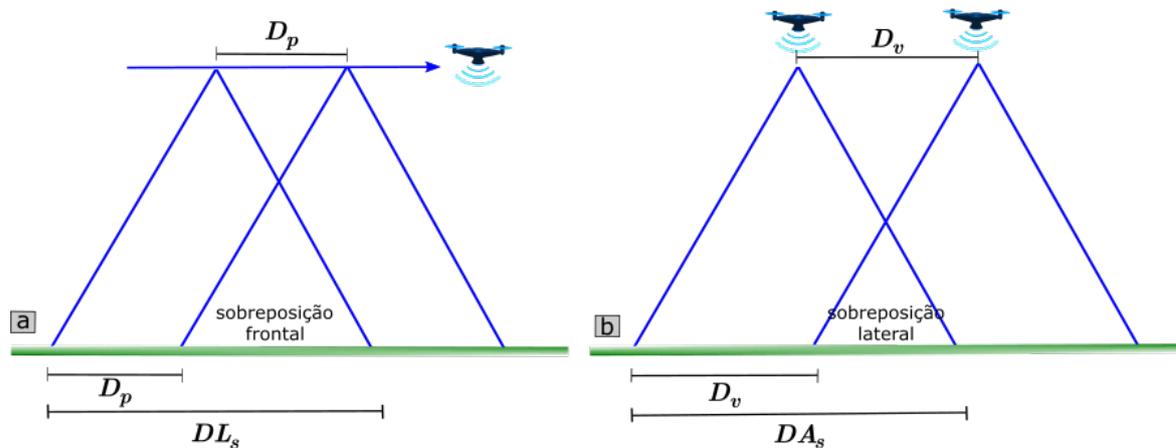
onde L_s é a largura do sensor, e L_i é a largura da imagem. É interessante destacar que, tanto H como f mudam as distâncias da imagem em um fator linear. Por exemplo, se dobrar H , a imagem resulta com a metade das distâncias, e se triplicar f , aumentaria as distâncias da imagem por um fator de três. Por tanto, o GSD se torna em um fator determinante na qualidade das imagens.

2.4.3 Sobreposição de imagens aéreas

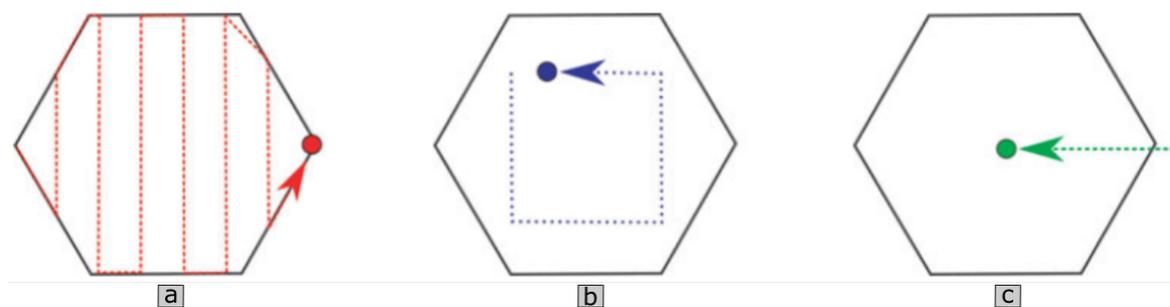
O olho humano percebe a geometria de um objeto no espaço tridimensional (3D) através de sobreposição dos campos de visão, os quais são formados por diferentes ângulos, dito de outra forma os humanos possuem a visão estereoscópica. Por conta disso, o mapeamento aéreo reproduz este fenômeno através de sobreposição de múltiplas imagens verticais, os quais são adquiridas em blocos e com múltiplas linhas durante a etapa de voo (ABER et al., 2016).

Frazier e Singh (2021), definem a sobreposição de imagens aéreas, como a quantidade de cobertura do solo a partir de duas imagens consecutivas ou adjacentes, e sua medição esta expresso em porcentagem. Na sobreposição aérea, dois tipos são considerados: frontal (também chamado de longitudinal) e lateral.

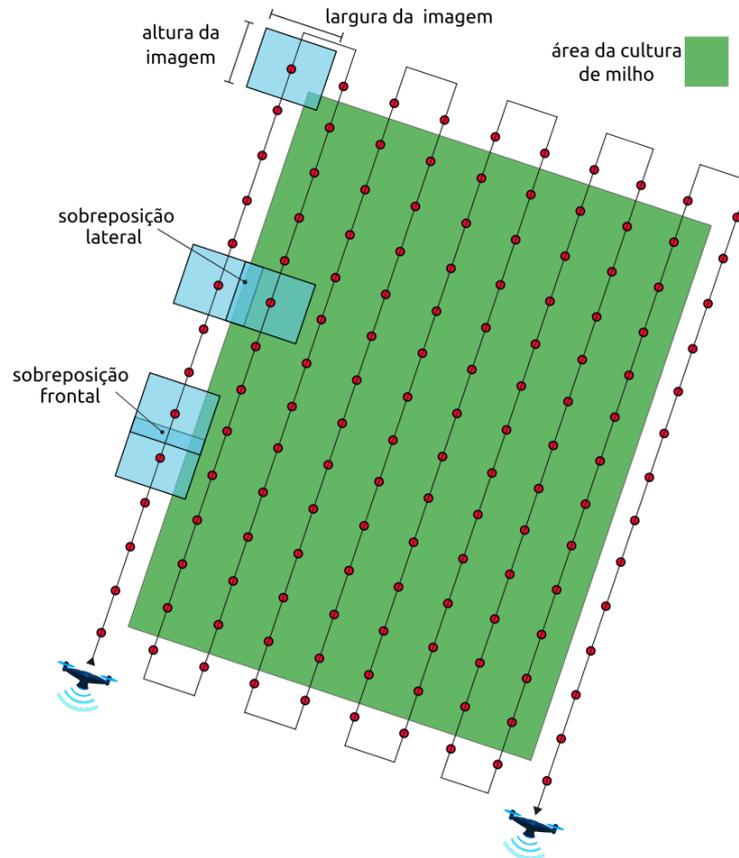
A sobreposição frontal, descreve a quantidade de sobreposição introduzidas intencionalmente entre imagens sucesivas capturadas ao longo de uma linha de voo. A Figura 10-a ilustra a sobreposição frontal, com a distância entre dos pontos centrais (D_p), e sua distância da largura coberta no chão (DL_s). Enquanto, a sobreposição lateral, descreve a quantidade de sobreposição entre imagens de linhas de voo adjacentes. A Figura 10-b ilustra a sobreposição frontal, com a distância entre as duas linhas de voo (D_v), também chamada de espaçamento de linha, e da mesma forma se ilustra a distância de altura coberta no chão (DA_s).

Figura 10 – Sobreposição frontal e lateral de imagens aéreas.

Nesse contexto, há múltiplas opções para a sobreposição de imagens aéreas, os quais estão fortemente ligados com o número de linhas de voo, assim como as rotas utilizadas durante o planejamento do voo. [Almadhoun et al. \(2019\)](#), descrevem três padrões que evitam colisões em diferentes altitudes, destacando que a abordagem dos autores tem por objetivo minimizar o tempo de mapeamento, além de ser adaptável para diferentes Drones e resoluções. A cobertura padrão do cortador de grama lateral (*side-lawnmower pattern*) é ilustrado na Figura 11-a, o padrão quadrado ilustrado na Figura 11-b, e o padrão centroide hexagonal é ilustrado na Figura 11-c.

Figura 11 – Padrão de sobreposição para cobertura do solo. Fonte: Adaptado de [Almadhoun et al. \(2019\)](#)

Para [Frazier e Singh \(2021\)](#), os valores recomendados para sobreposição frontal e lateral são considerados com porcentagens mínimos de 70% e 40% respectivamente, dessa forma os autores equilibram os efeitos da instabilidade da plataforma durante a etapa do voo. Enquanto [Aber et al. \(2016\)](#), afirma que para terrenos montanhosos, a sobreposição lateral e frontal devem ser mais densos, onde o objetivo incrementar as amostras na cobertura, e dessa forma evitar as lacunas entre as imagens capturadas. Com tudo isso, para a presente pesquisa desenvolvida, a sobreposição lateral e frontal foram de 80% respectivamente, seguindo um padrão de voo de linhas paralelas, ao longo de um conjunto de *waypoints*, como são ilustradas na Figura 12.

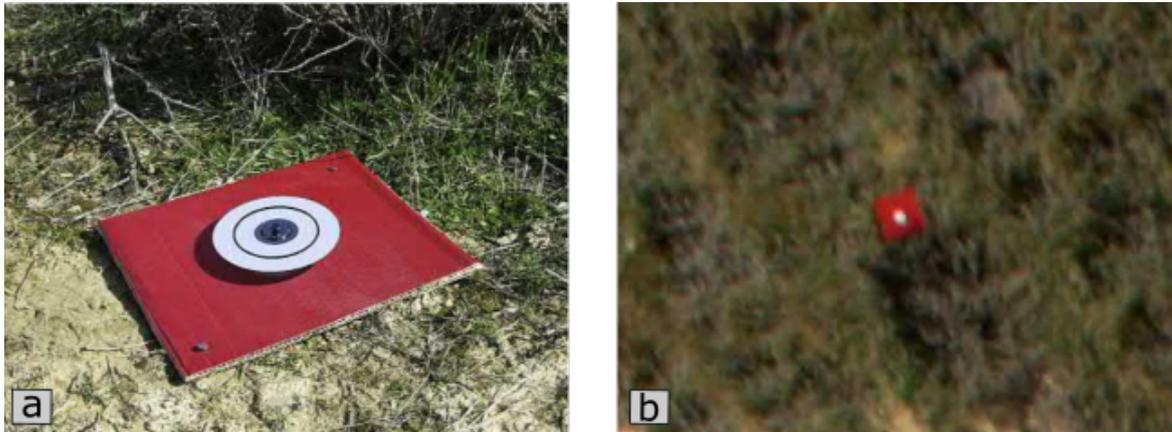
Figura 12 – Mapeamento durante a sobreposição das imagens aéreas.

2.4.4 Pontos de controle em campo

Drones possuem capacidades de voar livremente, do mesmo modo que podem ser pilotadas remotamente por um piloto no solo, ou programadas para voar de forma autônoma ao longo de rotas especificadas com pontos de passagens (do inglês *waypoints*) ou com pontos de controle em campo (GCP, do inglês *Ground Control Points*). Assim, de acordo com [Frazier e Singh \(2021\)](#), os *waypoints* são localizações referenciais nas vias aéreas especificados durante a missão de voo. Enquanto [Aber et al. \(2016\)](#), argumentam que os GCP representam dados do solo referenciados e marcados no campo de estudo, geralmente com alvos de cores de alto contraste (preto e branco) que podem ser de diferentes materiais (papelão, metal, etc) além de ser facilmente visíveis a partir de um Drone. A Figura 13-a exibe um exemplo de GCP feito com material de papelão e metal, enquanto a Figura 13-b ilustra um exemplo de GCP registrado a partir de um Drone.

Para esses pontos referenciais, dito de melhor forma os *waypoints* e GCP, podem ser localizados com certo grau de precisão a partir de diversos sistemas de navegação por satélite (GNSS, do inglês *Global Navigation Satellite System*). Nesse quesito, destaca-se o sistema de posicionamento global (GPS, do inglês *Global Positioning System*), o qual esta composto por 24 constelações de satélites americanos e esta operado pelo departamento de defesa americano. Desta forma [Jensen \(2015\)](#), descreve que as tecnologias GPS podem atingir acurácias até um

Figura 13 – Marcado de pontos de controle em campo. Fonte: Adaptado de Aber et al. (2016).



intervalo de ± 20 cm. Por outro lado, Barkóczy et al. (2018), descrevem que melhores precisões podem ser conseguidos com técnicas de posicionamento cinemático em tempo-real (RTK, do inglês *Real Time Kinematic*), que são utilizados em conjunto com sistemas GNSS, o qual pode atingir até uma acurácia de ± 2 cm.

Por conta disso, neste trabalho os pontos de controle em campo foram coletados por GPS de alta precisão em conjunto com um receptor RTK, com objetivo de registrar coordenadas geográficas até uma precisão de ± 1 cm. Desta forma os GCP coletados no levantamento topográfico, foram importados na estação de controle do Drone DJI® integrando-se com o sensor Micasense® e com o GPS 3DR® embarcado. Assim, finalmente os GCP marcados foram utilizados como os *waypoints* nas etapas de aquisição de imagens.

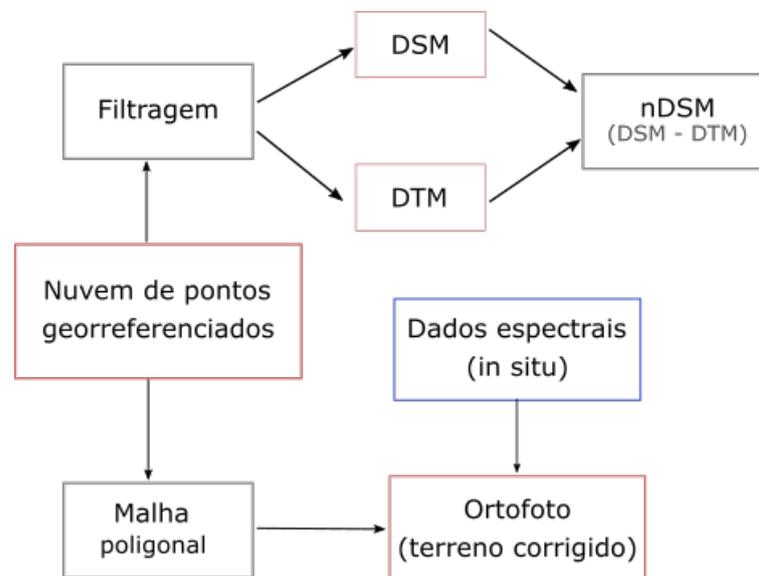
2.5 Caracterização das imagens multiespectrais

Segundo Frazier e Singh (2021), imagens adquiridas são apenas o primeiro passo em um fluxo de trabalho de SR baseado em Drone, e para produzir produtos de dados utilizáveis (modelos digitais e ortofotos) é necessário o uso de processos de estrutura de movimento (SfM do inglês, *Structure from Motion*) (ULLMAN, 1979). Nesse sentido, SfM consiste em uma técnica de visão computacional que cria modelos 3D a partir de de uma serie de imagens bidimensionais (2D), identificando características únicas nas imagens adquiridas, e compare-as com as áreas de sobreposição com a finalidade de gerar pontos chaves em um espaço 3D, na forma de nuvem de pontos (ABER et al., 2016).

A nuvem de pontos produzidos, não possuem coordenadas do mundo real, e para a correta geração dos produtos digitais é necessário de correções nos sistemas de coordenadas de referência. Desta forma, no processo SfM é necessário localizar e escalar as coordenadas geodéticas, os mesmos que estão baseados em latitude e longitude (LIU; MASON, 2016). Nesse quesito, imagens aéreas que foram coletadas com GPS de alta precisão, são registrados com

sistema geodético ou melhor definidos pelo modelo matemático DATUM WGS 84 ⁹, os mesmos que precisam ser transformados no sistema de referência geocêntrico para as américas 2000 (SIRGAS 2000 ¹⁰). Dito isso, o processo SfM gera uma nova nuvem de pontos georreferenciadas de acordo as coordenadas coletadas com GPS, em formato de arquivo raster ou também chamado de modelo digital de elevação (DEM do inglês *Digital Elevation Model*). A partir desse novo modelo retificado é possível obter novos produtos de acordo ao estudo da aplicação, como são ilustrados na Figura 14.

Figura 14 – Produtos derivados apartir do processo SfM, incluem modelos digitais de superfície (DSM, do inglês *Digital Surface Model*), modelos digitais de terreno (DTM, do inglês *Digital Surface Terrain*), modelos digitais de superfície normalizados (nDSM, do inglês *normalized DSM*) e ortofotos. Fonte: Adaptado de Frazier e Singh (2021)



2.5.1 Ortorretificação e ortomosaicos

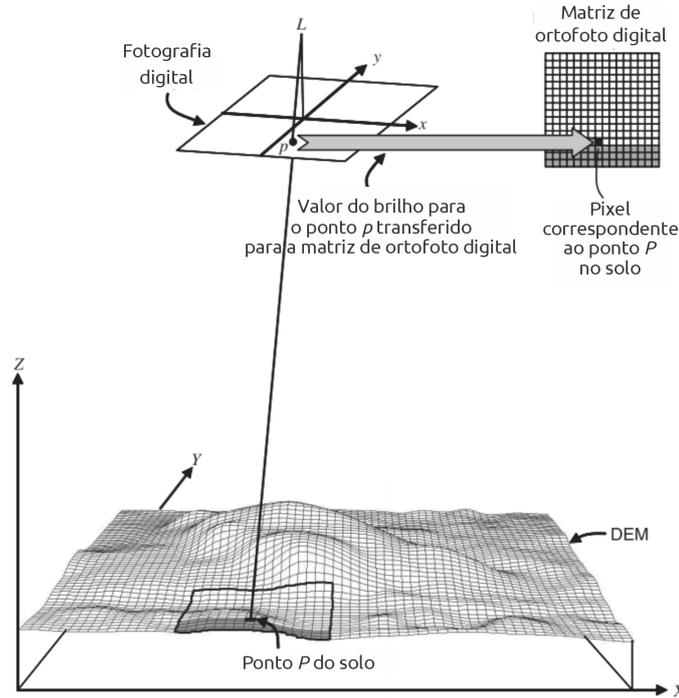
Como foi dito, objetos no solo possuem diferentes elevações (disparidades) na posição de suas imagens e de acordo com Schowengerdt (2006), essas disparidades podem ser usadas para estimar diferenças em elevações entre pontos e dessa forma poder gerar as imagens ortográficas, onde cada ponto em sua correta localização relativa se une a outros pontos. Para sua produção são necessárias a perspectiva de uma fotografia digital e um DEM.

Segundo Lillesand et al. (2015), existem vários meios de criar ortofotos, um dos processos é conhecido de *backward projection*, onde começa com os pontos no DEM e localiza esses pontos correspondentes na fotografia original. A Figura 15 ilustra esse processo, onde pontos da posição do solo no DEM (X_p, Y_p, Z_p) e os pontos associados à imagem (x_p, y_p) podem ser calculados através de equações colineares, definidos pela Equações 2.4 e 2.5 respectivamente.

⁹ WGS 84: *World Geodetic System 1984*

¹⁰ SIRGAS 200: estabelecido pelo instituto brasileiro de geografia e estatística desde 2005

Figura 15 – Processo de produção de uma ortofoto digital. Fonte: Lillesand et al. (2015).



$$x_p = -f \left[\frac{m_{11}(X_P - X_L) + m_{12}(Y_P - Y_L) + m_{13}(Z_P - Z_L)}{m_{31}(X_P - X_L) + m_{32}(Y_P - Y_L) + m_{33}(Z_P - Z_L)} \right] \quad (2.4)$$

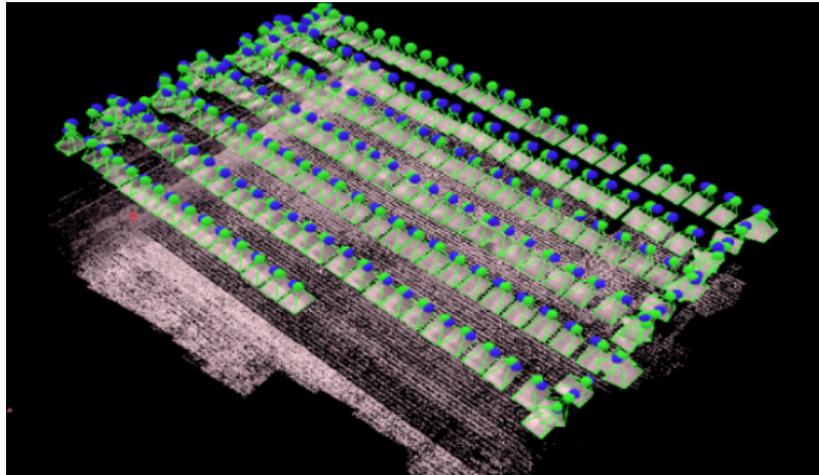
$$y_p = -f \left[\frac{m_{21}(X_P - X_L) + m_{22}(Y_P - Y_L) + m_{23}(Z_P - Z_L)}{m_{31}(X_P - X_L) + m_{32}(Y_P - Y_L) + m_{33}(Z_P - Z_L)} \right] \quad (2.5)$$

onde x_p e y_p correspondem às coordenadas da imagem aérea em qualquer ponto p ; X_P , Y_P , e Z_P correspondem às coordenadas do chão no ponto P ; X_L , Y_L , e Z_L correspondem às coordenadas do centro óptico L ; m_{11}, \dots, m_{33} correspondem aos coeficientes de uma matriz de rotação de 3×3 , os mesmos que transformam o sistema de coordenadas do chão para o sistema de coordenadas da imagem.

Muitos softwares SfM com os pontos localizados e retificados convertem a nuvem esparsa ou densa de pontos para uma malha triangular (do inglês *mesh*) ou também chamada de malha TIN (do inglês, *Triangulated Irregular Network*), onde é possível realizar operações básicas como a filtragem ou amostragem da malha. A partir disso, uma ortofoto pode ser costurada (do inglês *stitching*) juntamente com muitas das imagens aéreas coletadas, corrigindo seus respectivos deslocamentos das superfícies registradas. Nesse quesito, a união de imagens contíguas para imagem completa pode ser também chamado de ortomosaicos (ABER et al., 2016). A Figura 16 exibe um exemplo de nuvem de pontos esparsos gerados a partir de pacotes de software SfM.

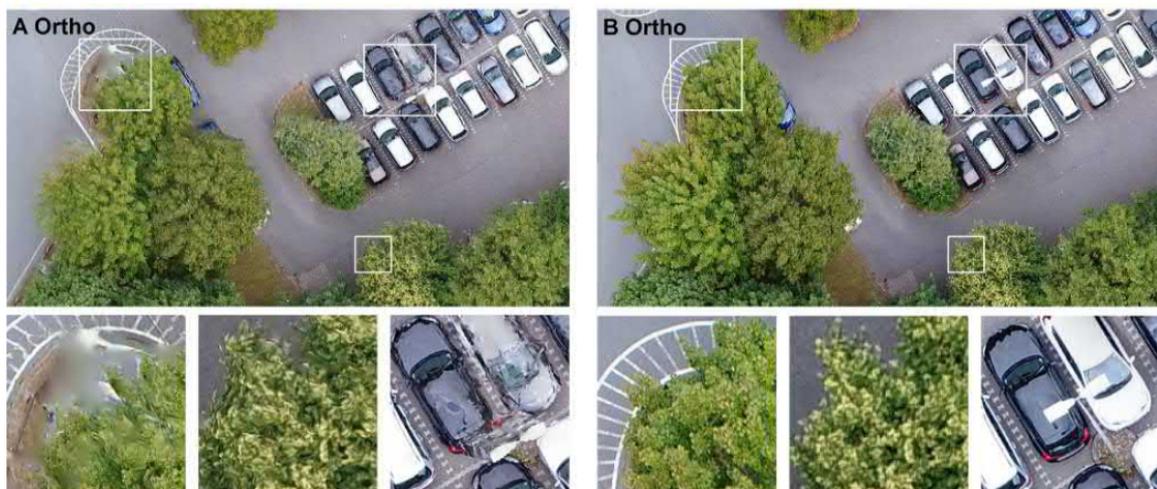
Muitas vezes a presença de vegetação é um desafio para o processo da ortorretificação. Por exemplo, o dossel de uma vegetação pode ser mais ou menos compacto e pode produzir superfícies bastante acidentadas e fragmentadas (artefatos) na produção de uma ortofoto final.

Figura 16 – Exemplo de nuvem de pontos esparsos a partir de canais RGB. Fonte: Yadav et al. (2019)



Por outro lado, dificuldades também podem surgir para objetos abruptamente altos e superfícies altamente refletivas. Por conta disso, dependendo da finalidade da ortofoto, pode ser aconselhável calcular um modelo de superfície mais suave para uso na ortorretificação ou reduzir o número de imagens de entrada para o mosaico, desta forma evitar muitos efeitos de borda nas linhas de costura. A Figura 17 ilustra um exemplo com dois fluxos de trabalhos diferentes.

Figura 17 – Exemplo de ortomosaico gerado com dois fluxos de trabalho diferentes, a partir de uma (A) malha com alta contagem de nuvem de pontos ultradensa levemente filtrada correspondente de 6 imagens, e (B) malha fortemente suavizada com baixa contagem de nuvem de pontos esparsa correspondente de 3 imagens. Fonte: Frazier e Singh (2021)



Embora, muitos pacotes de software clássicos e SfM agora integram a criação de mosaicos ortofotográficos de forma automática, considerando fatores de correção, possíveis artefatos, etc. O que possibilita a etapa final do fluxo de trabalho na caracterização de imagens aéreas, além de guardar as georeferências dos ortomosaicos em formato rasterizado (ABER et al., 2016; LIU; MASON, 2016).

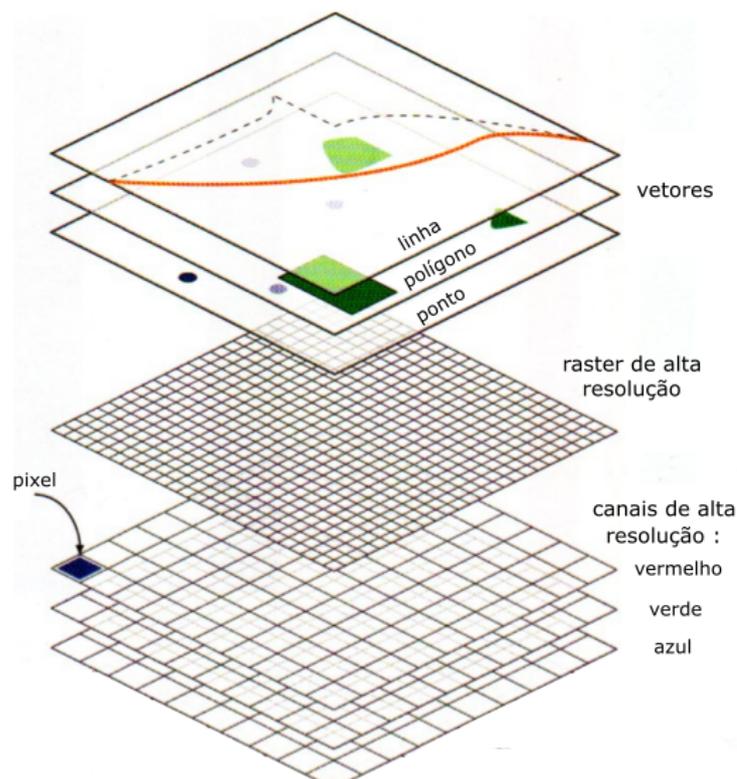
2.5.2 Representação de imagens aéreas

Originalmente pacotes de software SfM produzem nuvem de pontos em formato de arquivo LAS, mas produtos finais como modelos digitais e ortofotos, são frequentemente utilizados com estrutura de dados *raster* ou vetoriais, além de ser frequentemente utilizado por especialistas e pesquisadores de dados geospaciais (FRAZIER; SINGH, 2021).

De acordo com Konecny (2019), dados raster representam uma grade regular ou matriz 2D de números digitais (pixels), sendo caracterizada por: origem (geocodificada), orientação (geocodificada) e tamanho da célula. Claramente, objetos podem ser representados de acordo al tamanho del pixel da grade, os quais são definidos pela resolução: espacial e radiométrica. Além disso, o raster é utilizado mais apropriadamente, para imagens com alta variabilidade espacial.

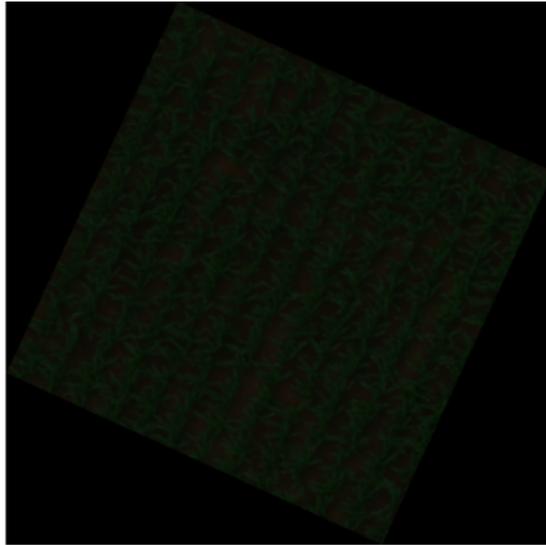
Estrutura de dados vetoriais incorporam discretamente, a geometria e a localização de características geográficas, mas são representados por gráficos vetoriais, ou seja são compreendidos por vértices, e arestas em forma de rotas. Além disso, suas informações são geralmente visualizadas e acessadas em forma de tabela, dito de outra forma, a característica vetorial tem um número de identificação único, assim como atributos numéricos ou textuais, de modo que as características possam ser simbolizadas (WEGMANN et al., 2016). A Figura 18 ilustra as imagens multiespectrais de canais RGB, representando os objetos da cena por um raster e vetores compostos como camadas.

Figura 18 – Imagens multiespectrais RGB com seus respectivas representações em formato raster e vetor, para objetos como pontos, linhas e áreas. Fonte: Adaptado de Wegmann et al. (2016)



No contexto da pesquisa, os ortomosaicos georreferenciados com formato raster foram trabalhados com auxílio de software QGIS. A seguir, foram cortadas em um ortomosaico geral sobre uma camada vetorial. Logo após, foram realizados recortes em blocos, os quais são referentes aos locais de análise. Finalmente esses blocos foram guardados em formato raster como os dados georreferenciados, conforme é exibido pela Figura 19.

Figura 19 – Ortomosaico final recortado em bloco.



2.5.3 Formato de armazenamento de imagens aéreas

Selecionar um formato ótimo para a imagem poderia evitar artefatos provocados pela compreensão de dados como acontece nos formatos de imagem JPEG, os mesmos que poderiam degradar a precisão nas medições da informação espectral. Desta forma, formatos de arquivo sem compreensão como RAW (formato cru) ou TIFF (do inglês, *Tagged Image File Format*) são eficientes para armazenamento de informação espectral, além de disponibilizar a customização no pós-processamento.

Por conta disso, o GeoTIFF (RITTER et al., 2000) introduziu o conceito dos metadados, onde apenas seis rótulos TIFF são suficientes para transportar todas as informações do georreferenciamento, i.e, projeção cartográfica, datum geodésico, tamanho do pixel, coordenadas espaciais da imagem e qualquer informação adicional, como as coordenadas projetadas. Todas essas informações, são salvos no formato de arquivo TIFF padrão sem destruir sua estrutura de dados (GAO, 2009).

Nesse contexto, e de acordo com as características atuais do formato de imagem GeoTIFF, esta pesquisa utilizou o GeoTIFF como formato de arquivo para armazenamento. Portanto, as principais características dos ortomosaicos multiespectrais são especificadas na Tabela 5.

Tabela 5 – Imagens Multiespectrais.

Descrição	Ortomosaico por Bloco
Tipo de arquivo	GeoTIFF
Bits por pixel	32 Bit (ponto flutuante)
Resolução	96 ppi (pixels per inch)
Total de imagens	40
Largura	546 pixels
Altura	546 pixels

2.6 Processamento Digital de Imagens

Atualmente, existem uma diversidade de conceitos que podem definir o processamento digital de imagens (PDI), o que implica algum entendimento complexo. Alguns autores propõem limitações entre processamento digital de imagens (GONZALEZ; WOODS, 2018), visão de máquina (DAVIES, 2017) e visão computacional (BHUYAN, 2019). Segundo Gonzalez e Woods (2018), sugerem o uso de paradigmas úteis para separá-los em três tipos de processos: (i) baixo nível, que envolve operações primitivas; (ii) médio nível, que envolve segmentação, descrição e (iii) alto nível, processamento que envolve "dar sentido" a um conjunto de objetos reconhecidos, abordando as funcionalidades cognitivas associadas à visão humana.

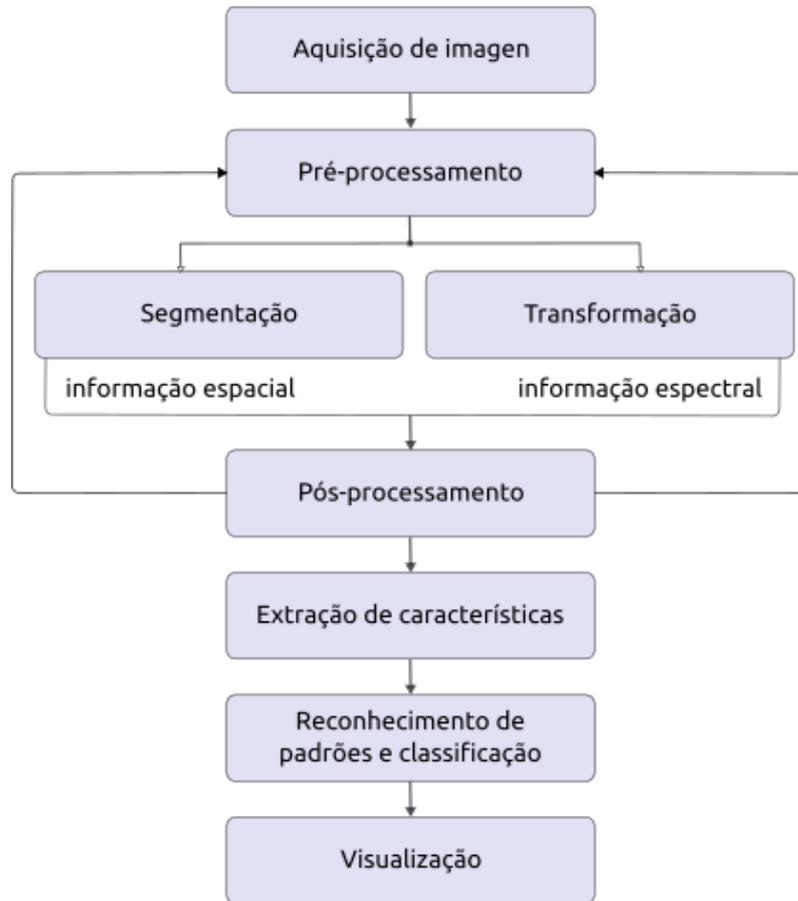
Baseado no paradigma acima mencionado, o PDI é dividido em uma série de fases que podem englobar uma grande variedade de técnicas existentes ou em evolução. Nesse contexto, a Figura 20 ilustra uma estruturação de etapas do PDI, começando da aquisição da imagem e finalizando até a visualização. Cada etapa não segue necessariamente a sequência apresentada, possibilitando adaptações às necessidades do problema.

2.6.1 Pré-processamento das imagens da cultura de milho

Como dito anteriormente, a imagem rasterizada contém pixels em cada um de suas posições em relação à grade regular. Dessa maneira Gonzalez e Woods (2018), representam a imagem como uma matriz 2D composta por funções de duas variáveis espaciais $F(x, y)$ (Equação 2.6), sendo que f corresponde aos pixels de um certo canal multiespectral.

$$F(x, y) = \begin{bmatrix} f(0, 0) & f(0, 1) & \dots & f(0, N - 1) \\ f(1, 0) & f(1, 1) & \dots & f(1, N - 1) \\ \vdots & \vdots & & \vdots \\ f(M - 1, 0) & f(M - 1, 1) & \dots & f(M - 1, N - 1) \end{bmatrix} \quad (2.6)$$

onde $(x, y) \in \mathbb{Z}^2$ são as coordenadas das intensidades do pixel, M e N representam o número de linhas e colunas da matriz da imagem para $x = 0, 1, 2, \dots, M - 1$ e $y = 0, 1, 2, \dots, N - 1$.

Figura 20 – Etapas do processamento digital de imagens. Fonte: Umbaugh (2017).

De acordo com a Equação 2.6, as imagens requerem decisões feitas pelos valores de M, N em conjunto do número de níveis de intensidade (L_*). Por conta do armazenamento e à quantização do hardware, comumente esse níveis de intensidade são de potencia dois: $L_* = 2^k$, $k \in \mathbb{Z}$.

Canais selecionados nesta pesquisa correspondem para canais vermelho, verde e azul. Por conta disso, a representação matricial de imagens pode ser denotada como F_r, F_g e F_b em relação aos canais anteriormente mencionados. Dito isso, processamento de imagens multiespectrais utilizam rotineiramente operações matriciais e vetoriais. Portanto, para formar uma imagem RGB com operações vetoriais, podem ser selecionados pixels de posições (x, y) de um certo canal, e desta forma a organização desses pixels podem ser denotados como um vetor coluna:

$$\mathbf{F}_{rgb} = \begin{bmatrix} \mathbf{F}_r(x, y) \\ \mathbf{F}_g(x, y) \\ \mathbf{F}_b(x, y) \end{bmatrix} \quad (2.7)$$

Dado que as imagens foram corrigidas radiometricamente na etapa de caracterização de imagens, algoritmos de alinhamento de canais não serão necessários (LIMA, 2020). Por outro lado, a combinação dos cores de canais mencionados (c_1, c_2, c_3) precisam da leitura (\mathcal{L}) de

imagens georreferenciadas com bibliotecas de abstração de dados geoespaciais (GDAL¹¹, do inglês *Geospatial Data Abstraction Library*) para seu correto pré-processamento. Nesse contexto, o Algoritmo 1 define o processo de combinação de canais (\cup).

Algorithm 1 Combinação de canais vermelho, verde e azul

Entrada: Matrizes imagens rasterizadas F_{c_1} , F_{c_2} , F_{c_3}

Saída: Matriz imagem RGB F_{rgb}

1 **início**

2 $F_r = \mathcal{L}(F_{c_1})$

3 $F_g = \mathcal{L}(F_{c_2})$

4 $F_b = \mathcal{L}(F_{c_3})$

5 $F_{rgb} = F_r \cup F_g \cup F_b$

6 **fim**

Transformações geométricas modificam a arranjo espacial dos pixels em uma imagem, alterando a posição, orientação e tamanho dos objetos na imagem. Entre suas operações básicas estão: transformação espacial de coordenadas, e interpolação de intensidade atribuídos aos pixels transformados espacialmente. A transformação espacial de coordenadas pode ser definido como:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{T} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t_{0,0} & t_{0,1} \\ t_{1,0} & t_{1,1} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2.8)$$

onde $t_{x,y} \in \mathbb{R}^{2 \times 2}$ representam os elementos da matriz transformação em coordenadas espaciais, e x', y' representam as coordenadas da imagem transformada.

De acordo com [Gonzalez e Woods \(2018\)](#), é interessante salientar que as características da transformação preservam pontos, linhas retas e planos. Embora, uma limitação desta representação é encontrada em relação à translação, para o qual é necessário que um vetor coluna 2D seja aderido ao lado direito da matriz da Equação 2.8. Desta forma, utilizando as coordenadas homogêneas é possível expressar todas as transformações geométricas (escala, translação, rotação e cisalhamento), as mesmas que são expressadas utilizando uma matriz de 3×3 na forma da equação geral:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{T} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} t_{0,0} & t_{0,1} & t_x \\ t_{1,0} & t_{1,1} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.9)$$

onde a submatriz superior esquerda 2×2 é ortogonal, e t_x, t_y permitem a translação. Dito isso, podemos formalmente definir a translação como:

¹¹ <<https://gdal.org/>>

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.10)$$

Rotacionar uma imagem com um ângulo θ em relação a sua origem, pode resultar em sentido anti-horário, conforme é definido pela equação:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.11)$$

Objetos de uma imagem podem estar no centro da cena, como foi ilustrado na Figura 19. Nesse quesito, as Equações 2.10 e 2.11 podem ser operados como uma sequência de transformações fundamentais. Desta forma, uma operação de rotação em torno ao centro do objeto (\mathbf{R}) é definido como:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -t_x \\ 0 & 1 & -t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.12)$$

onde $-t_x, -t_y$ correspondem à translação do objeto ao origem, e t_x, t_y deslocam o objeto a sua posição original.

Embora, dado que objeto da Figura 19 está centrada, rotada e com fundos pretos, é necessario extrair previamente a região de interesse dessa área rotacionada. Uma abordagem que minimize essa área e calcule com uma melhor aproximação seu ângulo de inclinação, pode ser alcançada com o código de cadeia e área retangular mínima.

O código de cadeia é uma técnica sugerida por Freeman (1961), que descreve uma sequência de segmentos de linha com um conjunto de orientações definidas, e esses segmentos de linha são originalmente definidos como uma sequência de dígitos. Embora a desvantagem desse código seja diferente para diverso pontos de partida, sua invariância pode ser alcançada por uma variante publicada por Tyagi (2018), decodificando as cadeias originais para números inteiros, sendo realizadas ciclicamente até obter os inteiros menores.

Para calcular a área retangular mínima, um dos algoritmos a considerar é o *convex hull*, pois os eixos de um retângulo são girados no sentido anti-horário pelo menor ângulo que faz o retângulo coincidir com outra aresta do polígono. Nesse contexto, os ângulos caem constantemente em um intervalo de $[-90, 0)$, porque o próximo vértice desse polígono é utilizado para calcular o ângulo a partir da coordenada horizontal do plano da imagem.

Nesse sentido, o algoritmo para extrair a região de interesse esta composto pelo cálculo de uma sequência de contornos, seguido da definação de um polígono retangular a partir dos

contornos calculados, obtendo sua coordenada em referência ao centro (tx, ty) da região, sua dimensão, e seu ângulo θ . Em seguida, calcula-se a matriz rotada \mathbf{R} definida acima, onde seu ângulo é adicionado 90° por conta de sua direção anti-horário. Finalmente, é recortada (\cap) a nova matriz levando em consideração informações de sua nova dimensão. Assim, o algoritmo 2 descreve os passos envolvidos na extração da região de interesse.

Algorithm 2 Extração da região de interesse

Entrada: Matriz imagem \mathbf{F}_{rgb}

Saída: Matriz sub-imagem \mathbf{I}_{rgb}

```

7 início
8   contornos = códigoCadeia( $\mathbf{F}_{rgb}$ )
9   para contornos faça
10    |   retângulo = áreaMínima(contornos)
11   fim
12   centro, dimensão,  $\theta$  = retângulo
13    $\mathbf{M}$  = zeros(dimensão) // cria uma matriz com fundo preto
14    $\mathbf{M}_r$  =  $\mathbf{R}$ (centro,  $\theta + 90^\circ$ )
15    $\mathbf{I}_{rgb}$  =  $\mathbf{M} \wedge \mathbf{M}_r$  // recorta  $\mathbf{M}$  com  $\mathbf{M}_r$ 
16 fim
  
```

Como dito na Seção 2.4.1, sensores multiespectrais embarcados em Drone frequentemente podem coletar imagens com pouco contraste ou brilho. Independentemente da correção radiométrica, espacial e espectral, é importante ajustar os valores dos pixels da imagem para melhorar a qualidade ou aparência da imagem. Por conta disso, nesta pesquisa foram utilizados operações pontuais como o ajuste de brilho e contraste, equalização de histogramas locais, simultaneamente com espaços de cores.

2.6.2 Ajustes de brilho e contraste

Em geral, operações de ponto como ajuste de brilho e contraste, mudam valores dos pixels sem alterar o tamanho, geometria ou estrutura local da imagem (BURGER; BURGE, 2009). Para manipulação desses ajustes, Nixon e Aguado (2019) sugerem, que incrementar brilho estica o contraste, implicitamente pela operação da multiplicação, mudando todos os valores dos pixels por um escalar α . Da mesma forma, para reduzir o constraste se divide por esse mesmo escalar. Por outro lado, para incrementar ou decrementar brilho uma operação de adição é realizado pelo escalar β . A equação utilizada para ajustar brilho e contraste é definida como:

$$G(x, y) = \alpha F(x, y) + \beta \quad (2.13)$$

onde $F(x, y)$ é a imagem de entrada, $G(x, y)$, é a imagem de saída, α e β são parâmetros de constraste e brilho respectivamente. Além disso, esses parâmetros são frequentemente chamados de ganho e viés.

2.6.3 Ajuste de histograma

Segundo (THANKI; KOTHARI, 2019), o histograma de uma imagem representa a frequência relativa de ocorrências de diversos níveis de cinza, fornecendo uma descrição total da aparência de uma imagem. Desta forma, é possível estabelecer uma ponte natural entre imagens e uma descrição probabilística.

Formalmente, define-se o histograma como uma função de probabilidade de primer ordem $p_r(r_k; x, y) = n_k$ para indicar a probabilidade do pixel (x, y) com r_k ocorrências do k -ésimo nível de intensidade, e um número de pixels n_k em intervalos de $[0, L_* - 1]$. Devido à natureza da probabilidade, ou seja $0 \leq r_k \leq 1$, o histograma define-se melhor normalizado como:

$$p_r(r_k; x, y) = \frac{n_k}{MN} \quad (2.14)$$

De acordo com Szeliski (2010), o histograma de uma imagem com diversas intensidades de cinza, podem resultar com excesso de valores escuros e valores claros. Por conta disso, uma abordagem interessante seria, simultaneamente clarear esse excessos na faixa dinâmica de tons de cinza. Nesse quesito, se utiliza uma equalização de histograma, que calcula uma função de mapeamento S_k a partir de amostras aleatórias de uma função de distribuição uniforme, o qual discretamente é definido como:

$$S_k = (L_* - 1) \sum_{j=0}^k \frac{n_j}{MN} \quad (2.15)$$

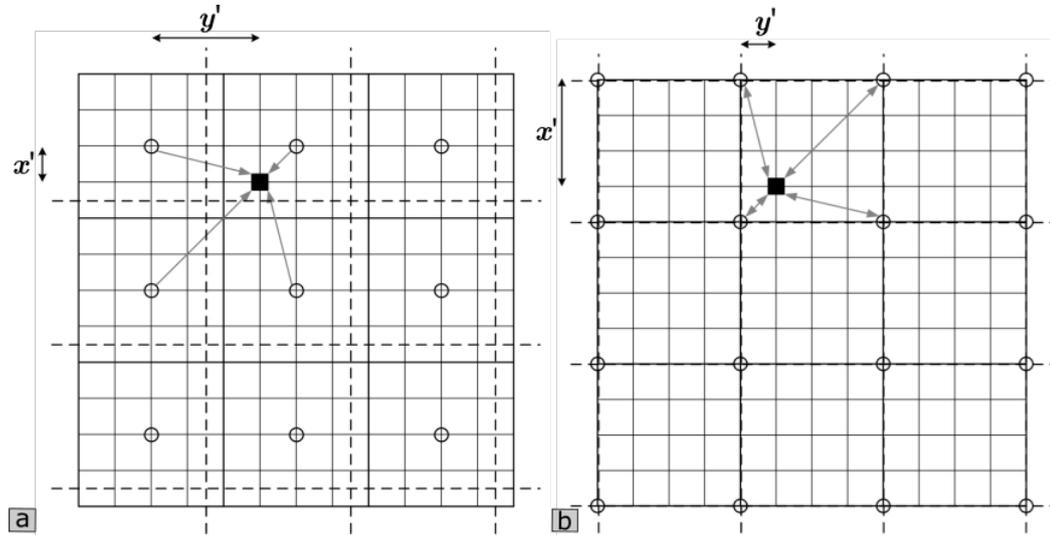
Em geral, operações de histograma anteriormente mencionadas são métodos globais, os quais não são apropriadas para realçar pequenos detalhes da imagem. Desse modo, que muitas vezes é desejável melhorar o contraste de regiões locais, isso poderia ser conseguido pela equalização adaptativa do histograma (SOLOMON; BRECKON, 2010).

Tal abordagem, foi ampliado por Pizer et al. (1987) conhecida como equalização de histograma adaptativa limitada por contraste (CLAHE, do inglês *Contrast Limited Adaptive Histogram Equalization*), o qual calcula apropriados histogramas de equalização baseadas em janela deslizante ou blocos não sobrepostos, seguindo uma sequência de passos: (i) sobrepõe uma grade de pontos na imagem, (ii) para cada ponto uma distância horizontal e vertical determina uma janela retangular, (iii) uma *lookup table* de histogramas equalizados é calculado para cada uma dessas janelas com vizinhança 4 (norte, sul, leste e oeste), e (iv) cada pixel do valor transformado é calculado com uma combinação de valores ponderados a partir dos valores da vizinhança. Assim, esses passos utilizam uma equação bilinear que é definida como:

$$S'_{x',y'} = (1 - x')(1 - y')F_1 + x'(1 - y')F_2 + (1 - x')y'F_3 + x'y'F_4 \quad (2.16)$$

onde x', y' são as distâncias ponderadas da grade de pontos de cada janela definida nos passos mencionados, conforme se ilustra na Figura 21-a.

Figura 21 – Interpolação de CLAHE utilizando coordenadas relativas (x', y'): (a) com pontos centrados nos blocos, (b) com pontos no cantos dos blocos. Fonte: Adaptado de Szeliski (2010)



De acordo com Szeliski (2010), uma pequena variante do algoritmo CLAHE coloca as *lookup tables* nos cantos de cada janela (Figura 21-b). Além de mezclar as quatro *lookup tables*, também distribui cada valor do pixel durante a fase de acumulação do histograma com um limiar de contraste nos passos mencionados anteriormente, os quais são definidos como:

$$h_{b_*, l_*}(F(x, y)) = \sum S'_{x,y}(x, y, b_*, l_*) \quad (2.17)$$

onde b_* é o tamanho do bloco, e l_* é o limiar do contraste aplicado em um determinado bloco.

2.6.4 Espaços de cores

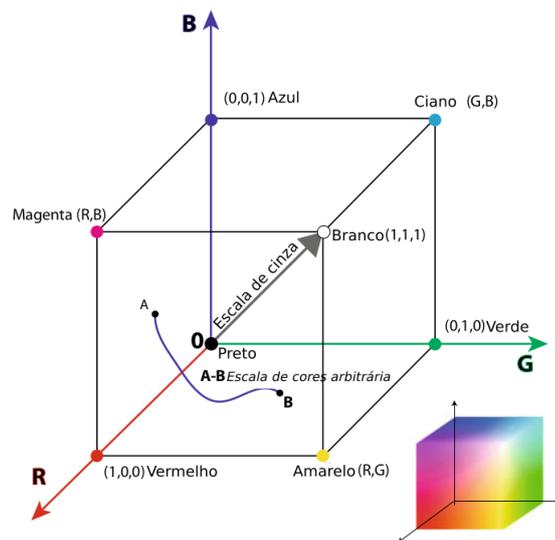
De acordo a teoria Young-Helmholtz, uma ampla gama de cores pode ser gerada pela mescla adequada de três feixes de luz visível com comprimentos de onda bem separados. Desta forma, três cores primários: vermelho, verde e azul, foram estabelecidos internacionalmente pelo CIE (*Commission International de L'eclairage*) em 1931 (DISTANTE; DISTANTE, 2020).

Forsyth e Ponce (2012), classificam os espaços de cores como: (i) lineares, composto pelo espaço RGB para monitores e câmeras de vídeo, CMY (*Cyan, Magenta, Yellow*) e CMYK (*Cyan, Magenta, Yellow, Black*) para impressoras; (ii) não lineares, composto pelo espaço HSV (*Hue, Saturation, Value*), HSI (*Hue, Saturation, Intensity*) e CIEL*a*b*. Em função disso, objetivo dos espaços de cores é facilitar a especificação de cores em alguma forma padrão, seja como um sistema de coordenadas ou como um subespaço desse sistema de coordenadas (GONZALEZ; WOODS, 2018).

2.6.4.1 Espaço de cor RGB

De acordo com [Solomon e Breckon \(2010\)](#), imagens RGB (cores puras) são matrizes 3D, que conceitualmente se compõe de planos 2D, cada um correspondentemente para cada canal de cores vermelho, verde e azul. Todas as cores podem ser representados em um espaço 3D (cubo) com eixos R, G e B (Figura 22), onde cada eixo esta em um mesmo intervalo de $0 \rightarrow 1$ (escalado com $0 - 255$ para 1 byte por canal de cor). Desta forma, a diagonal dos pontos entre a origem do cubo (posição $(0, 0, 0)$ – cor preta) e o canto oposto (posição $(1, 1, 1)$ – cor branca) representam tons de cinza (cores acromáticas).

Figura 22 – Espaço de cor RGB representado como um cubo. Fonte: Adaptado de [Distante e Distante \(2020\)](#).



2.6.4.2 Espaço de cor XYZ

Curvas relativas da distribuição espectral às coordenadas tricromáticas, tomam valores negativos e esse representa uma desvantagem na representação de cores em dispositivos eletrônicos. A CIE, definiu cores primárias virtuais **XYZ**, onde componente **Y** representa o valor tristímulo da luminância de cor a ser comparada, ou seja, **Y** coincide com a função da eficiência luminosa fotópica $V(\lambda)$. Portanto, o plano **XY** coincide com o plano da luminância de valor zero ([DISTANTE; DISTANTE, 2020](#)). Formalmente **Y** pode ser calculado como:

$$L_c(\lambda) = Y = 683 \int_{sv} V(\lambda) M(\lambda) d\lambda \quad (2.18)$$

onde $L_c(\lambda)$ é a luminância de uma cor com comprimento de onda (λ) , $M(\lambda)$ é a potência espectral. Dito isso, a transição das coordenadas tristímulo RGB, para as coordenadas de cores virtuais **XYZ** é dado pela seguinte transformação linear:

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} = \begin{bmatrix} 0,4125 & 0,3576 & 0,1804 \\ 0,2127 & 0,7152 & 0,0722 \\ 0,0193 & 0,1192 & 0,9503 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.19)$$

Da mesma forma a transição das coordenadas de cores virtuais \mathbf{XYZ} para coordenadas RGB é dado pela seguinte transformação linear:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3,2405 & -1,5371 & -0,4985 \\ -0,9693 & 1,8760 & 0,04156 \\ 0,0556 & 0,2040 & 1,0572 \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{bmatrix} \quad (2.20)$$

2.6.4.3 Espaço de cor YCbCr

De acordo com [Szeliski \(2010\)](#), este espaço de cor é extensivamente utilizado em vídeos digitais (NTSC e PAL). Por conta disso, informação de luminância esta representado pelo componente Y, entanto informações de cores são armanzenadas em componentes Cb e Cr respectivamente. É interessante salientar, que o componente Cb é uma diferença entre a cor azul e um valor de referência dado, enquanto componente Cr é a diferença entre a cor vermelha e um valor de referência dado. Dito isso, formalmente a transição do espaço de cor RGB para YCbCr é dada por:

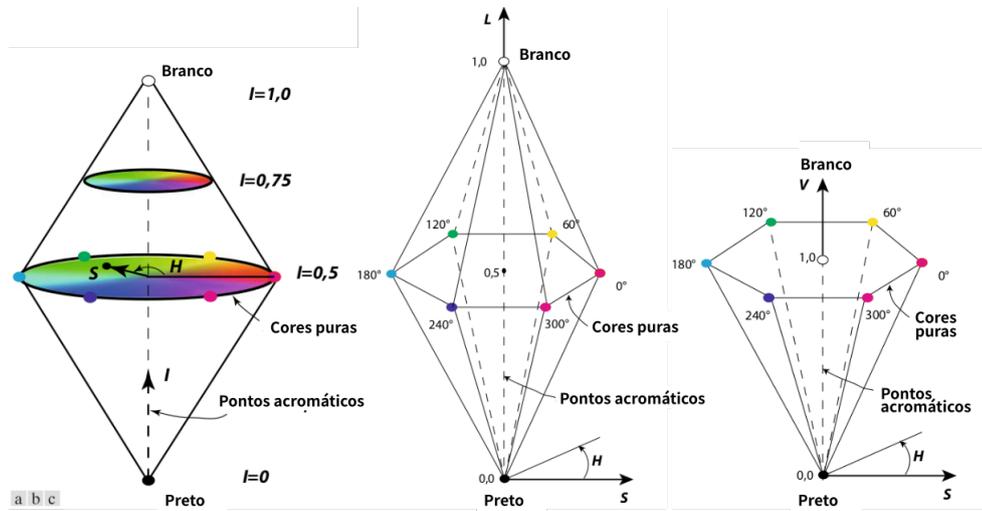
$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0,2990 & 0,5870 & 0,1140 \\ -0,1687 & -0,3313 & 0,5000 \\ 0,5000 & -0,4187 & 0,0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.21)$$

De acordo com [Solomon e Breckon \(2010\)](#), espaços de cores RGB tem algumas limitações por ser perceptualmente (estímulos) não lineares, devido a que sua representação cubica não necessariamente reproduz a cor consistentemente. Já [Distante e Distante \(2020\)](#), acrescenta que quantidades subjetivas e físicas de cores devem ser adaptativas às condições do entorno. Nesse quesito, espaços de cores RGB e XYZ podem definir outros espaços de cores uteis em várias aplicações, especialmente para reprodução de cores, PDI, e aquisição de imagens baseados em sensores.

2.6.4.4 Espaço de cor HSV

Apresentado por [Smith \(1978\)](#), como um modelo de cor que define três componentes: matiz (H, do inglês *Hue*), saturação (S, do inglês *Saturation*) e brilho (V, do inglês *Value*). Tais componentes, como a matiz determina uma cor ou tounalidade, saturação que determina a pureza da cor (esmaecida a intensa), e brilho que determina sua intensidade (cor mais clara ou escura). A representação é exibida pela Figura 23-c, na forma de um cone 3D invertido, com base hehagonal 2D para cores puras.

Figura 23 – Representações geométricas de espaços de cores (a) HSI (b) HSL e (c) HSV.
Fonte: Adaptado de Distante e Distante (2020).



De acordo com Blanchet e Charbit (2006), espaços de cores similares são definidos como: (i) modelo HSI, que representa cores em termos do matiz, saturação e intensidade (I, do inglês *Intensity*) que mede a luz associada a uma certa energia espectral percebida; e (ii) modelo HSL, definido pelos parâmetros matiz, saturação e luminosidade (L, de inglês *lightness*) que expressam a luminância de um objeto na forma da capacidade de sua refletância. Desta forma, a Figura 23-a e a Figura 23-b demonstram de maneira visual os espaços HSI e HSL.

Dito isso, uma transição do espaço RGB para espaço HSV faz uso de componentes R, G, B, a fim de encontrar seus respectivos valores máximos e mínimos, conforme é definido na equação:

$$H = \begin{cases} 60^\circ \left(\frac{G-B}{M'-m'} \right), & \text{se } M' = R \\ 60^\circ \left(\frac{B-R}{M'-m'} \right) + 120^\circ, & \text{se } M' = G \\ 60^\circ \left(\frac{R-G}{M'-m'} \right) + 240^\circ, & \text{se } M' = B \end{cases} \quad (2.22)$$

$$S = \begin{cases} \frac{M'-m'}{M'}, & \text{se } M' \neq 0 \\ 0 & \text{caso contrário} \end{cases}$$

$$V = M'$$

onde M' e m' representam o máximo e mínimo dos tristímulos R, G e B. Desta forma, a transição do espaço HSV para espaço RGB é definido como:

$$C = V \times S$$

$$X' = C(1 - |\frac{H}{60^\circ} \text{ mod } 2 - 1|)$$

$$(R_1, G_1, B_1) = \begin{cases} (C, X', 0), & \text{se } 0^\circ \leq H < 60^\circ \\ (X', C, 0), & \text{se } 60^\circ \leq H < 120^\circ \\ (0, C, X'), & \text{se } 120^\circ \leq H < 180^\circ \\ (0, X', C), & \text{se } 180^\circ \leq H < 240^\circ \\ (X', 0, C), & \text{se } 240^\circ \leq H < 300^\circ \\ (C, 0, X'), & \text{se } 300^\circ \leq H < 360^\circ \end{cases} \quad (2.23)$$

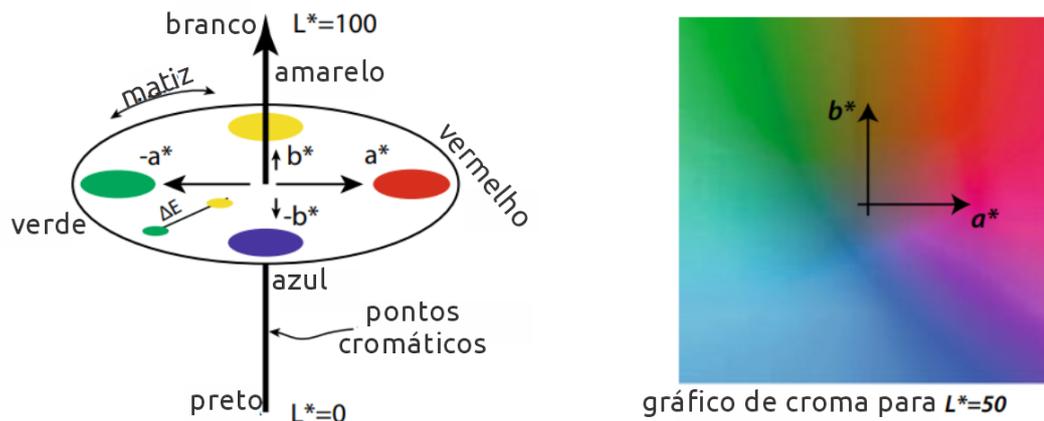
$$m_* = V - C$$

$$(R, G, B) = (R_1 + m_*, G_1 + m_*, B_1 + m_*)$$

2.6.4.5 Espaço de cor CIEL*a*b*

Adotado em 1976 para classificação de cores de acordo com o sistema Munsell e para estimar quantitativamente as diferenças de cores. Este espaço de cores é baseado na teoria das cores opostas, onde uma cor não pode ser verde e vermelha ao mesmo tempo, nem amarela ou azul. A Figura 24 ilustra uma representação geométrica 3D com um plano cromático e seus eixos a^* , b^* .

Figura 24 – Representação geométrica do espaço de cor CIEL*a*b*. Fonte: Adaptado de **Distante e Distante (2020)**.



Para a conversão do espaço RGB para o espaço CIEL*a*b*, são necessárias previamente as coordenadas de cores virtuais XYZ (Equação 2.20). Por conta disso, a conversão RGB para CIEL*a*b* é definida como:

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16, & \text{se } \frac{Y}{Y_n} > 0,0089 \\ 903,3 \left(\frac{Y}{Y_n} \right), & \text{se } \frac{Y}{Y_n} \leq 0,0089 \end{cases} \quad (2.24)$$

$$a^* = 500 \left[\left(\frac{X}{X_n} \right)^{\frac{1}{3}} - \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} \right]$$

$$b^* = 200 \left[\left(\frac{X}{X_n} \right)^{\frac{1}{3}} - \left(\frac{Z}{Z_n} \right)^{\frac{1}{3}} \right]$$

onde $\mathbf{X}_n, \mathbf{Y}_n, \mathbf{Z}_n$ descrevem um iluminador de referência acromático branco em \mathbf{XYZ} .

Pré-processamento de imagens com algoritmos para aprimoramento de brilho e contraste em canais RGB por separado, podem alterar algumas proporções das componentes RGB, além de causar alguns artefatos (distorções) de saturação ou mistura de cores (HITAM et al., 2013).

Neste contexto, para esta pesquisa foram utilizados algoritmos com operações pontuais, em forma global (Equação 2.13) e operações de vizinhança na forma local (Equação 2.17), em conjunto com os espaços de cores (\mathcal{S}^c) YCbCr, HSV e $L^*a^*b^*$. Assim, foram definidos os parâmetros α e β para controle de brilho e contraste globalmente, e parâmetros b_*, l_* para histograma local (CLAHE). Dito isso, define-se operadores de transformação para cada um dos espaços de cores (\mathcal{T}), para processamento de certos canais relacionados à luminância com algoritmo CLAHE (\mathcal{E}_{clahe}). Desta forma, imagem pré-processada é transformado para seu espaço RGB gradualmente, conseguindo assim o aprimoramento desejado. Esses processos, por sua vez são especificados no Algoritmo 3 de forma iterativa por uma sequência k vezes.

Algorithm 3 Ajuste brilho e contraste com espaços de cores

Entrada: Matriz imagem RGB: I_{rgb} , escalares: $k, \alpha, \beta, b_*, l_*$ e cadeia de cor: s

Saída: Matriz imagem R'G'B': I'_{rgb}

```

17 início
18    $I'_{rgb} = I_{rgb}$ 
19   para  $i = 0$  até  $i < k - 1$  faça
20      $I'_{rgb} = \alpha I'_{rgb} + \beta$ 
21     se  $s \in \mathcal{S}^c$  então
22       //  $I_t$ : imagem transformada
23        $I_t = \mathcal{T}_{\mathcal{S}^c}(I'_{rgb})$ 
24        $I_t = \mathcal{E}_{clahe}(I_t, b_*, l_*)$ 
25        $I'_{rgb} = \mathcal{T}_{rgb}(I_t)$ 
26     fim
27 fim
```

2.6.5 Filtragem no domínio espacial

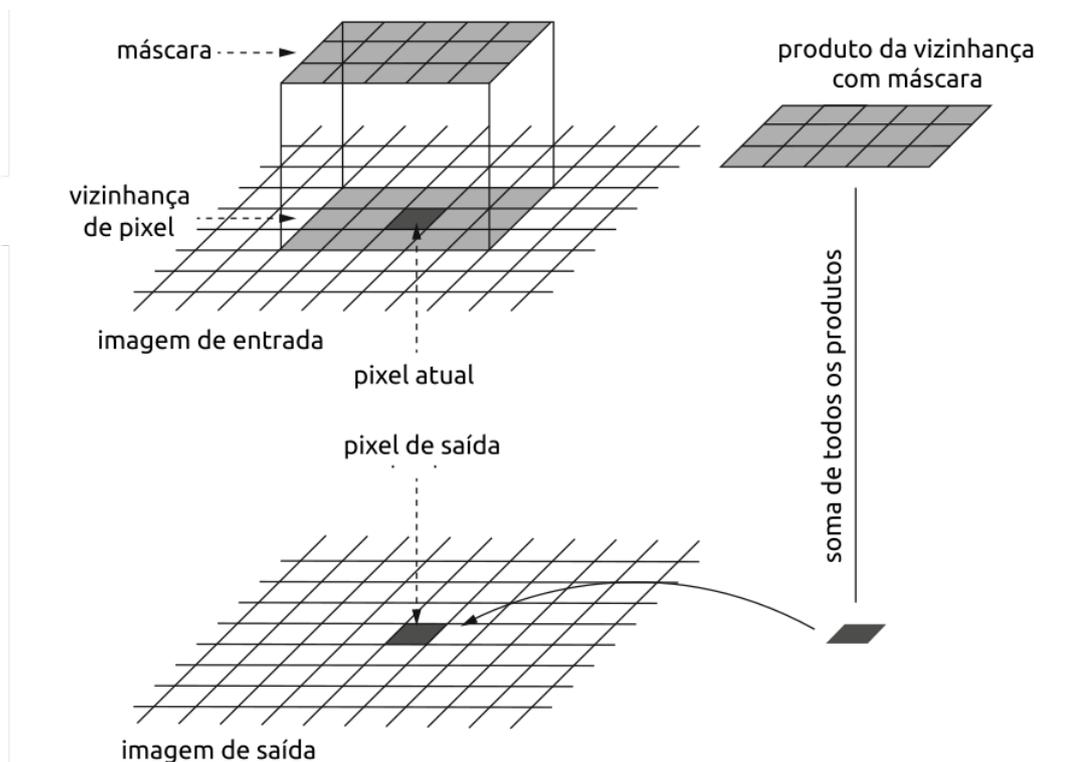
De acordo com Bhuyan (2019), a filtragem no domínio espacial executa operações para aprimorar ou suavizar características da imagem, mudando os valores dos pixels por meio de operações de vizinhança, melhorando assim a qualidade da imagem. É interessante destacar, que operações de filtragem podem ser classificadas como lineares ou não lineares, e também como filtragem de passa-baixa, passa-alta ou passa-banda (GONZALEZ; WOODS, 2018).

Uma expressão para filtragem, está definido formalmente pela convolução entre duas funções k e F , que produz uma nova função, o qual representa, o grau de solapamento da função k sobre a função F . Desta forma, a convolução discreta 2D é denotada como:

$$F(x, y) \star k(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b k(s, t) F(x - s, y - t) \quad (2.25)$$

onde, $F(x, y)$ representa a imagem; $k(x, y)$ representa o kernel, filtro ou máscara; a, b são os limites para percurso da máscara com $a = (m' - 1)/2$, $b = (n' - 1)/2$ assumindo que m' e n' são inteiros impares. Seu funcionamento se baseia na avaliação dos valores da máscara em (s, t) localizações, sobre a imagem completa de coordenadas (x, y) . Assim, a Equação 2.25 garante a filtragem refletida horizontal, vertical e girada em 180°. Por conta disso, a Figura 25 ilustra o processo de convolução.

Figura 25 – Processo de filtragem espacial por convolução. Fonte: Adaptado de (THANKI; KOTHARI, 2019).



2.6.5.1 Filtro da média

Esta classificada como um filtro de passa-baixa e linear, essencialmente esta filtragem, realiza a suavização de imagem, ou seja, frequências altas na imagem correspondentes às transições abruptas são atenuadas. Desta forma, é caracterizado por um leve borramento, minimizando o efeito de ruído. Assim, em relação à operação de convolução a máscara altera os valores dos pixels da imagem pela média de seus vizinhos. Na Figura 26 é ilustrado alguns exemplos de filtro média, cuja soma de pesos igual a 1.

Figura 26 – Exemplos de filtragem da média para suavização.

$$\frac{1}{9} \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad \frac{1}{10} \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 2 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad \frac{1}{16} \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

É importante salientar, que máscaras pequenas atenuam o ruído (sinal uniforme), máscaras maiores têm efeito de borrão e a perda de alguns detalhes fica cada vez mais evidente.

2.6.5.2 Filtro de Gaussiano

Similar à filtro da média, Gaussiana é um filtro de passa-baixa e linear, nesse sentido é utilizado como filtro de suavização na imagem. Seu uso é mais adequado para o tratamento de ruído gaussiano (ruído branco), tornando-se um processo estocástico com média (μ) zero. Os coeficientes da máscara são derivados a partir de uma função Gaussiana 2D (Figura 27-a). Por conta disso, sua definição discreta é dado por:

$$k(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2.26)$$

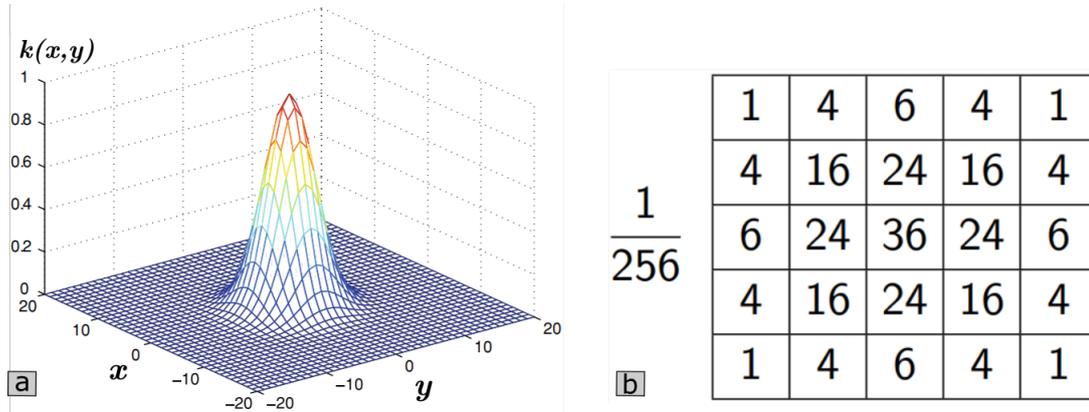
onde x, y representa as coordenadas da máscara, σ é desvio padrão e está associado à distribuição de probabilidade, que controla a quantidade de suavização.

De acordo com [Burger e Burge \(2009\)](#), vale ressaltar que funções Gaussianas são simétricas com relação à rotação, ou seja, suavização é realizado em todas as direções tornando-se um filtro isotrópico. Além disso, funções Gaussianas são separáveis, portanto a convolução Gaussiana pode ser realizada com um Gaussiano unidirecional, levando a reduzir o número de operações na convolução. A Figura 27-b ilustra a máscara Gaussiana.

2.6.5.3 Filtro da mediana

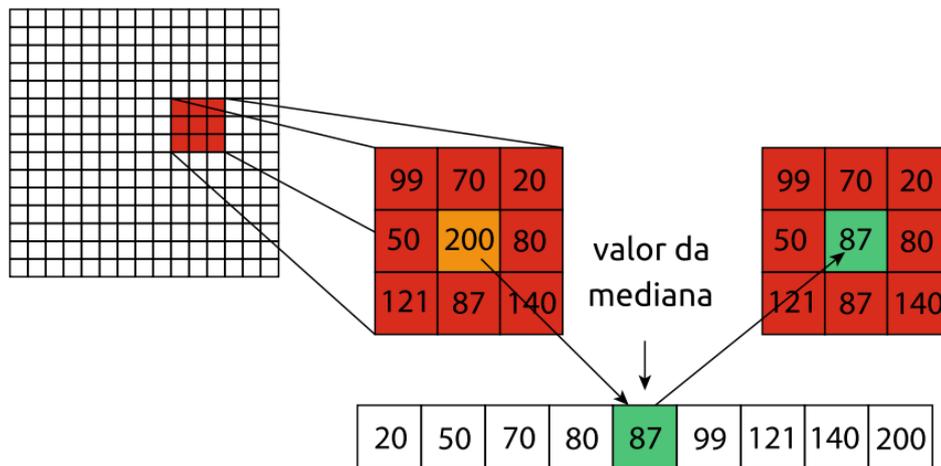
Diferente aos filtros anteriormente mencionados, a mediana é um filtro de passa-baixa, não linear e de ordem estatístico. O filtro consiste em ordenar os pixels da vizinhança, assim substituir o píxel central pela mediana da intensidade nessa vizinhança ([SUNDARARAJAN](#),

Figura 27 – Exemplo de um filtro Gaussiano, (a) representação bidimensional com $\sigma = 3$, e (b) máscara Gaussiano. Fonte: Adaptado de Distante e Distante (2020).



2017). Por conta disso, o processo minimiza bastante o efeito de borramento além de reduzir efeito de ruído impulsivo do tipo sal e pimenta. É interessante salientar que o filtro da mediana não introduz os valores de níveis de cinza diferentes daqueles que estão contidos na imagem original, nesse sentido afeta menos as bordas e sua aplicação pode ser utilizada iterativamente. Na Figura 28 é ilustrado os efeitos de aplicar filtro mediano para remover ruído impulsivo.

Figura 28 – Diagrama funcional do filtro da mediana. Fonte: Adaptado de Distante e Distante (2020).



2.7 Processamento morfológico

A morfologia teve seus inícios na biologia, para estudar a forma e a estrutura dos organismos desde o século XVII. Sua definição matemática foi proposta por Serra (1986) na teoria dos conjuntos. Desta forma, operadores morfológicos para PDI foram representados com operações da teoria de conjuntos, originalmente para imagens binárias (valores dos pixels em 0 – fundo, ou 1 – objeto), sendo estendidos para imagens de tons cinzas e cores. Viabilizando a primeira aplicação de operadores morfológicos na área da microscopia.

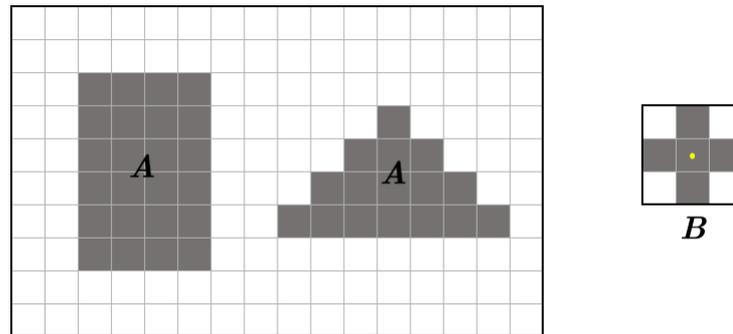
PDI, constantemente esta tratando formas ou objetos de uma imagem como um grupo de pixels, valores desses pixels podem ser representados como conjuntos (NIXON; AGUADO, 2019), descrevendo quantitativamente sua estrutura topológica (BHUYAN, 2019).

Nesse quesito, defina-se um conjunto de pares inteiros para imagem binária com coordenadas $(x, y) \in \mathbb{Z}^2$, onde os pontos pertencentes à imagem binária são representados em um conjunto A – objeto de pixels com valor 1, e pontos complementos são representados em um conjunto A^c – fundo com pixels de valor 0. Formalmente o objeto A é definido como:

$$A = \{(x, y) \mid (x, y) = 1\} \quad (2.27)$$

De acordo com Gonzalez e Woods (2018), operações morfológicas em PDI precisam de um elemento estruturante B especificados em termos de pixels de fundo e objeto, com diversas formas (quadrada, retangular, circular, etc.), tamanho e um ponto de origem. A Figura 29 ilustra objeto e elemento estruturante para PDI.

Figura 29 – Exemplo de objeto e elemento estruturante para PDI, onde o ponto amarelo representa a origem de B.



Similar às máscaras na convolução, elementos estruturantes processam pixels em vizinhança por meio de conceitos de reflexão e translação (SZELISKI, 2010). Desta forma, a reflexão de B sobre su origem é denotado por \hat{B} , sendo definido da seguinte forma:

$$\hat{B} = \{\hat{w} \mid \hat{w} = -b, \text{ para } b \in B\} \quad (2.28)$$

onde \hat{w} representam as coordenadas (x, y) . Desta forma B é substituído por $(-x, -y)$ com rotação de 180° sobre sua origem. Dito isso, a translação de B pelo ponto $\hat{z} = (z_1, z_2)$, denotado pelo $B_{\hat{z}}$ é definido como:

$$B_{\hat{z}} = \{\hat{c} \mid \hat{c} = b + \hat{z}, \text{ para } b \in B\} \quad (2.29)$$

onde as coordenadas de B são substituídas por \hat{z} , ou seja, por $(x + z_1, y + z_2)$, conseguindo deslizar B sobre uma imagem como a operação de convolução.

2.7.1 Dilatação e erosão

De acordo com Dougherty (2009), operação de dilatação aumenta o tamanho dos objetos (A) em primeiro plano, geralmente usando pixels brancos, embora em algumas implementações essa convenção seja invertida. Desta forma, o centro do elemento estruturante (B) é sobreposto à primeiro píxel da imagem, onde pelo menos um píxel pode coincidir com o píxel do objeto, e sua saída será o valor do objeto mais o píxel do elemento estruturante, assim o fundo se torna parte de um objeto de primeiro plano. Dito isso, podemos formalmente denotar a dilatação por \oplus , e defini-lo como:

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\} \quad (2.30)$$

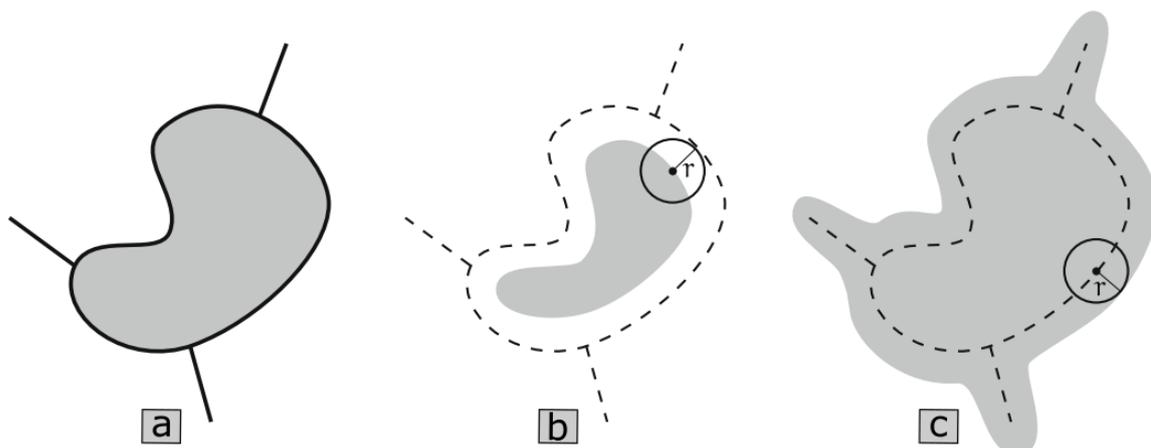
onde B esta refletindo sobre a origem, e por sua vez essa reflexão está sendo transladado por z .

Por outro lado, a erosão é uma operação que aumenta o tamanho do fundo e diminui os objetos do primeiro plano. Para isso, o elemento estruturante percorre a imagem de forma semelhante à dilatação, com a diferença de que os objetos vizinhos retornam parte do plano de fundo. Além disso, se houver fundos (buracos) dentro de objeto, a saída estará composto de fundos de maior tamanho (DOUGHERTY, 2009). Formalmente, a erosão pode ser denotado por \ominus , sendo definida como:

$$A \ominus B = \{z \mid B_z \subseteq A\} \quad (2.31)$$

onde objeto A esta sendo transladado por B em conjunto de todos os pontos z tais que B está contido em A . A Figura 30 exhibe um exemplo de operações morfológicas de erosão e dilatação, junto a um elemento estruturante com tamaho de rádio r centrado na origem:

Figura 30 – Exemplo de operações morfológicas da erosão e dilatação, (a) objeto dentro de uma imagem digital, sendo operados pela (b) erosão e (c) dilatação, junto a um elemento estruturante circular. Fonte: Adaptado de Velho et al. (2009)



2.7.2 Abertura e fechamento

De acordo com [Sonka et al. \(2015\)](#), erosão e dilatação não são transformações inversas, ou seja, se uma imagem é erodida e depois dilatada, a imagem original não é obtida novamente. Em vez disso, o resultado é uma versão simplificada e menos detalhada da imagem original, também chamada de cancelamento ([DAVIES, 2017](#)). Nesse quesito, outras operações em morfologia matemática que expressem o grau de cancelamento podem ser derivadas dos filtros de erosão e dilatação, entre as principais, citam-se a abertura e fechamento.

A abertura de um objeto A por B , é denota por $A \circ B$, ou seja, é definida como uma erosão seguida por uma dilatação, como segue:

$$A \circ B = (A \ominus B) \oplus B \quad (2.32)$$

Enquanto, o fechamento de um objeto A por B , é denotado por $A \bullet B$, ou seja, é definida como uma dilatação seguida por uma erosão, como segue:

$$A \bullet B = (A \oplus B) \ominus B \quad (2.33)$$

De acordo com [Sonka et al. \(2015\)](#), operações de abertura e fechamento com elemento estruturante isotrópico, são frequentemente utilizados para eliminar detalhes específicos da imagem, menores que o elemento estruturante, ou dito melhor, a forma global dos objetos não é distorcida, conseguindo em certo grau de invariância à translação do elemento estruturante.

2.8 Segmentação de imagen

Imagens podem conter diversas regiões, os quais estão conectados e apresentam características distintas em relação a alguma medida, como proximidade espacial (histograma, distância euclidiana) e similaridade de atributos (cor, textura) ([SUNDARARAJAN, 2017](#)). Assim, após realce dessas características desejadas, é possível segmentar uma imagem em diferentes regiões homogêneas. De acordo com [Distante e Distante \(2020\)](#), regiões homogêneas são agrupamentos de pixels, mas essas regiões podem ser hipotéticas. Por conta disso, a segmentação não necessariamente implica uma classificação das regiões. Além, de não fornecer uma teoria físico-matemática específica. Consequentemente, existem diversos algoritmos, que podem funcionar melhor ou pior dependendo do tipo da imagem.

2.8.1 Binarização

Um processo fundamental na segmentação de imagens, a binarização utiliza informação espacial das intensidades de imagem $F(x, y)$, desta forma torná-la em uma imagem binária $G(x, y)$ a partir de um limiar global. Esta imagem binária mapeada é definido por:

$$G(x, y) = \begin{cases} 1, & \text{se } F(x, y) \geq L \\ 0, & \text{se } F(x, y) < L \end{cases} \quad (2.34)$$

onde L representa o valor limiar global das intensidades de imagem, que pode ser obtido por diversos métodos. Alguns dos mais utilizados se baseiam em histogramas de limiar global (GONZALEZ; WOODS, 2018), ou limiar automático de Otsu (OTSU, 1979).

De acordo com Distante e Distante (2020), a binarização de imagen com intervalos de intensidades $[L_1, L_2]$ associados às regiões de imagem, pode se formalmente obtida por:

$$G(x, y) = \begin{cases} 1, & \text{se } F(x, y) \in [L_1, L_2] \\ 0, & \text{caso contrário} \end{cases} \quad (2.35)$$

Com objetivo de segmentar imagens de plantas de milho, neste trabalho foi considerado o algoritmo de segmentação baseado em intervalos de intensidades sobre espaço cor HSV. Assim, foi utilizado uma filtragem Gaussiana para matriz do Algoritmo 3. Em sequência, uma transformação para o espaço de cor HSV (\mathcal{T}_{hsv}) foi realizado, extraíndo valores de limiares mínimos e máximos para cada canal (H, S e V) em relação aos cores de plantas de milho. Para binarização de imagem HSV é comparado cada um dos limiares L_1, L_2 a cada canal, onde limiares pertencentes ao objeto de planta de milho, são representados com cor branco (valor 1), caso contrário em fundo (valor 0). Dito isso, para garantir uma segmentação global com bons resultados optou-se por utilizar operação morfológica de fechamento com elemento estruturante elíptico isotrópico para fornecer invariância na translação. Com tudo, o Algoritmo 4 destaca os passos envolvidos para a segmentação de planta de milho baseado espaço HSV.

Algorithm 4 Segmentação de planta milho baseado em espaço de cor HSV

Entrada: Matriz imagem rgb: I_{rgb} , escalares tamanho máscara e desvio padrão: b_*, σ

Saída: Matriz binária: I_b

28 **início**

29 $I_H, I_S, I_V : 45 - 80, 100 - 255, 100 - 255$

30 $I_{rgb} = \text{Gaussiana}(b_*, \sigma)$

31 $I_t = \mathcal{T}_{hsv}(I_{rgb})$

32 $L_1, L_2 = \min(I_t, I_H, I_S, I_V), \max(I_t, I_H, I_S, I_V)$

33 $I_b = \text{zeros}(I_t.\text{dimensão})$

34 **para** $i, j = 0$ **até** $i, j < I_t$ **faça**

35 **se** $I_t(i, j) \in [L_1(i, j), L_2(i, j)]$ **então**

36 $I_b(i, j) = 1$

37 **fim**

38 **fim**

39 $B \leftarrow \text{criar elemento estruturante circular de tamanho } b_*$

40 $I_b = (I_b \oplus B) \ominus B$

41 **fim**

2.8.2 Transformada de distância

Em relação as imagens aéreas que registram as formas do dossel das plantas de milho, quando são processados pela etapa de segmentação estas informações não necessariamente representam a topologia da forma do dossel, sendo separadas ou juntadas por regiões de pixels segmentadas em forma bruta. Nesse quesito, é necessário de métodos que representem as regiões segmentadas em forma topológica ou geométrica, como tamanho, orientação ou distância.

De acordo com [Szeliski \(2010\)](#), a distância de uma região segmentada pode ser calculada utilizando medidas de distâncias. Nesse sentido, o algoritmo da transformada de distância proposto por [Rosenfeld e Pfaltz \(1966\)](#), calculam a distância de um determinada região que se encontra à um contorno mais próximo do objeto, ou seja, os pontos de valores mais altos em forma de informação local. Dito isso, medidas de distância frequentemente utilizadas, são formalmente definidas como:

$$d_1(x, y) = |x - s| + |y - t| \quad (2.36)$$

$$d_2(x, y) = \sqrt{(x - s)^2 + (y - t)^2} \quad (2.37)$$

onde s, t indicam o pixel vizinho, d_1 representa a distância Manhattan e d_2 representa a distância Euclidiana. Assim, calcular distâncias de regiões para selecionar o menor valor deles, pode ser computacionalmente ineficiente, se algumas regiões fossem grandes ([MCANDREW, 2016](#)). A transformada de distância minimiza esse calculo, pela seguinte definição:

$$D(x, y) = \min_{s,t:\mathbf{k}(s,t)} d(x - s, y - t) \quad (2.38)$$

onde $D(x, y) \in \mathbb{R}$, \mathbf{k} é uma máscara adequada para ser usada na transformada de distância.

A transformada de distância, calcula informação local de uma região baseado em uma sequência de passos: (i) cada pixel é rotulado em relação à distância de região, se pixel pertencer à região é rotulado com 0, caso contrario com ∞ ; (ii) calcular distância de cada pixel em vizinhança (Equação 2.38) em conjunto de sua máscara \mathbf{k} , o valor mínimo encontrado é salvo no pixel central, considerando somas com ∞ ; (iii) repetir passo 2 até que todos os rótulos dos pixels visitados sejam valores finitos; (iv) finalmente dividir cada pixel pelo tamanho da máscara.

Embora, os passos mencionados sejam mais utilizados com distância Manhattan, para generalizar sua aplicação uma abordagem mais eficiente foi publicado por [Borgefors \(1988\)](#), conhecido como algoritmo de chamfer. Esse algoritmo, inclui máscaras propagados duas vezes como uma onda. A primeira começa no canto superior esquerdo, e a segunda começa na direção oposta de baixo para cima, ambos são propagados em uma direção diagonal ([BURGER; BURGE, 2009](#)). Nesse sentido, neste trabalho é definido as máscaras de distância euclidiana:

$$\mathbf{k}_{ezq} = \begin{bmatrix} \sqrt{2} & 1 & \sqrt{2} \\ 1 & \times & . \\ . & . & . \end{bmatrix}, \mathbf{k}_{der} = \begin{bmatrix} . & . & . \\ . & \times & 1 \\ \sqrt{2} & 1 & \sqrt{2} \end{bmatrix} \quad (2.39)$$

onde “×” indica localização central e não é tomado em conta no cálculo da distância. A Figura 31 exibe um exemplo do cálculo da transformada de distância de objeto retangular.

Figura 31 – Exemplo de resultado da aplicação da transformada de distância, sobre uma região retangular com distância Euclidiana. Fonte: Adaptado de Gonzalez e Woods (2018)

0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	1	1	1	1	1	1	1	0	0
0	0	1	2	2	2	2	2	1	0	0
0	0	1	2	3	3	3	2	1	0	0
0	0	1	2	2	2	2	2	1	0	0
0	0	1	1	1	1	1	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Com objetivo de obter a localização do dossel das plantas de milho, os quais estão dispostos em linhas pelas imagens aéreas. A etapa de segmentação conseguiu separar essas plantas de milho de fundo. Porém, a máscara global retornada não providenciaram localizações de cada possível planta. Por conta disso, um algoritmo que calcule a distância à localização central foi necessariamente desenvolvido com auxílio do algoritmo de chamfer.

Nesse contexto, para esta pesquisa foram utilizadas uma matriz de imagem binarizada para seu respectivo cálculo de informação geométrica local, considerando uma máscara de tipo euclidiana com tamanho de 3x3 para facilidade de cálculos além de no causar possíveis demoras. A sequência do algoritmo seguiu os passos recomendados pelo algoritmo de chamfer, mas o mapa de distâncias D' foi criado primeiramente com uma matriz de zeros para depois deixar só objeto, já preenchido com a recomendação do algoritmo original. Para os cálculos da distância, foi utilizado a vizinhança quatro com uma metodologia de desenvolvimento *top-down* e *bottom-up*. Finalmente, a sequência dos passos do Algoritmo 5 calcula o mapa de distâncias em relação à localização de possíveis dosséis da planta de milho.

2.9 Reconhecimento de padrões

Em 1960, como resultado de pesquisas teóricas na área de estatística, surgiu o reconhecimento de padrões (PR, *Pattern Recognition*), onde muitas das aplicações práticas foram utilizadas na indústria e na engenharia, construindo sistemas de tomada de decisão (BISHOP, 2006; THEODORIDIS; KOUTROUMBAS, 2008).

Algorithm 5 Mapa de distâncias por algoritmo de chamfer**Entrada:** Matriz segmentada I_b , escalar tamanho de máscara: b_* , cadeia tipo máscara: s **Saída:** Matriz mapa de distâncias D'

```

42  $M, N = \text{tamanho}(I_b)$  // M: linhas, N: colunas
43  $D' = \text{zeros}(M, N)$ 
44 para  $i, j = 0$  até  $i, j < M, N$  faça
45     se  $D'(i, j) \notin I_b(i, j)$  então
46          $D'(i, j) = \infty$ 
47     fim
48 fim
49  $k \leftarrow \text{criar máscara } s \text{ de tamanho } b_*$ 
50 para  $i, j = 0$  até  $M, N$  faça
51     se  $D'(i, j) > 0$  então
52          $d_1, d_2 = k_{\text{ezq},1} + D'(i-1, j), k_{\text{ezq},2} + D'(i-1, j-1)$ 
53          $d_3, d_4 = k_{\text{ezq},3} + D'(i, j-1), k_{\text{ezq},4} + D'(i+1, j-1)$ 
54          $D'(i, j) = \min(d_1, d_2, d_3, d_4)$ 
55     fim
56 fim
57 para  $i, j = M, N$  até  $0, 0$  faça
58     se  $D'(i, j) > 0$  então
59          $d_1, d_2 = k_{\text{der},1} + D'(i+1, j), k_{\text{der},2} + D'(i+1, j+1)$ 
60          $d_3, d_4 = k_{\text{der},3} + D'(i, j+1), k_{\text{der},4} + D'(i-1, j+1)$ 
61          $D'(i, j) = \min(d_1, d_2, d_3, d_4)$ 
62     fim
63 fim
64  $D' = D' \ominus k$ 

```

Segundo Dougherty (2013), o objetivo principal do PR é a classificação de objetos em uma série de categorias ou classes, onde uma das operações críticas envolve a extração e a seleção de características representativas expressas por meio de um vetor de características:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (2.40)$$

onde x_i , é o i -ésimo descritor da características, e n é o número total de descritores. Desta forma, vetor padrão pode representar um ponto no espaço Euclidiano n -dimensional, e classe padrão pode ser interpretada como hipernuvem de pontos desse espaço (GONZALEZ; WOODS, 2018).

Tarefas fundamentais como a detecção de formas ou padrões é realizada por meio de casamento de modelos (do inglês *template matching*), o qual esta baseado na correlação entre um template (modelo base) e a imagem (WILSON; RITTER, 2000). Assim, em relação ao mapa de distâncias de plantas de milho, nesta sessão foi de interesse procurar dosséis locais das plantas de milho com ajuda do template matching.

2.9.1 Template matching

Medidas de similaridade podem estar baseadas em valores de intensidades, imagens binárias, características ou proximidade de bordas, considerando que representações de objeto são descritas em translação, rotação com condições de luz constante ou variável além de imagens desfocadas. Porém, tarefas de localização de objetos pode ser tratada em domínios espaciais ou de frequência através de convolução ou correlação (HORNBERG, 2017; NIXON; AGUADO, 2019).

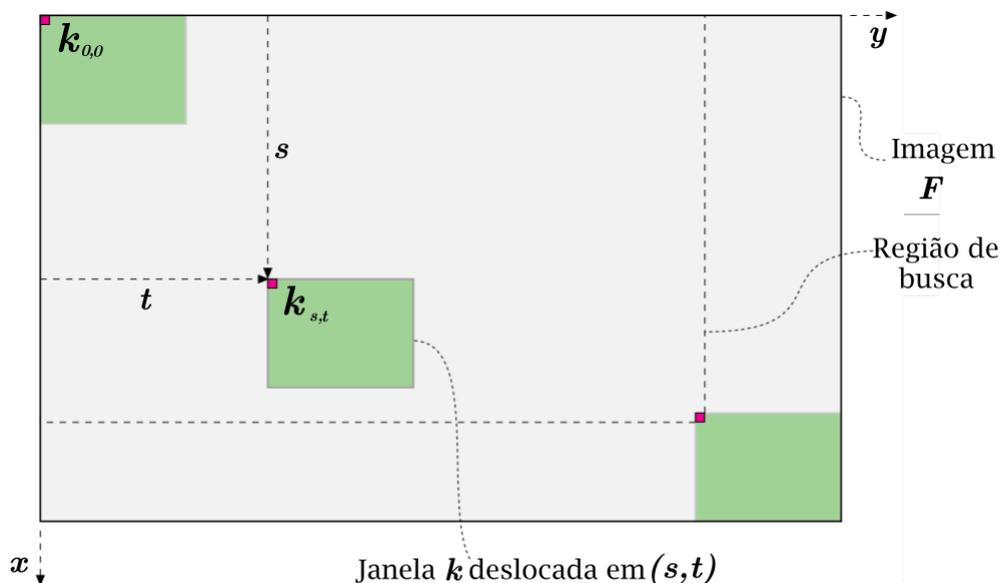
Como dito na Seção 2.6.5, operação de convolução usa operação de correlação invertido em 180° para aprimareamento ou suavização de imagens. Por conta disso, a correlação pode ser efetivamente usada para comparação de imagens ou padrões (GONZALEZ; WOODS, 2018). Assim, formalmente a correlação 2D no domínio espacial, é definida como:

$$(\mathbf{F} \star \mathbf{G})(x, y) = \sum_s \sum_t \mathbf{k}(s, t) \mathbf{F}(x + s, y + t) \quad (2.41)$$

onde $\mathbf{F}(x, y)$ representa a imagem, $\mathbf{k}(x, y)$ representa uma janela a ser utilizada sobre imagem, (x, y) e (s, t) são coordenadas da imagem \mathbf{F} e da janela \mathbf{k} respectivamente.

De acordo com Gonzalez e Woods (2018), correlação tem valores mais altos nas regiões onde \mathbf{F} e \mathbf{k} são iguais ou quase iguais, nesse sentido a Equação 2.41 procura localizações onde \mathbf{k} coincide com a região \mathbf{F} . Por outro lado, Burger e Burge (2009), descrevem que localização de objetos é possível obter com medidas de distância expressando graus de similaridade entre \mathbf{F} e \mathbf{k} . Por conta disso, o processo consiste em mover \mathbf{k} em \mathbf{F} , medindo as diferenças correspondentes em cada posição e ir registrando essas posições, onde valores máximos de similaridade são encontradas, conforme se ilustra na Figura 32.

Figura 32 – Geometria do template matching. Fonte: Adaptado de Burger e Burge (2009).



Uma destas medidas de comparação, está relacionado com a distância Euclidiana. Assim, em conjunto com a correlação, pode ser definida como:

$$d_E(x, y) = \sum_{(s,t) \in \mathbb{R}} [(\mathbf{F}(x + s, y + t) - \mathbf{k}(s, t))^2]^{1/2} \quad (2.42)$$

Propriedades da Equação 2.42 com espaço n -dimensional, é de especial importância na otimização e estatística. Por tanto, para encontrar localizações com melhores posições entre a \mathbf{F} e \mathbf{k} , é suficiente minimizar o quadrado da d_E (sempre positiva), o qual pode ser expandido como:

$$d_E^2(x, y) = \sum_{s,t} \mathbf{F}^2(x + s, y + t) + \sum_{s,t} \mathbf{k}^2(s, t) - 2 \sum_{s,t} \mathbf{F}(x + s, y + t) \mathbf{k}(s, t) \quad (2.43)$$

Para simplicidade da Equação 2.43, convencionalmente define-se $d_E^2 = \mathcal{A} + \mathcal{B} - 2\mathcal{C}$.

De acordo com [Burger e Burge \(2009\)](#), é interessante ressaltar que \mathcal{B} representa soma quadrática dos pixels em \mathbf{k} , ou seja, é uma constante e pode ser ignorada. Enquanto, \mathcal{A} é a soma quadrática dentro da sub imagem de \mathbf{F} no deslocamento de (s, t) . Finalmente, \mathcal{C} é chamada de correlação linear cruzado definido na Equação 2.41.

Enquanto equações 2.41 e 2.42, expressem métodos de comparação entre imagem \mathbf{F} e sub-imagem \mathbf{k} , há alguns problemas associados quando valores de \mathcal{A} se tornam grandes sobre uma sub-imagem em uma particular localização, levando que não aconteça similaridade. Por conta disso, outros métodos normalizados podem ser utilizados ([WILSON; RITTER, 2000](#)).

2.9.2 Template matching com correlação cruzada normalizada

Template matching (TM), de acordo com [Burger e Burge \(2009\)](#), para sair das dependências mencionadas acima, uma normalização deve ser levado em consideração junto a energia da imagem (\mathcal{A}) de referência e da subimagem atual, deixando a equação definida como:

$$\dot{\mathcal{C}}_N(x, y) = \frac{\sum_{s,t} \mathbf{F}(x + s, y + t) \mathbf{k}(s, t)}{\left[\sum_{s,t} \mathbf{F}^2(x + s, y + t) \right]^{1/2} \left[\sum_{s,t} \mathbf{k}^2(s, t) \right]^{1/2}} \quad (2.44)$$

onde $\dot{\mathcal{C}}_N$, é resultado de calcular correlação cruzada normalizada em intervalo de $[0, 1]$. Dito isso, $\dot{\mathcal{C}}_N \approx 1$ indica uma correlação alta, caso contrário se $\dot{\mathcal{C}}_N \approx 0$ indica que não tem semelhança.

Em relação à Equação 2.44, é possível ressaltar que a medida se trata de uma distância local. Embora, é importante identificar uma limitação na medida da distância absoluta entre imagem \mathbf{F} e a sub-imagem \mathbf{k} , ou seja, a intensidade geral da imagem se for alterada, $\dot{\mathcal{C}}_N$ poderia mudar drasticamente ([BURGER; BURGE, 2009](#)).

2.9.3 Template matching com coeficiente de correlação

De acordo com (LEWIS, 1995), uma solução para possíveis limitações com \check{C} é utilizar a média dos valores de \mathbf{k} e a média da sub-imagem atual \mathbf{F} . De forma que normaliza a imagem e as características para o comprimento da unidade, produzindo um coeficiente de correlação semelhante ao cosseno, o qual é definido como:

$$\check{C}_N(x, y) = \frac{\sum_{s,t} (\mathbf{F}(x+s, y+t) - \overline{\mathbf{F}})(\mathbf{k}(s, t) - \overline{\mathbf{k}})}{\left[\sum_{s,t} (\mathbf{F}(x+s, y+t) - \overline{\mathbf{F}})^2 \right]^{1/2} \left[\sum_{s,t} (\mathbf{k}(s, t) - \overline{\mathbf{k}})^2 \right]^{1/2}} \quad (2.45)$$

onde \check{C}_N é resultado de calcular coeficiente de correlação em intervalo de $[-1, 1]$. Por conta disso, se $\check{C}_N \approx 1$ indica uma correlação alta, caso contrário se $\check{C}_N \approx -1$ indica que não tem semelhança. Além disso, \check{C}_N descreve uma correlação local por partes entre \mathbf{k} e a sub-imagem atual.

De acordo com (BURGER; BURGE, 2009), \check{C}_N é recomendado para medir TM com imagens de intensidades em condições de luz natural, conseguindo robustez contra mudanças de brilho e contraste e tolera pequenos desvios padrão a partir de padrão referencia (\mathbf{k}). Embora, vale destacar algumas limitações como a invariância em relação à escala ou rotação (LEWIS, 1995). Por conta disso, uma possibilidade é utilizar em conjunto com a transformada de distância (seção 2.8.2) e algoritmo de chamfer.

2.9.4 Template matching com algoritmo de chamfer

Como dito, algoritmo de chamfer utiliza mascaras adequadas para calcular a distância local ponderadamente. Nesse quesito, a máscara de chamfer segue uma definição generalizada 2D de uma função de distância octogonal (ROSENFELD; PFALTZ, 1968) definido como:

$$d(x, y; \mathbf{B}') = \max \{d_1(x, y), d_2(x, y)\} \quad (2.46)$$

onde $\mathbf{B}' = \{b(i) : 1, 2, \dots, m; b(i) \in \{1, 2, \dots, n\}\}$, d_1 e d_2 são definidos como:

$$d_1(x, y) = p_* + \sum_{j=1}^{p_*} \left[\frac{x - y - f'(y-1) - 1}{f'(p_*)} \right], d_2(x, y) = \max(x, y) \quad (2.47)$$

onde $p_* = |\mathbf{B}'|$ é uma sequência periódica de comprimento \mathbf{B}' , e $f'(i) = \sum_{j=1}^i b(j) \quad \forall i 1 \leq i \leq p_*$.

Desta forma a mascara de chamfer é uma vizinhança estendida que enumera os valores de distância para seu centro. Assim, equações 2.44 e 2.45 aplicadas com a transformada de distância permite a possibilidade de invariância rotacional em TM. Por fim, o pesos desta máscara segue um padrão de tamanho $(2p_* + 1) \times (2p_* + 1)$ (MUKHOPADHYAY et al., 2016). A Figura 33 ilustra uma máscara de chamfer típica formada pela distância octogonal.

Figura 33 – Máscara de chamfer com distância octogonal

*	*	3	3	3	*	*
*	3	3	2	3	3	*
3	3	2	1	2	3	3
3	2	1	0	1	2	3
3	3	2	1	2	3	3
*	3	3	2	3	3	*
*	*	3	3	3	*	*

Nesse contexto, neste trabalho foi utilizado o TM com algoritmo de chamfer, onde a máscara de chamfer é o padrão k . Por conta disso, a máscara de chamfer pode ser de diferentes tamanhos p_* segundo a equação 2.46, útil para sua comparação em relação à mapa de distâncias das plantas de milho. Da mesma forma, modos de correlação podem ser especificados com a notação $\hat{C}_N = \{\dot{C}_N, \ddot{C}_N\}$. Os resultados das correlações são salvos junto a suas localizações, para ser avaliados com um limiar l_* . Assim, é retornado só correlações altas junto a suas localizações. O Algoritmo 6 especifica os passos para obter os máximos locais de planta de milho utilizando TM com algoritmo chamfer.

Algorithm 6 Template matching com algoritmo chamfer

Entrada: Matriz mapas de distância: D' , escalares tamanho e limiar: p_*, l_* , cadeia: \hat{C}_N

Saída: Matriz 2D de máximos locais de planta de milho M_{loc}

```

65 início
66    $M, N = tamanho(D')$ 
67    $d \leftarrow criar\ máscara\ chamfer\ com\ tamanho\ p_*$ 
68   para  $i, j = 0$  até  $M, N$  faça
69     se  $\hat{C}_N \in \{\dot{C}_N, \ddot{C}_N\}$  então
70        $V_{match} = \hat{C}_N(D'(i, j), d)$ 
71        $V_{loc} = (i, j)$ 
72     fim
73   fim
74    $M_{match} = [V_{match}, V_{loc}]$ 
75    $M_{loc} = max(M_{match} \times l_*)$ 
76 fim
  
```

2.10 Descritores de objetos e extração de características

De acordo com Szeliski (2010), determinar quais características vêm de locais correspondentes a diversos algoritmos de *matching* é estudado pelos descritores de algoritmos. Assim, grupos de pixels correspondentes a regiões devem descrever algumas formas geométricas em relação a sua geometria local (DISTANTE; DISTANTE, 2020).

Localizações de máximos locais obtidas na etapa de TM (Seção 2.9.1), levam informações de características que podem ser avaliadas por meio de descritores de contorno ou por descritores regionais. Nesse contexto, para a presente pesquisa, foram utilizados código de cadeia além de sua extração de características em prol de verificar seu desempenho na Seção 2.12.

2.10.1 Código da cadeia

O contorno de um objeto pode ser representado especificando um ponto inicial (x, y) que segue uma sequência ao redor do contorno de um objeto no sentido horário ou anti-horário. Desta forma, código da cadeia (FREEMAN, 1974) representa contornos de objetos, por uma sequência de segmentos de linha de componentes conectados condicionados a algum comprimento e direção específicos. Este processo de deslocamento é realizado para cada ponto até chegar ao ponto de partida, normalmente o código da cadeia é baseado em conectividades de 4 ou 8, e segmentos dessas conectividades codificam a direção em um esquema de numeração adequado.

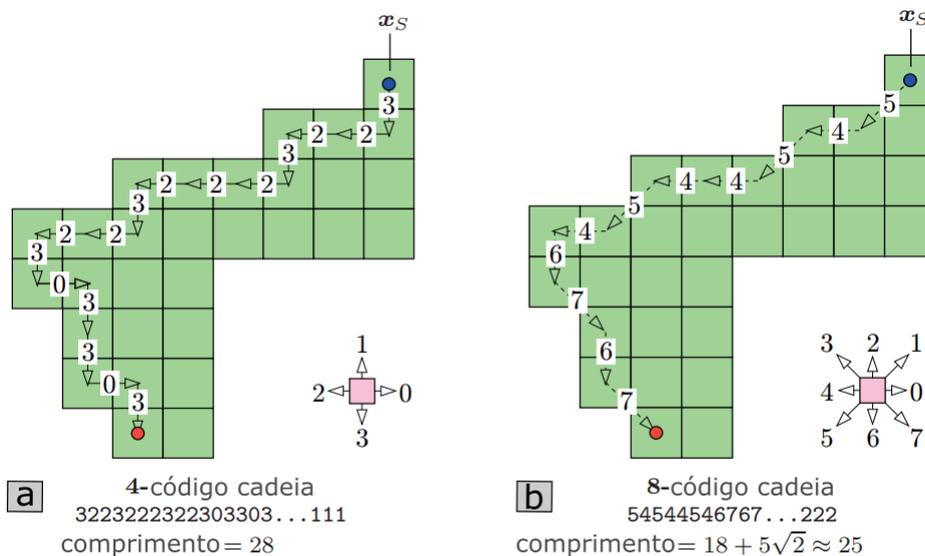
De acordo com Burger e Burge (2009), formalmente código cadeia começa com um ponto x_s , representado uma sequência de mudanças direcionais. Para um contorno de região A , descreve-se uma sequência de pontos $P_A = [x_0, x_1, \dots, x_{M-1}]$ com $x_i = \langle s_i, t_i \rangle$ criando assim os elementos do código cadeia: $P'_A = [P'_0, P'_1, \dots, P'_{M-1}]$, onde M é número total de pontos. Finalmente, cada ponto é calculado conforme:

$$P'_i = H(\Delta s_i, \Delta t_i)$$

$$(\Delta s_i, \Delta t_i) = \begin{cases} (s_{i+1} - t_i, s_{i+1} - t_i) & 0 \leq i < M - 1 \\ (s_0 - t_i, s_0 - t_i) & i = M - 1 \end{cases} \quad (2.48)$$

A Figura 34 exibe o método de extração de código em cadeia.

Figura 34 – Exemplo de representação de região por código de cadeia, em (a) vizinhança 4 e (b) vizinhança 8. Fonte: Adaptado de Burger e Burge (2009)



De acordo com [Sundararajan \(2017\)](#), a sequência numérica do código de cadeia depende de uma sequência circular, de modo que os números que se formam nessa sequência podem ser expressos com uma magnitude mínima. Além disso, vale ressaltar que o código da cadeia é invariante à translação, mas se uma diferença de adjacências for feita, ela será invariante para rotação.

Dito isso, nesta seção do trabalho foram utilizados mapa de distâncias da Seção 2.8.2 e matriz de máximos locais de TM (Seção 2.9.1). Desta forma, cada um dos pontos centrais encontrados nos máximos locais são usados para binarizar a região localizada (P_k). Em sequência, cada um dos contornos dessas regiões (A') se calculam com código de cadeia (Equação 2.48) como é especificado no Algoritmo 7.

Algorithm 7 Contorno de plantas de milho

Entrada: Matriz de mapa de distâncias e de máximos locais: D' , M_{loc}

Saída: Matriz de contornos com regiões locais: A'

```

77 início
78    $M, N = tamanho(D')$ 
79    $P_k = zeros(M, N)$ 
80   para cada ponto  $(x, y) \in M_{loc,i}$  faça
81      $P_k(x, y) = binarizar(D'(x, y))$ 
82      $A'(x, y) = H(P_k(x, y))$ 
83   fim
84 fim

```

2.10.2 Características geométricas

Uma vez que as regiões de interesse são representadas pelo código de cadeia, é necessário extrair características geométricas. Uma das principais características que foi de interesse para esta pesquisa, esta relacionada com os raios das regiões encontradas, que por sua vez representam a abertura de dosséis da planta milho.

Em relação às regiões obtidas (A') que são binárias e os contornos dessas regiões, propriedades binárias são calculadas como uma caixa delimitadora mínima, que consiste em um retângulo paralelo ao eixo mínimo que envolve todos os pontos de uma região ([BURGER; BURGE, 2009](#)). Por tanto, define-se pontos x, y para cada $x_{min} < x < x_{max}$ e $y_{min} < y < y_{max}$, desta forma caixa delimitadora é definido como:

$$C'_d = \langle x_{min}, x_{max}, y_{min}, y_{max} \rangle \quad (2.49)$$

onde x_{min}, x_{max} e y_{min}, y_{max} são valores mínimos e máximos de todos os pontos $(x_i, y_i) \in A'$ nas direções (x, y) respectivamente. Com tudo, o raio é calculado utilizando informação de diâmetro definido como:

$$\dot{d} = \max_{i,j} [d_e(x_i, y_j)] \quad (2.50)$$

onde d_e é a distância euclidiana, e x_i, y_i são pontos mínimos e máximos do contorno da região.

Nesse contexto, para o trabalho desenvolvido informação do diâmetro foi multiplicado por uma constante de um meio, além de uso da função teto. Desta forma, os cálculos representaram vetores de raios com V_r . Esses vetores por sua vez, são compostos para visar o processo de estimativa de altura (Seção 2.11) e classificação (Seção 2.12) de plantas de milho. Além disso, um mapa de características composto de raios circulares são sobrepostos para representar abertura do dossel da planta de milho. Assim, algoritmo 8 descreve o processo de extração do vetor de características, e da geração de mapas com círculos com a finalidade de visualização e detecção de plantas de milho.

Algorithm 8 Extração do vetor características

Entrada: Matriz mapa de distâncias e de contornos locais: D', A'

Saída: Matriz mapa de características: M_c , e vetor de características: V_r

```

85 início
86    $M, N = tamanho(D')$ 
87    $M_c = zeros(M, N)$ 
88   para cada ponto  $(x, y) \in A'$  faça
89      $x_{min}, x_{max}, y_{min}, y_{max} = C'_d(A'(x, y))$ 
90      $V_r = 1/2 \times \lceil \max(d_e(x_{min}, x_{max}, y_{min}, y_{max})) \rceil$ 
91      $M_c(x, y) = criarMapa(V_r)$ 
92   fim
93 fim
```

Para o presente trabalho foi elaborado uma análise espectral dos raios por meio de histogramas. Dito isso, uma análise das ocorrências em relação aos histogramas, pode estimar a realidade das sementes emergidas. Nesse quesito, foi normalizado o cálculo da abertura do raio, por meio de raio mínimo e máximo respectivamente de todos os blocos em geral. O Algoritmo 9 especifica o cálculo do histograma e a estimação da contagem de plantas de milho.

Algorithm 9 Cálculo de histograma e contagem de plantas de milho

Entrada: Vetor de características: V_r

Saída: Vetor de histogramas e de contagem: H', N'

```

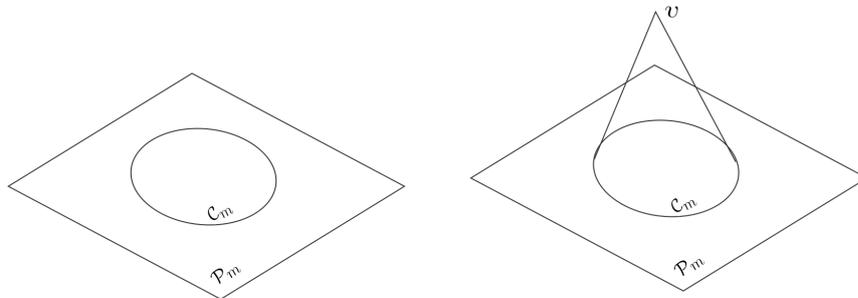
94 início
95    $r_{min}, r_{max} = \min(V_r), \max(V_r)$ 
96   para cada  $V_r$  faça
97      $H'_i, r_i, f_i = numero\ ocorrências(V_{r,i}) // \text{Calcula histograma, } f_i: \text{ocorrências}$ 
98      $N'_i = \sqrt{r_i \times 1,04 \times f_i}$ 
99   fim
100 fim
```

2.11 Modelo baseado em secções cónicas

As medidas que normalmente são feitas às plantas de milho são dadas por comprimento, largura e altura. Esses três parâmetros são alcançáveis para o modelo tridimensional em que vivemos, mas para imagens digitais, eles se projetam no nosso mundo tridimensional em um plano bidimensional. Portanto, uma forma de calcular a altura de uma planta de milho a partir de uma perspectiva bidimensional (imagem digital) pode ser feita através da relação entre as secções cónicas e sua base circular (VERAMENDI; CRUVINEL, 2021).

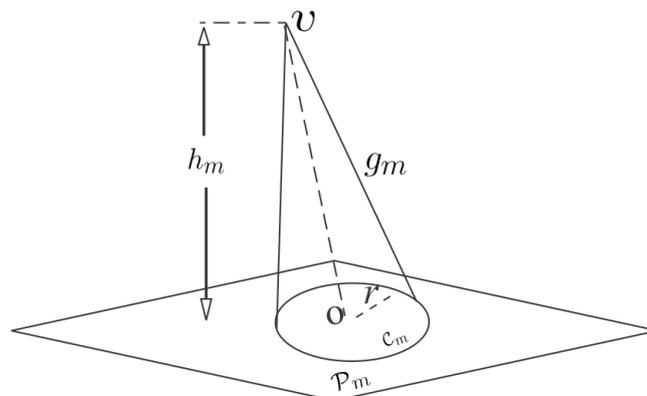
Relação entre círculo e um cone é dada pela seguinte definição formal. Seja um plano \mathcal{P}_m que inclui um círculo \mathcal{C}_m , e um ponto qualquer v (vértice) fora desse plano, uma figura pode ser formada a partir do círculo e do vértice, então podemos chamar essa figura de cone circular, conforme é exibido na Figura 35.

Figura 35 – Geração de um cone a partir de um círculo.



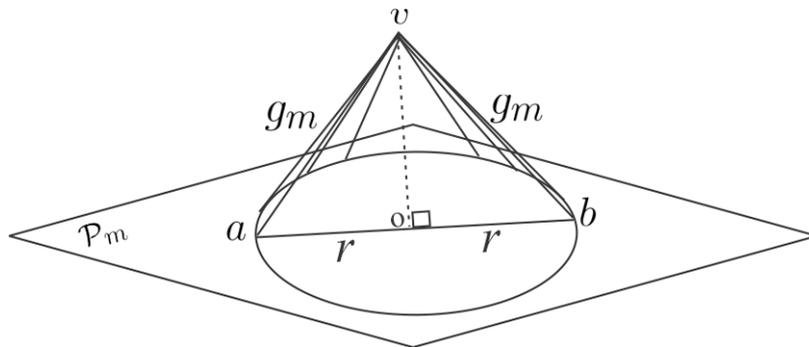
Alguns elementos (Figura 36) que podem ser destacados são dados pela altura h_m , que será a distância do vértice v ao plano \mathcal{P}_m , a geratriz g_m , que é um segmento de reta entre um ponto da circunferência e o vértice v , o eixo de rotação definido pela reta \overline{vO} .

Figura 36 – Elementos de um cone circular.



Uma seção interessante a ser estudada e que estima a altura com base em um cone reto, é a seção meridiana cuja superfície é determinada por um plano \mathcal{P}_m e seu eixo de rotação, formalmente chamada de seção meridiana, conforme é ilustrada na Figura 37.

Figura 37 – Seção meridiana.

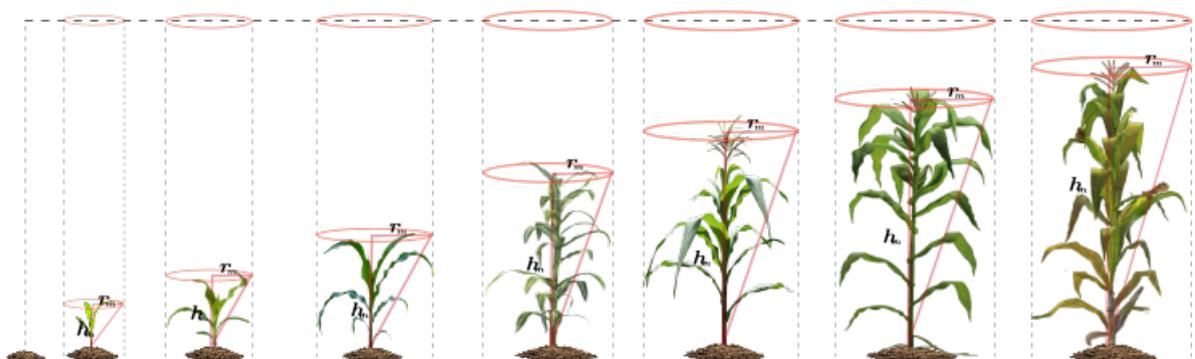


Assim, o triângulo conformado pelos segmentos de reta avb representa um triângulo equilátero, com essa premissa o cone obtido segue uma definição de cone equilátero. Nesse sentido, vale ressaltar propriedades da geratriz e altura do cone, os quais são definidos como:

$$g_m = 2r, \quad h_m = r\sqrt{3} \quad (2.51)$$

De acordo com a Equação 2.51, há possibilidade de modelar a estimativa da altura de milho. Por conta disso, com a detecção dos raios é possível projetar uma base circular definida acima, a mesma, que representa a abertura do dossel da planta de milho. Nesse contexto, medidas reais de planta foram definidas pelo comprimento da folha, e altura é definido pela medida a partir da superfície do solo até a base da última folha em relação aos estádios vegetativos. A Figura 38 ilustra a relação do modelo de seções cónicas com as plantas de milho.

Figura 38 – Relação entre modelo de seções cónicas e estádios fenológicos de planta de milho.



Nesse sentido, para esta seção do trabalho foram utilizadas a Equação 2.51 para calcular as alturas da planta de milho, a partir dos raios extraídos na forma de características da Seção 2.10.2. Dito isso, o Algoritmo 10 descreve o processo de estimativa de altura das plantas de milho, baseado no modelo de seções cónicas.

Algorithm 10 Estimativa de altura para plantas de milho

Entrada: Vetor de características de raios: V_r **Saída:** Altura da planta de milho: h_p

```
101 início
102   para cada raio  $r_i \in V_r$  faça
103      $h_p = r_i \times \sqrt{3}$ 
104   fim
105 fim
```

2.12 Classificação de padrões por aprendizado de máquina

De acordo com [Awad e Khanna \(2015\)](#), o aprendizado de máquina (ML, do inglês *Machine Learning*) é considerado um ramo da inteligência artificial que aplica algoritmos sistematicamente para sintetizar as relações entre dados e informações. Com objetivo de prever eventos ou cenários futuros que são desconhecidos para o computador.

Diversos conceitos para ML foram surgindo, ressaltando alguns delas: "computadores que aprendem sem serem explicitamente programados" ([SAMUEL, 1959](#)), "Experiências obtidas por algum tipo de tarefa, medida em desempenho" ([MITCHELL, 1997](#)) até o teste de Turing ([TURING, 1950](#)) para demonstrar máquinas inteligentes usando um benchmark padrão. Nesse quesito, também ML pode ser considerado como um framework algorítmico unificado ([DEISENROTH et al., 2020](#); [WATT et al., 2020](#)) para modelos computacionais que descrevem com precisão dados empíricos e fenômenos subjacentes com pouca ou nenhuma participação humana.

De acordo com [Braga-Neto \(2020\)](#), ML tem atualmente uma sobreposição substancial entre si e uma configuração matemática comum. Por conseguinte, para abordar o amplo grupo de problemas, uma taxonomia apresentada por [Raschka et al. \(2016\)](#) agrupam três categorias principais de aprendizado: (i) aprendizado supervisionado, com regras computacionais que envolvem um relacionamento entre entrada e saída, salientando problemas de *regressão* e a *classificação*; (ii) aprendizado não supervisionado, com regras computacionais usando apenas dados de entrada, ressaltando problemas fundamentais de *redução de dimensões* e *agrupamento (clustering)*; (iii) aprendizado por reforço, com um sistema (agente) que melhora desempenho a partir das interações com o ambiente ou circunstância (recompensa).

Com base nas diferentes abordagens mencionadas do ML, esta pesquisa utilizou o aprendizado supervisionado, com a tarefa de classificação. Por conta disso, o modelo que visou classificar as características encontradas na Seção 2.10.2, foi optado por usar o modelo máquina de vetores de suporte. A escolha do modelo, esta sujeita a vantagens, como a possibilidade de utilizar um número de dimensões menores, natureza de otimização convexa e tratamento com problema não linearmente separáveis.

2.12.1 Máquina de vetores de suporte

De acordo com [Beyerer et al. \(2018\)](#), máquina de vetores de suporte (SVM, do inglês *Super Vector Machine*) é um dos classificadores mais versáteis, relativamente simples, mas extremamente poderoso, e fornece boa generalização mesmo com um pequeno número de amostras de treinamento.

Originalmente publicado por [VAPNIK e CHERVONENKIS \(1964b\)](#) para otimizar os modelos baseados em separações lineares (margem) influenciadas na época pelo aprendizado estatístico ou a teoria de Vapnik-Chervonenkis (teoria VC), conseguindo superar os modelos baseados do perceptron ou regressão. Posteriormente [Boser et al. \(1992\)](#), abordaram problemas de separação não linear aplicando o truque do kernel e para casos práticos a proposta de margem suave ([CORTES; VAPNIK, 1995](#)) permitiu que o SVM fosse generalizado.

2.12.2 SVM linear

Segundo [Mohri et al. \(2018\)](#), um dos problemas iniciais que se realizam na classificação são as separações lineares com pequenos erros de generalização, que podem ser definidos como:

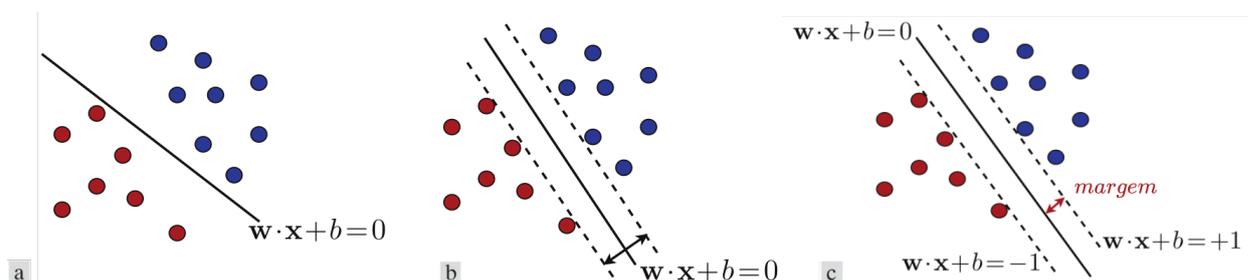
$$\varepsilon_n(h) = [h(x^{(i)}) \neq y^{(i)}] \quad (2.52)$$

onde y é a saída ou um espaço de rótulos (classes), h é a hipótese que pertence a um conjunto diferente de hipóteses \mathcal{H} que podem ser selecionadas usando a dimensão VC ([VAPNIK; CHERVONENKIS, 1964a](#)) ou a complexidade de Rademacher, alcançando melhores garantias de aprendizagem. Assim, um conjunto de hipóteses naturais com menor complexidade, é dado pelos classificadores lineares ou hiperplanos que são definidos por:

$$h(\mathbf{x}; \mathbf{w}, b) = \text{sgn}(\mathbf{w} \cdot \mathbf{x} + b) \quad \mathbf{w} \in \mathbb{R}^D, b \in \mathbb{R} \quad (2.53)$$

onde \mathbf{w} é um vetor normal diferente de zero (pesos) ao hiperplano, \mathbf{x} é um espaço de entradas, b é um bias escalar e D é a dimensão dos hiperplanos. Desta forma, é definido a equação da reta geral com rótulos positivos e negativos, ilustrados na Figura 39-a.

Figura 39 – Exemplos de hiperplanos de separação, (a) ingênua, (b) margem máxima, e (c) margem otimizada. Fonte: Adaptado de [Mohri et al. \(2018\)](#).



Nesse sentido, para uma escolha de um hiperplano ótimo a partir dos D -hiperplanos, uma noção de margem geométrico é aplicada na Equação 2.53. Assim, no ponto \mathbf{x} a distância Euclidiana desse hiperplano é definido como:

$$\rho_h = \frac{|\mathbf{w} \cdot \mathbf{x} + b|}{\|\mathbf{w}\|_2} \quad (2.54)$$

onde ρ_h é margem geométrico de um classificador linear h para uma amostra $\mathcal{S} = (\mathbf{x}_1, \dots, \mathbf{x}_d)$.

Por sua vez, ρ define uma distância mínima sobre cada um dos pontos mais próximos da amostra com $\rho_h = \min_{i \in [d]} \rho_h(x_i)$, onde d é um sub-espaco de \mathcal{D} . Assim, a solução SVM aplicando otimização primal com margem geométrico máximo (Figura 39 – b), é definido por:

$$\rho = \max_{\mathbf{w}, b} \min_{i \in [d]} \frac{y_i (\mathbf{w} \cdot \mathbf{x} + b)}{\|\mathbf{w}\|} \quad (2.55)$$

Essa igualdade decorre do fato, que amostra é linearmente separável maximizando (\mathbf{w}, b) , $y_i(\mathbf{w} \cdot \mathbf{x} + b)$ sujeitas para serem negativas para todo $i \in [d]$. Desta expressão, é interessante ressaltar a invariância para multiplicação de (\mathbf{w}, b) por um positivo escalar. Assim, uma restrição para os pares escalados (\mathbf{w}, b) de tal modo que $\min_{i \in [d]} y_i(\mathbf{w} \cdot \mathbf{x}_i + b) = 1$ é definido como:

$$\rho = \max_{\mathbf{w}, b} \frac{1}{\|\mathbf{w}\|} \quad (2.56)$$

onde (\mathbf{w}, b) esta sujeito à $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \forall i \in [d]$.

Desta forma, resultado de maximizar (\mathbf{w}, b) , torna-se em um mínimo de $y_i(\mathbf{w} \cdot \mathbf{x}_i + b)$ seja 1, conforme é ilustrado pelos hiperplanos marginais paralelos na Figura 39 – c. Com tudo, o problema de maximizar $1/\|\mathbf{w}\|$ é equivalente a minimizar $1/\|\mathbf{w}\|^2$. Além disso, o resultado de SVM separável se torna em uma otimização convexa:

$$\rho = \min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \quad (2.57)$$

Entanto, a restrição seja $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \forall i \in [d]$, é possível identificar que otimização convexa, há somente um mínimo que é global, porque o problema além de ser quadrático tem restrições lineares. Nesse sentido, a função objetivo é infinitamente diferenciável explicitamente é uma família da programação quadrática (MOHRI et al., 2018).

De acordo com Beyerer et al. (2018), é preciso transformar a formulação primal em dual, fazendo uso das otimizações de técnicas de programação quadrática introduzindo uma Lagrangiana na função de custo (CORTES; VAPNIK, 1995), conforme é definido por:

$$\mathcal{J}(\mathbf{w}, b, \boldsymbol{\alpha}^*) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^d \alpha_i^* [y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1], \quad \boldsymbol{\alpha}^* \in \mathbb{R}_+^d \quad (2.58)$$

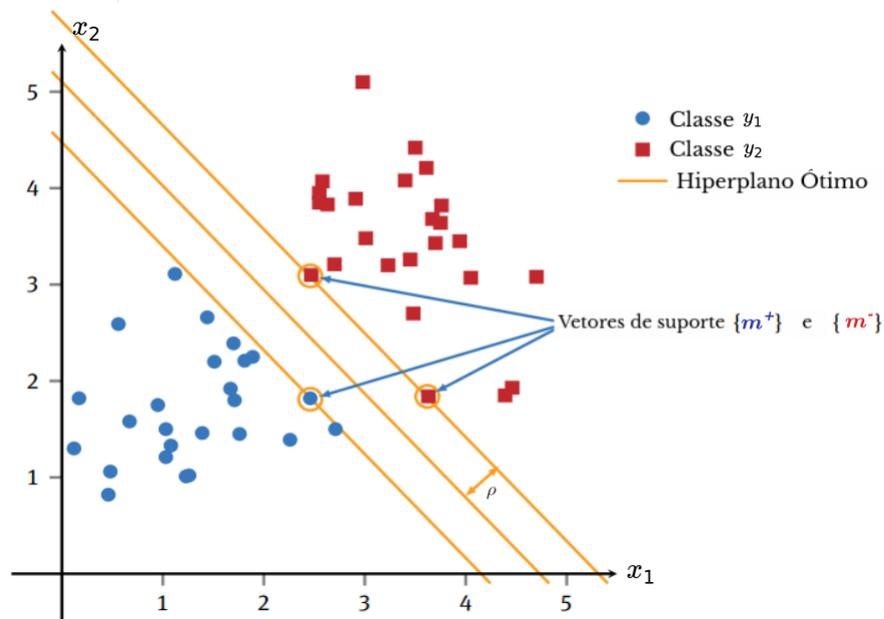
onde α^* é um vetor associado aos multiplicadores de Lagrange.

De acordo a esta função Lagrangiana \mathcal{J} , a resolução da minimização em relação a (\mathbf{w} e b), ou seja, $(\frac{\partial \mathcal{J}}{\partial \mathbf{w}}, \frac{\partial \mathcal{J}}{\partial b})$ permite a solução da forma primal para dual, apresentando restrições mais simples. Dito isso, vetores de suporte para definição do hiperplano precisam ser calculados com auxílio as condições de Karush-Kuhn-Tucker (KKT), assim b^* (uma das soluções duais) é definido como:

$$b^* = \frac{1}{N_{VS}} \sum_{x_j \in VS} \left[\frac{1}{w_j} - \sum_{x_i \in VS} \alpha^* w_i (x_i x_j) \right] \quad (2.59)$$

onde N_{VS} é o número de vetores de suporte (VS), x_i, x_j são os vetores características, e y_i, y_j são as classes. A Figura 40 ilustra o calculo dos vetores de suporte otimizados.

Figura 40 – Vetores de suporte com hiperplano otimizado. Fonte: Beyerer et al. (2018).



2.12.3 SVM com margem suaves

SVM linear até agora mencionado acima tem uma limitação que é interessante ressaltar, o qual esta relacionado com a esparsidade dos dados, ou seja, existe possibilidade de alguns vetores de suporte estejam em hiperplanos marginais muito pequenos e não estejam sujeitos as restrições discutidas acima (BEYERER et al., 2018). Por conta disso, algumas características atípicas (*outliers*) podem ser mal classificadas, e relaxar as restrições na otimização primal, pode dar a separabilidade, segundo a equação:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i \quad \forall i = 1, \dots, d \quad (2.60)$$

onde ξ_i são variáveis de folga ou de relaxamento, e $\xi_i \geq 0$.

Objetivo das variáveis de folga consiste em selecionar hiperplano que minimize o erro empírico. Por conta disso, possíveis erros mínimos podem maximizar a margem de separação entre as classes. Assim, classificador linear com máximo margem na otimização é definida como:

$$\rho = \min_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C^* \sum_{i=1}^d \xi_i \quad (2.61)$$

onde C^* é uma constante de peso para minimização de erro no treinamento. Assim, C^* define quanta ênfase deve ser colocada na classificação correta (C^* é grande) versus uma margem grande (C^* é pequena). Por conta disso, a formulação dual com a função de Lagrange, resulta na seguinte equação:

$$\max_{\alpha^*} = \sum_{i=1}^{\mathcal{D}} \alpha_i^* - \frac{1}{2} \sum_{i,j=1}^{\mathcal{D}, \mathcal{D}} \alpha_i^* \alpha_j^* y_i y_j (x_i \cdot y_j) \quad (2.62)$$

sujeito as seguintes restrições:

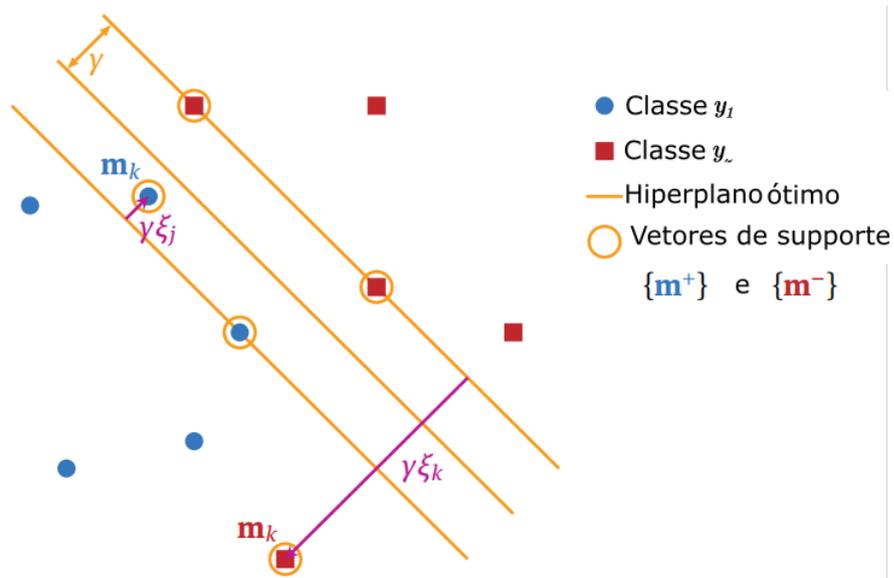
$$\sum_{i=1}^{\mathcal{D}} \alpha_i^* y_i = 0 \quad e \quad 0 \leq \alpha_i^* \leq C^* \quad (2.63)$$

onde $\alpha_i^* \neq 0$, e $\hat{\xi}_i$ são calculadas de acordo com:

$$\hat{\xi}_i = \max \left\{ 0, 1 - y_i \sum_{i=1}^{\mathcal{D}} \alpha_i^* \alpha_j^* y_i y_j x_i \cdot y_j + b \right\} \quad (2.64)$$

Desta forma, na Figura 41 é exibido graficamente a localização das variáveis de folga.

Figura 41 – Interpretação geométrica das variáveis de folga. Fonte: Beyerer et al. (2018).



2.12.4 SVM não linear

Distribuição do espaço de características não necessariamente são linearmente separáveis, e uma solução para garantir separabilidade é aplicar mapeamento dos dados para espaços de maior dimensão (BRAGA-NETO, 2020). Portanto seja as características originais mapeadas desde $\mathbb{M} \subseteq \mathbb{R}^d$ ao espaço do vetores de altas dimensões $\Phi \subseteq \mathbb{R}^{d^*}$ é dado por:

$$\phi: \begin{cases} \mathbb{M} \rightarrow \Phi \\ x \rightarrow \Phi(x) = [\Phi_1(x), \dots, \Phi_{d^*}(x)]^T \end{cases} \quad (2.65)$$

onde Φ_i são as funções não lineares. Por tanto, a função de decisão é definido como:

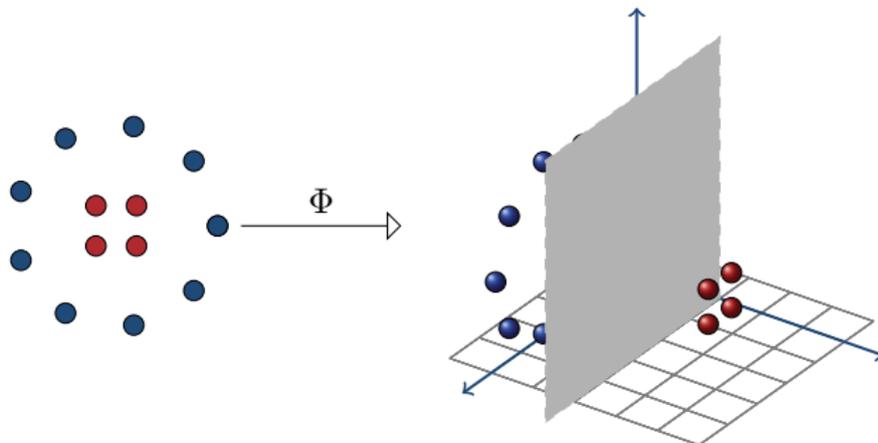
$$g(x) = \text{sgn} \left\{ \sum_{x_j \in sv} \alpha_j^* w_j \Phi(x_i)^T \Phi(x_j) + b^* \right\} \quad (2.66)$$

Assim, uma possível escolha de Φ no conjunto de dados \mathbf{x} , em seu espaço de entrada \mathbb{R}^2 , é projetado para um espaço de características \mathbb{R}^3 (Figura 42), sendo suas equações definidas como:

$$\Phi(x) = \Phi(x_1, x_2) = (x^2, \sqrt{2}x_1x_2, x_2^2) \quad (2.67)$$

$$h(x) = \mathbf{w} \cdot \Phi(x) + b = w_1x_1^2 + w_2\sqrt{2}x_1x_2 + w_3x_2^2 + b = 0 \quad (2.68)$$

Figura 42 – Exemplo de mapeamento não linear com função kernel, a partir de \mathbb{R}^2 para \mathbb{R}^3 . Fonte: Adaptado de Mohri et al. (2018)



Da Equação 2.66 é necessário definir Φ (possivelmente ∞ -dimensões), uma solução para essa definição é usar funções de kernel no espaço das funções que consideram as matrizes de autovalor. Portanto, partindo da matriz kernel é possível encontrar:

$$K(x_i, x_j) = \Phi(x_i)^T \Phi(x_j) \quad (2.69)$$

onde a função kernel é definido, de acordo a seguinte equação:

$$K(x_i, x_j) = (x_i^T x_j)^2 \quad (2.70)$$

Segundo o teorema de [Mercer e Forsyth \(1909\)](#), um kernel pode gerar uma matriz positiva na forma: polinomial, Gaussiana ou sigmoide. As aplicações desses kernels vai depender dos conjuntos de dados a serem classificados. A Tabela 6 apresenta esses tipos de kernel, os mesmos que são mais utilizados.

Tabela 6 – Tipos de Kernels

Kernel	Função	Parâmetros
Polinomial	$(\gamma x_i^T \cdot x_j + k)^p$	γ, k, p
Gaussiano	$\exp(-\gamma \ x_i - x_j\ ^2)$	γ
Sigmoide	$\tanh(-\gamma x_i^T \cdot x_j + k)$	γ, k

Para realizar os experimentos desta seção, foi necessário a elaboração de classificadores baseados nos vetores de raios características, correspondentes aos histogramas encontrados da Secao 2.10.2. Nesse quesito, vetores formados foram explorados com: SVM de margens suaves e SVM não linear. Além disso, foram testados e avaliados com kernels: Sigmoide e Gaussiano.

2.12.5 Conjunto de classificadores SVM

Como dito anteriormente, a elaboração dos classificadores foram separados por uma primeira etapa, que consistiu em cálculo de medianas e desvios padrão em relação aos histogramas encontrados. Já uma segunda etapa, consistiu em calcular um limiar de acordo ao Algoritmo 11. Esse limiar separa cada um dos histogramas em vetores de raio e ocorrências, além de um, rotulo que identifica a qual classe pertence. Por fim, para visar elaboração do conjunto de classificadores SVM, duas faixas consolidam proporcionalmente esses classificadores. Assim, o Algoritmo 11 especifica a separação dos classificadores.

Nesse quesito, para classificar as plantas de milho quanto ao seu estágio de pós-emergência, foram compostas dos conjuntos de classificadores calculados pelo Algoritmo 11. Lembrando, que cada classificador processa histogramas de raios correspondentes ao estágio V2. Desta forma, cada classificador sera responsável por classificar um determinado padrado da abertura do dossel em relação com o histograma descrito anteriormente. A Figura 43 exhibe o processo de treinamento de classificadores SVM responsáveis pela classificação da abertura do dossel das plantas milho em pós-emergência V2, explicitamente o classificador 1 e 2.

Algorithm 11 Separação de conjunto de classificadores

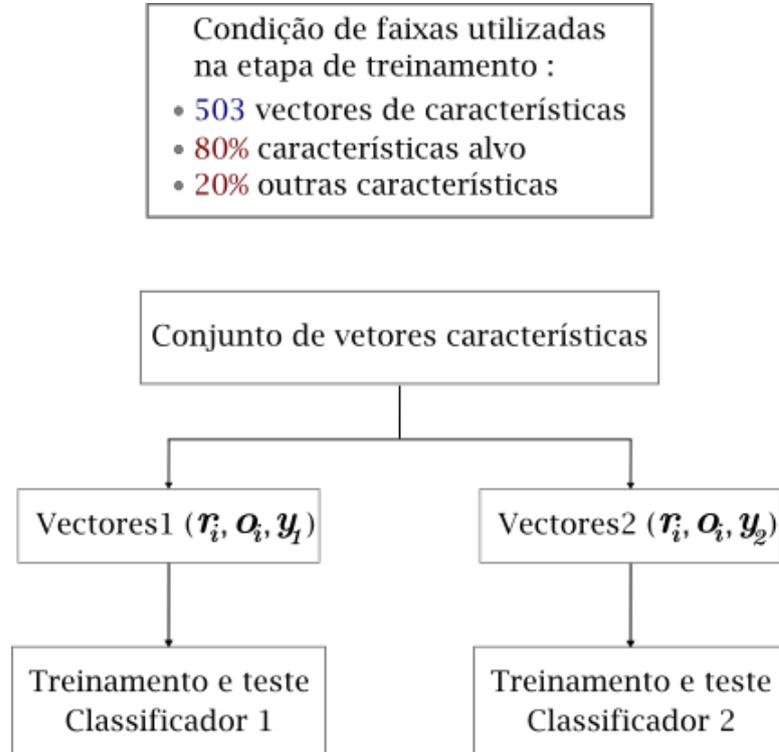
Entrada: Vetor de histogramas: H' , escalar de faixas: ψ_1, ψ_2

Saída: Matriz conjunto de classificadores: M_{c1}, M_{c2}

```

106 início
107   para cada  $H'_{r,i}$  faça
108     |  $\mu_i, \sigma_i = media(H'_{r,i}), std(H'_{r,i})$ 
109     fim
110      $\mu_{min}, \mu_{max} = min(\mu_i), max(\mu_i)$ 
111      $\sigma_{min}, \sigma_{max} = min(\sigma_i), max(\sigma_i)$ 
112      $l_* = \frac{\mu_{max} - \mu_{min}}{\sigma_{max} - \sigma_{min}}$ 
113     se  $\sigma_i \geq l_*$  então
114       |  $V_{c1} = raios(H'_{r,i}) + ocorrências(H'_{r,i}) + y_1$ 
115       fim
116     se  $\sigma_i < l_*$  então
117       |  $V_{c2} = raios(H'_{r,i}) + ocorrências(H'_{r,i}) + y_2$ 
118       fim
119      $M_{c1} = V_{c1} \times \psi_1 + V_{c2} \times \psi_2$ 
120      $M_{c2} = V'_{c1} \times \psi_2 + V'_{c2} \times \psi_1$ 
121 fim
    
```

Figura 43 – Diagrama de blocos para o conjunto de classificadores SVM.



Como dito, os classificadores são levados aleatoriamente para treinamento e teste em conjunto com os kernels anteriormente mencionados, ressaltando a utilização do parâmetro de regularização (C^*) de SVM com margem suaves e um parâmetro escalar γ . Assim, o Algoritmo 12 descreve o processo de classificação dos padrões das plantas de milho.

Algorithm 12 Classificação de características das plantas de milho**Entrada:** Matriz de características: M_c ; escalar gamma e regularização: γ, C^* ; modo kernel: k^* **Saída:** Acurácia, precisão, revocação, f1-score

```

122 início
123   para cada classificador faça
124      $V_{tr}$  = vetor de treinamento
125      $V_{ts}$  = vetor de teste
126     se  $k^* \in \text{Sigmoide}$  então
127        $\mathcal{K} = \tanh(-\gamma x_i^T \cdot x_j)$ 
128       treinar classificador SVM( $V_{tr}, M_c, C, \mathcal{K}$ )
129     fim
130     se  $k^* \in \text{Gaussiano}$  então
131        $\mathcal{K} = \exp(-\gamma \|x_i - x_j\|^2)$ 
132       treinar classificador SVM( $V_{tr}, M_c, C, \gamma, \mathcal{K}$ )
133     fim
134     testar classificador SVM( $V_{ts}$ )
135   fim
136 fim

```

Para as etapas de treinamento e teste do conjunto de classificadores SVM, foram realizados experimentos com as proporções 80% teste e 20% treinamento, 70% teste e 30% treinamento e 50% teste e 50% treinamento. Enquanto, para validação e qualificação do conjunto de classificadores foram utilizadas: (i) Matriz de confusão e (ii) Curva de característica de operação do receptor (ROC, do inglês *Receiving Operating Characteristics*).

Capítulo 3

RESULTADOS E DISCUSSÕES

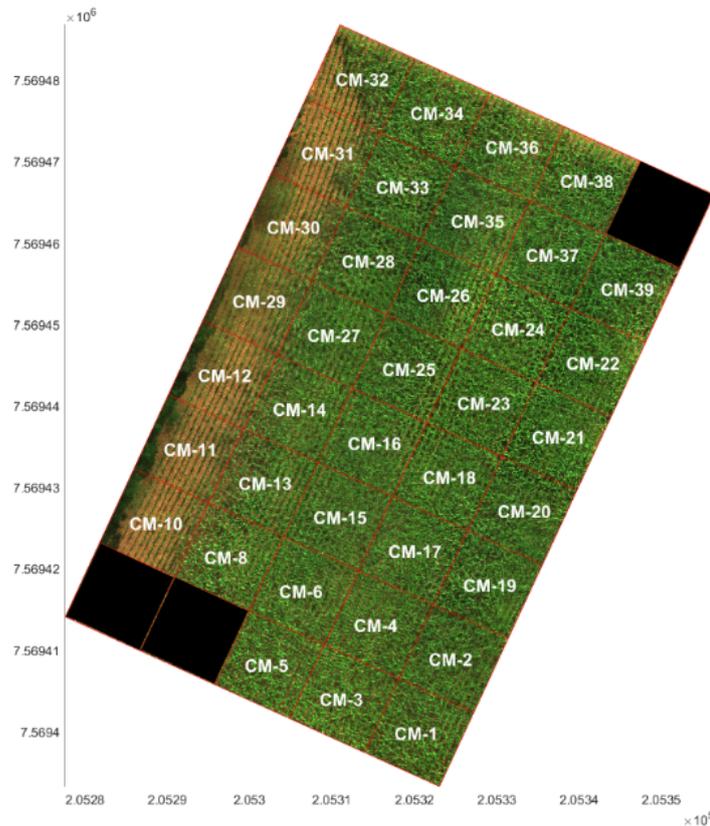
Este capítulo, tem como objetivo descrever os resultados dos experimentos a partir do método desenvolvido, bem como dos experimentos que foram realizados para validação da contagem de plantas de milho (*Zea mays. L.*) na fase de pós-emergência, considerando imagens georreferenciadas obtidas com um sensor multiespectral embarcado em Drone. Também, são apresentadas as avaliações dos ortomosaicos envolvendo a utilização de técnicas de processamento de imagens assim como a avaliação obtida na classificação dos resultados com uso de classificadores baseados em máquina de vetores de suporte (SVM).

3.1 Resultados em campo experimental

O arranjo experimental organizado para a avaliação dos riscos agrícola e logístico em área da cultura de milho (*Zea mays. L.*), foi estabelecido em uma área experimental de 4000 m^2 ($50 \times 80\text{ m}$), localizada no Laboratório de Referência Nacional de Agricultura de Precisão (LANAPRE) da Embrapa Instrumentação, com grade de amostragem de $10 \times 10\text{ m}$ em relação a sua largura e comprimento respectivamente. A Figura 44 exibe cada um dessas amostras divididos em blocos (40 blocos), relacionando a contagem de plantas com quadrante (Seção 2.1.1) por uma grade digital. Nota-se alguns desses blocos com cor preta, os mesmos indicam que não foram utilizados nos experimentos deste trabalho, sendo considerados para contagem manual por especialista. Deste modo, os experimentos foram viabilizados por 37 blocos de plantas de milho pertencentes ao estágio pós-emergência V2.

3.2 Resultados para aquisição das imagens

Nesta seção, são identificados parâmetros correspondentes à aquisição de imagens (Seção 2.4) com sensor multiespectral embarcado em Drone, com objetivo de analisar e organizar possíveis influências na qualidade de imagens. Começando listando parâmetros customizados do Drone e sensor multiespectral, em seguida é discutido planejamento de voo, e finalmente relata-se algumas práticas para aquisição de imagens em culturas de milho.

Figura 44 – Definição da grade de amostragem e dos blocos da área experimental.

3.2.1 Parâmetros para fase pré-voo

Um parâmetro do Drone DJI® Matrice 100, que é importante ressaltar esta relacionada a sua velocidade com GPS, por conta disso, para o presente trabalho essa velocidade foi de 11 m/s , lembrando que velocidade máxima é de 17 m/s . Dito isso, parâmetros do sensor multiespectral Micansense® RedEdge M destaca-se sua tecnologia CMOS, apresentando menores ruídos na imagem em comparação de um CCD; o tipo de opturador é global, minimizando o desfoque; o lente é de tipo retilínea, o mesmo tem pouca distorção almofada. Já parâmetros para definição de GSD estão baseados em distância focal, tamanho do sensor e resolução espacial, além da altura de voo do Drone. Por conta disso, altura de voo foi definida em 30 m o qual em consequência definiu um GSD com $2,08\text{ cm/s}$. A Tabela 7 descreve os parâmetros definidos na fase de pré-voo.

Tabela 7 – Parâmetros identificados e utilizados para fase pré-voo

Descrição	Valores	Unidades
Distância focal	5,4	mm
Taxa de captura	1	captura/seg
Tamanho do sensor	$4,8 \times 3,6$	mm
Resolução espacial	1280×960	píxeis
Velocidade máxima de voo para Drone	11	m/s
Altura de voo para Drone	30	m
GSD (Ground Sample Distance)	2,08	cm/pixel

3.2.2 Parâmetros para planejamento de voo

Como dito, parâmetros para geometria e sobreposição de imagens garantem qualidade de imagens em razão a sua densidade. Nesse quesito, áreas capturadas seguiram uma orientação de $\pm 0^\circ$, ou seja, a partir de nadir ou vertical. Enquanto, sobreposição lateral e frontal foram de 80% respectivamente. A orientação para os lentes da câmera Micanse seguiu uma orientação horizontal. Enquanto, o padrão de voo seguiu linhas paralelas em forma retangular com auxílio de GPS 3DR® com receptor RTK de georreferenciamento em torno de $\pm 1\text{ cm}$ de precisão. Dito isso, descreve-se resultados da missão, como tempo de voo do Drone em 12 minutos, o numero de imagens registradas por banda foram de 300, os que resultaram para as 5 bandas em um total de 1 500 imagens; a capacidade de armazenamento resultou em 3, 69 GB (*Gigabyte*); já a largura e altura de superfície de foram de $27 \times 20\text{ m}$ respectivamente; e a distância entre cada captura frontal e lateral foram de 4 e 5 m respectivamente. Finalmente, é importante ressaltar que os resultados do planejamento foram utilizados constantemente durante 8 fases de aquisição das imagens, enumeradas como Voo 01, até Voo 08, correspondentes aos estádios da planta de milho.

3.2.3 Calibração para câmera em cada voo

Aquisição de imagens para cada um das fases mencionadas anteriormente, precisam realizar a calibração radiométrica com objetivo de converter os dados brutos da imagem digital para uma escala física. Nesse quesito, foi utilizada uma base de medidas de um painel de reflectância de $6'' \times 6''$, que fornece propriedades Lambertianas e respostas de refletância em todo o espectro visível e infravermelho próximo. Desta forma, o Drone foi primeiramente posicionado com a câmera multispectral no painel, conforme é exibido na Figura 45.

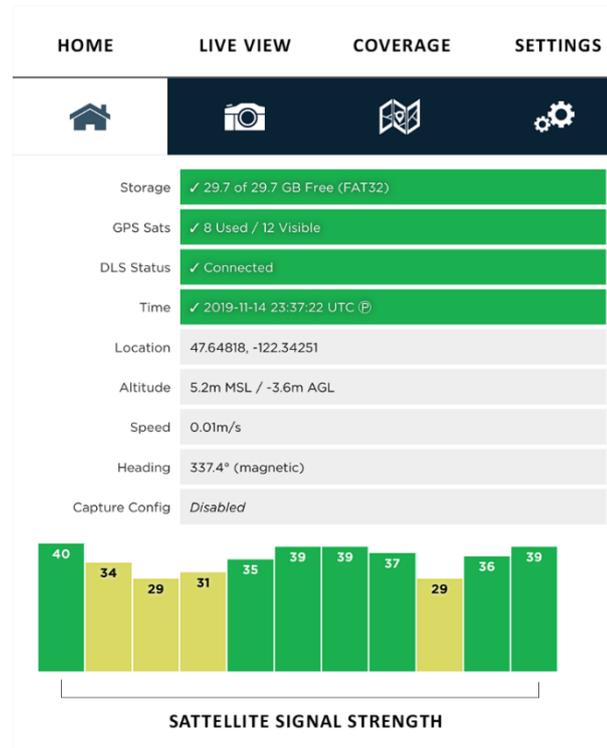
Figura 45 – Instruções recomendadas para calibração.



A condição para o posicionamento da calibração da câmera exigiu que nenhuma sombra estivesse sobre o painel, assim como os raios do sol estivessem com alguma angulação sobre o painel de reflectância, garantindo no mínimo uma reflectância de 60% para cada uma de suas bandas. Em seguida, foi utilizado um computador portátil para acessar a interface de configuração da câmera Micasense®, verificando seus estados principais, como do dispositivo

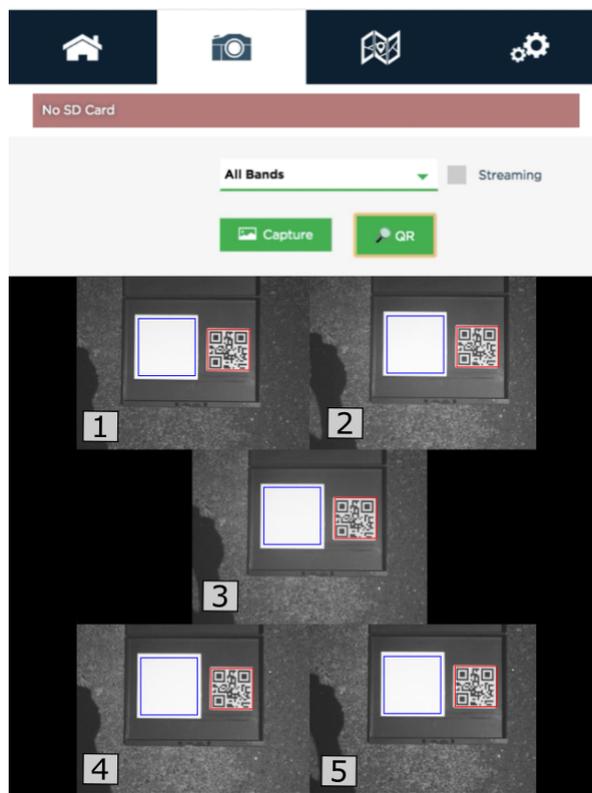
de armazenamento inserido no sensor, sinais do GPS 3DR®, estado do sensor DLS, todo eles representados pelos indicadores de barras verdes de acordo a ilustração da Figura 46.

Figura 46 – Interface principal para configuração da câmera Micasense®.



Uma vez, posicionando a câmera no painel de refletância e com as condições mencionadas, foi possível colocar a interface em operação. Desta forma, quando raios de luz del sol caim sobre o painel de refletância (de acordo ao entorno: nuvens, mais soleado, etc.), o sensor faz a calibração de níveis escuros, sensibilidade, ganho e exposição do sensor. Além disso, as intensidades coletadas são normalizadas de acordo ao formato da profundidade de bits (16 bits para este experimento), são calibrados e salvados. Já esses dados são guardados em forma de metadados na fase de registro de cada imagem que sera coletada. Dito isso, a Figura 47 exhibe em rótulos vermelhos a leitura por código QR, assim como os rótulos em azul que leem os raios refletidos do sol pelo material de refletância, sendo coletados para sua calibração em cada uma das bandas: (1) azul, (2) verde, (3) vermelho, (4) NIR e (5) RedEdge.

Finalmente, é importante ressaltar que os experimentos para a obtenção de imagens de alta qualidade foram realizados com tempo de referência horaria entre as 10h e 12h na parte da manhã. Assim, para cada uma das 8 fases comentadas anteriormente, as aquisições de imagens passam por uma fase previa de calibração do painel de refletância.

Figura 47 – Processo de calibração com captura manual da imagem de referência.

3.3 Calibração de medidas reais em campo

Para análise de crescimento, assim como para estimar altura de plantas de milho (modelo de secções cônicas), foram coletadas manualmente e de forma aleatória com cinco medidas referentes da abertura do dossel e da altura de milho. Desta forma, as medidas da abertura do dossel, foram consideradas a partir da folha mais desenvolvida, ou seja, seu comprimento total; e para altura, as medidas foram coletadas com referência ao caule da planta de milho, a partir desde o chão até o ponto central da folha com a abertura mais desenvolvida. As unidades de medida para comprimento e altura foram em centímetros. Deste modo, para viabilizar essa análise, uma calibração quantitativa foi realizada nas medidas obtidas das imagens e a realidade do campo, considerando exclusivamente as medidas obtidas no estádio V2 (correspondetes ao Voo 02) para avaliação pós-emergência. Mesmo com medidas do estádio V2, medidas do estádio V5 (Voo 05) e V7 (Voo 07) foram coletadas, conforme se descreve nas Tabelas 8, 9 e 10 respectivamente.

Tabela 8 – Dados coletados manualmente no campo para Voo 02

Abertura de dossel (cm)	Raio (cm)	Altura (cm)
35	17,5	27
33	16,5	32
29	14,5	29
32	16,0	33
31	15,5	28

Tabela 9 – Dados coletados manualmente no campo para Voo 05

Abertura de dossel (cm)	Raio (cm)	Altura (cm)
88	44,0	48
90	45,0	50
86	43,0	45
91	45,5	53
86	43,0	47

Tabela 10 – Dados coletados manualmente no campo para Voo 07

Abertura de dossel (cm)	Raio (cm)	Altura (cm)
104	52,0	165
84	42,0	177
99	49,5	168
83	41,5	184
104	52,0	198

Sabendo os dados manualmente coletados, é possível estimar um modelo de crescimento, levando uma confiança estabelecida entre as medidas de imagens e esses dados coletados. Nesse sentido, a partir dos dados coletados V2, V5 e V7, são calculados a média dos raios, média das altura (μ_r, μ_h), assim como o desvio padrão das alturas (σ_h), os mesmos servem de base para posterior análise do modelo de crescimento das plantas de milho. Dito isso, os resultados calculados das médias e dos desvios padrão são descritos na Tabela 11.

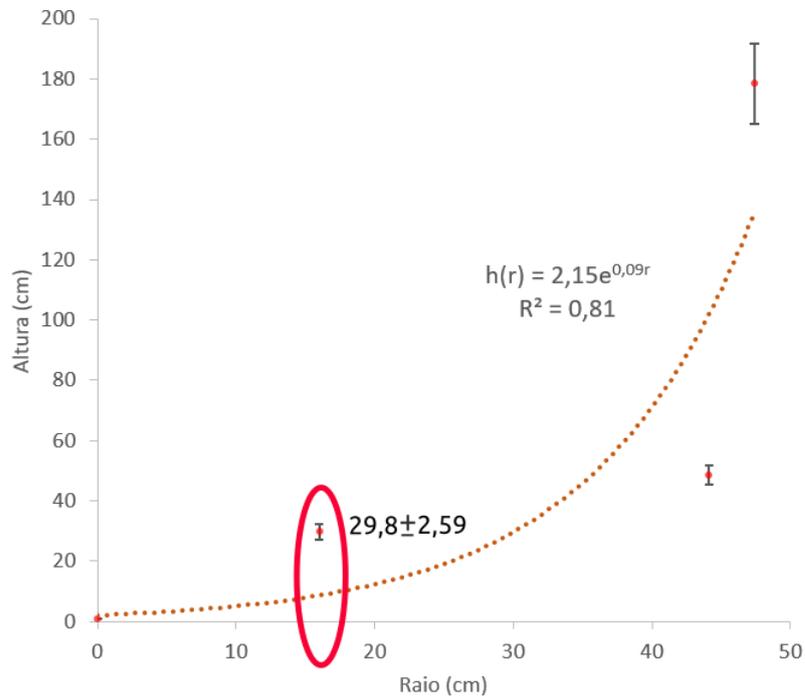
Tabela 11 – Médias e desvios padrão dos dados coletados

μ_r (cm)	μ_h (cm)	σ_h (cm)
16,00	29,08	2,59
44,10	48,60	3,05
47,40	178,40	13,28
0,00	1,00	0,00

Nota-se na Tabela 11, os valores da linha quatro, ou seja, as médias $\mu_r = 0$, $\mu_h = 1$ e o desvio padrão $\sigma_h = 0$ justificam a não presença de plantas de milho na área experimental. Dito isso, é possível identificar uma relação com o modelo de Müller et al. (2005), onde análise do comportamento de crescimento do milho apresenta aumento exponencial no período de instalação da cultura, linear e positivo (após V7) até o pendoamento, seguido de variação linear negativa até o início da rápida senescência das folhas.

Nesse contexto, os dados de campo que foram utilizados de forma a encontrar o modelo projetado para o crescimento das plantas de milho e verificar sua concordância, com os estudos que foram apresentados por Müller e colaboradores. Por conta disso, é apresentado o modelo calculado na Figura 48 o qual incluiu os desvios padrões referentes às alturas coletadas.

Figura 48 – Modelo de estimativa para crescimento das plantas de milho, obtidos a partir de medidas reais em campo experimental realizado para validação do método.



Os resultados deste modelo projetado, demonstram que o coeficiente de determinação R^2 tem uma correlação forte, com valor de 0,89. Destacando o análise para esta pesquisa com circulo vermelho em relação ao estádio de pós-emergência V2, onde seus valores da média e desvio padrão são $29,8 \pm 2,59$ unidades centímetro respectivamente.

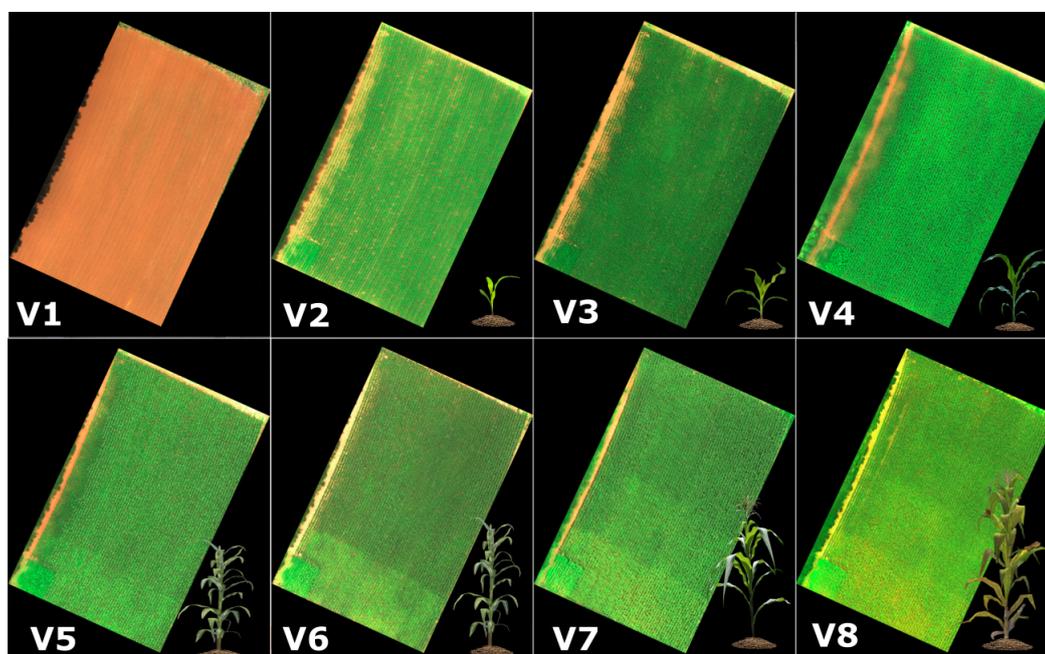
3.4 Resultados para caracterização de imagens

Como dito na Seção 2.5, produtos derivados a partir das imagens coletadas para cada fase do voo foram otimizadas para produzir as ortofotos, com informação dos dados espectrais assim como sua correção de terreno. Nesse sentido, as imagens coletadas foram amostradas, filtradas e costuradas com auxílio de software SfM Pix4Dmapper®, o qual permitiu a geração dos ortomosaicos para cada banda espectral de interesse. A Figura 50 ilustra produtos finais gerados por Pix4Dmapper® em forma de ortomosaicos nas bandas multiespectrais vermelho ($\lambda = 668 \text{ nm}$), azul ($\lambda = 475 \text{ nm}$) e verde ($\lambda = 560 \text{ nm}$), ressaltando que para propositos de visualização cada um desses ortomosaicos seguem um padrão de pseudo cor R, G e B respectivamente, lembrando que cada canal contém informação das intensidades em relação à luminância do campo experimental.

Após a geração dos ortomosaicos do campo experimental, foram recortadas a área experimental (área onde estão localizadas a cultura de milho) de cada um dos estádios fenológicos, com auxílio do software QGIS. Representando o ortomosaico da área experimental em formato rasterizado. Desta forma, cada um desses ortomosaicos recortados descrevem propriedades

Figura 49 – Ortomosaicos gerais para bandas vermelho, verde e azul respectivamente.

de resolução espacial em 3935×5019 píxeis, assim como resolução radiométrica de 32 bits e finalmente sua resolução digital de 96 dpi. Nesse sentido, os ortomosaicos em relação dos estádios fenológicos, são exibidos na Figura 50. Vale ressaltar, que para propósitos de visualização cada um desses ortomosaicos estão com cores compostas das bandas espectrais RGB, além de ajustes de brilho e contraste. Assim, cada estágio fenológico da área experimental, é salientado numericamente a partir de V1 até V7 em conjunto com uma planta de milho correspondentes aos estádios que foram registrados.

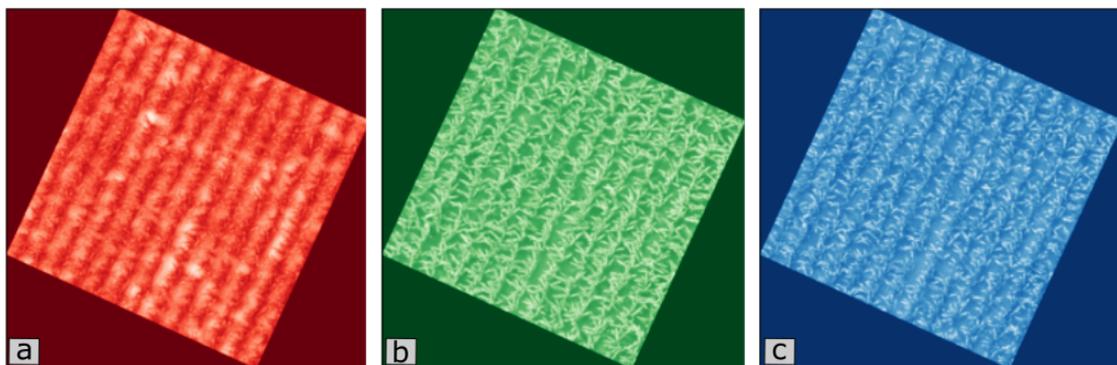
Figura 50 – Ortomosaicos da área experimental das plantas de milho.

Tanto feito a caracterização das imagens, o trabalho prosseguiu os experimentos considerando o estágio pós-emergência V2 conforme foi especificado no objetivo desta pesquisa. Desta forma, em cada canal R, G e B são cortadas com software QGIS da área experimental em blocos de 10×10 m.

3.5 Qualificação do banco de imagem RGB V2

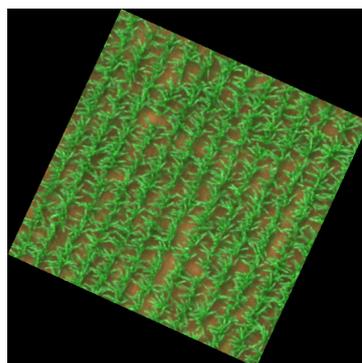
Como dito na Seção 2.4.1, curvas de refletância em canais RGB conseguem coletar informação da refletância da planta de milho. Nesse sentido, na seção anterior foi possível observar que os resultados detalham as características do estágio de pós-emergência V2. Assim, bandas espectrais R, G e B foram escolhidas como unidade de estudo para exploração da contagem de plantas de milho. Para isso, o banco de imagens foram avaliados através das unidades de sítio específico ou também chamado de bloco de imagem, conforme se ilustra na Figura 51 onde as bandas espectrais correspondem ao bloco 30 com pseudo cores para propósito de visualização.

Figura 51 – Bloco 30 com suas respectivas bandas espectrais vermelho, verde, azul.



De acordo com às três bandas espectrais selecionadas, para cada um dos blocos correspondentes das imagens coletadas, foram compostas as cores verdadeiras RGB com o Algoritmo 1, com objetivo de ler os metadados apropriadamente. Dito isso, o conjunto de imagens foram compostas por 37 blocos de imagens da área experimental. A Figura 52 ilustra as cores compostas das bandas espectrais que formam o bloco 30 RGB. Ressaltando suas propriedades em largura e altura de 546×546 píxeis, com 32 bits de resolução radiométrica e resolução digital de 96 dpi.

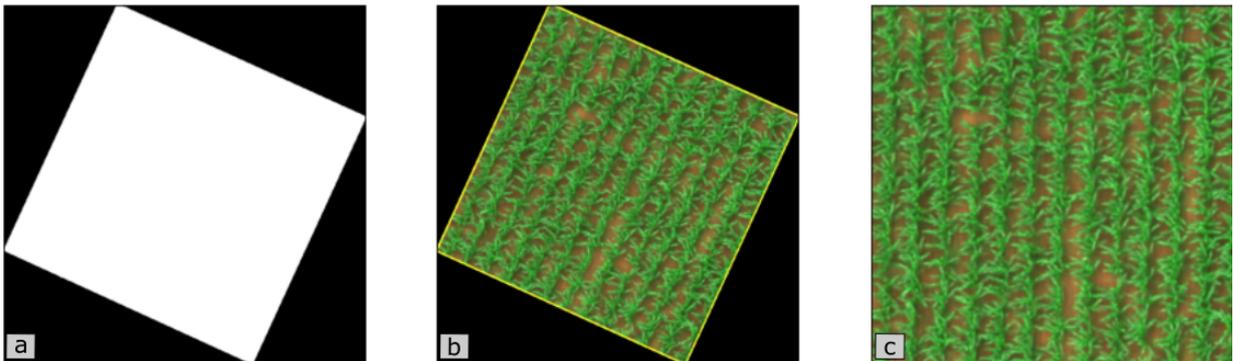
Figura 52 – Bloco número 30 multispectral RGB utilizado para validação do método.



3.6 Resultados obtidos através da transformação geométrica

Nesta seção, são apresentados resultados relacionados à extração da região de interesse referente à cultura do milho. Este processo se justifica pela necessidade de estudar estritamente apenas a cultura do milho, sem a necessidade de informações adicionais de intensidade de outros elementos tais como solo, pedras entre outros ao redor. Nesse contexto, identificou-se que para o corte da área foi necessário calcular o ângulo de inclinação, onde foram organizados três subprocessos: filtragem, limiarização, detecção de contornos e rotação em torno ao centro do objeto. Dito isso, os subprocessos envolvidos seguem a execução do Algoritmo 2. A Figura 53 mostra o processo da transformação geométrica em torno ao objeto.

Figura 53 – Transformação geométrica em torno ao objeto, por etapas de (a) filtragem e limiarização, (b) detecção de contornos, (c) rotação e extração da área de interesse.



Para a etapa de filtragem, foi utilizado um filtro de tipo Gaussiano, visando suavizar as bordas da área de interesse. A escolha dos parâmetros para este filtro foi selecionado um tamanho de *kernel* b_* de 3×3 , bem como da abertura com valor desvio padrão $\sigma = 0$. Em relação, a etapa de limiarização utilizou-se um grau de intensidade de valor $l_* = 10$, como é observado esse valor indica pouca informação de intensidade o qual sera processado apropriadamente na Seção 3.7.

Com imagem binária, contornos são calculados com um descritor de objetos de código de cadeia, onde os números de píxeis que definem os contornos tem um valor de 1 499. Dito isso, a área em valores de píxeis é calculado em 168 100, definindo a área de interesse. Finalmente, com a área encontrada á estimação do ângulo encontrado é de $\theta \approx -65^\circ$, e para rotacionar a área de interesse foi preciso realizar um soma complementar (somado em 90°) devido a sua natureza anti horária, sendo o ângulo real utilizado de $\theta \approx 25^\circ$. Assim, o Algoritmo 2 rotaciona a região de interesse baseado em seu centro. Por fim, a região de interesse é cortada da imagem original, é a nova dimensão fica definida em 410×410 píxeis em relação à cultura do milho.

Cada um dos parâmetros, com seus respectivos valores anteriormente mencionados, garantiram a extração da região a ser analisada para prosseguir os experimentos desta pesquisa. Nesse sentido é resumido na Tabela 12 os valores dos parâmetros encontrados nesta seção.

Tabela 12 – Principais parâmetros para a transformação geométrica.

Descrição	Parâmetros	Valores
Tamanho do <i>kernel</i>	b_*	3×3
Abertura para filtragem	σ	0
Limiar para binarização	l_*	10
Ângulo real para rotação	θ	$\approx 25^\circ$

3.7 Resultados obtidos para ajustes de brilho e contraste

Como dito anteriormente na Seção 2.4.1, imagens para Drone podem apresentar mudanças na coleta de imagens, resultando com mudanças de brilho e contraste. Além disso, conforme foi observado na Seção 3.6, o limiar da intensidade foi baixa. Por conta disso, uma inspeção no histograma evidenciou a necessidade de pre-processar imagem com operações pontuais de ajuste para brilho e contraste. Enquanto os experimentos foram executados, foram identificados a necessidade de operações locais devido à variabilidade de luminância em regiões pequenas. Nesse sentido, foi utilizada uma equalização de histograma adaptativa limitada por contraste (CLAHE). Além disso, como as imagens coletadas tem informação de luminância e refletância em forma de intensidades, foram utilizados espaços de cores HSV, YCbCr e $L^*a^*b^*$. Toda abordagem comentada anteriormente foram feitas de forma iterativa pelo Algoritmo 3.

Resultados para os parâmetros de brilho e contraste foram $\alpha = 1,25$ e $\beta = 5$ processadas no espaço RGB. Parâmetros para CLAHE foram de um tamanho para a janela de $b_* = 16$, em conjunto do limiar de contraste para essa janela local $l_* = 2$, a diferença que esses parâmetros são aplicados localmente nos espaços de cores mencionados. Assim, são transformados novamente para seu espaço original RGB. Finalmente, essa abordagem foi aplicado um número iterativo de vezes $k = 2$. A Tabela 13 descreve os parâmetros utilizados nesta seção.

Tabela 13 – Parâmetros utilizados para ajustes de brilho, contraste e CLAHE

Descrição	Parâmetros	Valores
Contraste	α	1,25
Brilho	β	5
Tamanho da janela CLAHE	b_*	16
Limiar local CLAHE	l_*	2

A Figura 54 ilustra pré processamento no espaço de cor YCbCr, destacando a utilização do canal Y. A Figura 55 exibe o pré processamento no espaço de cor HSV, destacando a utilização do canal H. Finalmente a Figura 56 mostra o pré processamento no espaço de cor $L^*a^*b^*$, destacando a utilização do canal L^* . Assim, os canais mencionados em cada espaço de cor estão relacionados com informação de luminância.

Figura 54 – Exemplo de ajustes para brilho, contraste e CLAHE com espaço de cor YCbCr. A linha superior demonstra resultados para espaço YCbCr. A linha inferior demonstra os resultados finais em RGB, após pré processamento do YCbCr.

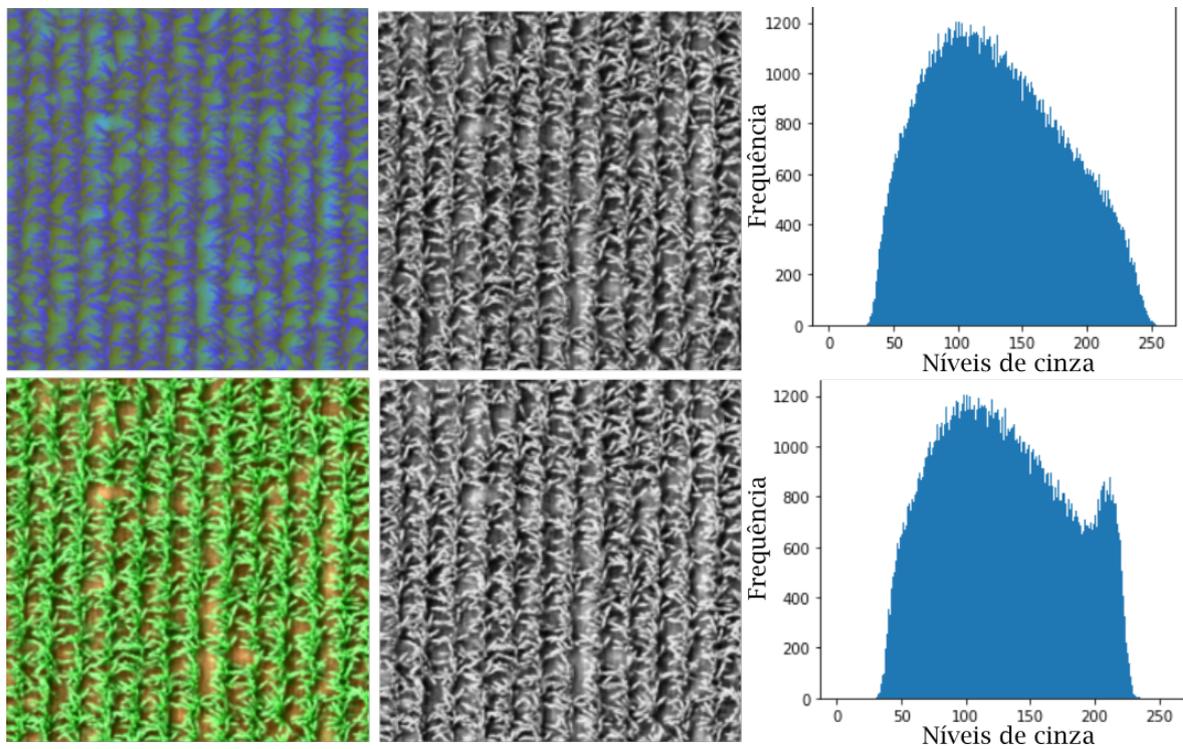


Figura 55 – Exemplo de ajustes para brilho, contraste e CLAHE com espaço de cor HSV. A linha superior demonstra resultados para espaço HSV. A linha inferior demonstra os resultados finais em RGB, após pré processamento do HSV.

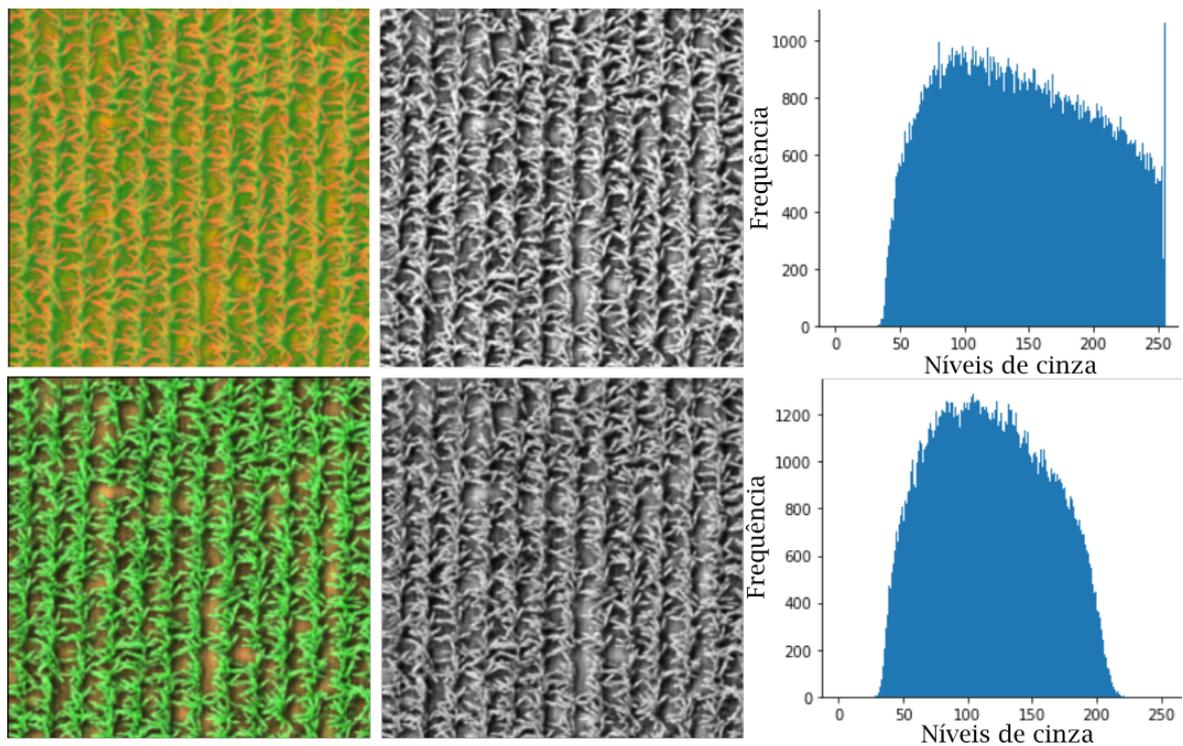
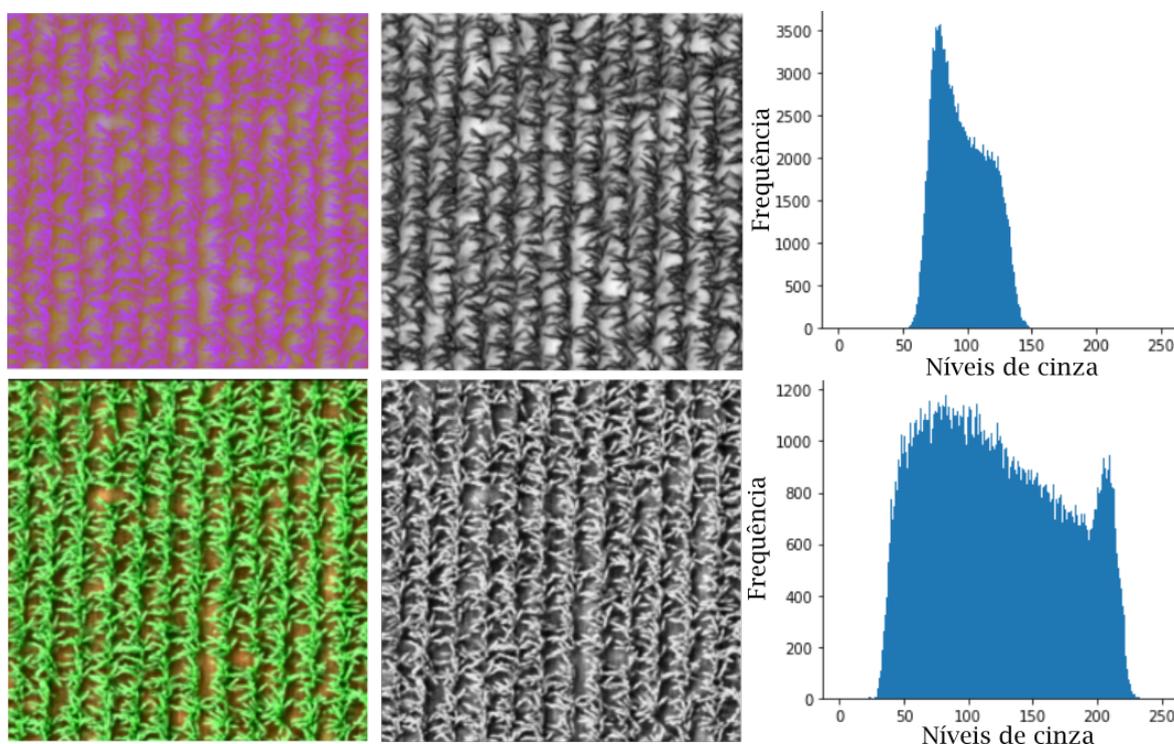


Figura 56 – Exemplo de ajustes para brilho, contraste e CLAHE com espaço de cor $L^*a^*b^*$. A linha superior demonstra resultados para espaço $L^*a^*b^*$. A linha inferior demonstra os resultados finais em RGB, após pré processamento do $L^*a^*b^*$.

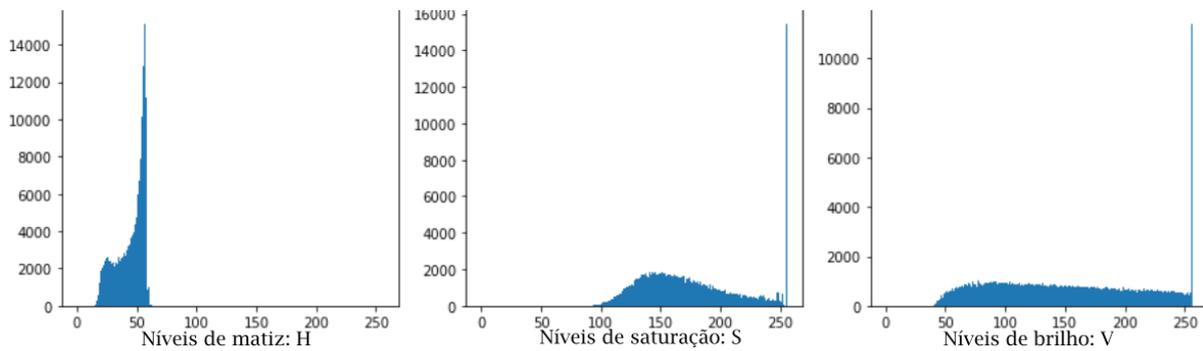


Em relação aos ensaios realizados para melhorar o brilho, contraste e equalização CLAHE, foi possível observar que os espaços de cores HSV e $L^*a^*b^*$ possuem as informações necessárias para os detalhes de contraste e brilho, respectivamente. Dito isso, a utilização iterativa com os espaços de cores envolvidos no processamento digital de imagens aplicadas à agricultura, demonstrou que o uso do espaço de cor HSV possui um histograma mais uniforme em relação à distribuição de intensidades. Assim, HSV pode viabilizar experimentos de segmentação em relação a aplicações de ambiente natural.

3.8 Resultados obtidos para segmentação de imagens

Nesta seção, são apresentados os resultados obtidos durante a etapa de segmentação baseado no espaço de cor HSV e para sua implementação foi utilizado o Algoritmo 4. Nesse sentido, é ilustrado histogramas dos canais H, S e V na Figura 57, esse histograma por sua vez é referente à imagem do bloco número 30. Vale ressaltar, que histograma em análise é resultado de aplicar os processos da realce de brilho e contraste (Seção 3.7).

A partir de análise dos histogramas, intervalos de intensidades foram definidos com valores de mínimos e máximos em referência aos cores refletidas pela planta de milho. Para a escolha sintética desses valores, foi utilizado um mapa de cores HSV como referência. Assim, cores da planta de milho em imagens digitais RGB são transformados ao mapa HSV, que

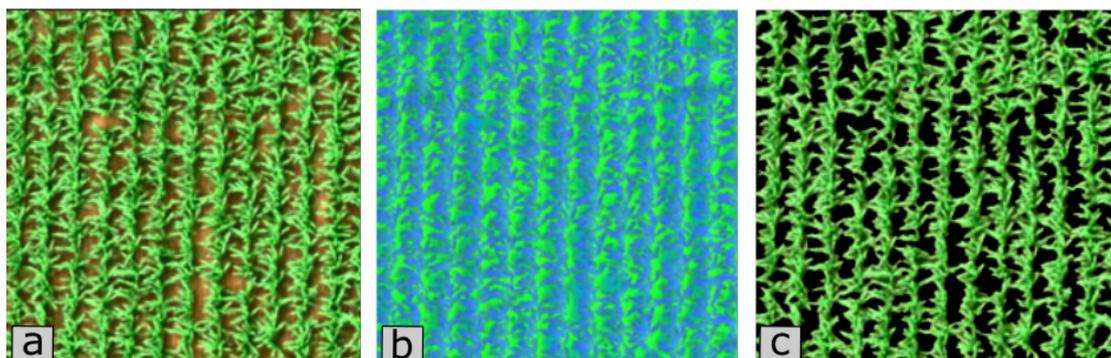
Figura 57 – Histograma do espaço de cor HSV.

resultaram em intervalos de valores para os canais H: 45 – 75, S: 100 – 250 e V: 50 – 250. Nesse contexto, a Tabela 14 apresenta parâmetros mínimos e máximos dos canais H, S e V em referência à planta de milho.

Tabela 14 – Parâmetros mínimos e máximos para plantas de milho no espaço de cor HSV

Parâmetros	Canal H	Canal S	Canal V
Mínimo	45	100	50
Máximo	75	250	250

Durante a aplicação da técnica de segmentação baseada em espaço de cores (Algoritmo 4), uma filtragem Gaussiana de tamanho 5×5 com desvio padrão $\sigma = 0$ foi aplicada na imagem. Após, as faixas identificadas da imagem HSV foram operadas utilizando um operador morfológico de fechamento ao lado de um elemento estruturante isotrópico com tamanho 5×5 de forma circular, o que forneceu invariância na translação. Além disso, elemento estruturante removeram alguns ruídos acontecidos na execução da segmentação. Desta forma, a Figura 58 mostra exemplos de resultados obtidos durante a segmentação de imagens com padrão de cores HSV. Pode-se observar que a segmentação aplicada resolveu de forma satisfatória a discriminação da planta de milho em relação ao solo, palha, e outros elementos possivelmente existentes na região considerada.

Figura 58 – Resultado da segmentação pelo espaço de cor HSV, (a) imagem aprimorada na etapa de brilho e contraste, (b) imagem HSV e (c) resultados da segmentação para a planta de milho.

3.9 Resultados obtidos para transformada de distância

Como dito na Seção 2.8.2, a etapa de segmentação podem ser utilizadas para localização da presença das plantas de milho, descrevendo informação geométrica com medidas de distâncias. Por conta disso, nesta seção, são apresentados os resultados obtidos pela aplicação da transformada de distância, cuja implementação segue o Algoritmo 5 baseado em sua variação espacial e forma das regiões. Esse processo forneceu informações suficientes para o reconhecimento de padrões, especificamente da abertura dos dosséis da planta de milho.

A execução do Algoritmo 5 precisou como entrada a imagem das plantas em forma binária (Seção 3.8). Dito isso, um elemento estruturante de forma quadrada e de tamanho 3×3 foi aplicado com operador morfológico de erosão, com objetivo de diminuir bordas inessárias para o cálculo da transformada de distância. Em sequência, as distâncias utilizadas para calcular a transformada foi executada com uma Euclidiana. Assim, o resultado da transformada da distância em referência à imagem segmentada, originou um mapa de distâncias com intensidades locais representando o objeto de interesse em relação a suas distâncias mais próximas da abertura central do dossel. A Figura 59 ilustra exemplo de resultados obtidos com a aplicação da técnica de transformada de distância.

Figura 59 – Resultado da transformada de distância, utilizando (a) imagem segmentada, onde (b) máscara é operado por erosão, (c) mapa de distâncias com Euclidiana.



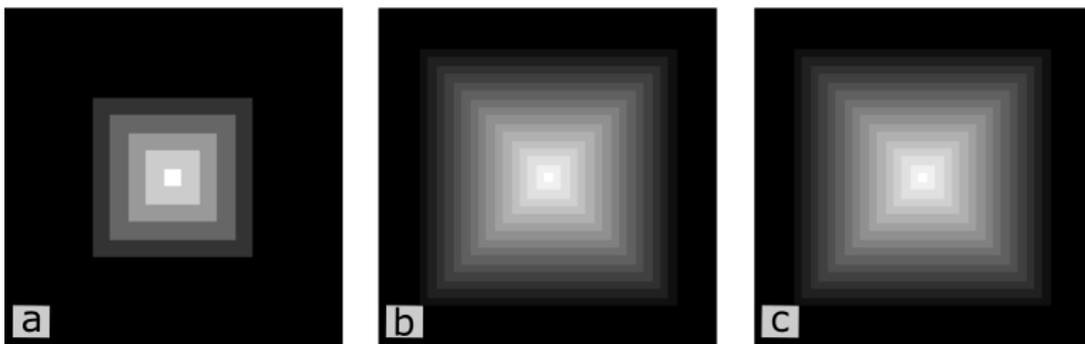
Analisando o resultado, é possível observar que a técnica de transformada de distância localizou satisfatoriamente a presença de plantas de milho em relação à sua distância central de forma geométrica.

3.10 Resultados obtidos para a etapa de template matching

Após de obter o mapa de distâncias para abertura dos dosséis das plantas de milho, nesta seção é apresentada a localização espacial desses dosséis com uso da técnica de TM com máscara de chamfer, conforme é executado o Algoritmo 6. Assim, o TM é utilizado para avaliar a localização dos máximos locais do objeto de interesse.

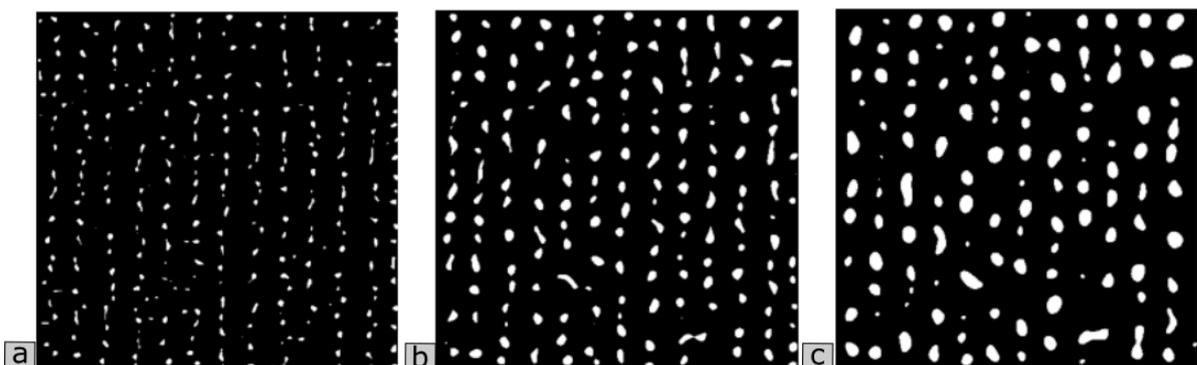
Nesse contexto, foram definidos três máscaras de distância octogonal conforme é especificado pelo algoritmo de chamfer na Equação 2.47. Distâncias dessas máscaras, são influenciados pelo parâmetro de tamanho p_* , e nesse sentido, foram definidos $p_1 = 8, p_2 = 13$ e $p_3 = 18$. Em seguida, um mapa de distâncias (transformada de distância) é calculado para a máscara de chamfer. Desta forma o mapa de distância da máscara chamfer é analisado como padrão sub-imagem do TM em conjunto do mapa de distâncias das plantas de milho, cada uma destas máscaras por é utilizado por separado nos experimentos. A Figura 60 ilustra os padrões utilizados para seu uso em conjunto com a técnica de TM.

Figura 60 – Mapa de distâncias para máscara padrão de chamfer, com (a) 8×8 , (b) 13×13 e (c) 18×18 píxeis.



Para prosseguir, o algoritmo de TM foi utilizado em conjunto com correlação normalizada, assim como o coeficiente de correlação. A abordagem, resulta em uma faixa de valores de baixa e alta correlação, salvando localizações espaciais a fim de garantir presença da planta de acordo à correlação. Para avaliar esses valores na faixa de correlação, foi utilizado um limiar no mínimo de um 50% desses valores altos. Nesse sentido, o coeficiente de correlação viabilizaram melhores resultados, lembrado que valores de -1 correspondem a ausência de máximos locais, e $+1$ representa a presença dos máximos locais. A Figura 61 ilustra os resultados de localizações utilizando as máscaras de chamfer como sub imagem no algoritmo de TM. Vale ressaltar que essas localizações representam informações de picos na presença de plantas de milho.

Figura 61 – Picos localizados após o uso do algoritmo de template matching com coeficiente de correlação e máscara de chamfer, para (a) 8×8 , (b) 13×13 , e (c) 18×18 píxeis.



Como dito, resultados do TM mostraram a presença de máximos locais representados pelos picos das regiões encontradas. Por conta disso, foi possível observar que a presença de máximos locais mais próximos (Figura 61 – a), da mesma forma esses resultados mostraram a presença de máximos locais um pouco mais uniformes (Figura 61 – b), embora a existência de proximidade se referem à variabilidade de crescimento das plantas de forma aleatória em relação à localização dos caules. Também foi possível observar a existência de máximos locais mais separadas (Figura 61 – c), sendo que também ocorreu ausência de correlação de alguns máximos locais. Por fim, para a avaliação em relação aos resultados obtidos, o padrão chamfer constituído pelo parâmetro $p_* = 13$ possibilitou a obtenção de informações locais de forma mais uniforme em relação aos demais padrões.

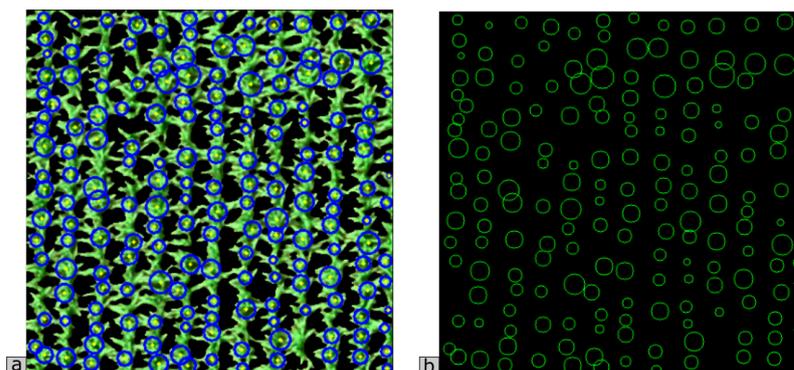
3.11 Resultados para a etapa de extração de características

Como foi observado na Seção 3.10, regiões de abertura dos dosséis foram localizados de acordo às altas correlações de TM. Nesse sentido, para esta seção são apresentados os resultados alcançados com uso do descritor de contorno de código em cadeia, bem como a representação das características dos padrões circulares encontrados na seção anterior.

Aplicação do descritor de código cadeia foi executada de acordo com o Algoritmo 7, que define os contornos dessas regiões para serem binarizados em relação a um área mínima da abertura do dossel. Já execução do Algoritmo 8, viabilizou extrair as características de contorno dos picos encontrados na etapa de TM a fim de representarem a presença das plantas de milho referentes à cultura no estágio de pós-emergência V2.

A Figura 62 ilustra um exemplo de resultado obtido por descrição geométrica circular, os mesmos, representam a localização das plantas de milho na área da cultura. Nesse sentido, esses locais indicam que houve emergência das sementes referente ao estágio fenológico V2.

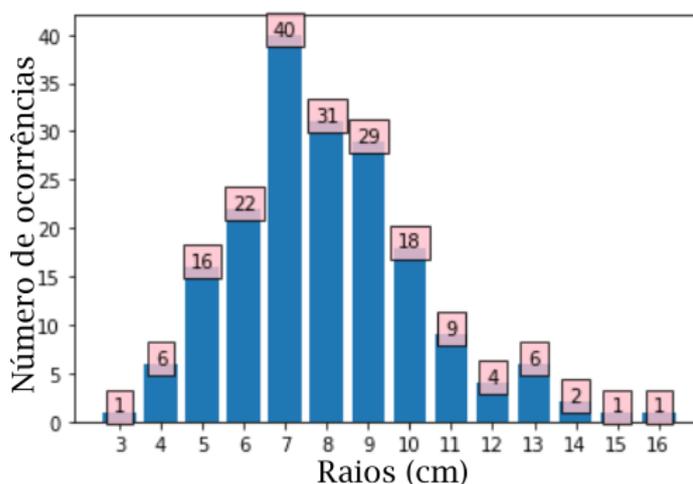
Figura 62 – Localizações pós-emergência V2 com descritores geométricos, (a) localizações do mapa circular na cultura do milho, (b) mapa de características circulares.



Uma vez obtidos os mapas de descrição geométrica, também foi possível extrair vetores de características para os raios circulares, correspondentes aos blocos da cultura do milho que foram analisados para validação de método. Nesse contexto, 92,5% dos blocos de imagens foram analisados para validação do método, do conjunto, três blocos foram analisados apenas por contagem manual de plantas a fim de se ter uma avaliação baseada em verdade de campo sobre o número de plantas por bloco. Assim, de cada um desses vetores que foram constituídos de raios do dossel das plantas, foi observado que suas dimensões variam de acordo com os raios detectados ao longo das plantas de milho. Nesse sentido, foram encontrados tanto um vetor com uma dimensão menor 116 raios (correspondente ao bloco 7), bem como um vetor com maior dimensão, ou seja com 186 raios (correspondente ao bloco 39).

Outra característica avaliada para esta seção foi a elaboração do número de ocorrências de raios em relação aos blocos analisados (Algoritmo 9), a fim de conhecer o dossel das plantas detectadas (abertura dos raios) e, por sua vez, a inferência da taxa de emergência da cultura do milho na área considerada (Algoritmo 10). A Figura 63 ilustra o histograma obtido a partir do extrator de características para o bloco número 30 do estádio de pós-emergência V2. Enquanto a Figura 64, exibe os histogramas para blocos analisados neste trabalho.

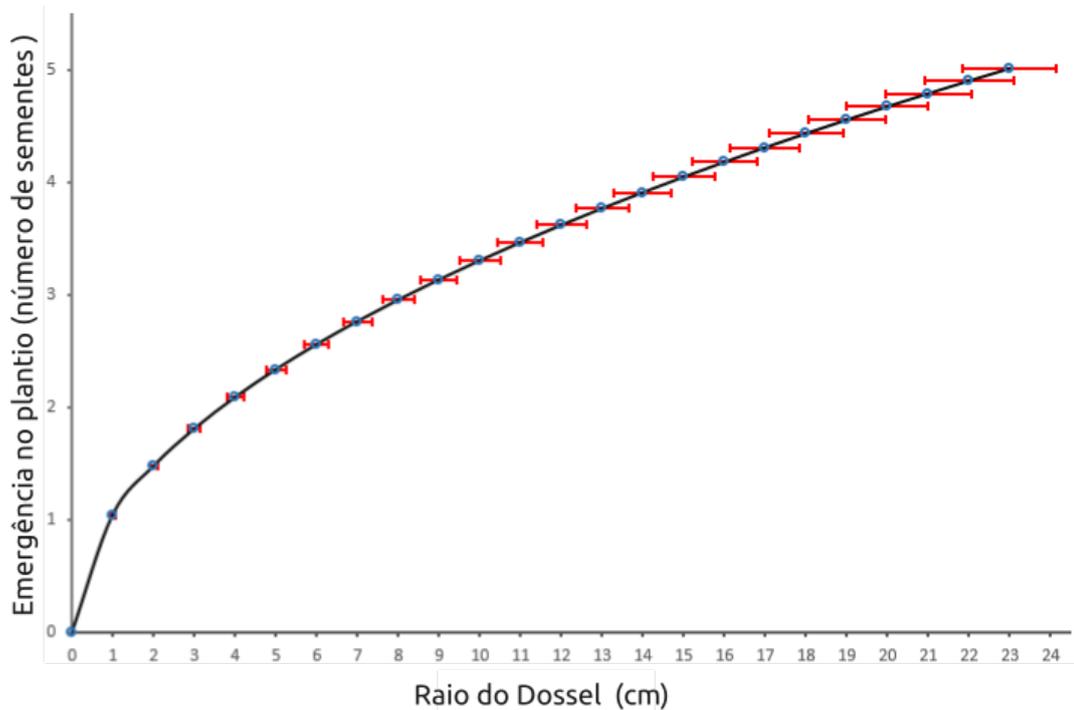
Figura 63 – Histograma para bloco número 30, com eixo x de raios em cm e, com eixo y de número de ocorrências.



Para contagem da presença de plantas de milho, foi necessário realizar uma normalização com raio mínimo ($r_{min} = 1$) e máximo ($r_{max} = 23$). Considerando que para o menor raio, a menor unidade da semente emergida correspondeu em 1, enquanto para o maior raio correspondeu de 5 sementes emergidas. Dito isso, a quantidade das plantas de milho foram estimadas com o número de ocorrências mencionadas anteriormente, os quais resultaram para o bloco 7 (menor bloco) em 247, entanto para o bloco 39 (maior bloco) resultaram em 527, já o bloco 30 utilizado como análise central contabilizou 507 plantas de milho. Assim, com o Algoritmo 9 utilizado o total foi de 16 598 plantas de milho emergidas no estádio V2.

Contudo, os resultados de contagem das plantas de milho anteriormente mencionados, estão baseados na calibração de um modelo projetado, o qual é mostrado na Figura 65. Além disso, o modelo incluiu os desvios padrões referentes aos raios de abertura do dossel.

Figura 65 – Modelo de estimativa para contagem das plantas de milho, obtidos a partir do método computacional.



Nesse sentido, vale resaltar que o uso do Algoritmo 9 baseado no modelo calibrado da Figura 65, correspondeu na contagem da emergência do plantio, ou seja, o número das sementes emergidas no estágio V2 da área experimental, e esses números são exibidos pela Tabela 15.

Com tudo isso, é importante destacar que com uso do modelo estabelecido entre o raio da abertura do dossel da planta e altura (Figura 48), assim como dos dados obtidos com a curva de crescimento do milho e dos histogramas de distribuição foi possível inferir uma relação entre as medidas reais (verdade de campo) e as medidas computacionais apresentadas nos histogramas relacionados ao tamanho de raio e número de ocorrências (Figura 65). Nesse contexto, foi observado que para o estágio V2 a altura real das plantas amostradas foi igual a $29,80 \pm 2,59$ cm. Logo considerando o valor médio dos raios obtidos computacionalmente, foi igual à 12 unidades de medida o que corresponde a um altura de 20,78 unidades de medidas, valor obtido com uso de modelo de seções cônicas, foi possível encontrar um coeficiente $\alpha = 0,69$ entre essas unidades. Finalmente, é possível observar que com os modelos desenvolvidos existe oportunidade de avaliar automaticamente a emergência e altura do milho no estágio V2, importante como indicador de crescimento para os seguintes estádios vegetativos do milho.

Tabela 15 – Quantidade de sementes emergidas para cada um dos sítios específicos (blocos) na cultura de milho.

Número do bloco (sítio específico)	Emergências (sementes de milho)
Bloco 01	414
Bloco 02	328
Bloco 03	297
Bloco 04	264
Bloco 05	268
Bloco 06	266
Bloco 07	247
Bloco 09	434
Bloco 10	448
Bloco 11	470
Bloco 12	467
Bloco 13	467
Bloco 14	475
Bloco 15	440
Bloco 17	401
Bloco 18	441
Bloco 19	465
Bloco 20	487
Bloco 21	501
Bloco 22	498
Bloco 23	474
Bloco 24	490
Bloco 25	476
Bloco 26	500
Bloco 27	519
Bloco 28	504
Bloco 29	474
Bloco 30	507
Bloco 31	516
Bloco 32	467
Bloco 34	526
Bloco 35	534
Bloco 36	500
Bloco 37	494
Bloco 38	497
Bloco 39	527
Bloco 40	515

3.12 Resultados obtidos através da etapa de treinamento dos classificadores SVM

Nesta seção, são apresentados o desempenho dos resultados, considerando o uso de classificadores SVM com funções de *kernel* Sigmoide e Gaussiano. Para a validação de cada

classificador SVM, foram considerados parâmetros de acurácia, precisão, revocação e F1-score, visando classificar a distribuição da ocorrência pós-emergência das plantas de milho (*Zea mays. L.*), no estágio fenológico V2. Nesse sentido, os histogramas analisados descrevem um padrão que se encontra próximo às distribuições Gaussianas (Figura 64), o que levou a organização de vetores das características que acompanham esse padrão considerado (Algoritmo 11). Logo, foi possível organizar 503 vetores, cada um deles constituído de raio e número de ocorrências.

Dois conjuntos de dados foram organizados de acordo com intervalo de valores do desvio padrão e média dos histogramas obtidos (Figura 64). Dito isso, foram encontrados valores sigma $\sigma_{min} = 1,86$ e $\sigma_{max} = 3,42$, assim como, valores média $\mu_{min} = 5,06$ e $\mu_{max} = 9,09$, o que possibilitou estabelecer um limiar que relaciona uma maior e menor dispersão de observações das ocorrências das distribuições observadas nesses histogramas. Portanto, um limiar $\sigma_t = 2,5$ foi calculado de acordo ao Algoritmo 11. Desta forma, histogramas com $\sigma \geq 2,5$ formaram a classe 1, e histogramas com $\sigma < 2,5$ formaram a classe 2. Logo, a classe 1 foi composta por 230 vectores e a classe 2 por 273 vectores.

Para prosseguir os ensaios para treinamento e teste dos Classificadores, foram constituídos proporções de faixas ψ de acordo ao Algoritmo 11. Nesse sentido, faixas ψ_1 e ψ_2 de 80% e 20% respectivamente, organizaram o Classificador 1 em 184 e 55 vectores da classe 1 e classe 2 respectivamente. Por outro lado, o Classificador 2 foi constituído de faixas complementarias em 20% e 80% respectivamente, contendo assim, 46 e 218 vectores para classe 1 e 2 respectivamente. Desta forma, para validação dos classificadores SVM, os mesmos foram treinados e testados considerando as proporções de 50% de treinamento e 50% de teste, 70% de treinamento e 30% de teste, bem como 80% de treinamento e 20% de teste.

A Tabela 16 apresenta os resultados obtidos pelos testes realizados em referência à classificação dos padrões do estágio de pós-emergência V2, utilizando *kernel Sigmoide* considerando a relação 50%-50% para treinamento e teste respectivamente.

Tabela 16 – Resultados para a classificação dos padrões do estágio de pós-emergência V2 com *kernel Sigmoide* utilizando a relação de 50%-50%.

Classificadores	Precisão	Revocação	F1-Score	Vetores de suporte	Acurácia (%)
Classificador 1	0,81	0,92	0,86	92	0,78
Classificador 2	0,82	0,94	0,88	109	0,79

As Figuras 66 e 68, ilustram as matrizes de confusão. Enquanto, as Figuras 67 e 69 exibem suas curvas ROC, referentes aos resultados da Tabela 16.

Figura 66 – Matriz de confusão do Classificador 1 com *kernel Sigmoide* e relação 50%-50%.

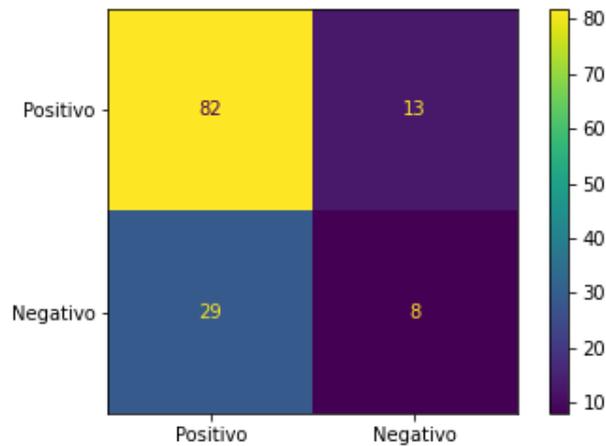


Figura 67 – Curva ROC do Classificador 1 com *kernel Sigmoide* e relação 50% - 50%.

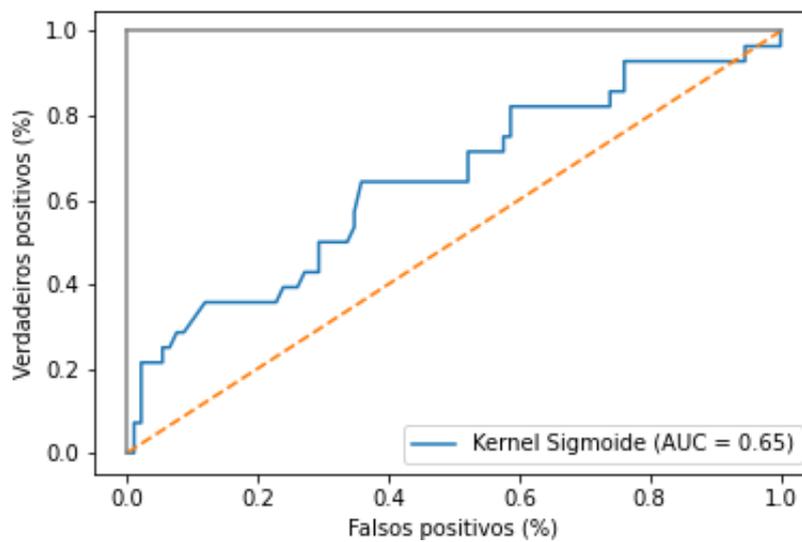


Figura 68 – Matriz de confusão do Classificador 2 com *kernel Sigmoide* e relação 50% - 50%.

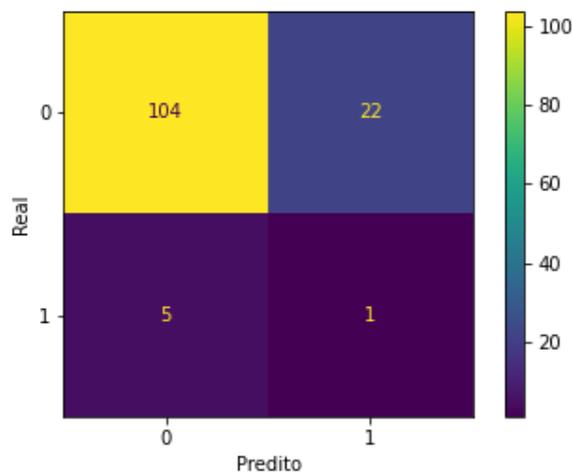
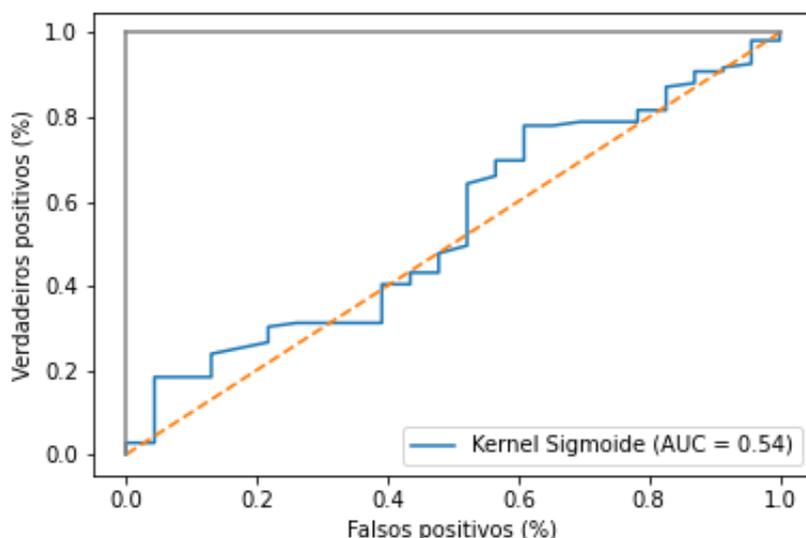


Figura 69 – Curva ROC do Classificador 2 com *kernel Sigmoide* e relação 50% - 50%.



A Tabela 17 apresenta os resultados obtidos pelos testes realizados em referência à classificação dos padrões do estádio de pós-emergência V2, utilizando *kernel Sigmoide* considerando a relação 70%-30% para treinamento e teste respectivamente.

Tabela 17 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com *kernel Sigmoide* utilizando a relação de 70%-30%.

Classificadores	Precisão	Revocação	F1-Score	Vetores de suporte	Acurácia (%)
Classificador 1	0,74	0,87	0,80	55	0,67
Classificador 2	0,78	0,76	0,77	66	0,62

As Figuras 70 e 72, ilustram as matrizes de confusão. Enquanto, as Figuras 71 e 73 exibem suas curvas ROC, referentes aos resultados da Tabela 17.

Figura 70 – Matriz de confusão do Classificador 1 com *kernel Sigmoide* e relação 70%-30%.

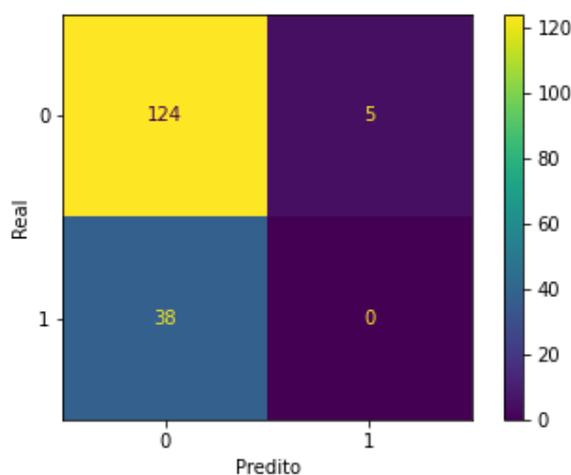


Figura 71 – Curva ROC do Classificador 1 com *kernel Sigmoide* e relação 70% - 30%.

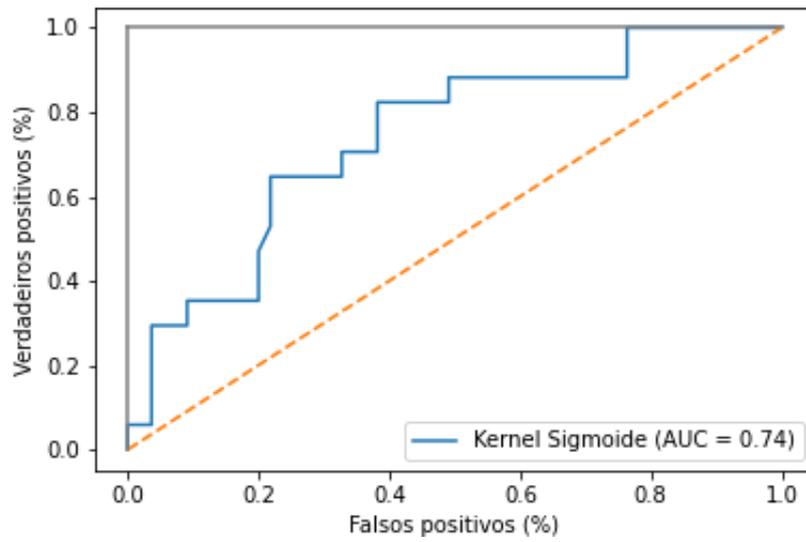


Figura 72 – Matriz de confusão do Classificador 2 com *kernel Sigmoide* e relação 70%-30%.

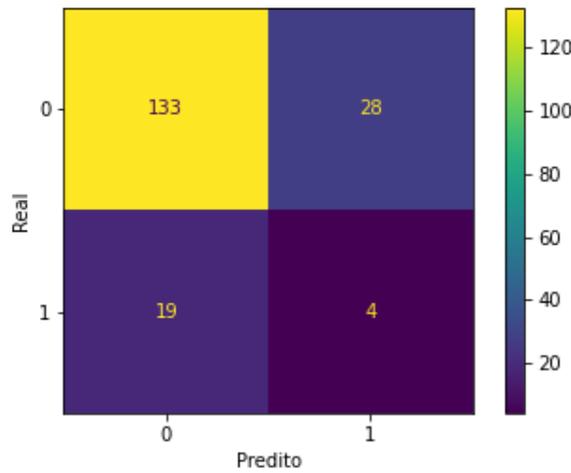
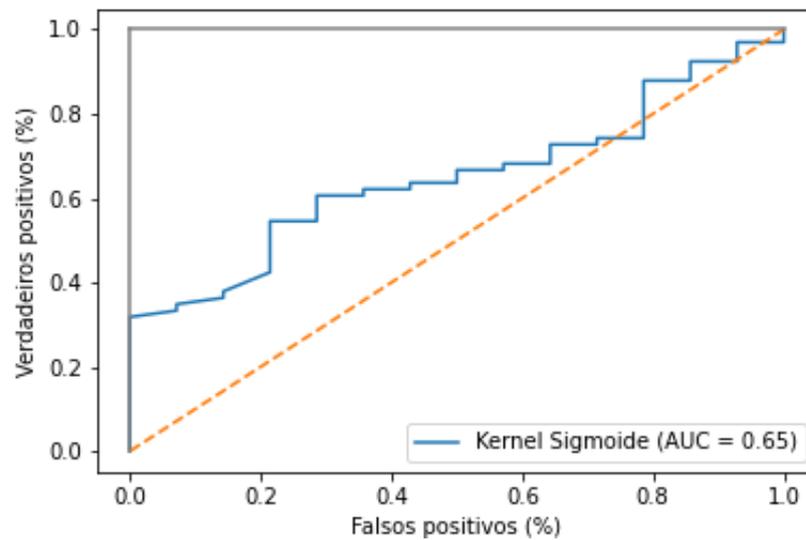


Figura 73 – Curva ROC do Classificador 2 com *kernel Sigmoide* e relação 70% - 30%.



A Tabela 18 apresenta os resultados obtidos pelos testes realizados em referência à classificação dos padrões do estádio de pós-emergência V2, utilizando *kernel Sigmoide* considerando a relação 80%-20% para treinamento e teste respectivamente.

Tabela 18 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com *kernel Sigmoide* utilizando a relação de 80%-20%.

Classificadores	Precisão	Revocação	F1-Score	Vetores de suporte	Acurácia (%)
Classificador 1	0,84	0,97	0,90	37	0,83
Classificador 2	0,83	0,98	0,90	44	0,81

As Figuras 74 e 76 ilustram as matrizes de confusão. Enquanto, que as Figuras 75, 77 exibem suas curvas ROC, referentes aos resultados da Tabela 18.

Figura 74 – Matriz de confusão do Classificador 1 com *kernel Sigmoide* e relação 80%-20%.

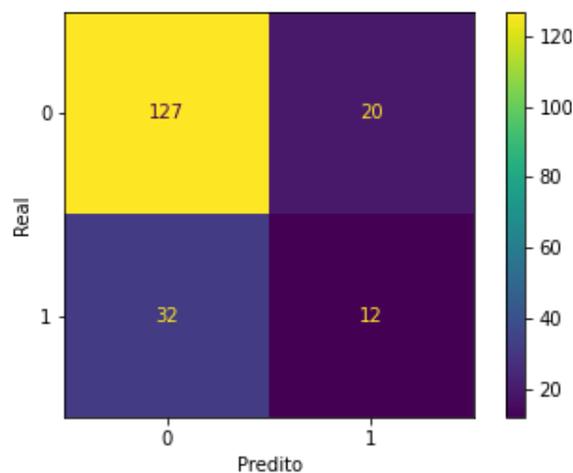


Figura 75 – Curva ROC do Classificador 1 com *kernel Sigmoide* e relação 80% - 20%.

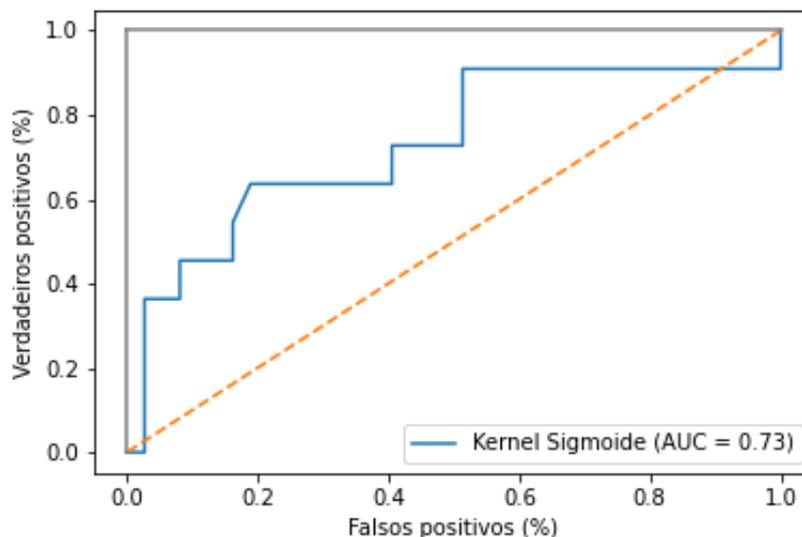


Figura 76 – Matriz de confusão do Classificador 2 com *kernel Sigmoide* e relação 80%-20%.

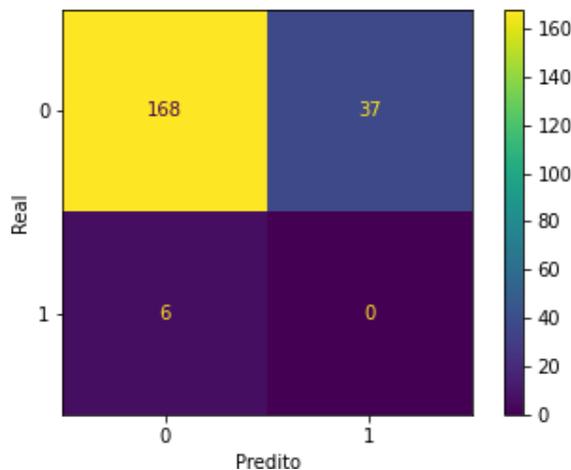
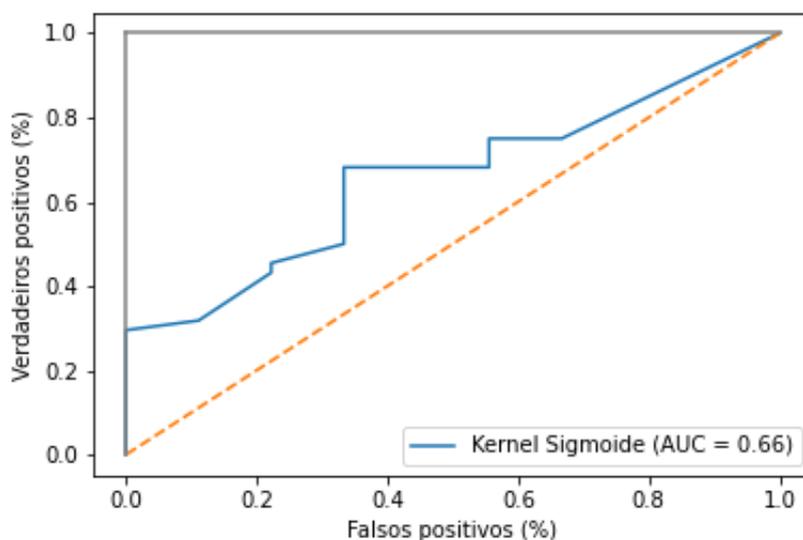


Figura 77 – Curva ROC do Classificador 2 com *kernel Sigmoide* e relação 80% - 20%.



A Tabela 19 apresenta os resultados obtidos pelos testes realizados em referência à classificação dos padrões do estádio de pós-emergência V2, utilizando *kernel Gaussiano* considerando a relação 50%-50% para treinamento e teste respectivamente.

Tabela 19 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com *kernel Gaussiano* utilizando a relação de 50%-50%.

Classificadores	Precisão	Revocação	F1-Score	Vetores de suporte	Acurácia (%)
Classificador 1	0,82	0,97	0,89	92	0,82
Classificador 2	0,84	0,94	0,88	198	0,80

As Figuras 78 e 80 ilustram as matrizes de confusão. Enquanto, as Figuras 79 e 81 exibem suas curvas ROC, referentes aos resultados da Tabela 19.

Figura 78 – Matriz de confusão do Classificador 1 com *kernel Gaussiano* e relação 50%-50%.

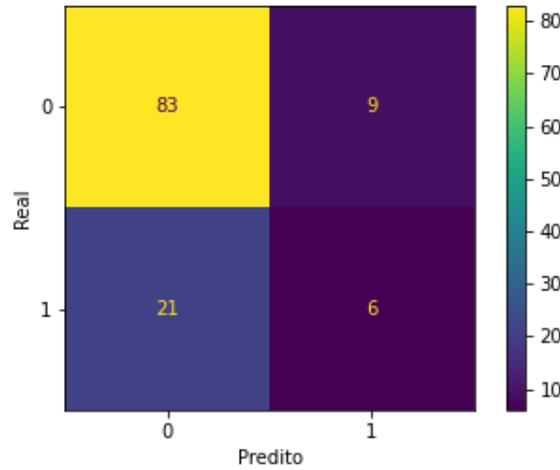


Figura 79 – Curva ROC do Classificador 1 com *kernel Gaussiano* e relação 50% - 50%.

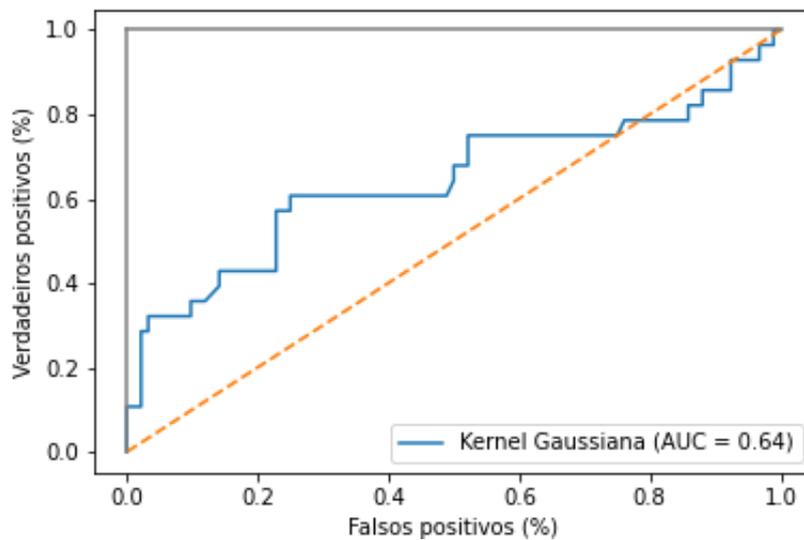


Figura 80 – Matriz de confusão do Classificador 2 com *kernel Gaussiano* e relação 50%-50%.

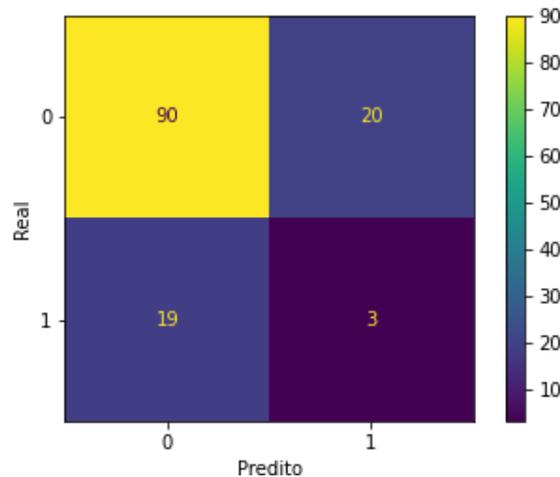
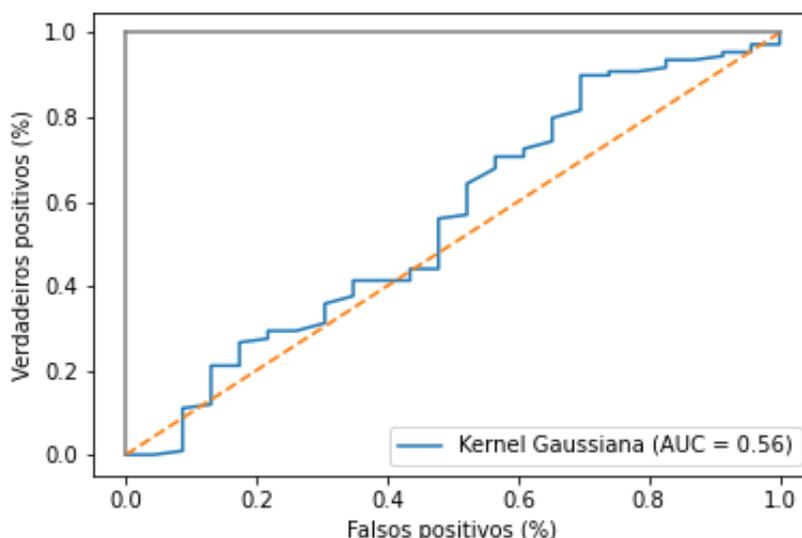


Figura 81 – Curva ROC do Classificador 2 com *kernel Gaussiano* e relação 50% - 50%.



A Tabela 20 apresenta os resultados obtidos pelos testes realizados em referência à classificação dos padrões do estádio de pós-emergência V2, utilizando *kernel Gaussiano* considerando a relação 70%-30% para treinamento e teste respectivamente.

Tabela 20 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com *kernel Gaussiano* utilizando a relação de 70%-30%.

Classificadores	Precisão	Revocação	F1-Score	Vetores de suporte	Acurácia (%)
Classificador 1	0,83	1,00	0,91	55	0,85
Classificador 2	0,84	1,00	0,91	66	0,84

As Figuras 82 e 84 ilustram as matrizes de confusão. Enquanto, as figuras 83 e 85 exibem suas curvas ROC, referentes aos resultados da Tabela 20.

Figura 82 – Matriz de confusão do Classificador 1 com *kernel Gaussiano* e relação 70%-30%.

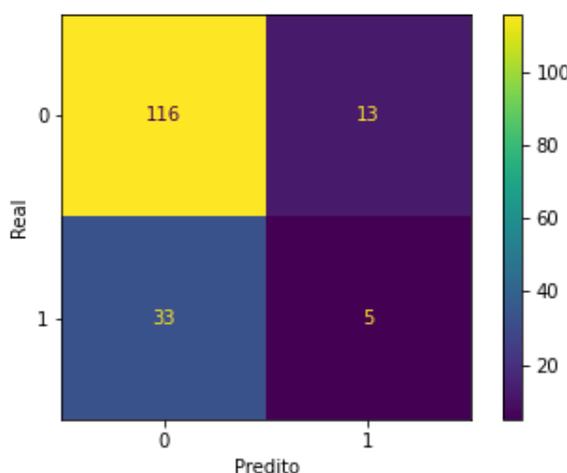


Figura 83 – Curva ROC do Classificador 1 com *kernel Gaussiano* e relação 70% - 30%.

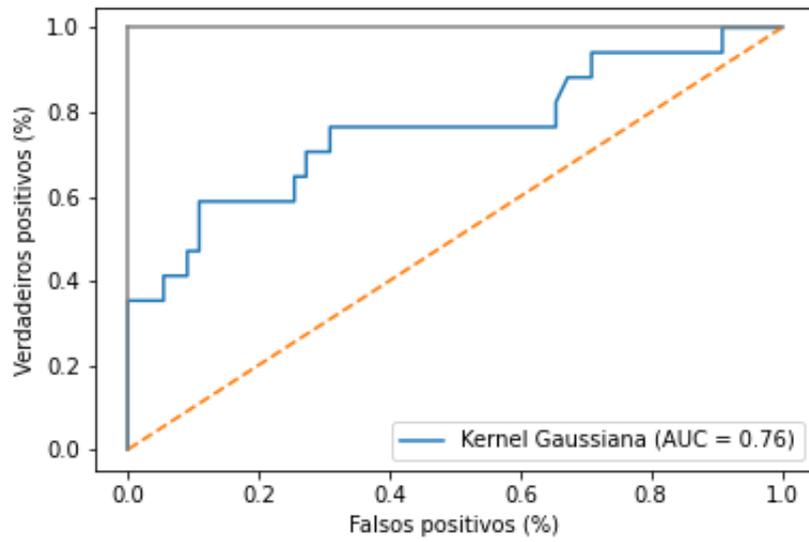


Figura 84 – Matriz de confusão do Classificador 2 com *kernel Gaussiano* e relação 70%-30%.

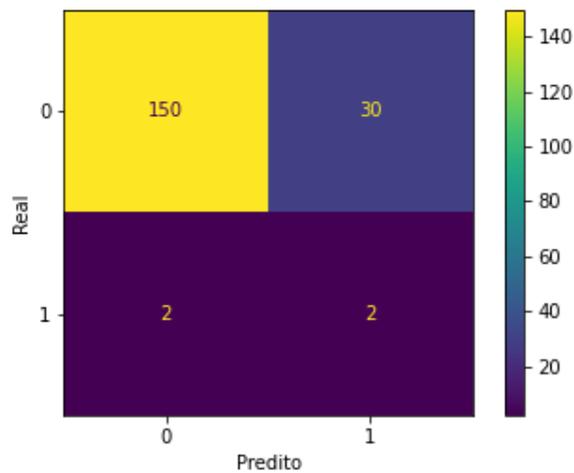
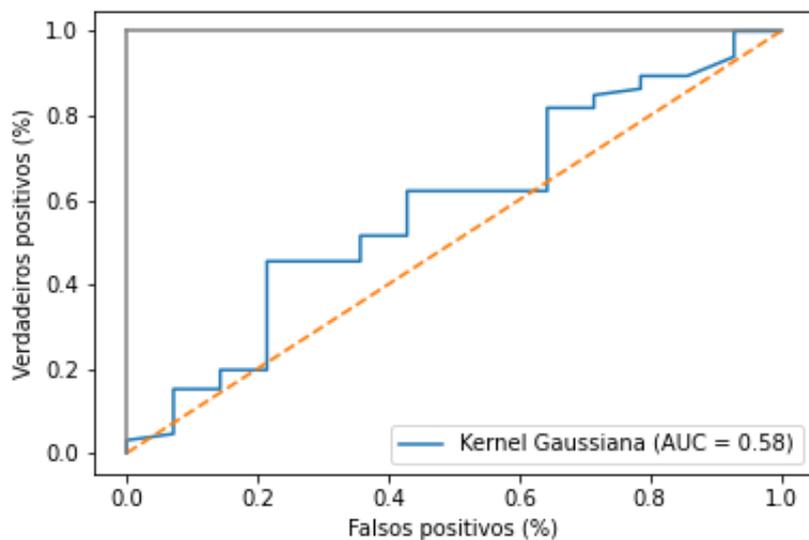


Figura 85 – Curva ROC do Classificador 2 com *kernel Gaussiano* e relação 70% - 30%.



A Tabela 21 apresenta os resultados obtidos pelos testes realizados em referência à classificação dos padrões do estádio de pós-emergência V2, utilizando *kernel Gaussiano* considerando a relação 80%-20% para treinamento e teste respectivamente.

Tabela 21 – Resultados para a classificação dos padrões do estádio de pós-emergência V2 com *kernel Gaussiano* utilizando a relação de 80%-20%.

Classificadores	Precisão	Revocação	F1-Score	Vetores de suporte	Acurácia (%)
Classificador 1	0,82	1,00	0,90	37	0,83
Classificador 2	0,84	0,95	0,89	44	0,81

As Figuras 86, 88, ilustram as matrizes de confusão referentes aos resultados da Tabela 21, da mesma forma que as figuras 87, 89 ilustram suas curvas ROC.

Figura 86 – Matriz de confusão do Classificador 1 com *kernel Gaussiano* e relação 80%-20%.

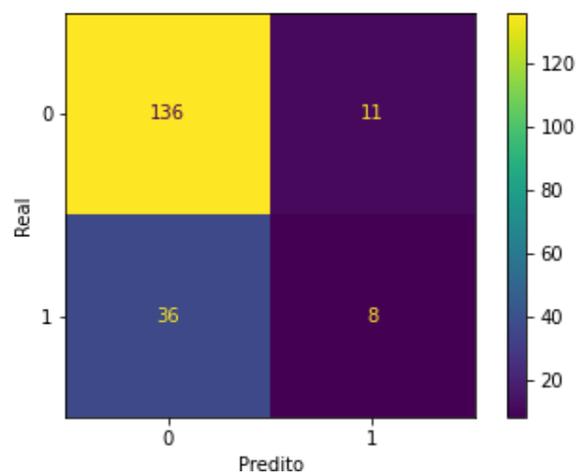


Figura 87 – Curva ROC do Classificador 1 com *kernel Gaussiano* e relação 80% - 20%.

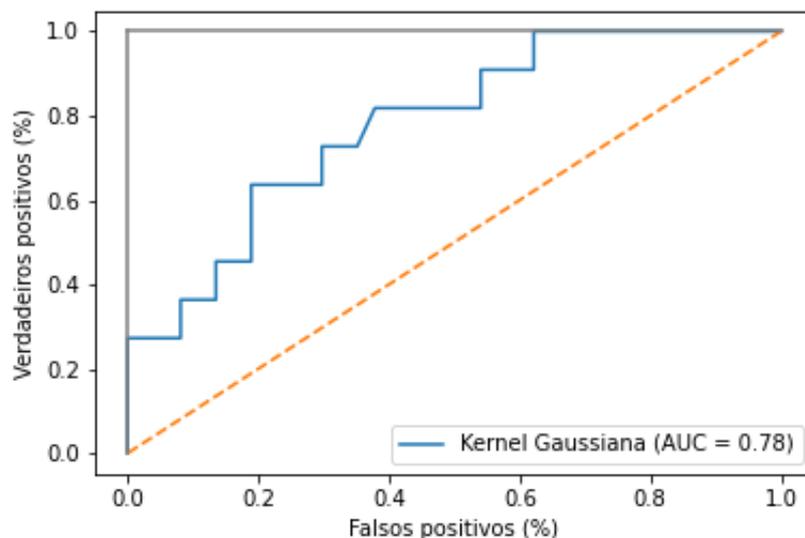


Figura 88 – Matriz de confusão do Classificador 2 com *kernel Gaussiano* e relação 80%-20%.

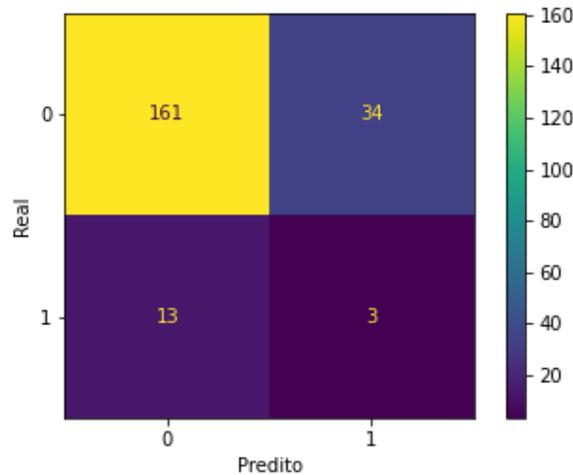
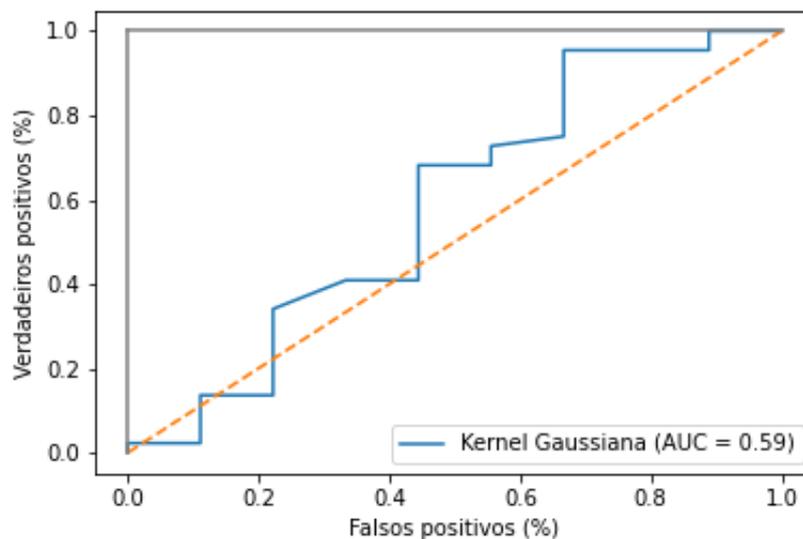


Figura 89 – Curva ROC do Classificador 2 com *kernel Gaussiano* e relação 80% - 20%.



3.12.1 Avaliação da etapa de treinamento do conjunto de classificadores SVM

Analisando os resultados dos experimentos realizados para treinamentos e testes dos classificadores SVM, foi possível avaliar que a taxa de precisão se manteve com um valor considerado satisfatório (acima do 74%) para a classificação de padrões pós-emergência V2 em relação à abertura do dossel das plantas de milho (*Zea mays L*). Além disso, foi observado que para avaliação da taxa de acurácia dos Classificadores a função de *kernel* Sigmoide com proporção 70%-30% de treinamento e teste, foram de percentagem baixa, ou seja, em 67% e 62% para Classificador 1 e 2 respectivamente.

Dito isso, a proporção de 50% utilizada para treinamento e teste, destaca-se o desempenho dos Classificadores SVM com função *kernel* Sigmoide, no que se refere a taxa de acurácia,

com resultados de 78% e 79% para Classificador 1 e 2, respectivamente. Da mesma forma, desempenho dos Classificadores SVM com função *kernel* Gaussiano obtiveram resultados altos em 82% e 80% para Classificador 1 e 2, respectivamente. Quanto analisada a taxa de precisão, os Classificadores com função Gaussiano obtiveram os melhores resultados de 82% e 84% para Classificador 1 e 2, respectivamente.

Para a proporção de 70% e 30% referentes a treinamento e teste, respectivamente, é interessante ressaltar que desempenho dos Classificadores SVM com função *kernel* Gaussiano, obtiveram melhores resultados para a classificação dos padrões do estádio pós-emergência V2, com taxa de acurácia de 85% e 84% para Classificador 1 e 2, respectivamente. Enquanto, análise da taxa de precisão, Classificadores com função *kernel* Sigmoide obtiveram resultados de 74% e 78%, da mesma forma, Classificadores com função *kernel* Gaussiano resultaram com taxa de precisão de 83% e 84%, respectivamente.

Para a proporção de 80% e 20% referentes a treinamento e teste, respectivamente, o desempenho dos Classificadores SVM com função *kernel* Sigmodal e Gaussiano obtiveram os mesmos melhores resultados, com taxa de acurácia de 83% e 81% para Classificador 1 e 2, respectivamente para a classificação dos padrões do estádio pós-emergência V2. Enquanto, avaliação da taxa de precisão, os Classificadores SVM *kernel* Sigmoide obtiveram melhores resultados com taxas de 84% e 83%, da mesma forma, os Classificadores SVM *kernel* Gaussiano obtiveram resultados para acurácia de 82% e 83%, respectivamente.

Capítulo 4

CONCLUSÕES

Neste capítulo é apresentado as conclusões correspondentes do trabalho desenvolvido, em resposta aos objetivos específicos e ao objetivo geral. Contribuindo no estabelecimento de um método para contagem de plantas de milho, além da classificação de padrões com abertura do dossel na cultura de milho (*Zea mays L.*), utilizando imagens multiespectrais RGB registradas por Drone.

4.1 Principais conclusões

A contagem de plantas de milho é de grande importancia para o ciclo da agricultura, tornando-a relevante na tomada de decisões de acordo à aplicação do especialista na área. Porém, o trabalho desenvolvido apresentou um método para contagem de plantas de milho baseado no processamento digital de imagens multiespectrais utilizando Drones em ambiente de campo.

A validação do método foi estabelecida considerando a aquisição de imagens com voos de Drones em uma cultura de milho, bem como estruturando ortomosaicos de canais RGB obtidos por sensores multiespectrais embarcados. Essa abordagem, permitiu verificar a qualidade das imagens multiespectrais coletadas pelo planejamento do Drone em área de cultura, o que visou em um GSD de $2,08 \text{ cm/s}$. Além disso, resultados da calibração de medidas reais em campo demonstraram que estadio de pós-emergência V2 tem relação com a realidade de campo.

Para resultados obtidos na etapa de pré-processamento das imagens, foi possível concluir que a transformação geométrica da rotação em torno ao objeto com relação aos ortomosaicos possibilitou as etapas iniciais de processamento das imagens, encontrando-se o ângulo de inclinação com valor de $\theta \approx 25^\circ$. Da mesma forma, a etapa de melhora de brilho, contraste e equalização do histograma adaptativo estabeleceu melhorias nas intensidades da imagem para diversos espaços de cores, levando a concluir que espaço de cores HSV representa melhor viabilidade para as etapas seguintes de processamento da imagem.

Para a etapa de segmentação da imagem, as técnicas utilizadas foram baseadas em uso de e espaço de cores HSV estabelecida dentro de intervalos de cores validados em conjunto com

operadores morfológicos, o que viabilizou a separação entre as plantas de milho e outros objetos presentes na imagem.

Em relação à localização da planta de milho nas culturas, foi possível concluir que o uso integrado da transformada de distância, bem como o coeficiente de correlação normalizado do *template matching*, permitiram localizar os máximos locais em relação à distância central de planta de milho. Onde foi possível identificar que a inclusão de templates baseados em máscaras de Chamfer viabiliza possíveis candidatos para a localização de máximos locais em relação à transformada de distância.

A partir da extração de características foi possível concluir que os códigos de cadeia descreviam os contornos para representar características circulares. Da mesma forma, foram encontrados padrões de histograma espectrais que descrevem larguras de comprimento mais abertas ou fechadas, ou ambas relacionadas ao estágio de crescimento das plantas.

Com relação aos resultados obtidos para a classificação do estágio de pós-emergência V2 de plantas de milho (*Zea mays L.*), a partir dos resultados obtidos foi possível concluir que para sua classificação, classificador 1 com função de *kernel* Sigmoide quando operado de 80% treinamento e 20% de testes, os mesmos foram mais precisos apresentando taxa de acerto de 84%. Da mesma forma, para a classificação do classificador 2 com uma função de *kernel* Gaussiano com proporção de 70% de treinamento e 30% de testes, foi mais precisa, com taxa de acerto de 84%.

Finalmente, foi possível concluir que a organização das técnicas avaliadas possibilitaram a validação da contagem das plantas de milho (*Zea mays L.*) em estágio de pós-emergência ou estágio fenológico V2, o que é de interesse para agricultura 4.0, pesquisas e atendimento ao setor produtivo.

4.2 Sugestões para trabalhos futuros

Como contribuição para trabalhos futuros foram considerados:

- Estudar e avaliar algoritmo de separação de diversas espécies de plantas, garantido que planta de milho seja separado ainda mais confiavelmente;
- Evoluir métodos de segmentação considerando segmentação semântica ou por instâncias;
- Estudar e avaliar o método desenvolvido no domínio de frequência;
- Estudar e avaliar o uso de modelos não supervisionados para reconhecimento de padrões;
- Estudar e avaliar a oportunidade de embarcar o algoritmo desenvolvido para análise em tempo real durante a realização do tempo dos voos

REFERÊNCIAS

ABER, J. S.; MARZOLFF, I.; RIES, J. *Small-format aerial photography: Principles, techniques and geoscience applications*. San Diego, CA, USA: Elsevier Science, 2016. ISBN 0444638237. Citado 6 vezes nas páginas 41, 42, 43, 44, 46 e 47.

AHMAD, I.; SAEED, U.; FAHAD, M.; ULLAH, A.; RAHMAN, M. H. ur; AHMAD, A.; JUDGE, J. Yield forecasting of spring maize using remote sensing and crop modeling in faisalabad-punjab pakistan. *Journal of the Indian Society of Remote Sensing*, v. 46, p. 1701–1711, 2018. ISSN 0974-3006. Citado na página 16.

ALMADHOUN, R.; TAHA, T.; SENEVIRATNE, L.; ZWEIRI, Y. A survey on multi-robot coverage path planning for model reconstruction and mapping. *SN Applied Sciences*, Springer, v. 1, n. 8, p. 1–24, 2019. Disponível em: <<https://doi.org/10.1007/s42452-019-0872-y>>. Citado na página 42.

AWAD, M.; KHANNA, R. *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. 1st. ed. USA: Apress, 2015. ISBN 1430259892. Citado na página 81.

BAH, M. D.; HAFIANE, A.; CANALS, R. CRoWNet: Deep network for crop row detection in UAV images. *IEEE Access*, v. 8, p. 5189–5200, 2020. ISSN 2169-3536. Citado na página 17.

BARKÓCZI, N.; BERTALAN, L.; SZABÓ, G.; DEÁK, M.; KARI, S.; VIZI, J. C.; ZAGORÁ CZ, M.; SIK, A.; RIEDEL, M.; KOHÁN, B. et al. Examples from the boundaries of geographic survey: Architecture and flood modeling. In: *Small Flying Drones*. Cham: Springer, 2018. p. 127–156. Disponível em: <https://doi.org/10.1007/978-3-319-66577-1_5>. Citado na página 44.

BENDIG, J.; YU, K.; AASEN, H.; BOLTEN, A.; BENNERTZ, S.; BROSCHEIT, J.; GNYP, M. L.; BARETH, G. Combining uav-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, v. 39, p. 79–87, 2015. ISSN 1569-8432. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0303243415000446>>. Citado na página 15.

BENJAMIN, L. Growth analysis, crops. In: THOMAS, B.; MURRAY, B. G.; MURPHY, D. J. (Ed.). *Encyclopedia of Applied Plant Sciences (Second Edition)*. Second edition. Oxford: Academic Press, 2017. p. 23–28. ISBN 978-0-12-394808-3. Citado na página 31.

BEYERER, J.; RICHTER, M.; NAGEL, M. *lassifiers and principles*. Berlin, Boston: Walter de Gruyter GmbH, 2018. (De Gruyter textbook). ISBN 978-3-11-053793-2. Disponível em: <<https://doi.org/10.1515/9783110537949>>. Citado 4 vezes nas páginas 82, 83, 84 e 85.

BHUYAN, M. K. *Computer vision and image processing: fundamentals and applications*. Boca Raton, FL: CRC Press/Taylor & Francis Group, 2019. ISBN 978-0-8153-7084-0 978-0-367-26573-1. Citado 3 vezes nas páginas 50, 62 e 65.

BISHOP, C. M. *Pattern recognition and machine learning*. Berlin, Heidelberg: Springer, 2006. (Information science and statistics). ISBN 978-0-387-31073-2. Citado na página 70.

BLANCHET, G.; CHARBIT, M. *Digital Signal and Image Processing Using MATLAB (Digital Signal and Image Processing Series)*. London: ISTE, 2006. ISBN 1905209134. Citado na página 59.

BORGEFORS, G. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 10, n. 6, p. 849–865, 1988. Citado na página 69.

BOSE, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*. New York, NY, USA: Association for Computing Machinery, 1992. (COLT '92), p. 144–152. ISBN 089791497X. Citado na página 82.

BRAGA-NETO, U. *Fundamentals of Pattern Recognition and Machine Learning*. 1st ed. 2020 edição. ed. Switzerland: Springer, 2020. ISBN 978-3-030-27656-0. Citado 2 vezes nas páginas 81 e 86.

BURGER, W.; BURGE, M. J. *Principles of Digital Image Processing: Core Algorithms*. 2009ª edição. ed. London: Springer, 2009. ISBN 978-1-84800-194-7. Citado 8 vezes nas páginas 54, 63, 69, 72, 73, 74, 76 e 77.

CHANG, C.-C.; LIN, C.-J. Libsvm: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, Acm New York, NY, USA, v. 2, n. 3, p. 1–27, 2011. Citado na página 27.

CHEN, Y.; HOU, C.; TANG, Y.; ZHUANG, J.; LIN, J.; HE, Y.; GUO, Q.; ZHONG, Z.; LEI, H.; LUO, S. Citrus tree segmentation from uav images based on monocular machine vision in a natural orchard environment. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 19, n. 24, p. 5558, 2019. Citado na página 27.

CLELAND, E. E.; CHUINE, I.; MENZEL, A.; MOONEY, H. A.; SCHWARTZ, M. D. Shifting plant phenology in response to global change. *Trends in ecology & evolution*, Elsevier, v. 22, n. 7, p. 357–365, 2007. Citado na página 32.

COMER, R.; KINN, G.; LIGHT, D.; MONDELLO, C. Talking digital. *Photogrammetric Engineering and Remote Sensing*, [Falls Church, Va.] American Society of Photogrammetry., v. 64, n. 12, p. 1139–1142, 1998. Citado na página 41.

CORTES, C.; VAPNIK, V. Support-Vector Networks. *Machine learning*, v. 20, n. 3, p. 273–297, 1995. ISSN 08859000. Citado 2 vezes nas páginas 82 e 83.

CRUVINEL, P.; MINATEL, E. Image processing in automated pattern classification of oranges. In: AMERICAN SOCIETY OF AGRICULTURAL AND BIOLOGICAL ENGINEERS. *World Congress of Computers in Agriculture and Natural Resources, Proceedings of the 2002 Conference*. 2002. p. 56. Disponível em: <<https://elibrary.asabe.org/abstract.asp?aid=8312&t=1>>. Citado na página 16.

CUI, Z.; QI, W.; LIU, Y. A Fast Image Template Matching Algorithm Based on Normalized Cross Correlation. *Journal of Physics: Conference Series*, v. 1693, p. 012163, dez. 2020. ISSN 1742-6596. Publisher: IOP Publishing. Disponível em: <<https://doi.org/10.1088/1742-6596/1693/1/012163>>. Citado na página 23.

DAVIES, E. R. *Computer Vision: Principles, Algorithms, Applications, Learning*. 5th edition. ed. London, United Kingdom ; Cambridge, MA, United States: Academic Press, 2017. ISBN 978-0-12-809284-2. Citado 2 vezes nas páginas 50 e 67.

DEISENROTH, M. P.; FAISAL, A. A.; ONG, C. S. *Mathematics for Machine Learning*. Cambridge: Cambridge University Press, 2020. Disponível em: <<https://mml-book.github.io/book/mml-book.pdf>>. Citado na página 81.

DISTANTE, A.; DISTANTE, C. *Handbook of image processing and computer vision. Volume 1, Volume 1*. Cham: Springer, 2020. OCLC: 1157084891. ISBN 978-3-030-38148-6. Disponível em: <<http://public.ebib.com/choice/PublicFullRecord.aspx?p=6213178>>. Citado 9 vezes nas páginas 56, 57, 58, 59, 60, 64, 67, 68 e 75.

DORJ, U.-O.; LEE, M.; YUN, S.-s. An yield estimation in citrus orchards via fruit detection and counting using image processing. *Computers and Electronics in Agriculture*, Elsevier, v. 140, p. 103–112, 2017. Citado na página 19.

DOUGHERTY, G. *Digital image processing for medical applications*. Cambridge: Cambridge University Press, 2009. Citado na página 66.

DOUGHERTY, G. *Pattern recognition and classification: an introduction*. Cambridge: Springer Science & Business Media, 2013. Citado na página 71.

ELZINGA, C. L.; SALZER, D. W. *Measuring & monitoring plant populations*. USA: US Department of the Interior, Bureau of Land Management, 1998. Citado na página 32.

FAO; IFAD; UNICEF; WFP; WHO. *The State of Food Security and Nutrition in the World 2021*. Rome, Italy: The Publishing Group - FAO, 2021. 240 p. Citado na página 15.

FERNANDEZ, R.; MONTES, H.; SURDILOVIC, J.; SURDILOVIC, D.; GONZALEZ-DE-SANTOS, P.; ARMADA, M. Automatic detection of field-grown cucumbers for robotic harvesting. *IEEE Access*, IEEE, v. 6, p. 35512–35527, 2018. Citado na página 20.

FORSYTH, D.; PONCE, J. *Computer vision: a modern approach*. 2nd ed. ed. Boston: Pearson, 2012. ISBN 978-0-13-608592-8. Citado na página 56.

FRAZIER, A. E.; SINGH, K. K. *Fundamentals of Capturing and Processing Drone Imagery and Data*. Boca Raton: CRC Press, 2021. ISBN 9780429283239. Citado 11 vezes nas páginas 36, 38, 39, 40, 41, 42, 43, 44, 45, 47 e 48.

FREEMAN, H. On the encoding of arbitrary geometric configurations. *IRE Transactions on Electronic Computers*, IEEE, n. 2, p. 260–268, 1961. Citado na página 53.

FREEMAN, H. Computer processing of line-drawing images. *ACM Computing Surveys (CSUR)*, ACM New York, NY, USA, v. 6, n. 1, p. 57–97, 1974. Citado na página 76.

GAO, J. *Digital analysis of remotely sensed imagery*. New York: McGraw-Hill Education, 2009. ISBN 0071604650. Citado na página 49.

GARCÍA-MARTÍNEZ, H.; FLORES-MAGDALENO, H.; KHALIL-GARDEZI, A.; ASCENCIO-HERNÁNDEZ, R.; TIJERINA-CHÁVEZ, L.; VÁZQUEZ-PEÑA, M. A.; MANCILLA-VILLA, O. R. Digital Count of Corn Plants Using Images Taken by Unmanned Aerial Vehicles and Cross Correlation of Templates. *Agronomy*, v. 10, n. 4, p. 469, abr. 2020. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute. Citado na página 23.

GHAOUI, L. E.; VIALON, V.; RABBANI, T. Safe feature elimination in sparse supervised learning technical report no. *Technical report, UCB/EECS-2010-126, EECS Department, University of California, Berkeley*, 2010. Citado na página 26.

GNÄDINGER, F.; SCHMIDHALTER, U. Digital Counts of Maize Plants by Unmanned Aerial Vehicles (UAVs). *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 9, p. 544, jun. 2017. Citado na página 16.

GONZALEZ, R. C.; WOODS, R. E. *Digital image processing*. New York, NY: Pearson, 2018. ISBN 978-0-13-335672-4. Citado 10 vezes nas páginas 37, 50, 52, 56, 62, 65, 68, 70, 71 e 72.

HALL, O.; DAHLIN, S.; MARSTORP, H.; BUSTOS, M. F. A.; ÖBORN, I.; JIRSTRÖM, M. Classification of maize in complex smallholder farming systems using UAV imagery. *Drones*, Multidisciplinary Digital Publishing Institute, v. 2, n. 3, p. 22, 2018. Citado na página 16.

HASITUYA; ZHONGXIN, C.; FEI, L.; YUNCAI, H. Mapping plastic-mulched farmland by coupling optical and synthetic aperture radar remote sensing. *International Journal of Remote Sensing*, v. 41, p. 7757–7778, 2020. Citado na página 16.

HERNÁNDEZ-HERNÁNDEZ, J.; RUIZ-HERNÁNDEZ, J.; GARCÍA-MATEOS, G.; GONZÁLEZ-ESQUIVA, J.; RUIZ-CANALES, A.; MOLINA-MARTÍNEZ, J. A new portable application for automatic segmentation of plants in agriculture. *Agricultural Water Management*, Elsevier, v. 183, p. 146–157, 2017. Citado na página 21.

HITAM, M. S.; AWALLUDIN, E. A.; YUSSOF, W. N. J. H. W.; BACHOK, Z. Mixture contrast limited adaptive histogram equalization for underwater image enhancement. In: *IEEE. 2013 International conference on computer applications technology (ICCAT)*. Sousse, Tunisia, 2013. p. 1–5. Citado na página 61.

HONG, B.; ZHANG, W.; LIU, W.; YE, J.; CAI, D.; HE, X.; WANG, J. Scaling up sparse support vector machines by simultaneous feature and sample reduction. *J. Mach. Learn. Res.*, v. 20, p. 121–1, 2019. Citado na página 26.

HORNBERG, A. (Ed.). *Handbook of Machine and Computer Vision: The Guide for Developers and Users*. 2ª edição. ed. Weinheim, Germany: Wiley, 2017. ISBN 978-3-527-41339-3. Citado na página 72.

HOSHMAND, R. *Design of Experiments for Agriculture and the Natural Sciences Second Edition*. New York: CRC Press, 2006. Citado na página 30.

JENSEN, J. R. *Introductory Digital Image Processing: A Remote Sensing Perspective*. 4th. ed. USA: Prentice Hall Press, 2015. ISBN 013405816X. Citado na página 43.

JIN, X.; LIU, S.; BARET, F.; HEMERLÉ, M.; COMAR, A. Estimates of plant density of wheat crops at emergence from very low altitude uav imagery. *Remote Sensing of Environment*, v. 198, p. 105–114, 2017. ISSN 0034-4257. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425717302651>>. Citado na página 18.

JOHANSEN, K.; RAHARJO, T.; MCCABE, M. F. Using multi-spectral uav imagery to extract tree crop structural properties and assess pruning effects. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 10, n. 6, p. 854, 2018. Citado na página 17.

JR, F. F. S.; ELLIS, J. M. *Remote Sensing: Principles, Interpretation, and Applications*. Long Grove: Waveland Press, 2020. Citado na página 36.

KALANTAR, B.; IDREES, M.; MANSOR, S.; HALIN, A. A. Smart counting – oil palm tree inventory with uav. *Coordinates e-Navigation Assistance System for Cooperative Collision Avoidance at Sea*, XIII, p. 17–22, 05 2017. Citado na página 22.

KESTUR, R.; ANGURAL, A.; BASHIR, B.; OMKAR, S.; ANAND, G.; MEENAVATHI, M. Tree crown detection, delineation and counting in uav remote sensed images: A neural network based spectral–spatial method. *Journal of the Indian Society of Remote Sensing*, Springer, v. 46, n. 6, p. 991–1004, 2018. Citado na página 16.

KHAN, P. W.; XU, G.; LATIF, M. A.; ABBAS, K.; YASIN, A. Uav’s agricultural image segmentation predicated by clifford geometric algebra. *Ieee Access*, IEEE, v. 7, p. 38442–38450, 2019. Citado na página 22.

KITANO, B. T.; MENDES, C. C.; GEUS, A. R.; OLIVEIRA, H. C.; SOUZA, J. R. Corn plant counting using deep learning and uav images. *IEEE Geoscience and Remote Sensing Letters*, IEEE, 2019. Citado na página 16.

KOC-SAN, D.; SELIM, S.; ASLAN, N.; SAN, B. T. Automatic citrus tree extraction from UAV images and digital surface models using circular Hough transform. *Computers and Electronics in Agriculture*, v. 150, p. 289–301, jul. 2018. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169918302928>>. Citado 2 vezes nas páginas 16 e 17.

KONECNY, G. *Geoinformation: remote sensing, photogrammetry and geographical information systems*. Boca Raton: cRc Press, 2019. ISBN 9780429162480. Citado na página 48.

KRISHNA, K. R. *Agricultural drones: a peaceful pursuit*. Oakville, Canada: CRC Press, 2018. ISBN 9781771885959. Citado na página 36.

KUMAR, J. P.; DOMNIC, S. Image based leaf segmentation and counting in rosette plants. *Information Processing in Agriculture*, v. 6, n. 2, p. 233–246, jun. 2019. ISSN 2214-3173. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2214317318301562>>. Citado na página 16.

LABORDE, D.; MARTIN, W.; SWINNEN, J.; VOS, R. Covid-19 risks to global food security. *Science*, American Association for the Advancement of Science, v. 369, n. 6503, p. 500–502, 2020. ISSN 0036-8075. Citado na página 15.

LAL, R.; STEWART, B. A. *Soil-specific farming: precision agriculture*. Boca Raton: CRC Press, 2015. v. 22. ISBN 9780429172496. Citado na página 30.

LEWIS, J. Fast normalized cross-correlation. *Vision Interface, 1995*, 1995. Citado na página 74.

LILLESAND, T.; KIEFER, R. W.; CHIPMAN, J. *Remote sensing and image interpretation*. New York: John Wiley & Sons, 2015. ISBN 1118919459, 9781118919453. Citado 2 vezes nas páginas 45 e 46.

- LIMA, D. C. d. Métodos para estimativa de imagens nir a partir de imagens de câmeras rgb. Universidade Federal de São Carlos, 2020. Citado na página 51.
- LIU, J. G.; MASON, P. J. *Image processing and GIS for remote sensing: techniques and applications*. United Kingdom: John Wiley & Sons, 2016. ISBN 9781118724200. Citado 2 vezes nas páginas 44 e 47.
- LIU, S.; BARET, F.; ANDRIEU, B.; BURGER, P.; HEMMERLÉ, M. Estimation of wheat plant density at early stages using high resolution imagery. *Frontiers in Plant Science*, v. 8, p. 739, 2017. ISSN 1664-462X. Citado na página 15.
- LOTTE, P.; KHANNA, R.; PFEIFER, J.; SIEGWART, R.; STACHNISS, C. Uav-based crop and weed classification for smart farming. In: IEEE. *2017 IEEE International Conference on Robotics and Automation (ICRA)*. Singapore, 2017. p. 3024–3031. Citado na página 24.
- LU, B.; DAO, P. D.; LIU, J.; HE, Y.; SHANG, J. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sensing*, v. 12, p. 2659, 2020. Citado na página 16.
- LU, J.; LEE, W. S.; GAN, H.; HU, X. Immature citrus fruit detection based on local binary pattern feature and hierarchical contour analysis. *Biosystems Engineering*, Elsevier, v. 171, p. 78–90, 2018. Citado na página 24.
- MAGALHAES, P. C.; DURÃES, F. O. Fisiologia da produção de milho. *Embrapa Milho e Sorgo-Circular Técnica (INFOTECA-E)*, Sete Lagoas: Embrapa Milho e Sorgo, 2006., 2006. Citado na página 32.
- MALAMBO, L.; POPESCU, S. C.; MURRAY, S. C.; PUTMAN, E.; PUGH, N. A.; HORNE, D. W.; RICHARDSON, G.; SHERIDAN, R.; ROONEY, W. L.; AVANT, R. et al. Multitemporal field-based plant height estimation using 3d point clouds generated from small unmanned aerial systems high-resolution imagery. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, v. 64, p. 31–42, 2018. Citado na página 17.
- MALIK, M.; ZHANG, T.; LI, H.; ZHANG, M.; SHABBIR, S.; SAEED, A. *Mature Tomato Fruit Detection Algorithm Based on improved HSV and Watershed Algorithm*. *IFAC-PapersOnLine 51 (17) pp. 431-436*. 2018. Citado na página 21.
- MARINELLO, F. Last generation instrument for agriculture multispectral data collection. *Agricultural Engineering International: CIGR Journal*, v. 19, n. 1, p. 87–93, jun. 2017. ISSN 1682-1130. Citado na página 16.
- MARQUES, O. *Practical image and video processing using MATLAB*. Hoboken, N.J: Wiley/IEEE Press, 2011. OCLC: ocn711777140. ISBN 978-0-470-04815-3. Citado na página 37.
- MCANDREW, A. *A computational introduction to digital image processing*. New York: CRC Press Boca Raton, 2016. v. 2. ISBN 9780367783334. Citado na página 69.
- MENESATTI, P.; ANGELINI, C.; PALLOTTINO, F.; ANTONUCCI, F.; AGUZZI, J.; COSTA, C. Rgb color calibration for quantitative image analysis: The “3d thin-plate spline” warping approach. *Sensors, Molecular Diversity Preservation International*, v. 12, n. 6, p. 7063–7079, 2012. Citado na página 19.

MERCER, J.; FORSYTH, A. R. XVI. functions of positive and negative type, and their connection the theory of integral equations. v. 209, n. 441, p. 415–446, 1909. Publisher: Royal Society. Disponível em: <<https://royalsocietypublishing.org/doi/10.1098/rsta.1909.0016>>. Citado na página 87.

MITCHELL, T. M. *Machine Learning*. 1. ed. USA: McGraw-Hill, Inc., 1997. ISBN 0070428077. Citado na página 81.

MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. *Foundations of Machine Learning*. 2nd. ed. [S.l.]: The MIT Press, 2018. ISBN 0262039400. Citado 3 vezes nas páginas 82, 83 e 86.

MUKHOPADHYAY, J.; DAS, P. P.; CHATTOPADHYAY, S.; BHOWMICK, P.; CHATTERJI, B. N. *Digital geometry in image processing*. New York: CRC Press, 2016. ISBN 9780367380212. Citado na página 74.

MÜLLER, A. G.; BERGAMASCHI, H.; BERGONCI, J. I.; RADIN, B.; FRANÇA, S.; SILVA, M. d. Estimativa do índice de área foliar do milho a partir da soma de graus-dia. *Revista Brasileira de Agrometeorologia*, v. 13, n. 1, p. 65–71, 2005. Citado na página 95.

NGUYEN, T. T.; HOANG, V.-N.; LE, T.-L.; TRAN, T.-H.; VU, H. A vision based method for automatic evaluation of germination rate of rice seeds. In: IEEE. *2018 1st International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*. Ho Chi Minh, 2018. p. 1–6. Citado na página 19.

NIXON, M.; AGUADO, A. *Feature Extraction and Image Processing for Computer Vision*. USA: Elsevier Science, 2019. ISBN 9780128149775. Citado 3 vezes nas páginas 54, 65 e 72.

OLIVER, F. W.; TANSLEY, A. G. Methods of surveying vegetation on a large scale. *New Phytologist*, JSTOR, v. 3, n. 9, p. 228–237, 1904. Citado na página 16.

OTSU, N. A threshold selection method from gray-level histograms. *IEEE Trans. on Systems, Man and Cybernetics*, v. 9, n. 1, p. 62–66, 1979. Citado na página 68.

OWEN, D. C.; BENSI, M. T.; DAVIS, A. P.; AYDILEK, A. H. Measuring soil coverage using image feature descriptors and the decision tree learning algorithm. *Biosystems Engineering*, Elsevier, v. 196, p. 112–126, 2020. Citado na página 25.

PALLOTTINO, F.; FIGORILLI, S.; CECCHINI, C.; COSTA, C. Light drones for basic in-field phenotyping and precision farming applications: Rgb tools based on image analysis. In: *Crop Breeding*. New York, NY: Springer, 2021. p. 269–278. Disponível em: <https://doi.org/10.1007/978-1-0716-1201-9_18>. Citado na página 18.

PEREIRA, T. M.; GASPAR, P. D.; SIMÕES, M. P. Fruit recognition and classification based on svm method for production prediction of peaches-preliminary study. In: *IV Balkan Symposium on Fruit Growing 1289*. Leuven, Belgium: [s.n.], 2019. p. 141–150. Disponível em: <<https://doi.org/10.17660/ActaHortic.2020.1289.21>>. Citado na página 28.

PIZER, S. M.; AMBURN, E. P.; AUSTIN, J. D.; CROMARTIE, R.; GESELOWITZ, A.; GREER, T.; ROMENY, B. ter H.; ZIMMERMAN, J. B.; ZUIDERVELD, K. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, Elsevier, v. 39, n. 3, p. 355–368, 1987. Citado na página 55.

POPP, J.; LAKNER, Z. Global socio-economic and environmental dimensions of gm maize cultivation. Scientific Research Publishing, 2013. Citado na página 15.

- POUND, R.; CLEMENTS, F. E. *A method of determining the abundance of secondary species*. Minnessota, USA: [s.n.], 1898. Citado na página 16.
- PUSCH, M.; MACHADO, T. P.; AMARAL, L. R. Introdução a agricultura de precisão. 03 2019. Disponível em: <www.feagri.unicamp.br/gitap>. Acesso em: 20/05/2022. Citado na página 35.
- RASCHKA, S.; JULIAN, D.; HEARTY, J. *Python deeper insights into machine learning: leverage benefits of machine learning techniques using Python : a course in three modules*. Birmingham, UK: Packt Publishing, 2016. ISBN 978-1-78712-854-5. Citado na página 81.
- RIEHLE, D.; REISER, D.; GRIEPENTROG, H. W. Robust index-based semantic plant/background segmentation for rgb-images. *Computers and Electronics in Agriculture*, Elsevier, v. 169, p. 105201, 2020. Citado na página 22.
- RITCHIE, S. W.; HANWAY, J. J. *How a corn plant develops*. USA, 1989. Disponível em: <<https://dr.lib.iastate.edu/handle/20.500.12876/90185>>. Citado 2 vezes nas páginas 17 e 32.
- RITTER, N.; RUTH, M.; GRISSOM, B. B.; GALANG, G.; HALLER, J.; STEPHENSON, G.; COVINGTON, S.; NAGY, T.; MOYERS, J.; STICKLEY, J. et al. Geotiff format specification geotiff revision 1.0. *SPOT Image Corp*, v. 1, p. 154–172, 2000. Citado na página 49.
- ROSENFELD, A.; PFALTZ, J. L. Sequential operations in digital picture processing. *Journal of the ACM (JACM)*, ACM New York, NY, USA, v. 13, n. 4, p. 471–494, 1966. Citado na página 69.
- ROSENFELD, A.; PFALTZ, J. L. Distance functions on digital pictures. *Pattern recognition*, Elsevier, v. 1, n. 1, p. 33–61, 1968. Citado na página 74.
- SAMUEL, A. L. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.*, IBM Corp., USA, v. 3, n. 3, p. 210–229, jul. 1959. ISSN 0018-8646. Disponível em: <<https://doi.org/10.1147/rd.33.0210>>. Citado na página 81.
- SCHOWENGERDT, R. A. *Remote sensing: models and methods for image processing*. Burlington: Elsevier, 2006. ISBN 0080480586. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780123694072500048>>. Citado na página 45.
- SCHWARTZ, M. D. et al. *Phenology: an integrative environmental science*. Springer, 2003. ISBN 9789400769250. Disponível em: <<https://doi.org/10.1007/978-94-007-6925-0>> Citado na página 32.
- SERRA, J. Introduction to mathematical morphology. *Computer vision, graphics, and image processing*, Academic press, v. 35, n. 3, p. 283–305, 1986. Citado na página 64.
- SHALEV-SHWARTZ, S.; ZHANG, T. Accelerated proximal stochastic dual coordinate ascent for regularized loss minimization. In: *International conference on machine learning*. Beijing, China: PMLR, 2014. p. 64–72. Disponível em: <<https://proceedings.mlr.press/v32/shalev-shwartz14.html>>. Citado na página 27.
- SHANNON, D. K.; CLAY, D. E.; KITCHEN, N. R. *Precision Agriculture Basics*. USA: John Wiley & Sons, 2020. ISBN 978-0-89118-366-2. Citado na página 35.

- SHIBAGAKI, A.; KARASUYAMA, M.; HATANO, K.; TAKEUCHI, I. Simultaneous safe screening of features and samples in doubly sparse modeling. In: BALCAN, M. F.; WEINBERGER, K. Q. (Ed.). *International Conference on Machine Learning*. New York, New York, USA: PMLR, 2016. p. 1577–1586. Disponível em: <<https://proceedings.mlr.press/v48/shibagaki16.html>>. Citado na página 27.
- SHIFERAW, B.; PRASANNA, B. M.; HELLIN, J.; BÄNZIGER, M. Crops that feed the world 6. past successes and future challenges to the role played by maize in global food security. *Food Security*, Springer, v. 3, n. 3, p. 307, 2011. Citado na página 15.
- SMITH, A. R. Color gamut transform pairs. *ACM Siggraph Computer Graphics*, ACM New York, NY, USA, v. 12, n. 3, p. 12–19, 1978. Citado na página 58.
- SOLOMON, C.; BRECKON, T. *Fundamentals of Digital Image Processing: A practical approach with examples in Matlab*. John Wiley Sons, Ltd, 2010. 1-28 p. ISBN 9780470689776. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470689776>>. Citado 3 vezes nas páginas 55, 57 e 58.
- SONKA, M.; HLAVAC, V.; BOYLE, R. *Image processing, analysis, and machine vision*. Fourth edition. Stamford, CT, USA: Cengage Learning, 2015. ISBN 978-1-133-59360-7. Citado na página 67.
- SRESTASATHIERN, P.; RAKWATIN, P. Oil palm tree detection with high resolution multi-spectral satellite imagery. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 6, n. 10, p. 9749–9774, 2014. Citado na página 16.
- SRINIVASAN, A. *Handbook of precision agriculture: principles and applications*. Boca Raton: CRC press, 2006. ISBN 9780429180309. Citado 2 vezes nas páginas 35 e 36.
- SU, J.; COOMBES, M.; LIU, C.; ZHU, Y.; SONG, X.; FANG, S.; GUO, L.; CHEN, W.-H. Machine Learning-Based Crop Drought Mapping System by UAV Remote Sensing RGB Imagery. *Unmanned Systems*, v. 08, 2019. ISSN 2301-3850. Citado na página 16.
- SUN, S.; LI, C.; PATERSON, A. H.; CHEE, P. W.; ROBERTSON, J. S. Image processing algorithms for infield single cotton boll counting and yield prediction. *Computers and electronics in agriculture*, Elsevier, v. 166, p. 104976, 2019. Citado na página 25.
- SUNDARARAJAN, D. *Digital image processing: a signal processing and algorithmic approach*. Singapore: Springer, 2017. ISBN 978-981-10-6113-4 978-981-10-6112-7. Citado 3 vezes nas páginas 64, 67 e 77.
- SUTHERLAND, W. J. *Ecological census techniques: a handbook*. 2. ed. Cambridge: Cambridge university press, 2006. ISBN 9780511790508. Citado na página 31.
- SYLVESTER, G. *E-agriculture in action: Drones for agriculture*. Bangkok: Food and Agriculture Organization of the United Nations (FAO), 2018. ISBN 978-92-5-130246-0s. Disponível em: <<http://www.fao.org/3/I8494EN/i8494en.pdf>>. Citado na página 30.
- SZELISKI, R. *Computer Vision: Algorithms and Applications*. 2011^a edição. ed. London ; New York: Springer, 2010. ISBN 978-1-84882-934-3. Citado 6 vezes nas páginas 55, 56, 58, 65, 69 e 75.

TESTA, G.; REYNERI, A.; BLANDINO, M. Maize grain yield enhancement through high plant density cultivation with different inter-row and intra-row spacings. *European Journal of Agronomy*, Elsevier, v. 72, p. 28–37, 2016. Citado na página 16.

THANKI, R. M.; KOTHARI, A. M. *Digital Image Processing using SCILAB*. 1st ed. 2019. ed. Cham: Springer International Publishing : Imprint: Springer, 2019. ISBN 978-3-319-89533-8. Citado 2 vezes nas páginas 55 e 62.

THEODORIDIS, S.; KOUTROUMBAS, K. *Pattern Recognition, Fourth Edition*. 4th. ed. USA: Academic Press, Inc., 2008. ISBN 1597492728. Citado na página 70.

TUPIN, F.; INGLADA, J.; NICOLAS, J.-M. *Remote Sensing Imagery*. John Wiley & Sons, 2014. ISBN 9781118899106. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118899106>>. Citado na página 37.

TURING, A. M. Computing machinery and intelligence. *Mind*, Oxford University Press, v. 59, n. October, p. 433–60, 1950. Citado na página 81.

TYAGI, V. *Understanding Digital Image Processing*. 1ª edição. ed. Boca Raton: CRC Press, 2018. ISBN 978-1-138-56684-2. Citado na página 53.

ULLMAN, S. The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, The Royal Society London, v. 203, n. 1153, p. 405–426, 1979. Citado na página 44.

UMBAUGH, S. E. *Digital image processing and analysis: with MATLAB and CVIPtools*. Third edition. Boca Raton: Taylor & Francis, CRC Press, 2017. ISBN 978-1-4987-6602-9. Citado na página 51.

VAPNIK, V. N.; CHERVONENKIS, A. Y. On a class of learning algorithms for pattern recognition. *Avtomat. and telemekh.*, v. 25, n. 6, p. 937–945, 1964. Citado na página 82.

VAPNIK, V. N.; CHERVONENKIS, A. Y. On a class of perceptrons. *Avtomat. and telemekh.*, v. 25, n. 6, p. 112–120, 1964. Citado na página 82.

VELHO, L.; FRERY, A. C.; GOMES, J. *Image processing for computer graphics and vision*. [S.l.]: Springer Science & Business Media, 2009. Citado na página 66.

VERAMENDI, W. N. C.; CRUVINEL, P. E. Algorithm for the countering maize plants based on uav, digital image processing and semantic modeling. In: IEEE. *2021 IEEE 15th International Conference on Semantic Computing (ICSC)*. Laguna Hills, CA, USA, 2021. p. 393–397. Citado 2 vezes nas páginas 17 e 79.

WANG, Y.; ZHU, X.; WU, B. Automatic detection of individual oil palm trees from uav images using hog features and an svm classifier. *International Journal of Remote Sensing*, Taylor & Francis, v. 40, n. 19, p. 7356–7370, 2019. Citado na página 28.

WANG, Z.; WANG, K.; YANG, F.; PAN, S.; HAN, Y. Image segmentation of overlapping leaves based on chan–vese model and sobel operator. *Information processing in agriculture*, Elsevier, v. 5, n. 1, p. 1–10, 2018. Citado na página 21.

WATT, J.; BORHANI, R.; KATSAGGELOS, A. K. *Machine learning refined: foundations, algorithms, and applications*. Second edition. Cambridge: Cambridge University Press, 2020. ISBN 978-1-108-69093-5. Citado na página 81.

WEAVER, J. E. The quadrat method in teaching ecology. *The plant world*, JSTOR, v. 21, n. 11, p. 267–283, 1918. Citado na página 16.

WEGMANN, M.; LEUTNER, B.; DECH, S. *Remote sensing and GIS for ecologists: using open source software*. UK: Pelagic Publishing Ltd, 2016. ISBN 9781784270223. Citado na página 48.

WILSON, J. N.; RITTER, G. X. *Handbook of Computer Vision Algorithms in Image Algebra*. 2nd edition. ed. Boca Raton: CRC Press, 2000. ISBN 978-0-8493-0075-2. Citado 2 vezes nas páginas 71 e 73.

WINTERHALTER, L.; MISTELE, B.; SCHMIDHALTER, U. Evaluation of active and passive sensor systems in the field to phenotype maize hybrids with high-throughput. *Field Crops Research*, Elsevier, v. 154, p. 236–245, 2013. Citado na página 16.

XIE, X.; ZHANG, Y.; LING, X.; WANG, X. A novel extended phase correlation algorithm based on Log-Gabor filtering for multimodal remote sensing image registration. *International Journal of Remote Sensing*, v. 40, n. 14, p. 5429–5453, jul. 2019. ISSN 0143-1161. Disponível em: <<https://doi.org/10.1080/01431161.2019.1579941>>. Citado na página 23.

YADAV, P. K.; THOMASSON, J. A.; ENCISO, J.; SAMANTA, S.; SHRESTHA, A. Assessment of different image enhancement and classification techniques in detection of volunteer cotton using uav remote sensing. In: SPIE. *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV*. 2019. v. 11008, p. 152–165. Disponível em: <<https://doi.org/10.1117/12.2518721>>. Citado na página 47.

ZACHARIAH, A. B. *Precision Agriculture and the Future of Farming*. Canada: Delve Publishing, 2018. ISBN 978-1-77361-283-6. Citado na página 36.

ZHANG, Q. *Precision agriculture technology for crop farming*. Boca Raton: Taylor & Francis, 2016. ISBN 10.1201/b19336. Disponível em: <<https://library.oapen.org/handle/20.500.12657/40083>>. Citado na página 35.

ZHOU, C.; LIANG, D.; YANG, X.; YANG, H.; YUE, J.; YANG, G. Wheat ears counting in field conditions based on multi-feature optimization and twsvm. *Frontiers in plant science*, Frontiers, v. 9, p. 1024, 2018. Citado na página 26.