

JOSUÉ LEAL MOURA DANTAS

**UTILIZAÇÃO DE PADRÃO DISCRIMINATIVO, LÓGICA
FUZZY E COMPONENTES CONEXOS PARA
IDENTIFICAÇÃO DE PLANTA EM TEMPO REAL**

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção
do título de Doutor em Ciências.

São Paulo
2023

JOSUÉ LEAL MOURA DANTAS

**UTILIZAÇÃO DE PADRÃO DISCRIMINATIVO, LÓGICA
FUZZY E COMPONENTES CONEXOS PARA
IDENTIFICAÇÃO DE PLANTA EM TEMPO REAL**

Versão Corrigida

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção do
título de Doutor em Ciências.

Área de concentração:

Engenharia de Computação, Sistemas de In-
formação e Aplicações Computacionais

Orientador:

Prof. Dr. Bruno de Carvalho Albertini

São Paulo
2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo 25 de outubro de 2023

Assinatura do autor: *Josué Lival Moura Dantas*

Assinatura do orientador: 

Catlogação-na-publicação

Dantas, Josué

UTILIZAÇÃO DE PADRÃO DISCRIMINATIVO, LÓGICA FUZZY E COMPONENTES CONEXOS PARA IDENTIFICAÇÃO DE PLANTA EM TEMPO REAL / J. Dantas -- versão corr. -- São Paulo, 2023.

109 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Computação e Sistemas Digitais.

1.Processamento de Imagens 2.Visão Computacional
3.Reconhecimento de padrões 4.Reconhecimento de plantas
I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Computação e Sistemas Digitais II.t.

A todos que ajudaram nessa caminhada!

AGRADECIMENTOS

Primeiramente, preciso agradecer a Deus por ter permitido estudar na USP, reconhecidamente, melhor universidade do nosso país. Pelas pessoas que ele colocou ao meu redor. Minha mãe e meu irmão, que sempre me apoiaram em minha jornada. Em especial a minha esposa por todo o suporte e carinho. Ao saudoso professor André com quem aprendi coisas que vão além do conhecimento acadêmico e também ao professor Bruno, pela paciência e auxílio, fundamentais para conclusão desta etapa. Aos colegas de laboratório que também foram de suma importância na caminhada. Sinto-me rico de tudo que o dinheiro não pode comprar.

RESUMO

Reconhecimento de padrões é uma área que tem sido utilizada sobretudo no contexto do desenvolvimento de sistemas ou dispositivos dotados de inteligência artificial. No processo de reconhecimento, diversas técnicas têm sido utilizadas, dentre elas o reconhecimento utilizando padrões discriminativos. De uma maneira genérica, esses padrões buscam analisar similaridades em uma imagem utilizando estruturas simples. Uma de suas vantagens é a possibilidade de desenvolvimento da técnica sem a necessidade de uma base de dados extensa. Este trabalho busca analisar a utilização de um padrão discriminativo para reconhecimento de uma erva daninha, *Ipomoea* sp., em uma plantação de cana-de-açúcar. Folhas de plantas têm sido utilizadas para sua identificação e, no caso de *Ipomoea* sp. que é uma planta dicotiledônea, pode-se utilizar para isso a diferença entre a venação reticulada e a paralenívea da cana-de-açúcar, que é uma monocotiledônea. Em ambiente não controlado, o reconhecimento da planta pelas folhas apresenta problemas como: sobreposição, intensidade de luz e oclusão. Diante desse cenário, esta pesquisa propõe um método de identificação precisa e rápida usando templates do tipo Haar-like, lógica fuzzy e componentes conectados para diferenciar monocotiledôneas e dicotiledôneas em um ambiente real. O Haar-like features usa templates para identificar padrões. Usamos esta técnica com o conceito de Imagem Integral para reduzir o tempo de processamento. O objetivo da Imagem Integral é reduzir a quantidade de operações aritméticas no processamento de imagens. Já a lógica fuzzy é utilizada neste trabalho para auxiliar na definição do tamanho do template a ser utilizado e, dadas as correspondências, classificar a folha. Em um ambiente não controlado, geralmente são encontradas diversas folhas de dicotiledôneas na mesma imagem. O objetivo da utilização dos componentes conexos é identificar pixels que compõem objetos em uma imagem. Nossa proposta é usar os componentes conexos, lógica fuzzy, Haar-like e Imagem Integral para encontrar a localização correta das folhas dicotiledôneas em uma imagem de um ambiente real. Testes foram realizados com o intuito de verificar o comportamento do método em relação algumas condições verificadas no ambiente não controlado. Testamos a invariância a rotação e a oclusão das folhas. O método apresentou bons resultados em relação a esses aspectos. Em relação à acurácia, o método apresenta 90,0% de acurácia no melhor caso, levando-se em consideração o total de folhas de dicotiledôneas com folhas pouco sombreadas e algumas folhas com nível de oclusão baixo ou médio.

Palavras-chave: Cana-de-açúcar, Visão Computacional, Reconhecimento de padrões, Padrões discriminativos, Reconhecimento de plantas, Detecção de Ervas Daninhas.

ABSTRACT

Pattern recognition is an area that has been widely used, especially in the context of developing systems or devices with artificial intelligence. In the recognition process, several techniques have been used, among them the recognition using discriminative patterns. Generically, these patterns seek to analyze similarities in an image using simple structures. One of their advantages is the possibility of developing the technique without the need for an extensive database. This paper seeks to analyze the use of a discriminative pattern for the recognition of a weed, *Ipomoea* sp. Plant leaves have been used for its identification, and in the case of *Ipomoea* sp. which is a dicotyledonous plant, the difference between the reticulate and the parallel venation of sugarcane, which is a monocotyledonous plant, can be used for this. In an uncontrolled environment, the recognition of the plant by the leaves presents problems such as overlapping, light intensity, and occlusion. Given this scenario, this research proposes an accurate and fast identification method using Haar-like templates, fuzzy logic, and connected components to differentiate monocots and dicots in a real environment. Haar-like features use templates to identify patterns. We use this technique with the concept of Integral Imaging to reduce processing time. The goal of the Integral Image is to reduce the number of arithmetic operations in image processing. Fuzzy logic, on the other hand, is used in this work to help define the size of the template to be used and, given the matches, to classify the sheet. In an uncontrolled environment, usually, several dicot leaves are found in the same image. The goal of using the connected components is to identify pixels that make up objects in an image. We propose to use the connected components, fuzzy logic, Haar-like, and Integral Image to find the correct location of dicot leaves in an image of a real environment. Tests were performed to verify the behavior of the method concerning some conditions in the uncontrolled environment. We tested the invariance to rotation and occlusion of the leaves. The method presented good results regarding these aspects. Regarding accuracy, the method presents 90.0% accuracy in the best case, taking into account the total number of dicot leaves with little shaded leaves and some leaves with low or medium occlusion levels.

Keywords: Sugarcane, Computer Vision, Pattern recognition, Discriminative patterns, Plant recognition, Weeds Detection.

LISTA DE FIGURAS

1	Infestação de corda de viola em plantação de cana-de-açúcar.	5
2	Etapas realizadas no processo de reconhecimento de padrões.	12
3	Identificação de caracteres de placas de automóveis.	14
4	Utilização da técnica Haar-like features para reconhecimento de faces.	21
5	Exemplo do cálculo da imagem integral.	23
6	Extração da venação de uma folha utilizando detecção pelo método de Canny e transformada de Hough.	29
7	Saída do sistema proposto por Codizar e Solano (2016).	30
8	Identificação da venação das folhas.	31
9	Robô Shrimp.	32
10	Ladybird.	33
11	AgBotII.	34
12	Exemplo de como é calculado o valor do LBP. Neste caso, o valor de raio utilizado é 2 e o $P = 8$. Cada pixel é comparado com o central, começando no P_0 e um número binário é gerado e depois convertido em um valor inteiro.	34
13	Robô que identifica ervas daninhas.	36
14	Robôs desenvolvidos pela Deepfield Robotics para reconhecimento e remoção de ervas daninhas.	37
15	a) Trator com o equipamento utilizado para reconhecimento das ervas daninhas acoplado. b) Visão interna da estrutura projetada para captura de imagens.	38
16	Robô utilizado para remoção de ervas daninhas em plantações de arroz.	39

17	a) Vista do artefato utilizado para identificação da erva e utilização do spray. b) Vista lateral do sistema desenvolvido.	40
18	Gráficos das principais funções de pertinência utilizadas na lógica Fuzzy.	42
19	Representação dos elementos envolvidos em um Sistema baseado em lógica Fuzzy.	43
20	Representação do conceito utilizado pelos componentes conexos. . . .	46
21	Esquema de representação de um pixel e sua vizinhança.	47
22	a) Imagem de entrada. b) Após o processo de dilatação usando lógica Fuzzy c) Imagem Segmentada. d) Imagem de saída com veículos encontrados.	49
23	Diagrama de estados do Autômato Finito Determinístico.	50
24	Transições de estados a partir do estado A.	50
25	<i>Templates</i> propostos para reconhecimentos do padrão de venação. . .	53
26	Etapas propostas no processo de identificação das ervas daninhas. . .	55
27	Processo de entrada.	56
28	Síntese dos processos utilizados na etapa de pré-processamento. . . .	57
29	Resumo das etapas de pré-processamento. Primeiro a imagem de entrada, depois a mudança para o HSI e por fim o resultado da aplicação do algoritmo de Canny.	57
30	Relação entre o tempo de processamento e o tamanho da imagem utilizando a transformada de Hough.	58
31	Síntese da etapa de extração de características.	60
32	Relação entre o tamanho da imagem utilizada e a acurácia.	61
33	Quantidade de correspondências encontradas de cada template em ângulos diferentes nas folhas monocotiledôneas.	62
34	Quantidade de correspondências encontradas de cada template em ângulos diferentes nas folhas dicotiledôneas.	63

35	Erva Daninha e cana-de-açúcar em ambiente real.	65
36	Gráfico que relaciona a área em pixels do quadrado e o percentual da imagem que não será analisada.	66
37	Imagem dividida em quadrados mostrando o percentual de verde em cada um deles.	67
38	Gráfico da função de pertinência referente ao percentual de pixels verdes existentes em uma determinada área da folha.	68
39	Gráfico da função de pertinência referente ao tamanho do template a ser utilizado.	69
40	Gráfico da função de pertinência referente ao tipo da folha a qual pertence uma determinada área.	70
41	Folha segmentada artificialmente para simular a oclusão.	71
42	Técnicas utilizadas na etapa de processamento.	71
43	Relação entre o percentual da folha e a quantidade de intersecções obtidas.	72
44	a) Imagem de entrada b) Imagem de saída final do algoritmo proposto.	74
45	a) Imagem de entrada com maioria da área das folhas sombreadas. b) Imagem de saída após a aplicação da técnica.	77
46	Relação entre a taxa de acerto e o sombreamento das folhas.	78
47	Representação do Sistema RGB em coordenadas cartesianas.	90
48	Modelo de cores HSI.	91
49	Comparação de técnicas de interpolação da imagem. (a) Imagem Original (b). Vizinho mais próximo (c) Bi-linear (d) Bi-cúbica.	94
50	Aplicação do método proposto por Canny.	95
51	Identificação de pontos colineares.	97

LISTA DE TABELAS

1	Comparativo entre insumos utilizados na plantação de cana-de-açúcar orgânica e inorgânica	9
2	Acurácia relacionada com o tamanho do template.	60
3	Comparação entre técnicas para reconhecimento de folhas, em ambiente controlado.	63
4	Comparação entre técnicas para reconhecimento de folhas em ambiente real.	73

LISTA DE ABREVIACOES

BBI	Bayes Belief Integration
CUDA	Compute Unified Device Architecture
PCA	Factor Analysis
FRTCA	Fuzzy Real Time Classification Algorithm
GPU	Graphics Processing Unit
GLCM	Gray Level Co-occurrence Matrix
HOG	Histograms of Oriented Gradients
HSI	Hue Saturation Intensity
HSV	Hue, Saturation and Value
LDA	Linear Discriminant Analysis
LBP	Local Binary Pattern
MCS	Multiple Classifiers System
NIR	Near Infra-Red
NVDI	Normalised Difference Vegetation Index
OCR	Optical Character Recognition
k-FLBPCM	Filtered LBP method with Countor Mask and coefficient k
PCA	Principal Componente Analysis
RGB	Red Green Blue
RGNIR	Red Green NIR
SVM	Support Vector Machine

SUMÁRIO

1	Introdução	1
1.1	Objetivo	4
1.2	Justificativa	4
1.3	Problema	7
1.4	Hipóteses	10
1.5	Estrutura da Tese	11
2	Reconhecimento de Padrões	12
2.1	Correspondência	13
2.2	Probabilidade	14
2.3	Aprendizado de máquina	15
2.4	Técnicas para reconhecimento de objeto em tempo real	16
2.4.1	Modelos de subespaço linear	17
2.4.2	Filtros Locais baseados na intensidade dos pixels	18
2.5	Processamento dos Descritores	19
2.6	Haar Wavelets	20
2.6.1	Haar-like <i>Features</i>	20
2.6.1.1	Imagem Integral	21
3	Reconhecimento de Plantas	24
3.1	Reconhecimento de ervas daninhas	25
3.2	Venação das folhas	28
3.3	Artefatos aplicados no reconhecimento de padrões	31

4	Lógica Fuzzy	41
5	Componentes Conexos	46
6	Metodologia	51
6.1	Método proposto para extração da venação das folhas	51
7	Resultados	56
7.1	Entrada	56
7.2	Pré-processamento	57
7.3	Utilização da Transformada de Hough	58
7.4	Utilizando o Haar-like <i>features</i>	59
7.5	Tamanho dos templates	59
7.6	Tamanho da Imagem	60
7.7	Taxa de acerto com as imagens rotacionadas	60
7.8	Desempenho da metodologia	62
7.9	Dataset de imagens em ambiente não controlado	64
7.10	Varredura da imagem	64
8	Discussões	76
9	Conclusão	80
9.1	Contribuições do trabalho	81
9.2	Trabalhos futuros	81
	Referências	82
	Apêndice A - Considerações sobre o processamento de imagens deste trabalho	89

A.1	Pré-processamento	89
A.1.1	Tamanho da Imagem - redução, transformação em escala de cinza	89
A.2	Sistema RGB	89
A.3	Sistema HSI	90
A.3.1	A.1.3.1 Conversão de RGB para HSI	91
A.4	Interpolação	92
A.5	Segmentação	93
A.5.1	Detector de Bordas Canny	93
A.5.2	Transformada de Hough	96

1 INTRODUÇÃO

Atualmente, os robôs são utilizados em diversas atividades em diferentes áreas. Há, por exemplo, robôs colaboradores que auxiliam na troca de pneus (LEVRATTI et al., 2019) e outros que auxiliam em atividades de alta precisão como cirurgias (HU et al., 2019). Além disso, robôs estão sendo projetados para ajudar nas atividades domésticas, como o Turtlebot (WILSON et al., 2019), que busca auxiliar idosos em seu cotidiano dentro de seus lares. Assim, a tendência é que os robôs estejam presentes nas mais diversas áreas, inclusive nas residências, nos próximos anos.

Na agricultura, a tendência também é a aplicação de novas tecnologias, como descreve, por exemplo, Polymeni et al. (2023). Os autores citam o campo de pesquisa intitulado Agricultura 5.0 que utilizará as técnicas como sensores quânticos, robótica, redes sem fio avançadas e Inteligência Artificial com os conceitos existentes relacionados a Agricultura de Precisão. Essas tecnologias serão utilizadas para a implantação de fazendas conectadas, produção de novos equipamentos, tratores e máquinas conectadas. Busca-se, com a aplicação dessas tecnologias, o aumento da produtividade, qualidade e proteção do meio ambiente.

No contexto de inovações tecnológicas utilizadas na Agricultura, Lytridis et al. (2021) ressaltam o uso da robótica no contexto do que se chama de Agricultura 5.0. Ressaltam que a introdução dos robôs tem os seguintes objetivos: melhorar a eficiência e a produtividade, combater a escassez de mão de obra sazonal e realizar tarefas difíceis ou perigosas.

A importância de aplicação de novas tecnologias na produção de alimentos é justificada, dentre outros fatores, pela demanda crescente na produção de alimentos. Roshanianfard (2018) descreve a preocupação que se têm relacionada a quantidade de alimentos para os próximos anos, pois estima-se que em 2100 a população mundial será de 11,8 bilhões de habitantes. O autor ainda destaca que a utilização de robôs na agricultura será crucial para aumentar a produtividade e atender à necessidade mundial de alimentos.

Na agricultura atual os robôs já têm sido utilizados em diversas tarefas. Um exem-

plo disso é o Shrimp (HUNG et al., 2016), robô utilizado para estimar a produção de frutas de um determinado pomar. A utilização de robôs visa a diminuição de custos, aumento da precisão, e a realização de práticas mais sustentáveis, como o manejo da cultura sem a utilização de produtos químicos. Solomon, Sharma e Pathak (2019) corroboram essa ideia e citam a mecanização como um meio de tornar a produção de cana-de-açúcar mais sustentável.

Dentre as várias possibilidades de atuação de um robô em um campo agrícola, ressalta-se o combate a ervas daninhas, que é uma atividade comum nas plantações.

O Brasil é um importante produtor de diversas culturas. Dentre elas destaca-se a cana-de-açúcar, onde o Brasil é o responsável por 25% da produção mundial (RIGHETTO et al., 2019). Como não poderia ser diferente, a cana-de-açúcar também disputa nutrientes com certas ervas daninhas. Desta forma, é importante o estudo de técnicas que possibilitem a construção de artefatos que possibilitem a modernização da produção da cana-de-açúcar e outras culturas nesse cenário de Agricultura 4.0. Righetto et al. (2019) destacam que as ervas daninhas provocam distúrbios na área de cultivo, o que pode resultar em diminuição da qualidade da cana-de-açúcar cultivada, já que as ervas competem por águas e nutrientes com a mesma.

Dentre as diversas espécies que podem causar perdas no cultivo da cana, destacamos a corda de viola, *Ipomoea* sp. Bhullar et al. (2012) descrevem a *Ipomoea* sp. como dicotiledôneas perenes ou anuais de verão, que sobe e envolve o caule da cana. Por ser uma erva daninha trepadeira, sua remoção é mais difícil. Righetto et al. (2019) apresenta um dado interessante: em uma área predominantemente infestada pela planta em questão, há uma redução de cerca de 46% na produtividade. Os mesmos autores afirmam que, em média, 8,4% do custo total da produção é referente ao combate de ervas daninhas.

Monquero et al. (2011) descrevem a mudança do sistema de colheita da cana-de-açúcar que favoreceu a predominância de algumas espécies de ervas daninhas, incluindo a *Ipomoea* sp. No Estado de São Paulo, há uma previsão legal que proíbe a queimada em canaviais, seguindo um esquema de restrições progressivas que proibiu a prática totalmente em 2021 em áreas com possibilidade de mecanização total

da colheita. Para as demais áreas, o prazo é 2031¹. Os autores (MONQUERO et al., 2011) ressaltam que, sem o uso da queima, as folhas, bainhas, ponteiros, além de quantidade variável de pedaços de colmo, são cortados, triturados e lançados sobre a superfície do solo, formando o que é chamado de “cobertura de palha”. Essa cobertura, ocasiona diversas mudanças no solo, sejam elas físicas, químicas e biológicas, além de provocar uma seleção na comunidade biológica infestante.

Geralmente, o controle de ervas daninhas é realizado por defensivos agrícolas. Righetto et al. (2019) ressalta que a aplicação desses produtos é ineficiente, pois, segundo os autores, é difícil diferenciar em uma imagem de uma grande área se um pixel representa uma cana-de-açúcar ou uma erva daninha, para que a aplicação seja feita de forma mais adequada. Além disso, há danos ao meio ambiente devido a utilização de produtos químicos. Possebon e Polli (2020) ressaltam que a utilização de adubos solúveis acarreta a contaminação do solo e dos lençóis freáticos. Outro aspecto importante a ser ressaltado é que a aplicação desses produtos acabam também por fazer a seleção de espécies resistentes, e seu posterior proliferamento (ELSTONE et al., 2020).

Kshirsagar (2005) conceitua o cultivo orgânico como um sistema de gerenciamento de produção agrícola holístico que sustenta e melhora a saúde do ecossistema agrícola, abrangendo a biodiversidade, os ciclos biológicos dos nutrientes e as atividades bioquímicas e microbianas do solo. Assim, esta técnica não utiliza produtos químicos sintetizados, considerados perigosos, e promove o uso de recursos orgânicos renováveis, promovendo uma agricultura sustentável.

O Brasil, conforme evidenciado por Petrônio et al. (2022), é também o maior produtor de cana-de-açúcar orgânico do mundo, produzindo em torno de 181 mil toneladas por ano.

Há fazendas que se propõem a realizar o cultivo da cana-de-açúcar de forma orgânica, ou seja, sem utilizar produtos químicos que possam de alguma forma afetar o meio ambiente ou o consumidor final. Roberto e Junior (2015) realizaram um estudo sobre a diversidade biológica existente nesse tipo de cultivo. Eles identificaram 340 espécies de vertebrados silvestres, sendo que destas, 49 espécies são conside-

¹Decreto nº 47.700 de 11/3/2003, que regulamenta a Lei nº 11.241 de 19/9/2002

radas sob algum risco ou ameaça de extinção no estado de São Paulo. Os autores atribuem essa diversidade de espécies à colheita ser realizada sem queimada, a não utilização de agroquímicos, entre outros fatores relacionados ao cultivo e o manejo da cana-de-açúcar orgânica.

Nas plantações de cana-de-açúcar orgânica, o controle de ervas daninhas como a *Ipomoea* sp. é ainda mais custoso, pois ainda é feita por trabalhadores de forma manual, muitas vezes utilizando somente as mãos por conta da proximidade da planta com a cana-de-açúcar, conforme observado em visita *in loco*.

A proposta deste trabalho é investigar um método, que possa ser utilizado em um robô, para identificação de uma erva daninha comum às plantações de cana-de-açúcar. Nesse sentido, o trabalho busca aplicar o estado da arte para que se desenvolva um método de visão computacional que possa ser aplicado a um artefato que auxilie a aplicação dos conceitos de Agricultura 4.0 nas plantações de cana.

1.1 Objetivo

O objetivo da pesquisa é investigar um método que possibilite a identificação de erva daninha no cultivo da cana-de-açúcar utilizando um algoritmo de processamento de imagens de tempo real, baseado na modificação da transformada de Haar, e que se adapte às variações dos parâmetros nas imagens em cenários reais não controlados, possibilitando uma futura utilização da técnica em um artefato que remova essas ervas.

1.2 Justificativa

Andres et al. (2017) ressalta que a quantidade de herbicidas utilizados no processo de controle de ervas daninhas tem aumentado no Brasil. A dificuldade em encontrar mão de obra humana para realizar o controle manual e a destruição causada pelo controle mecânico, justificam o aumento na utilização de produtos químicos. Uma das consequências da utilização de herbicidas, destacadas por Andres et al. (2017), é a proliferação de ervas daninhas mais resistentes.

Há uma busca por técnicas que combatam ervas daninhas, mas que causem menos impactos ambientais (TODERO et al., 2018). Assim como em outras áreas,

uma maneira automatizada de remover essas plantas melhoraria a rentabilidade, bem como a qualidade da plantação, além de proporcionar uma maneira mais sustentável de controle dessa praga.

Dentre as diversas ervas daninhas existentes, nesse trabalho busca-se um método para combater especificamente ervas daninhas trepadeiras que atacam, entre outras culturas, a cana-de-açúcar. Fuzita (2009) destaca que a mudança no processo de colheita da cana, passando para o mecanizado sem queima, passou a deixar mais palha na área de plantio, o que provavelmente favorece o aumento na densidade de espécies de corda de viola na cultura. Fuzita (2009) aponta ainda uma dificuldade em combater essa praga com herbicidas, devido a necessidade de transportar a palha para eliminar a planta invasora.

Figura 1 - Infestação de corda de viola em plantação de cana-de-açúcar.



Fonte: Elaboração própria.

Fuzita (2009) ressalta que as espécies de corda de viola, segundo suas pesquisas, estão entre as ervas daninhas mais comuns nas plantações de cana-de-açúcar. Seu trabalho ainda caracteriza a corda de viola como uma planta trepadeira, em 74% das espécies. Ela compete por água, luz, nutrientes e espaço, além de causar danos a cana-de-açúcar no momento da colheita, dificultando o corte mecanizado pois

compromete o desempenho das máquinas e a qualidade do produto colhido. Na Figura 1 pode-se observar a ocorrência da corda de viola em uma plantação de cana-de-açúcar.

Nesse sentido, este trabalho investiga um método para identificação da corda de viola e suas espécies, utilizando conceitos de visão computacional. A técnica possibilitará o desenvolvimento de um robô que realize a remoção da erva e, conseqüentemente, resultará em ganhos para o produtor, além de promover a redução do impacto do uso dos herbicidas.

Técnicas de processamento de imagens e visão computacional têm sido utilizadas no processo de remoção de ervas daninhas, bem como em outras aplicações nos campos agrícolas (INÁCIO; RIEDER, 2018). Nesse sentido, as folhas das espécies têm sido estudadas como um dos meios para sua identificação (AMLEKAR et al., 2015).

Várias pesquisas buscam uma técnica eficiente para determinar uma espécie através da análise de sua folha. É recorrente o uso de uma técnica de aprendizado de máquina para identificação de plantas, as redes neurais (ZHU; ZHU; REN, 2018). Essas técnicas necessitam de um hardware robusto, com placa de vídeo, para o processamento mais rápido das informações, como apresentaram Milioto e Mar (2018) em um projeto de identificação de plantas e ervas daninhas em tempo real.

Utstumo et al. (2018) afirmam que as técnicas até então desenvolvidas para a classificação de folhas não apresentam robustez satisfatória quando lidam com sobreposição e reflexão. Há a necessidade de investigar qual técnica permite o reconhecimento da espécie em ambientes reais, tratando inclusive a oclusão das folhas e a reflexão da luz.

Wang, Zhang e Wei (2019) detalham várias dificuldades a serem transportadas no processo de identificação de plantas em um ambiente real. A sobreposição das folhas é uma delas, o que dificulta o desempenho de técnicas que utilizam características morfológicas. Além disso, as folhas podem sofrer total oclusão ou estarem danificadas por alguma praga ou ainda amareladas ou mortas alterando também as propriedades espectrais.

A variação da luz em ambientes abertos é uma dificuldade a ser levada em consi-

deração. Essa variação produz diferença nas cores, níveis de ruído, sombra, reflexão, contraste, saturação e brilho (WANG; ZHANG; WEI, 2019).

Outro aspecto importante nessa pesquisa é o tempo que a técnica leva em reconhecer a erva daninha, dado que o cultivo extensivo pode facilmente ocupar uma área de alguns quilômetros quadrados. Assim, esta pesquisa buscou analisar também, entre as possíveis soluções, a que apresenta menor tempo computacional.

Em relação ao processamento em tempo real, Sebastian e Ben-Tzvi (2019) apresenta a utilização de *Support Vector Machine* (SVM) para identificação de tipos diferentes de terrenos para facilitar o deslocamento de robôs autônomos. A velocidade utilizada no deslocamento dos robôs foi de 0.2-0.5 metro por segundo. Foi utilizado uma taxa de 100Hz para aquisição de dados da pose do robô, para auxiliar no seu deslocamento, e do contexto para estimar o tipo de solo. Dada essa frequência, temos para tempo real, a possibilidade de análise de cada imagem em cerca de 0,01 segundo. O trabalho é utilizado como referência por utilizar um robô que se assemelha ao que o algoritmo proposto seria embarcado, sendo que ele é adaptado para se locomover a tipos diferentes de solos, dentre eles o que se assemelha ao de uma plantação de cana-de-açúcar.

1.3 Problema

Dentre as propriedades que cultivam cana-de-açúcar há as que optam por não utilizar produtos químicos para o combate de ervas daninhas ou outras pragas. Em uma visita a fazenda São Francisco que fica localizada Sertãozinho, São Paulo, cuja latitude é -21.17442095978422 e a longitude é -48.11822006292568, que adota o cultivo orgânico da cana, pudemos entender melhor a necessidade do combate automatizado de ervas daninhas. Um dos agrônomos responsáveis relatou que para certas pragas, como insetos, já se desenvolveu uma forma biológica suficientemente eficaz para combatê-los. Outras ervas daninhas que nascem entre as fileiras também já são controladas através da utilização de uma máquina que percorre entre as linhas de plantio da cana, chamada de cultivador, e as remove. Porém, próximo à cana-de-açúcar, em local que o cultivador não alcança, desenvolvem-se ervas daninhas, como a corda de viola, e trepadeiras que impactam a produção da cana-de-açúcar de diversas formas.

Além de disputar nutrientes com a planta, o agrônomo também relatou que na fase de colheita a erva daninha enrosca na colheitadora e o seu acúmulo faz a máquina parar de funcionar, sendo necessário desligar a máquina e realizar a remoção desses cipós. Assim, o profissional relatou que a praga de mais difícil controle, hoje, são as ervas daninhas trepadeiras. Seu combate, nessa fazenda, tem sido feito através de funcionários que utilizam as mãos para extraí-las.

Outro aspecto importante a ressaltar é que a fazenda possui uma área vasta que não possui cobertura de sinal de dados móveis que permitiria testar uma solução utilizando computação em nuvem. Além disso, como deverá ser utilizado um artefato para fazer a remoção da erva, que provavelmente será um robô com uma estrutura considerável, não seria um grande problema adicional adicionar um hardware para detecção da erva daninha.

Atualmente, o controle da erva daninha dar-se, basicamente, de maneira manual, utilizando mão de obra humana que realiza a maioria da remoção das plantas utilizando as próprias mãos. Possebon e Polli (2020) endossam que esta forma de remoção das ervas daninhas é um dos gargalos para produção de cana-de-açúcar orgânica. Kshirsagar (2005) apresenta uma tabela comparativa que apresenta uma estimativa do trabalho humano em plantações de cana-de-açúcar orgânica e não orgânicas. Os dados mostram um acréscimo, esperado, na quantidade de trabalho humano realizado nas plantações orgânicas, mostrados na Tabela 1.

Partindo do pressuposto acima, investigamos um método para identificação dessas ervas daninhas utilizando visão computacional, de forma a identificá-las em tempo real com baixo custo computacional. No entanto, o reconhecimento de plantas, sobretudo em ambientes reais, apresenta diversas dificuldades.

Wang, Zhang e Wei (2019) citam os principais problemas relacionados ao processamento de imagens para identificação de ervas daninhas capturadas por um artefato que se locomove pela terra. O primeiro é a sobreposição das folhas das plantas com as das ervas daninhas, tornando as duas difíceis de identificar. Outro seria a oclusão ou a danificação das folhas, que também pode conter barro ou folhas mortas, o que modifica a morfologia da folha e ainda suas propriedades espectrais. Além desses, é enumerado a variação das condições de iluminação, a sombra da copa das plantas

Tabela 1 - Comparativo entre insumos utilizados na plantação de cana-de-açúcar orgânica e inorgânica

Sr No.	Input	Organic Sugarcane (OS)	Inorganic Sugarcane (IS)	% increase over Inorganic
1.	Human Labor (days)	247.80	206.15	20.20
2.	Bullock Labour (pair days)	9.72	8.51	14.22
3.	Tractor (hours)	6.42	5.96	7.72
4.	Seed (ton)	2.97	3.35	-11.34
5.	Organic Manures (ton)	11.40	6.36	79.25
6.	Bio-fertilizers (kg)	178.70	-	-
7.	Chemical Ferlitizers (kg)			
	Nitrogen (N)	-	31.37	-
	Phosphate (P)	-	110.25	-
	Potash (K)	-	77.42	-
8.	Insecticide / Pesticide (kg)	2.03	2.50	-18.80
9.	Number of Irrigations	21.45	26.51	-19.09

Fonte: Kshirsagar (2005)

ou o ângulo do sol, que influencia diretamente a cor da vegetação. Iqbal et al. (2018)) corrobora a dificuldade que a variação da luz causa no processamento de imagens de folhas, que na sua aplicação é para a identificação de doenças. Outro problema, ressaltado pelos autores, no processamento de imagens com folhas é a diferença entre os estágios de crescimento da planta que também representam uma alteração nas propriedades morfológicas, de textura e espectrais das folhas

Bakhshipour et al. (2017) utilizam *wavelet* para a identificação de ervas daninhas em uma plantação de beterrabas, porém ressalvam que a técnica não funciona bem se houver muitas folhas oclusas e sobrepostas.

Segundo Wang, Zhang e Wei (2019), no cenário no qual ocorre oclusão e sobreposição, a utilização de características morfológicas não possui bom desempenho, já que os formatos das folhas são imprevisíveis. Uma estratégia que tem sido adotada para diminuir o problema é dividir a imagem em sub imagens de tamanho $n \times n$, porém os autores ressaltam que em ambiente com muita sobreposição os resultados ainda são insatisfatórios. Outra abordagem que tem sido aplicada nesse sentido é a definição de pontos chaves na imagem, nos quais determinadas características da folha

são analisadas. Na pesquisa, no entanto, é ressaltado que essa técnica possui um custo computacional elevado.

Amend et al. (2019) apresenta ainda algumas dificuldades a serem superadas na tarefa de reconhecimento de ervas daninhas. Dentre elas, está a relação entre taxa de acerto e velocidade e a diversidade de ervas daninhas a ser reconhecida.

Neste trabalho, busca-se encontrar um método que apresente acurácia aceitável mesmo quando aplicada em ambiente externo real não controlado e com as dificuldades elencadas acima, e que também apresente baixo tempo de resposta de forma a viabilizar a identificação das plantas de forma automatizada em tempo real. No próximo capítulo, é apresentada a revisão da literatura dividida em tópicos que constituem a base na qual nossa pesquisa está fundamentada.

1.4 Hipóteses

A proposta de pesquisa é baseada nas seguintes hipóteses:

1. A cana-de-açúcar é monocotiledônea e a corda de viola é dicotiledônea. Há diferença na venação entre elas, o que pode ser utilizado para diferenciá-las.
2. O reconhecimento de padrões pode ser utilizado para diferenciar a venação das folhas da erva daninha e da cana.
3. O reconhecimento utilizando aprendizado de máquina através, geralmente, faz uso de uma grande base de imagens o que aumenta o tempo para desenvolvimento da solução e o custo computacional.
4. O reconhecimento de padrões utilizando estrutura morfológica pode apresentar um bom resultado em tempo real.
5. O Haar-like *features* pode ser utilizado, através da definição de *templates* que identifiquem os padrões contidos na folha como método morfológico.
6. O desempenho das técnicas é diferente em ambiente real e um controlado. É necessário investigar os ajustes necessários para que o método possa ser aplicado em ambientes reais.

7. Há a necessidade de estabelecer um método adaptativo de ajuste das configurações de processamento de imagem de acordo com as variações dos parâmetros.

1.5 Estrutura da Tese

Este documento está estruturado em 8 capítulos:

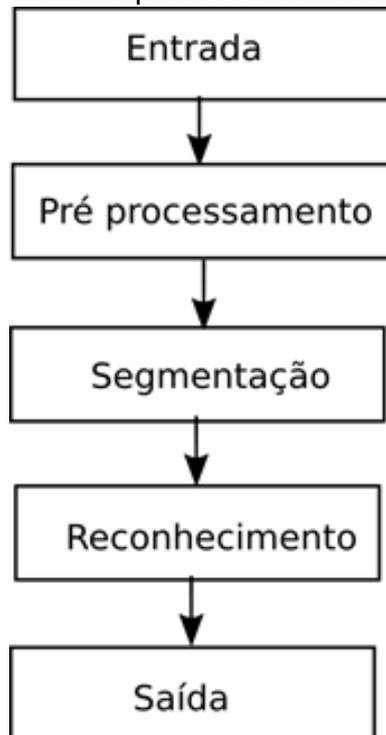
- Capítulo 2 - apresenta os conceitos e revisão da literatura referente ao conceito de reconhecimento de padrões.
- Capítulo 3 - discorre sobre o reconhecimento de plantas, analisando as diversas metodologias utilizadas para este fim, e apresenta robôs que utilizam visão computacional para realizar a remoção de ervas daninhas ou contagem de frutos utilizando.
- Capítulo 4 - descreve o conceito de lógica Fuzzy bem como algumas aplicações.
- Capítulo 5 - apresenta a metodologia utilizada neste trabalho.
- Capítulo 6 - apresenta os resultados, descrevendo os testes realizados a partir da metodologia proposta.
- Capítulo 7 - onde são apresentadas as discussões relacionadas aos testes e a técnica proposta.
- Capítulo 8 - dispõe as conclusões tendo em vista os resultados obtidos em face à revisão da bibliografia realizada e os outros trabalhos que se propõem a realizar tarefas similares.

2 RECONHECIMENTO DE PADRÕES

Gonzalez e Woods (2008) definem “padrão” como uma combinação de descritores. Um descritor pode ser entendido como uma característica. Uma classe padrão é definida como um conjunto de padrões que compartilha algumas propriedades comuns. O processo de reconhecimento de padrões envolve técnicas para atribuir padrões às suas respectivas classes.

As etapas no processo de reconhecimento de padrões são descritas por (SVYRYDOV; KUCHUK; TSIAPA, 2018) conforme a Figura 2. Na primeira etapa, temos o processo de aquisição da imagem. Depois a imagem é alterada de forma a ressaltar as regiões de interesse. Na segmentação é aplicada uma técnica que separa os pixels da(s) região(ões) de interesse das demais. A próxima etapa consiste na aplicação de uma técnica que consiga analisar as regiões segmentadas e classificá-las em uma determinada classe. Por fim, há a saída que mostra o resultado final do processo, indicando as áreas de interesse identificadas.

Figura 2 - Etapas realizadas no processo de reconhecimento de padrões.



Fonte: Svyrydov, Kuchuk e Tsiapa (2018).

Tendo em vista as diferentes áreas em que o reconhecimento de padrões pode ser aplicado, Svyrydov, Kuchuk e Tsiapa (2018) enfatizam que não existe um método universal que possa ser aplicado na resolução de todos os problemas de forma satisfatória. Por isso, existem diferentes categorias de técnicas utilizadas nessa tarefa.

Gonzalez e Woods (2008) dividem as técnicas de reconhecimento de padrões em dois grupos: reconhecimento baseado em métodos de decisão teóricos e métodos estruturais.

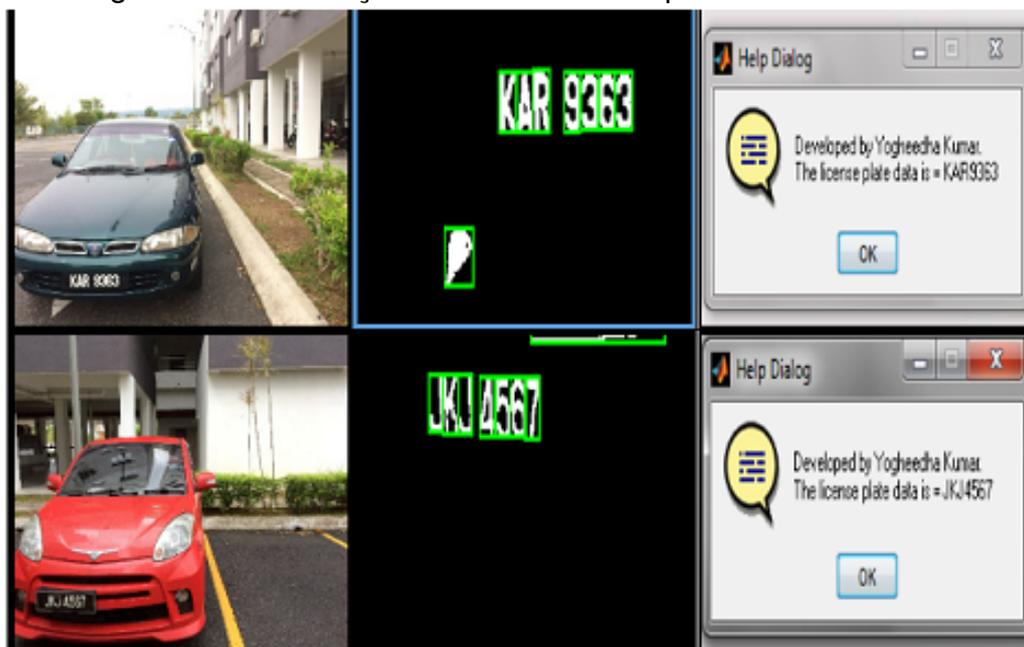
O primeiro realiza o reconhecimento baseado no uso de funções de decisão (ou discriminantes). Essas funções analisam cada padrão e retorna um valor que indica se a característica pertence a uma determinada classe ou não. Já o segundo, busca encontrar um relacionamento na estrutura inerente a forma do padrão a ser encontrado. Dentro do grupo de métodos de decisão teóricos, Gonzalez e Woods (2008) ainda os divide em métodos por correspondência, baseados em probabilidade e aprendizado de máquina.

2.1 Correspondência

As técnicas baseadas em correspondência representam cada classe por um protótipo de vetor de padrões. Um padrão desconhecido é atribuído à uma determinada classe se ele possui um valor próximo a uma métrica preestabelecida. Um exemplo de métrica simples a ser utilizada é a distância euclidiana entre os vetores da classe e os preestabelecidos.

Outro tipo de estratégia utilizada na correspondência é o uso da correlação de um determinado *template* com a imagem. Yogheedha (2018) usa a correspondência de padrões na criação de um sistema automático para a identificação de placas de veículos. Nesse processo foi utilizado o OCR (*Optical Character Recognition*) que analisa uma determinada janela na imagem e compara o caractere obtido com uma base utilizada no processo de treino. A Figura 3 mostra a identificação dos caracteres em placas dos automóveis.

Figura 3 - Identificação de caracteres de placas de automóveis.



Fonte: Yogheedha (2018)

2.2 Probabilidade

Nesse tipo de técnica é calculada a probabilidade de um determinado padrão pertencer à uma classe. Essa técnica se aplica, sobretudo, para áreas nas quais os padrões que cada classe gera são aleatórios (GONZALEZ; WOODS, 2008).

Dentre os métodos probabilísticos utilizados para o reconhecimento de padrões podemos citar o *Principal Componente Analysis* (PCA). O PCA utiliza o conjunto de treinamento $A = (A_1, A_2, \dots, A_n)$, o conjunto de treino tem N imagens, $j = 1, 2, \dots, N$, pertencentes a c classes, $i = 1, 2, \dots, c$, e o número de pixels na imagem é n . A matriz de dispersão entre classes é expressa da seguinte forma:

$$S_b = \sum_c^{i=1} P(\omega_i)(\mu_i - \mu_0)(\mu_i - \mu_0)^T \quad (2.1)$$

onde S_b é a matriz de dispersão entre as classes, $P(\omega_i)$ é a probabilidade a priori de ω_i , em geral, $P(\omega_i) = (1/c)$, μ_i é o vetor de médias da classe ω_i , $\mu_0 = (1/N) \sum_{j=1}^N A_j$ é o vetor de médias de todas as amostras. A matriz de dispersão interna de uma classe

é dada por:

$$S_w = \sum_{i=1}^C P(\omega_i) S_i \quad (2.2)$$

onde S_w é a matriz de dispersão interna de uma classe $S_i = E(A - \mu_i)(A - \mu_i)^T | A \in \omega_i$ é a matriz de covariância de ω_i . A matriz total de dispersão é definida como:

$$S_t = S_b + S_w = 1/N \sum_{j=1}^N (A_j - \mu_0)(A_j - \mu_0)^T \quad (2.3)$$

Uma das principais aplicações do PCA é no reconhecimento de faces. Para esta aplicação, o PCA utiliza principalmente a matriz de dispersão total ou a matriz de dispersão interna das classes para extrair as características de uma face. Se a matriz de dispersão total é usada para gerar a matriz, a matriz de projeção ótima é igual a calcular os máximos autovalores e os correspondentes auto vetores de S_t , e a matriz de projeção ótima (X_1, X_2, \dots, X_d) são os auto vetores associados com os maiores autovalores generalizados. Portanto, (X_1, X_2, \dots, X_d) descrevem a contribuição de cada autovalor na descrição da imagem de entrada da face, então é possível extrair as características faciais através dele (ZHOU; SUN; SHI, 2006).

2.3 Aprendizado de máquina

Em certos casos é difícil estimar uma função que receba um padrão e retorne sua correspondência com uma determinada classe. Assim como as propriedades estatísticas de uma classe podem ser desconhecidas ou não podem ser estimadas (GONZALEZ; WOODS, 2008). As técnicas de aprendizado de máquina utilizam um processo de treinamento ou aprendizado para obter as funções de decisão. Atualmente, existem várias técnicas e aplicações de reconhecimento de imagens que utilizam aprendizado de máquina.

Jin, Liang e Filter (2017) utilizam uma rede neural profunda para realizar o reconhecimento de imagens, em um caso de peixes. Segundo Jin, Liang e Filter (2017), para treinar uma rede neural profunda é necessária uma grande base de dados, porém adquirir imagens em baixo da água para formar uma grande base teria um custo elevado. Desta forma, foram utilizadas imagens de peixes da ImageNet, que possui mais de 15 milhões de imagens rotuladas em mais de 22.000 categorias, para treinar

uma rede chamada ConvNet. Depois, um ajuste fino nos parâmetros foi realizado com imagens de peixes em seu habitat natural, um processo conhecido como transferência de conhecimento.

Jin, Liang e Filter (2017) ressaltam que a ConvNet funciona como uma extratora de parâmetros. Desta forma, temos a aplicação da definição dessa categoria de técnicas para reconhecimento de padrões em imagens, na qual a própria técnica de aprendizado de máquina molda-se em uma função de maneira a encontrar o padrão na imagem alvo.

2.4 Técnicas para reconhecimento de objeto em tempo real

Gao et al. (2014) apresentam uma classificação para técnicas de detecção de objetos, que trabalham em tempo real, e necessitam de baixo recurso de hardware em sua execução. As técnicas são divididas em modelos generativos e modelos discriminativos.

Modelos generativos são construídos como modelos para a forma dos objetos alvo. Abordagens recentes tendem a incluir a textura no modelo da forma. O sucesso da abordagem generativa frequentemente está em quão bem o modelo se aproxima da aparência atual do objeto. Gao et al. (2014) ressaltam que este tipo de modelo serve bem para objetos que possuem formas bem definidas como humanos, mas para outros objetos como rochas, que não possuem forma bem definida, é mais difícil aplicar esse tipo de técnica.

Já os modelos discriminativos utilizam uma função discriminante, onde que não é necessário representar a aparência dos objetos alvo. Segundo Gao et al. (2014), esse tipo de abordagem envolve menos recursos computacionais do que a descrita anteriormente e também é mais fácil de implementar.

Os modelos discriminativos se baseiam em dois importantes aspectos chave. O primeiro é encontrar características que ajudem a definir a classe do objeto e que sejam rápidas de extrair. O segundo é o uso de classificadores dos vetores com as características (GAO et al., 2014). A seguir são descritas as técnicas discriminativas mais utilizadas para extração de características.

2.4.1 Modelos de subespaço linear

Os modelos de subespaços lineares assumem que todas as amostras positivas residem de forma linear, com uma dimensionalidade muito menor do que o espaço original. Os principais métodos que representam esse subespaço são: *Principal Component Analysis* (PCA), *Factor Analysis* (FA) e *Linear Discriminant Analysis* (LDA).

Principal Component Analysis - é um método que busca uma projeção linear que projete pontos de dados em um espaço de k dimensões, tal que a variação total dos pontos seja maximizada. O PCA é feito pela decomposição da matriz de covariância dos pontos de dados usando a decomposição de valores singulares ou deposição de auto vetores (GAO et al., 2014). Há várias aplicações do PCA, sobretudo, para reconhecimento facial.

Factor Analysis (FA) - é uma técnica semelhante ao PCA. A diferença entre elas é que o FA assume que as amostras de dados observadas são provenientes de um modelo bem definido como o representado na seguinte equação:

$$x = Df + u + \mu \quad (2.4)$$

onde o D é a matriz de constantes, f e u são vetores aleatórios, e μ é a média. O método busca D e a matriz de covariância de u que represente a estrutura de covariância de x . Caso u seja assumida como 0, o procedimento é equivalente ao PCA (GAO et al., 2014).

Linear Discriminant Analysis (LDA) - é um método que projeta pontos vindos de M classes em um espaço de dimensão $M - 1$, onde aparecem mais separados que em seus espaços originais. O LDA trabalha com a busca por uma matriz de projeção W tal que o seguinte quociente seja maximizado:

$$J(W) = WTS_bWWTS_wW \quad (2.5)$$

onde S_b é a matriz de dispersão entre classes que nos informa quão variado as M classes são em um espaço de alta dimensão, e S_w é uma matriz de dispersão interna que nos informa quão variados são os pontos dentro de sua própria classe. Uma característica do LDA é que ele maximiza o quociente $J(W)$ procurando os $M - 1$ mais significantes autovalores da matriz $S_w^{-1} S_b$ usando a deposição generalizado dos auto-

valores (GAO et al., 2014).

Gao et al. (2014) afirma que os modelos de subespaço linear requerem a extração da característica em tempo linear relacionado ao tamanho da janela a ser explorada. Para aplicações em tempo real, que podem analisar milhões de janelas em uma imagem, isso pode ser um gargalo. Gao et al. (2014) cita os filtros locais baseados na intensidade dos pixels e o Haar-like *features* como técnicas mais rápidas para extração de características, pois são modelos discriminativos.

Varatharajan, Manogaran e Priyan (2018) utilizam o LDA a fim de reduzir as características relevantes no processo de classificação de imagens de eletrocardiograma. Os autores utilizaram PCA e LDA para extração de características e algumas técnicas de aprendizado de máquina para classificação. O menor erro quadrático foi obtido utilizando o LDA com o SVM.

2.4.2 Filtros Locais baseados na intensidade dos pixels

Filtros locais que operam sobre intensidade de pixels são frequentemente usados como conjunto de características. Nessas abordagens, uma característica é simplesmente o sinal da diferença de intensidade entre dois pixels, sem necessidade de normalização pela média ou desvio padrão entre as amostras.

Haar-like *features*: é listada como uma das técnicas que utilizam modelos discriminativos por Gao et al. (2014). Como é a técnica aplicada neste trabalho, uma seção foi dedicada para melhor explicá-la.

Local receptive fields: é utilizada para selecionar extratores de características que se adaptam a tarefa, ocorrendo no processo de treinamento do modelo (GAO et al., 2014).

Coodebook feature patches: extrai características ao redor de pontos de interesse em uma imagem. Por exemplo, as características que representam um pedestre poderiam ser extraídas por essa técnica (GAO et al., 2014) Zhu, Zhu e Ren (2018) utilizaram o conceito desta técnica para selecionar as características mais discriminativas utilizando a minimização do erro de reconstrução para cada ponto central. A técnica foi aplicada em três datasets: Flower-17, Caltech-256 e PASCAL VOC 2007. O resultado

apresentado mostra uma taxa de acerto maior do que outras técnicas.

Histograms of Oriented Gradients (HOG): parte do pressuposto de que a aparência e a forma de um objeto podem ser caracterizadas pela distribuição local da intensidade dos gradientes ou pela direção das bordas. Nesta técnica, uma janela da imagem em uma região espacial pequena (células) acumula um histograma local de direção do gradiente ou orientação da borda para cada pixel na célula. A representação é formada através da combinação das entradas dos histogramas. A técnica também realiza o acúmulo da medida do histograma de energia local de pequenos blocos e utiliza os resultados para normalizar as células do bloco. Esse procedimento é útil para melhorar a invariância a iluminação (CHENG et al., 2019). Cheng et al. (2019) utilizam o HOG como descritor de imagens que contêm pés de laranja. O HOG é utilizado juntamente com o SVM para a classificação de troncos de pés de laranja. A pesquisa alcançou uma taxa de acerto de 98.96% utilizando as imagens da câmera e também um sensor ultrassônico. A localização dos troncos na plantação pode ser utilizada no processo de locomoção de um robô, sendo esta a justificativa da pesquisa.

Covariance Matrix Descriptor: Gao et al. (2014) descrevem a utilização da matriz de covariância como um descritor que apresenta resultados robustos em imagens com variação de iluminação (Gao et al apud Tuzel et al. (2015)). Essa técnica correlaciona a variação espacial local com as características do gradiente.

2.5 Processamento dos Descritores

Esta seção descreve as principais técnicas utilizadas no processo de aprendizado de máquina para reconhecimento de padrões. Essas técnicas, listadas abaixo, utilizam os dados que foram obtidos através da utilização de um processo anterior que extraiu características importantes da imagem.

Há uma divisão em duas categorias das técnicas de aprendizado de máquina, sendo elas: técnicas supervisionadas e as não supervisionadas. Nas técnicas supervisionadas há um conjunto de dados rotulados no qual os valores de entrada e saída são conhecidos. Já nas técnicas não supervisionadas não existe um conjunto de treino cujos valores de saída são conhecidos, apenas os de entrada.

A seguir teremos a descrição de algumas técnicas de aprendizado de máquina, bem como pesquisas que mostram como elas podem ser aplicadas.

2.6 Haar Wavelets

Papageorgiou, Oren e Poggio (1998) definem a Haar *wavelets* como um conjunto natural de funções bases que diferenciam regiões com diferentes médias de intensidade.

Oren et al. (1997) utilizam a Haar *wavelet* para o reconhecimento de pedestres. São utilizados três tipos de Haar *wavelets* de duas dimensões. Essas funções base capturam mudança de intensidade na direção horizontal, vertical e na diagonal.

Viola e Jones (2001), baseado no trabalho de Papageorgiou, Oren e Poggio (1998), utilizou o Haar-like *features* para detecção de faces. A próxima seção é dedicada a explicar melhor o Haar-like *features*.

2.6.1 Haar-like Features

Viola e Jones (2001) afirma que a análise de características da imagem é bem mais rápida do que a dos pixels. Além disso, quando a quantidade de dados para o processo de treino é reduzida, a obtenção de características singulares dos objetos pode ser uma alternativa para identificá-los. A velocidade de processamento de uma característica é mais rápida do que a de pixels.

A ideia do Haar-like é percorrer a imagem com templates, geralmente retângulos, que possuem duas ou mais regiões pretas e brancas. Os valores dos pixels na(s) região(s) branca(s) são somados e subtraídos da(s) região(s) preta(s), como na Equação 6. A diferença entre as regiões é utilizada para encontrar arestas, barras e outras estruturas simples na imagem (VIOLA; JONES, 2001).

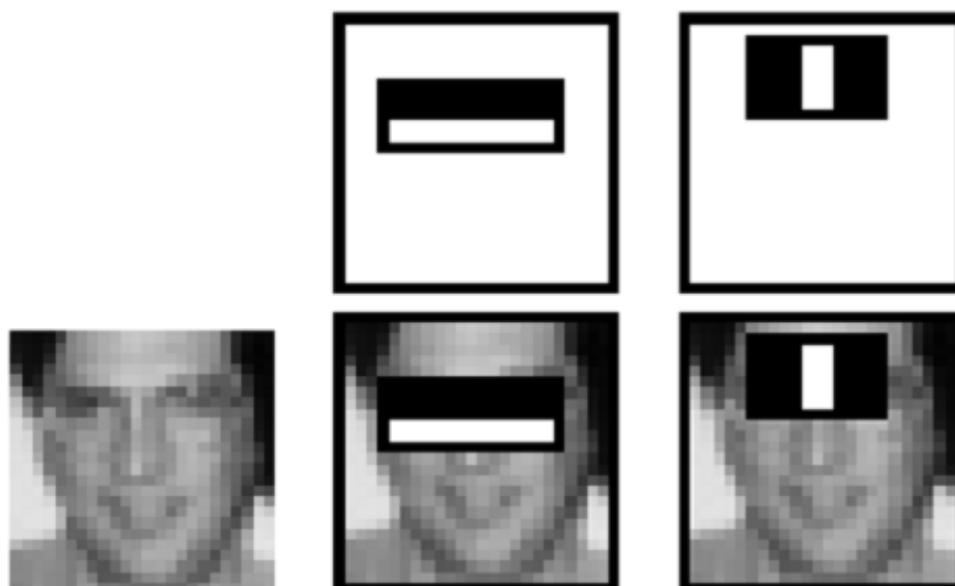
$$total = \frac{\sum \text{valores dos pixels brancos}}{nb} - \frac{\sum \text{valores dos pixels preto}}{np} \quad (2.6)$$

onde nb é a quantidade de pixels brancos e np a quantidade de pixels pretos.

O Haar-like tem sido utilizado em várias aplicações. Uma delas é o conhecimento

de faces, como mostrado na Figura 4, onde é possível entender o funcionamento da técnica. São definidos *templates* que são usados para varrer a imagem em busca de padrões. Por exemplo, na região dos olhos e sobrancelhas há uma região mais escura, que contrasta com a região mais branca das bochechas. Da mesma forma na região do nariz, que possui dos seus lados uma região de pixels mais pretos, correspondente a a região dos olhos. No artigo original em que se propõe a técnica (VIOLA; JONES, 2001), os autores a aplicaram justamente no reconhecimento de faces em conjunto com o Adaboost, técnica de aprendizado de máquina. Li et al. (2011) utiliza essa metodologia para identificação de pedestres.

Figura 4 - Utilização da técnica Haar-like features para reconhecimento de faces.



Fonte: Viola (2001)

Viola e Jones (2001) tem em vista uma diminuição no tempo de processamento das imagens, utilizam ainda o conceito de Imagem Integral para calcular os retângulos divididos em regiões.

2.6.1.1 Imagem Integral

A imagem integral é uma representação intermediária da imagem usada para um processamento mais rápido. Os pixels acima e a esquerda de um determinado ponto (x, y) são somados, juntamente com o valor armazenado na posição, seguindo a for-

mula:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (2.7)$$

onde $ii(x, y)$ é o ponto na imagem integral e o ponto $i(x', y')$ é o ponto da imagem original. O cálculo de um determinado ponto pode ser particionado como a soma da imagem integral do ponto acima, conforme a primeira equação abaixo, mais a soma dos elementos desde o ponto até os demais elementos a esquerda deste, na segunda equação abaixo.

$$s(x, y) = s(x, y - 1) + i(x, y) \quad (2.8)$$

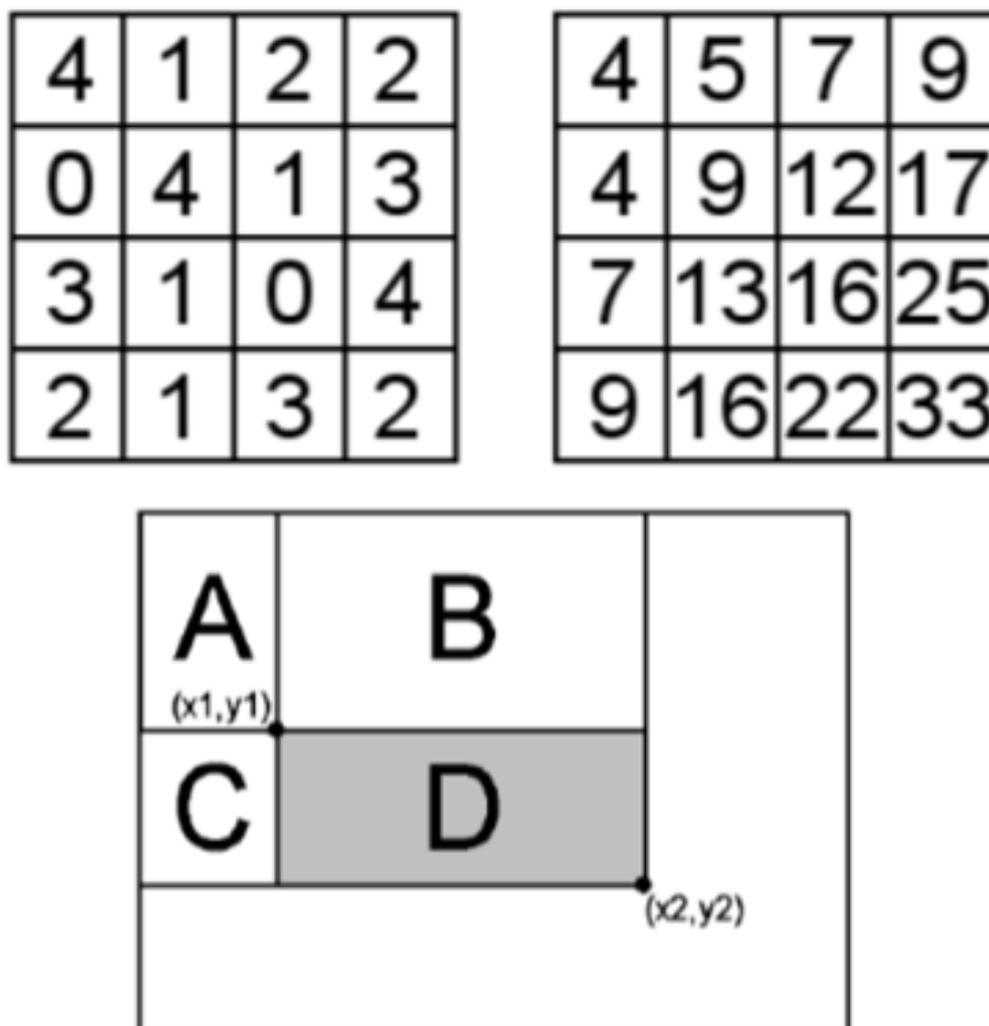
$$ii(x, y) = ii(x - 1, y) + s(x, y) \quad (2.9)$$

A Figura 5 mostra como o cálculo da imagem integral é realizado. Na imagem, à esquerda temos o que seriam os valores originais dos pixels da imagem. Na imagem ao centro temos o valor calculado da imagem integral e na imagem a direita um esquema ilustrativo de como o cálculo é feito. Por exemplo, para calcular o valor da região D , deve-se somar $C + B + D$ e diminuir pelo valor de A , assim a fórmula ficaria $C + B + D - A$. Vamos exemplificar calculando o valor na posição de linha 2 e coluna 2. O cálculo é feito somando os valores já calculados da célula acima (1,2) que é 5, da célula (2,1) que é 4, o valor da célula (1,1) que é 4 e próprio valor da célula (2,2) na tabela de entrada que é 4. A fórmula utilizada é a seguinte $(2,2) = (2,1) + (1,2) + (2,2) - (1,1)$, assim o resultado é 9, como visto na tabela ao centro.

A vantagem da utilização da imagem integral está na redução de operações a serem realizadas no cálculo. Por exemplo, para calcular o valor da célula (4,4) com a imagem integral são utilizados 4 acessos a matriz, portanto caso todas as células fossem somadas para achar o valor da tabela 16 acessos seriam feitos.

O Haar-like *features* utiliza *templates* que percorrem toda a imagem, realizado inúmeras somas para identificar uma região na qual há um determinado padrão. Suponhamos que um *template* de 16x4 fosse utilizado para percorrer uma imagem. Dividindo o *template* ao meio, uma para a região preta e outra para região branca, teríamos duas regiões de tamanho 8x2. Para calcular o valor em cada região somando todos os pixels, seriam utilizados 16 acessos a matriz e, utilizando o conceito de ima-

Figura 5 - Exemplo do cálculo da imagem integral.



Fonte: Peuwuan, Wararatpanaya e Pasupa(2016) .

gem integral, apenas 4 acessos seriam feitos. Dependendo do tamanho da imagem, essa técnica promove um ganho considerável no tempo de processamento.

No próximo Capítulo, apresenta como tem sido realizado o reconhecimento de plantas, bem como alguns robôs que aplicam o reconhecimento de padrões em sua visão computacional para diversas atividades na área da agricultura.

3 RECONHECIMENTO DE PLANTAS

A identificação de plantas é uma tarefa importante para várias áreas, tais como a medicina e a agricultura. Entretanto, este não é um processo simples, pois se for realizado manualmente requer um especialista e tempo (CODIZAR; SOLANO, 2016). Em virtude disso, várias pesquisas têm sido desenvolvidas para encontrar uma forma automatizada de identificação das plantas.

A análise dos padrões existente nas folhas das plantas é o parâmetro mais utilizado para a determinação de sua espécie. Zheng e Wang (2010) enumeram uma série de características das folhas que podem ser utilizadas para gerar parâmetros de modo a melhor realizar a classificação. Essas características podem ser divididas em três grupos: forma, textura e cor.

Wäldchen e Mäder (2018) apresentam uma divisão das características que podem ser utilizadas para reconhecimento de uma planta, utilizando a folha, bem parecida com a relatada por Wang, Zhang e Wei (2019). Ambas apresentam a divisão geral em três categorias: forma, cor e textura. Porém, Wäldchen e Mäder (2018) acrescentam outra divisão que é composta por estruturas específicas da folha: as veias e as bordas.

Wäldchen e Mäder (2018) realizam uma revisão da literatura de vários trabalhos e relatam que alguns pesquisadores utilizaram somente a venação como característica para diferenciar certas espécies. Por outro lado, Zheng e Wang (2010) propõem a classificação de diversas espécies de plantas utilizando diversas características das folhas para melhorar o processo de classificação. Lembramos que pesquisa desta tese precisa diferenciar somente duas espécies. Desta forma, procura-se uma característica ou a quantidade mínima possível dentre elas necessária para realizar a classificação das espécies com boa acurácia e no menor tempo possível. Como uma das diferenças marcante entre monocotiledôneas e dicotiledôneas é o formato das veias na folha (uma é a paralela e a outra reticulada, respectivamente) decidimos inicialmente utilizar essa característica e investigamos como extrair a venação de uma folha.

As próximas sub seções foram divididas de forma a apresentar as técnicas utilizadas para realizar o reconhecimento de plantas utilizando diversas características

e outra com os trabalhos que utilizam somente a venação da folha para realizar a classificação.

3.1 Reconhecimento de ervas daninhas

Ashraf e Khan (2020) propõem a identificação de ervas daninhas em plantações de arroz. Para isto, comparam duas técnicas, uma utilizando as características de textura e outra o formato e anatomia das plantas. Os autores ressaltam que essas características foram escolhidas por apresentarem invariância a rotação e a escala. Outro aspecto a destacar é que os autores também analisam o tempo de execução e o potencial de aplicação da técnica em tempo real. Para definir a textura da imagem, é utilizado a matriz de co-ocorrência de níveis de cinza (*Gray Level Co-occurrence Matrix* (GLCM)). Essas matrizes são utilizadas com um classificador SVM e *Random Forest*.

Para representar as características morfológicas das plantas um conjunto de momentos. Ashraf e Khan (2020) definem momentos como a média ponderada das intensidades de pixel. Os momentos utilizados são a média, variância, enviesamento e curtose. A partir dessas quatro características são calculados outros seis momentos, considerados complexos, que não são descritos explicitamente na pesquisa. Como classificadores usado para essas características, os autores utilizaram também o SVM e *Random Forest*. No comparativo, o melhor resultado obtido pelos autores foi utilizando o conjunto de características representado pelos momentos com o *Random Forest*. O valor obtido foi de 86% de taxa de acerto. Em relação ao tempo de processamento, o tempo médio para extrair a GLCM é de 11,231 segundos. Já extrair o conjunto de características representadas pelos momentos demora, em média, 30,965 segundos.

Kamath, Balachandra e Prabhu (2020) apresentam um técnica de visão computacional para identificação de ervas daninhas em plantações de arroz. Para isto, utilizam um sistema com múltiplos classificadores (MCS - *Multiple Classifiers System*). Utilizam 71 características de entrada, que foram selecionadas a partir de 93 por um método baseado na análise de variância. A taxa de acerto encontrada foi de 91,36%. Os autores não apresentam dados sobre o tempo de processamento da técnica proposta.

Por serem muitas características a serem processadas, provavelmente o tempo de processamento é elevado.

A abordagem utilizada por Lottes et al. (2018) é utilizada para o reconhecimento de ervas daninhas na plantação de beterrabas. Os autores descrevem uma sequência de passos usados a fim de fazer o reconhecimento. Primeiramente, a imagem é obtida e vegetação é separada das outras partes. Depois, é extraída uma série de características das regiões correspondentes a vegetação e depois é utilizado o classificador *Random Forest*. A pesquisa leva em conta duas variantes no problema de classificação. A primeira calcula as características locais de pontos chaves e classifica a área ao redor como ponto chave. A segunda variante é a classificação baseada em segmento ou objeto, cuja decisão é feita baseada nos pixels de um segmento pertencente a vegetação. Após a classificação usando o *Random Forest*, a informação sobre a vizinhança de cada ponto chave predito é obtida. Nesta tarefa, Lottes et al. (2018) utilizam campos aleatórios de Markov, o que resulta em melhoria na rotulação da imagem.

Lottes et al. (2018) ainda utilizam os pontos chaves de forma a contornar o problema de sobreposição de plantas que ocorre no ambiente real não controlado. Porém, o tempo de processamento das características utilizando pontos chaves ao invés de segmentos é maior. Na pesquisa, em cada ponto chave ou segmento são extraídas 18 características estatísticas como média, desvio padrão, entropia, entre outras, e relativas a forma como convexidade, retangularidade, solidez, entre outras. Um aspecto importante da pesquisa de Lottes et al. (2018) é que há uma análise do tempo de execução da técnica proposta. Como há várias características obtidas para classificar cada segmento ou ponto chave, é de se esperar um tempo elevado para classificação das imagens. No entanto, os pesquisadores utilizaram o processamento por placa de vídeo (GPU GeForce GTX 980) e o *framework* CUDA. No trabalho, ressalta-se que a GPU aumentou de 20-100 vezes velocidade da extração das características. O tempo médio de extração para todas as características é em torno de 0,4 ms para um ponto chave e de 9,75 ms para um objeto. Por fim, os autores apresentam um tempo médio de 0,2 segundo para a classificação usando objetos ou segmentos de vegetação e 0,8 segundos utilizando pontos chave. Testes foram realizados com três conjunto de dados diferentes e a maior taxa de acerto obtida foi de 92%.

Dyrmann, Christiansen e Midtiby (2017) utilizam uma técnica para classificação de plantas na qual as folhas são extraídas da imagem. São obtidas características da planta, como um todo, segmentadas as folhas e depois a classificação é realizada. A classificação das características é feita através do *Bayes Belief Integration* (BBI). O primeiro passo da estratégia proposta por Dyrmann, Christiansen e Midtiby (2017) é realizar a separação da planta do solo, etapa denominada de segmentação. Para isto, é realizado um cálculo que utiliza um conceito chamado de verde excessivo e vermelho excessivo. As plantas são definidas a partir de valores acima de um determinado limiar, após a subtração do vermelho excessivo do verde excessivo. Após esse passo, é realizada a separação da folha utilizando um algoritmo que identifica um ponto no qual começa o caule. Para a classificação da planta são extraídas 19 características da imagem e para classificar as folhas são extraídas 14. A maioria delas estão relacionadas a morfologia do objeto, tais como área, esfericidade, variância elíptica, entre outras. A taxa de acerto encontrada foi de 96,7%. Os autores também descrevem o tempo de processamento médio, por cada planta, sendo de 143 ms, desde a segmentação até a classificação. Ressaltam também, que esse valor é rápido suficiente para operações em tempo real. O conjunto de dados utilizado para os testes foi construído pelos próprios autores, que retiraram as imagens utilizando um *photo box* branco, o que ajudou o flash da câmera espalhar a luz diminuindo os efeitos de um possível sombreamento. No total, o conjunto de dados possui 1061 imagens de 8 espécies diferentes.

Ahderom e Apopei (2020) enfatizam as perdas ocasionadas pelas ervas daninhas e propõem uma técnica para reconhecimento destas. Na etapa de extração de características da imagem é utilizado o LBP (*Local Binary Pattern*), que será melhor explicado na seção 3.3. O LBP é uma das técnicas utilizadas para capturar da textura de uma imagem. No entanto, os autores propõem a utilização de um novo método chamado de k-FLBPCM (*Filtered LBP method with Countor Mask and coefficient k*). O método proposto é utilizado para realizar a classificação de duas plantas, a canola (cultura) e o rabanete selvagem (erva daninha). Como classificador é utilizado o SVM. A fim de testar o método proposto, Ahderom e Apopei (2020) construíram um conjunto de dados com imagens com 1920x1080 pixels. Ao todo foram utilizadas 19600 imagens coletadas em quatro diferentes estágios de desenvolvimento. A taxa de acerto

máxima obtida foi de 98,63%.

Há diversos trabalhos que utilizam aprendizado profundo com Redes Neurais Convolucionais para reconhecimento de plantas e ervas daninhas. Por exemplo, os trabalhos de Olsen et al. (2019), Wu (2019), Lottes et al. (2018), Wang et al. (2018) e Umamaheswari, Arjun e Meganathan (2018). Para utilizar esse tipo de abordagem, há a necessidade de geração de uma grande base de dados para treino e teste. Além disso, o hardware para reconhecimento em tempo real deve ser robusto. Chavan e Nandedkar (2018) treinaram uma rede neural convolucional para classificação das plantas cultivadas e das ervas daninhas. Para o treinamento e teste da solução foi utilizado um computador com um processador Intel Xeon E5-2695 v4 (64 GB RAM) e uma placa de vídeo NVIDIA Titan Xp (12 GB RAM). O hardware especificado é mais utilizado em servidores e possui um custo relativamente alto. O tempo aproximado para a classificação é cerca de 0,97 segundos.

3.2 Venação das folhas

Apesar do ser humano identificar as veias de uma folha com certa facilidade, por possuir um padrão de cor muito próximo ao do restante da folha, não consiste em uma tarefa computacionalmente simples. Em virtude disso, tão somente a aplicação de uma técnica de identificação de bordas não extrai com precisão a veia existente na folha.

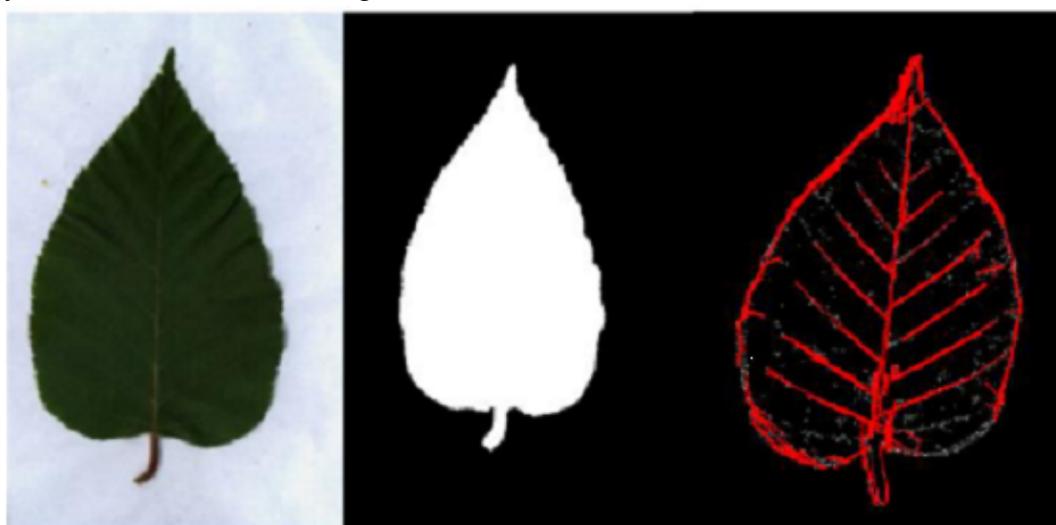
Para facilitar a identificação das veias, depois da aquisição da imagem, há autores que utilizam outro espaço de cores como o *Hue Saturation Intensity* (HSI) ao invés do *Red Green Blue* (RGB), pois o segundo possui limitações no processamento de certas imagens. O HSI separa os componentes pela intensidade, além de atenuar a influência causada pela iluminação (CAO et al., 2017).

Cao et al. (2017), após converterem a imagem para o domínio HSI, utilizam a técnica de detecção de Canny, que utiliza a diferença entre gradientes das imagens, para determinar as bordas. Porém, é necessário inserir dois valores de limiar para o algoritmo. Cao et al. (2017) utilizaram o algoritmo de Otsu para determinar esses limiares. O algoritmo de Otsu divide a imagem em duas classes utilizando o conceito de variância máxima entre *clusters*. Por fim, Cao et al. (2017) utilizam a transformada

de Hough para definir as linhas que formam a folha. Nesse processo, primeiramente é identificada a nervura principal e depois as secundárias.

A Figura 6 mostra a aplicação da técnica descrita por Cao et al. (2017). Primeiramente há uma imagem de uma folha, que é utilizada como entrada pelo algoritmo. Depois a imagem é binarizada com um valor de limiar que ignora as veias, mas possibilita a definição da venação central. Depois a imagem final mostra em vermelho as linhas que foram definidas utilizando a transformada de Hough.

Figura 6 - Extração da venação de uma folha utilizando detecção pelo método de Canny e transformada de Hough.

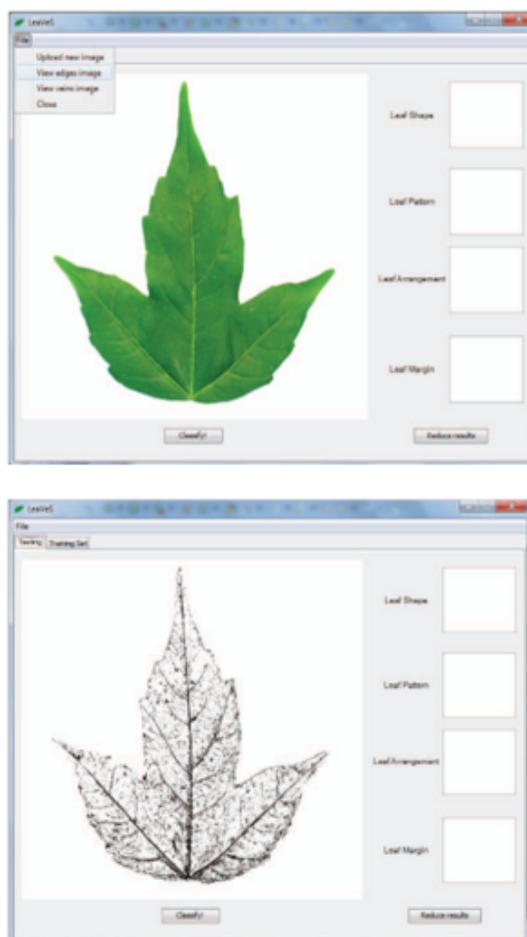


Fonte: Cao et al. (2017).

Codizar e Solano (2016) desenvolveram um sistema chamado LeaVeS para auxiliar na classificação de plantas utilizando o formato e a venação da mesma. Para a retirada da venação os autores convertem a imagem para escala de cinza e aplicam um operador morfológico matemático de abertura para ressaltar a parte de interesse. A saída do sistema pode ser observada na Figura 7. Na primeira tela do sistema, nota-se a imagem de entrada e na outra tela a imagem resultante que mostra a venação da folha.

Zheng e Wang (2010) também utilizam operações de morfologia matemática para obtenção das veias da folha. Antes, os autores transformam a imagem para o espaço HSV, realizam as operações e utilizam o algoritmo de Otsu para binarizar a imagem. Além disso, remove-se pixels pretos que possuem poucos pixels pretos vizinhos e curvas descontínuas são conectadas. Por fim, pequenos blocos de pixels pretos se-

Figura 7 - Saída do sistema proposto por Codizar e Solano (2016).

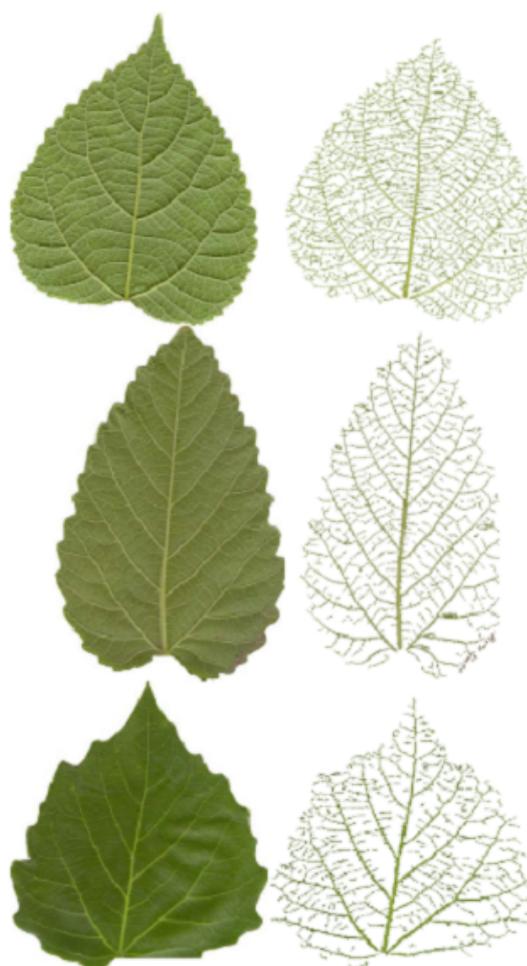


Fonte: Codizar e Solano (2016)

parados são removidos. O resultado da técnica aplicada por Zheng e Wang (2010) é mostrado na Figura 8, que mostra três tipos de folhas diferentes: com diferença em sua coloração, relacionando a imagem de entrada e a venação resultante da técnica proposta pelos autores.

Nas próximas seções são descritos conceitos essenciais para melhorar a compreensão dos procedimentos metodológicos necessários no processo de diferenciação entre as folhas da cana-de-açúcar e da erva daninha. Como o objetivo é desenvolver um método que seja utilizado na parte de visão computacional de um robô, começaremos descrevendo os conceitos relacionados a este tema, porém apresentamos também o uso da visão computacional em outros artefatos, como tratores, que também realizam o reconhecimento das plantas.

Figura 8 - Identificação da venação das folhas.



Fonte: Zheng e Wang (2010)

3.3 Artefatos aplicados no reconhecimento de padrões

Os dispositivos utilizados na visão computacional de determinados artefatos para a navegação também são utilizados para mapear ervas daninhas, contar frutos e outras atividades. Hung et al. (2016) utilizam um robô para teste de um algoritmo de detecção de frutas em geral. Para isso, eles utilizaram um robô de uso geral chamado Shrimp, visto na Figura 9, que possui uma câmera em seu topo, capturando imagens com resolução de 1232x1616 pixels. O processo de segmentação das imagens é feito utilizando uma rede neural convolucional que divide uma imagem original em 32 sub imagens de 308x202 pixels. A detecção da fruta é feita utilizando a transformada de Hough para círculos.

Hung et al. (2016) enfatizam que a maioria dos outros algoritmos são elaborados

Figura 9 - Robô Shrimp.

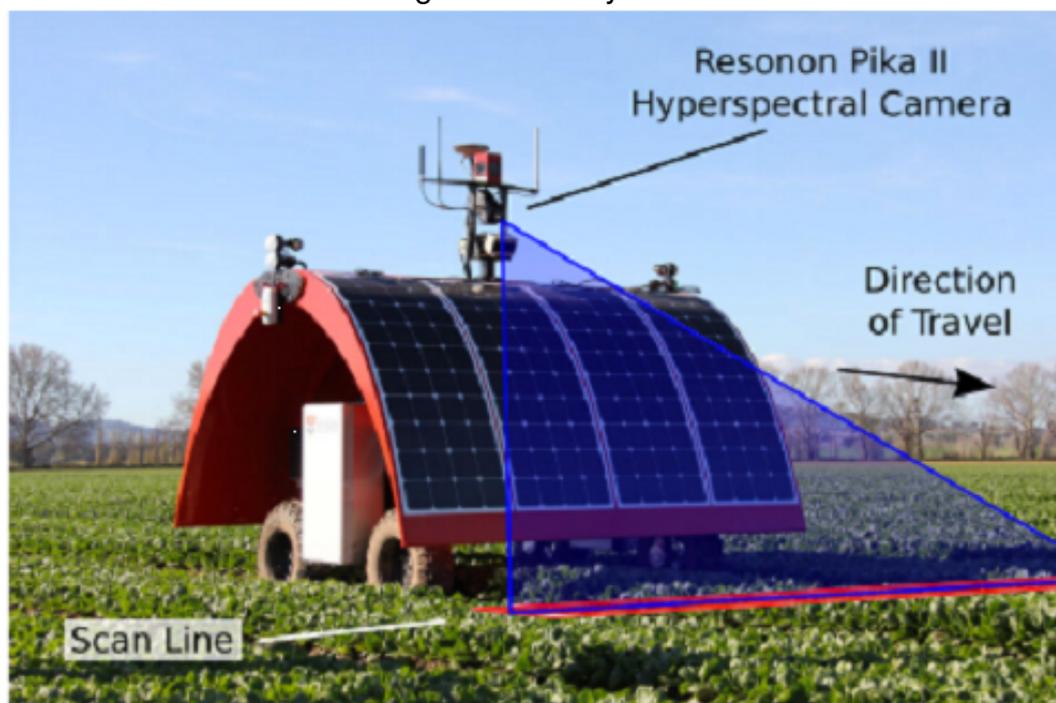


Fonte: Hung et al. (2016)

para a identificação de uma fruta específica, enquanto o testado por eles foi projetado para identificar e contar qualquer espécie de fruta. No entanto, a taxa de acerto do método proposto por eles é diferente para cada espécie de fruta. Assim, a taxa de acerto encontrada em plantações de maçãs foi cerca de 95,7%, amêndoas 73,9%, mangas 90,8% e lichia 75,2%. A pesquisa conclui que o desempenho do algoritmo depende do tamanho da fruta.

Outro robô autônomo que utiliza visão computacional é descrito por Wendel e Underwood (2016), o Ladybird, visto na Figura 10. Como pode ser visto na imagem, o robô é equipado com uma câmera hiperespectral para detectar ervas daninhas em uma plantação de algodão. Como a classificação de plantas utilizando características morfológicas pode ser afetada pela oclusão, deformações provocadas por pragas ou apresentar uma forma não identificável por estar dobrada, a pesquisa optou por utilizar imagens hiperespectrais, que se mostram invariante a esses efeitos na maioria dos casos.

Figura 10 - Ladybird.



Fonte: Wendel e Underwood (2016)

Wendel e Underwood (2016) resumem a pesquisa nos seguintes passos: primeiramente, os pixels da vegetação são separados dos pixels do fundo utilizando o índice de Diferença de Vegetação Normalizado (*Normalised Difference Vegetation Index - NVDI*); depois a média e o mínimo valor do espectro individual são normalizado para 1 e 0, e depois a extração dessa característica é feita utilizando o Análise do Componente Principal (*Principal Component Analysis - PCA*). Por fim, a classificação é feita pela Análise de Discriminante Linear (*Linear Discriminant Analysis - LDA*). A acurácia obtida ficou entre 85 e 91%.

Os robôs descritos acima não apresentam um mecanismo para remoção das ervas daninhas identificadas. Já o AgBotII, visto na Figura 11, além de identificar as plantas invasoras, remove-as (BAWDEN et al., 2017). No processo de identificação das plantas são utilizadas duas abordagens: o Padrão Binário Local (*Local Binary Pattern - LBP*) e a covariância de características.

O LBP é utilizado na referida pesquisa por ser, segundo Bawden et al. (2017), computacionalmente eficiente e robusto em relação a variações de iluminação, pois é calculado utilizando os valores de um conjunto de pixels. Nesta técnica, um pixel central é escolhido e comparado a outros pertencentes a região adjacente em um

Figura 11 - AgBotII.



Fonte: Bawden et al. (2017)

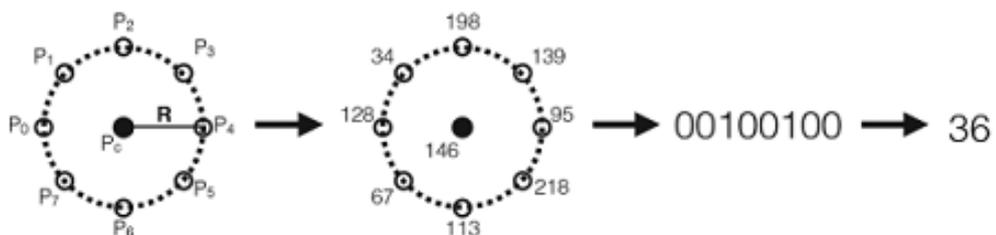
determinado raio escolhido previamente, conforme mostrado na Figura 12. O código LBP é calculado com a seguinte equação:

$$LBP_{(P,R)} = \sum_{p=0}^{P-1} h(P_c, P_p)2^p \tag{3.1}$$

onde P_c é o pixel central, P_p é o valor do p ésimo pixel e

$$h(x, y) = \{1 \text{ Se } y \geq x, 0 \text{ Se } y < x \tag{3.2}$$

Figura 12 - Exemplo de como é calculado o valor do LBP. Neste caso, o valor de raio utilizado é 2 e o $P = 8$. Cada pixel é comparado com o central, começando no P_0 e um número binário é gerado e depois convertido em um valor inteiro.



Fonte: Bawden et al. (2017)

A covariância de características é utilizada para obter o gradiente de uma área que é representada utilizando uma matriz de covariância. Para extrair o gradiente de cada

pixel é utilizada a seguinte equação:

$$y = [r, g, b, |I_x|, |I_y|, |I_x x|, |I_y y|, \sqrt{I_x^2 + I_y^2}, \arctang(|I_y|)/(|I_x|)] \quad (3.3)$$

onde $|I_x|, |I_y|, |I_x x|, |I_y y|$ são os gradientes de primeira e a segunda ordem da imagem em níveis de cinza, cujo pixel está localizado em (x, y) . A magnitude do gradiente é calculado pela fórmula $\sqrt{I_x^2 + I_y^2}$ e a orientação por $\arctang(|I_y|)/(|I_x|)$. As variáveis r, g e b correspondem aos valores do vermelho (red), verde (green) e ao azul (blue), no pixel localizado em (x, y) . Essa fórmula gera um vetor de nove dimensões para cada pixel. Cada região é representada utilizando a matriz de covariância ϕ ,

$$\phi = \sum_{n=1}^N (y - \mu_y)(y - \mu_y)^T \quad (3.4)$$

esta matriz de covariância é utilizada para identificar a área de vegetação na imagem. Na fórmula, o y representa o gradiente do pixel e o μ_y representa a média dos pontos Bawden et al. (2017). No mesmo sentido, Sujaritha et al. (2017) descrevem a construção de um robô para detecção de ervas daninhas em uma plantação de cana-de-açúcar. O robô é equipado com duas câmeras. A primeira captura imagens com resolução de 3264x2448, utilizada para navegação e a segunda 2592x1944, utilizada para reconhecimento das ervas daninhas. Também possui uma placa solar de 20-40 V que alimenta o veículo, que pode ser visto na Figura 13. Para o processo de treinamento foram adquiridas 300 imagens, em ambiente externo, com diferentes condições de luminosidade. Os experimentos foram conduzidos com 10 espécies de plantas, a cana-de-açúcar e 9 ervas daninhas, sendo 30 imagens de cada espécie.

No processo de reconhecimento das ervas daninhas, há uma extração da textura das folhas e depois as características obtidas são utilizadas em um Algoritmo de Classificação em Tempo Real utilizando Lógica Fuzzy (FRTCA). Os métodos utilizados para extração das características relacionadas a textura são: matriz de níveis de cinza de segunda ordem, características de textura de Laws, Gabor *wavelet* e um método de deposição por *wavelet* invariante a rotação. Embora sejam utilizadas 10 espécies na aplicação, no classificador FRTCA são reconhecidas apenas duas classes, uma é a cana-de-açúcar e a outra é a erva daninha. O algoritmo utilizado é similar a um classificador que utiliza distância euclidiana ponderada. O peso atribuído a cada elemento do vetor de distância é fuzzificado baseado na sua diferença. Quanto menor for a

Figura 13 - Robô que identifica ervas daninhas.



Fonte: Sujaritha et al. (2017)

diferença entre os valores de uma determinada característica entre a cana-de-açúcar e a erva daninha, maior será o seu peso. A acurácia média da técnica proposta por Sujaritha et al. (2017) foi de 92.9% e o tempo de processamento de cada imagem foi de 0.02 segundos. Vale ressaltar que o trabalho de Sujaritha et al. (2017) assume que há somente uma espécie de planta em cada imagem. Na nossa pesquisa buscamos classificar uma imagem que pode apresentar dois tipos diferentes de plantas, ervas daninhas e cana-de-açúcar.

Amend et al. (2019) apresenta uma abordagem desenvolvida pela Deepfield Robotics, que é uma divisão do grupo Bosch. A proposta consiste na aplicação de aprendizado profundo. Para isso utilizam o que chamam de “tremenda quantidade de dados de treino” para classificação de plantas. Eles utilizam 4,3 milhões de pixels rotulados com mais de 40 classes de plantações e ervas daninhas, em diferentes estágios de crescimento, condições de luminosidade, condições climáticas e tipos de solos. A técnica apresenta taxa de acerto média de 93% em ambientes reais. Os autores não informam o tempo de processamento médio ou a velocidade na qual os robôs podem

se deslocar. Além disso, analisando o projeto do robô na Figura 14, entende-se que ele é projetado para cultivos que possuem até determinada altura.

Figura 14 - Robôs desenvolvidos pela Deepfield Robotics para reconhecimento e remoção de ervas daninhas.



Fonte: Amend et al. (2019)

Elstone et al. (2020) utilizam uma solução de visão computacional para reconhecimento de ervas daninhas acoplada em um trator no cultivo de alface, como na Figura 15. Os autores definem uma velocidade constante para o trator de 5km/h. O artefato acoplado ao trator tem iluminação para garantir invariância nesse aspecto, independentemente das condições do ambiente. Além disso, também é projetado para não sofrer influência do vento, o que causaria a dispersão do herbicida utilizado, pois a proposta dos pesquisadores é fazer uso de herbicidas de forma eficiente, aplicando-o somente nas ervas daninhas.

Para a identificação das plantas, Elstone et al. (2020) utilizam os componentes vermelho, verde e o infravermelho próximo (NIR - *Near Infra-Red*), combinados e partindo da premissa que as alfaces são sempre maiores que as ervas daninhas. Os autores calculam um índice que chamaram de RGNIR (*Red Green NIR - Vermelho Verde Infravermelho Próximo*), a partir do qual é aplicado o algoritmo de Otsu para limiarização, e então diferenciadas as alfaces das ervas daninhas. O cálculo do RGNIR

segue a seguinte equação

$$RGNIR = \frac{\beta(G_c + g_c)}{(\alpha(R_c + r_c) + \beta(G_c + g_c) + \gamma(B_c + b_c) + L')} \quad (3.5)$$

Onde G_c, R_c , e B_c são os canais verde, vermelho e azul, respectivamente, e o g_c, r_c , e o b_c , são um valor para correção de fundo para cada canal (0-255), α, β e γ são um valor de peso para os canais (0-1) e L é uma constante para ajuste do solo. Os autores não descrevem, pelo menos de maneira clara, quais valores foram utilizados na obtenção dos resultados apresentados. A técnica possui uma taxa de acerto média entre 81% e 88%, no reconhecimento das ervas daninhas.

Figura 15 - a) Trator com o equipamento utilizado para reconhecimento das ervas daninhas acoplado. b) Visão interna da estrutura projetada para captura de imagens.



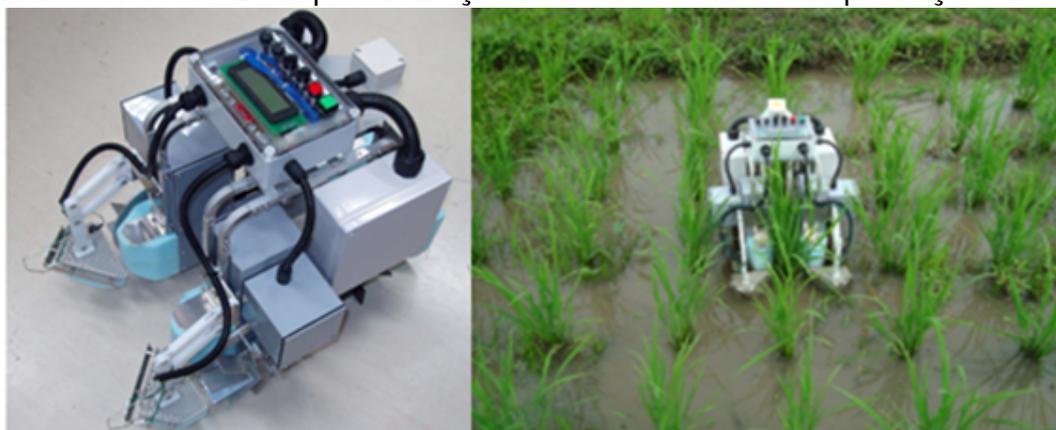
Fonte: Elstone et al. (2020).

Champ et al. (2020) utiliza um robô para remoção de ervas daninhas através da eletricidade. Para tanto é utilizada uma máscara de uma rede neural convolucional baseada em regiões para detecção das ervas daninhas. Os hiperparâmetros são reajustados para adaptar aos tipos de imagens e formatos encontrados na agricultura de precisão. O desempenho da técnica foi mensurado em plantações de milho e feijão, de forma semi controlada. Os autores ressaltam que o tamanho das plantas, ou seja, o seu grau de maturidade, influenciam no resultado. A depender disso, a taxa de remoção correta de ervas daninhas, sem o risco de remoção da outra planta, está entre 10% a 60%.

Sori et al. (2018) projetam um robô para remoção de ervas daninhas em plantações de arroz em áreas alagadas, mostrado na Figura 16. O Robô não realiza a identificação das ervas utilizando visão computacional. Ele realiza um percurso com

as plantas do arroz no centro e, no espaço entre as linhas, um mecanismo é utilizado para revirar o solo para que as ervas daninhas fiquem submersas e não realizem fotossíntese, o que evidentemente provoca a morte da planta.

Figura 16 - Robô utilizado para remoção de ervas daninhas em plantações de arroz.



Fonte: Sori et al. (2018)

O robô descrito por Raja et al. (2020), mostrado na Figura 17, é utilizado para detecção e remoção de ervas daninhas em plantações de tomate. O algoritmo utilizado para o reconhecimento das ervas é bem robusto, como os autores mesmo definem. São capturadas seis imagens, de diferentes posições, para que a possível oclusão da planta seja superada. Após a captura são realizados diversos passos, tais como: a definição de pontos de interesses, segmentação baseado no valor HSV dos canais e operações morfológicas. O robô possui um mecanismo, denominado de faca, que realiza a remoção da erva daninha. A acurácia apresentada, no caso para reconhecimento dos caules dos pés de tomate, é de 99,19%.

O próximo Capítulo descreve a Lógica Fuzzy seus fundamentos e como tem sido aplicada de forma a auxiliar na resolução de problemas reais.

Figura 17 - a) Vista do artefato utilizado para identificação da erva e utilização do spray.
b) Vista lateral do sistema desenvolvido.



(a)



(b)

Fonte: Raja et al. (2020)

4 LÓGICA FUZZY

Zadeh (1965) foi o precursor da lógica Fuzzy utilizada em aplicações computacionais. Em 1965, o autor publicou um artigo intitulado Fuzzy Sets, no qual se baseia na lógica matemática aplicada a conjuntos fuzzy para formular seus pressupostos. Zadeh (1965) parte da ideia de que algumas classes como "uma mulher bonita" ou "um homem alto" não constituem classes ou conjuntos no senso usual matemático. O autor ressalta ainda que essas classes, não definidas precisamente, possuem uma importância significativa no pensamento humano, sobretudo, em atividades como reconhecimento de padrões, comunicação e abstração. Zadeh (1965) parte de algumas definições matemáticas necessárias para aplicação da lógica fuzzy.

Uma primeira definição importante é a função que associa um determinado número real a um intervalo definido, como o intervalo $[0, 1]$, que é utilizado para associar um elemento em uma classe. Zadeh (1965) propõe o seguinte exemplo considerando X como um espaço de pontos e um elemento genérico de X expresso como x . Então, $X ::= \{x\}$. Um conjunto fuzzy ou classe A em X é caracterizado por uma função de pertinência $f_A(x)$ que associa cada ponto em X a um número real no intervalo $[0, 1]$, sendo o valor de $f_A(x)$ em x representando o grau de pertencimento de x em A . Assim, quanto mais próximo o valor de $f_A(x)$ a 1 maior é o grau de pertencimento de x em A .

Talpur, Salleh e Hussain (2017) discorrem sobre os tipos básicos das funções de pertinência. Abaixo, são apresentados esses tipos conforme a descrição dos autores.

Trapezoidal - possui 4 parâmetros escalares para definição de sua curva: a , b , c e d . A representação matemática é dada pela seguinte fórmula trapezoidal:

$$\text{trapezoidal}(x, a, b, c, d) = \max\left(\min\left(\frac{(x-a)}{(b-a)}, 1, \frac{(d-x)}{(d-c)}\right)\right) \quad (4.1)$$

Sino generalizada - possui uma forma similar a um sino. A função utiliza três parâmetros: a para a largura da curva do sino, b é um inteiro positivo e c ajusta o centro da curva. A fórmula é a seguinte:

$$\text{sino}(x, a, b, c) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}} \quad (4.2)$$

Triangular - apresenta o formato mais simples dentre as funções. É definida a partir de três pontos: a e c para base e b para o topo da curva. A fórmula é expressa conforme abaixo:

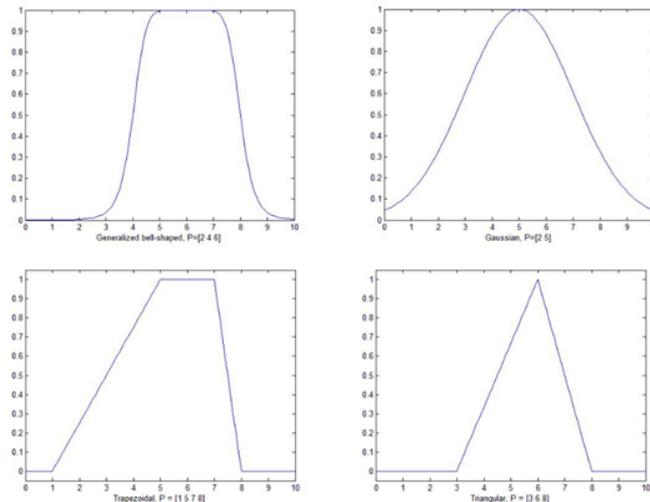
$$\text{triangular}(x, a, b, c) = \max\left(\min\left(\frac{(x-a)}{(b-a)}, 1, \frac{(c-x)}{(c-b)}\right), 0\right) \quad (4.3)$$

Gaussiana - é uma função que apresenta uma curva suave. Utiliza somente dois parâmetros: c para o centro e σ que determina a largura da curva. A fórmula utilizada é a seguinte:

$$\text{gaussiana}(x, c, \sigma) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2} \quad (4.4)$$

Na Figura 18, temos a representação das funções de pertinência descritas acima.

Figura 18 - Gráficos das principais funções de pertinência utilizadas na lógica Fuzzy.



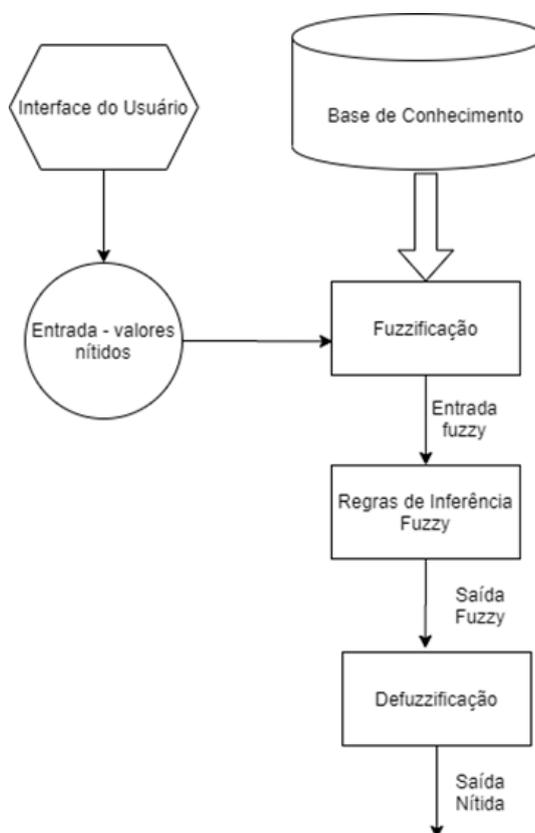
Fonte: Talpur et al. (2017)

Kaur e Singh (2019) enumera quatro componentes principais de um sistema baseado em lógica fuzzy: fuzzificação, motor de inferência, base de conhecimento e módulo de defuzzificação, vistos na Figura 19. Segue a definição dos autores para cada um dos componentes.

Módulo de fuzzificação - é utilizado para transformar um valor escalar real em um valor fuzzy. São utilizadas variáveis linguísticas fuzzy para representar determinados intervalos. Nesta etapa, os dados nítidos de entrada são recebidos e transformados em entradas fuzzy (variáveis linguísticas), que são utilizadas na próxima etapa.

Base de Conhecimento - contém o conhecimento relacionado ao domínio da aplicação. Armazena a base de dados e a base de regras de controle linguístico fuzzy. Todo

Figura 19 - Representação dos elementos envolvidos em um Sistema baseado em lógica Fuzzy.



Fonte: Adaptado de Kaur & Singh (2019)

o suporte de dados em relação as regras de controle linguístico está contido neste módulo.

Motor de inferência - é o modulo utilizado para tomada de decisões. Nesse módulo, cada entrada é mapeada para uma saída usando regras de especialistas que estão armazenadas na base de dados.

Módulo de Defuzzificação - a saída fuzzy é enviada para esse módulo para que seja convertida em valores nítidos.

Kaur e Singh (2019) apresentam algumas áreas em que a lógica fuzzy tem sido aplicada, fazendo referência a trabalhos publicados. É interessante notar que os autores citam o processamento de imagens, robótica e agricultura, entre outras, sendo áreas relacionadas a pesquisa que descrevemos neste documento.

Megahed e Mohammed (2020) propõem a utilização de redes neurais profundas e lógica fuzzy para o reconhecimento de expressões faciais. Os dados de entrada são

projetados para representar o nível de aprendizado, tempo de teste decorrido, taxa de validação de resposta e tuplas de sete classes de expressões faciais de estados emocionais agregados. Os autores reiteram que o fuzzificador realiza a conversão de valores de entrada de acordo com a linguagem fuzzy, usando a função de pertencimento. Os autores utilizaram dois tipos de funções de pertencimento, a triangular e a trapezoidal, para converter os valores de entrada para a linguagem fuzzy. No trabalho, os autores propõem 11 variáveis linguísticas no sistema de inferência fuzzy, categorizando-as em 4 tipos diferentes de variáveis que descrevem estados emocionais, calculadas através de testes e do gerenciamento no fim do exame.

No motor de inferência, Megahed e Mohammed (2020) usam regras codificadas condicionais para formar a base de conhecimento fuzzy. Abaixo um exemplo de uma regra que é mencionada no trabalho dos referidos autores:

```
If HappyExpr IS VeryHigh
AND AngryExpr IS VeryLow
AND DisgustExpr IS VeryLow
AND FearExpr IS VeryLow
AND SadExpr IS VeryLow
AND SurpriseExpr IS VeryLow
AND NeutralExpr IS VeryLow
AND AnswerValidity IS DefinitelyValid
AND TestElapsedTime IS VeryLow
AND CurrentLevel IS Intermediate
THEN NextLevel IS VeryAdvanced;
```

É interessante a observação dos autores que essas regras são definidas usando o conhecimento de especialistas, suas percepções e julgamentos pessoais. Na fase defuzzificação são obtidos valores de saídas a partir de valores fuzzy convertidos.

Como mencionado anteriormente, Megahed e Mohammed (2020) utilizam a lógica fuzzy juntamente com redes neurais profundas, que são utilizadas na etapa de extração de parâmetros. É comum a utilização da lógica fuzzy com outras técnicas de

inteligência artificial. Por exemplo, Li et al. (2017) utilizam árvores de decisão juntamente com o conceito da lógica fuzzy para realizar a anotação automática de imagens. Juneja e Rana (2020) utilizam o filtro Gabor, lógica fuzzy e SVM para a classificação de expressões faciais.

Liu et al. (2017) utilizam a lógica fuzzy para realçar imagens microscópicas de fitoterápicos chineses. Na metodologia definida pelos autores, uma imagem em níveis de cinza é gerada e cada pixel é relacionado com um nível de cinza utilizando uma função de pertinência. Então, é avaliado se o valor da função de pertencimento de cada pixels é maior do que 0,5 ou menor. A depender desse valor, é realizada uma correção.

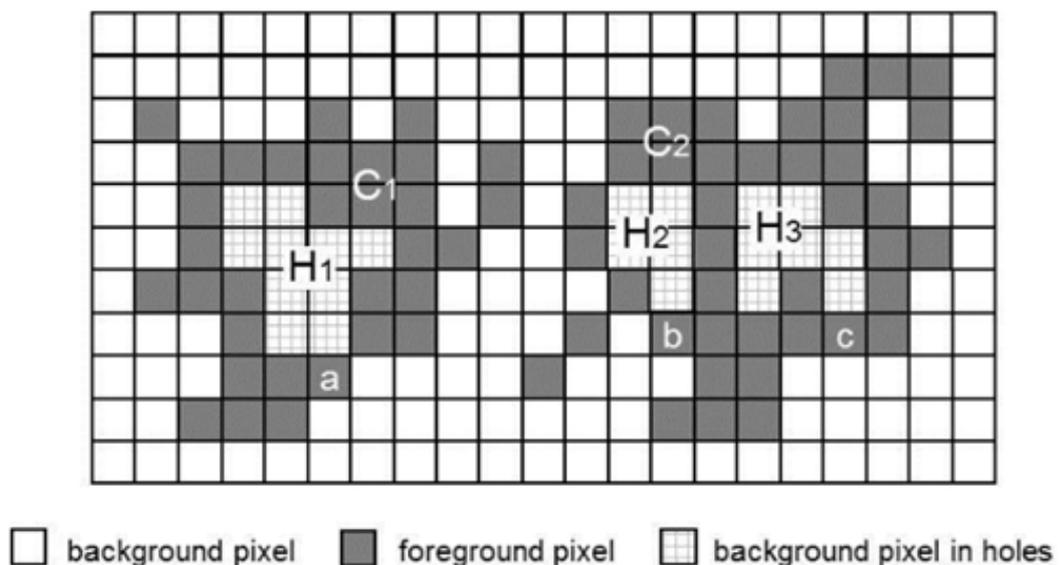
No próximo capítulo, trataremos do conceito de componentes conexos que é uma metodologia utilizada no processamento de imagens para identificar pixels pertencentes a um mesmo objeto.

5 COMPONENTES CONEXOS

Componentes conexos é uma abordagem amplamente utilizada no contexto de processamento de imagens e reconhecimento de padrões, conforme citado por He e Chao (2015) e Bataineh et al. (2012). A técnica é usada para identificar os pixels pertencentes a um determinado objeto, usando-o para gerar uma imagem binária. Bataineh et al. (2012) cita a utilização desta técnica em diversas aplicações no campo da visão computacional, envolvendo o reconhecimento de um alvo, reconhecimento de placas de veículos, análise de imagens médicas, reconhecimento de rostos e impressões digitais, entre outros.

He e Chao (2015) apresentam uma definição, considerada de clássica pelos autores, do processo de rotulagem dos componentes relacionados. Dada uma imagem binária de tamanho $N \times N$, o pixel coordenado (x, y) , onde $0 \leq x \leq N - 1$ e $0 \leq y \leq N - 1$, na imagem é denominado $b(x, y)$. Os pixels do primeiro plano também são chamados de pixels do objeto. Embora não seja declarado de outra forma, os valores dos pixels do objeto e pixels do fundo são considerados 1 e 0, respectivamente. Além disso, por conveniência, todos os pixels na borda de uma imagem são considerados pixels de fundo. A Figura 20 mostra um exemplo da técnica.

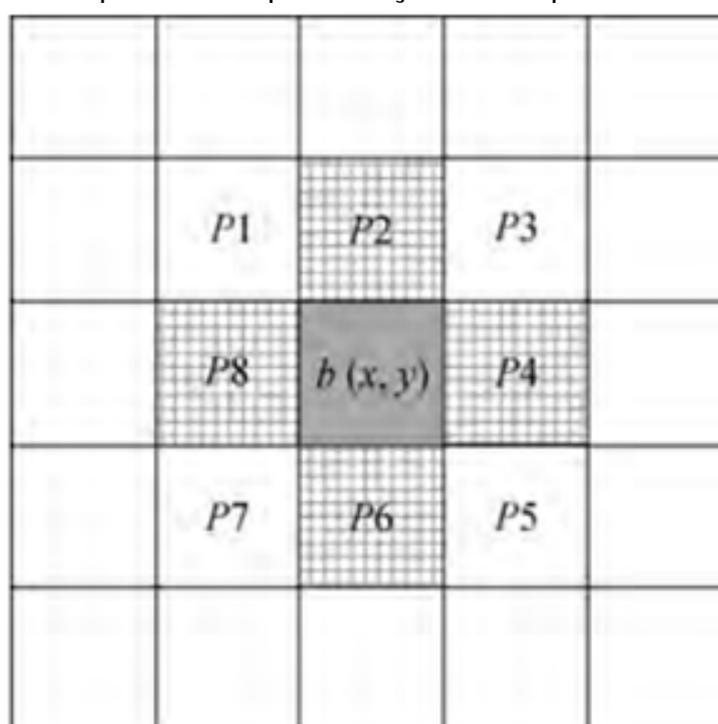
Figura 20 - Representação do conceito utilizado pelos componentes conexos.



Fonte: He& Chao (2015)

Para analisar a conectividade de um objeto, são usados os 4 pixels vizinhos de um pixel central, ou 8 pixels vizinhos. Os 4 e 8 vizinhos de um pixel são representados na Figura 21, onde os pares P_n representam 4 posições de vizinhos e de $P1$ a $P8$ compõem a representação de 8 vizinhos. Em geral, 8 conectividade é usada para componentes conectados (pixels de primeiro plano) e 4 conectividade para orifícios (pixels de fundo) (HE; CHAO, 2015). Em resumo, o objetivo da técnica é analisar os vizinhos de um pixel em busca de um determinado valor e, caso seja encontrado, um algoritmo é aplicado recursivamente.

Figura 21 - Esquema de representação de um pixel e sua vizinhança.



Fonte: He& Chao (2015)

Fatichah et al. (2016) utilizam o conceito de componentes conexos e lógica fuzzy para fazer a contagem de carros em uma determinada imagem. A lógica fuzzy é utilizada na etapa de segmentação da imagem. Depois da segmentação é feita rotulação dos objetos, no caso carros, encontrados na imagem. Fatichah et al. (2016) dividem em dois passos a busca por componentes conexos. O primeiro passo:

1. Percorrer os pixels da imagem através das linhas e colunas (*Scanning Raster*) ;
2. Se o pixel não pertencer ao fundo:

- (a) Armazene os pixels vizinhos ao pixel corrente;
- (b) Caso não exista nenhum pixel vizinho, o pixel é rotulado como único e o passo c é realizado;
- (c) Senão, procure um vizinho com o menor rótulo e selecione-o como o pixel atual;
- (d) Armazene a equivalência entre a vizinhança rotulada.

O segundo passo consiste nas seguintes etapas:

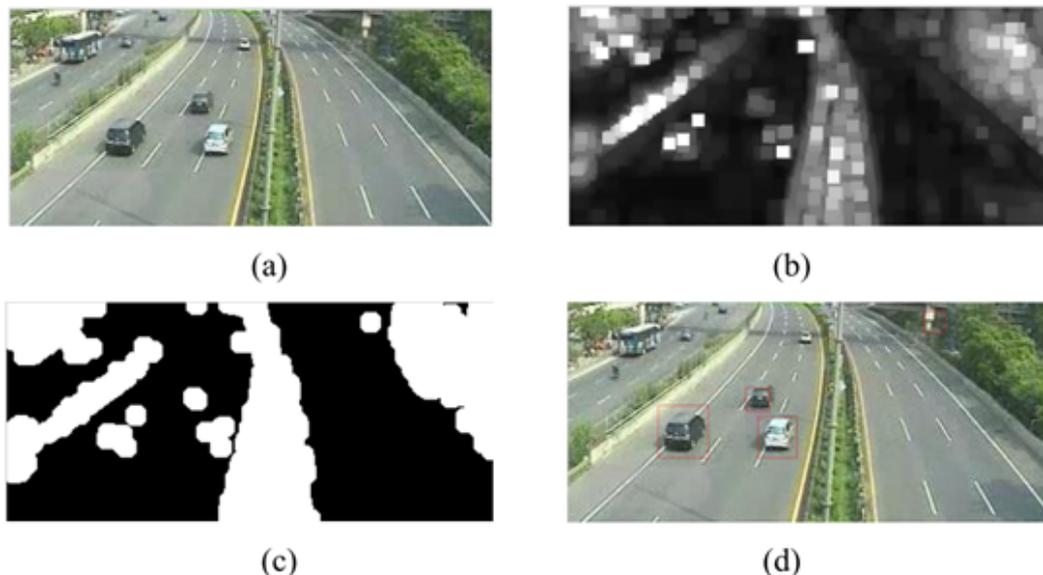
1. Percorrer cada pixel na imagem;
2. Se o pixel corrente não pertence ao fundo:
 - (a) Rotular novamente o pixels com o menor rótulo possível equivalente.

Após a verificação dos componentes conexos, há a verificação do tamanho da área dos possíveis veículos encontrados. Foi estimado a área mínima de um veículo em 120 pixels e a máxima em 800. Para ser considerado um veículo, o tamanho da área conexa deve estar dentro desse intervalo. A taxa de acerto obtida na pesquisa foi de 78,21%. Na figura 22 são apresentadas as imagens das respectivas etapas definidas por Fatichah et al. (2016).

Geralmente, uma estrutura de dados é utilizada para guardar a relação entre um determinado pixel e seus vizinhos, de acordo com um critério de similaridade. Depois os vizinhos dos vizinhos são percorridos, até o final da imagem. Liu e Gu (2014) propõem a utilização de um Autômato Finito Determinístico para rotular os componentes conexos de uma imagem. Na Figura 23 encontra-se o autômato proposto no trabalho e na Figura 24 a transição entre os estados do autômato.

Liu e Gu (2014) comparam com outras técnicas a metodologia por eles proposta em relação a complexidade e tempo de processamento, através de tabela comparativa. Conforme o comparativo a técnica possui menor complexidade e menor tempo de processamento. Em relação ao processamento os autores não deixam claro qual hardware foi utilizado no comparativo, se foi o mesmo das outras duas técnicas. Foram analisadas imagens de base de dados do site CMU com três tamanhos diferentes

Figura 22 - a) Imagem de entrada. b) Após o processo de dilatação usando lógica Fuzzy c) Imagem Segmentada. d) Imagem de saída com veículos encontrados.



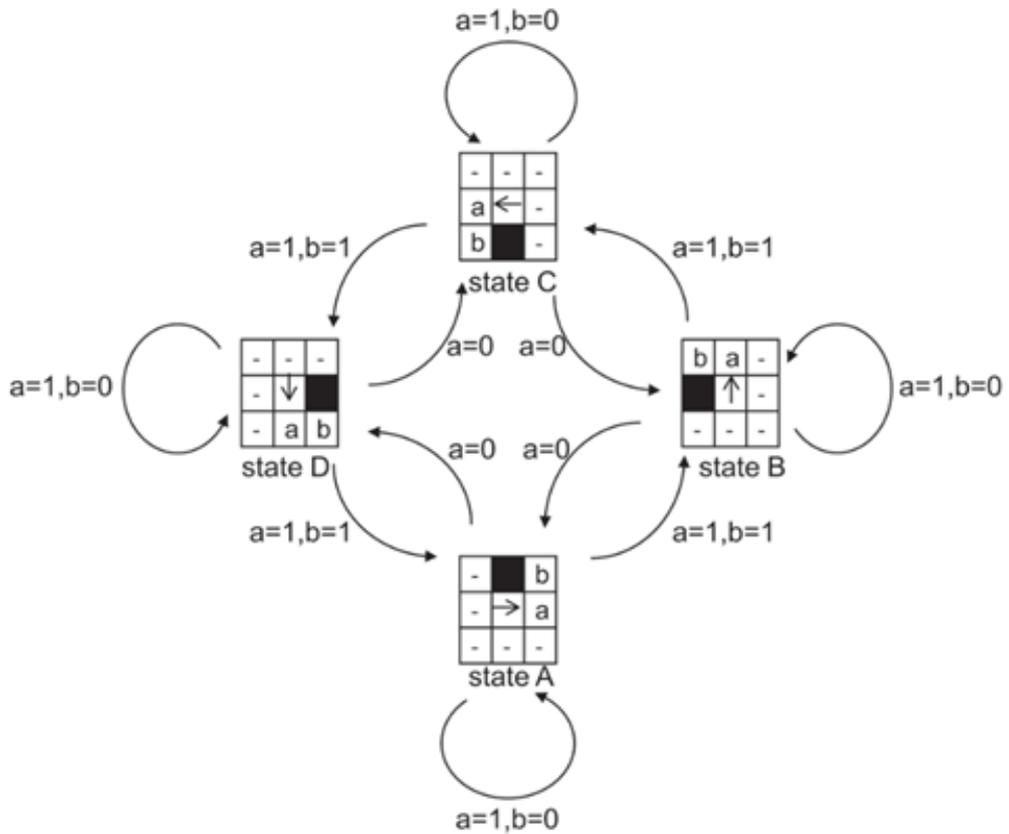
Fonte: Fatichah et al. (2016)

0,16Mpixels, 0,64Mpixels e 2,56Mpixels. O tempo de processamento fica entre 6,2 a 11 ms para as imagens de 0,16 M, 19,3 a 25 ms para imagens de 0,64M e 66,2 a 76,5 ms para imagens de 2,56M.

A técnica de Componentes Conexos é tão amplamente utilizada que Palaniappan e Natarajan (2014) realizam um estudo para aplicar o processamento paralelo dessa abordagem, tendo em vista a execução em tempo real. É utilizado um processador multicore XMOS para o teste. Por fim, os autores conseguem uma redução média de 51% na velocidade de processamento das imagens.

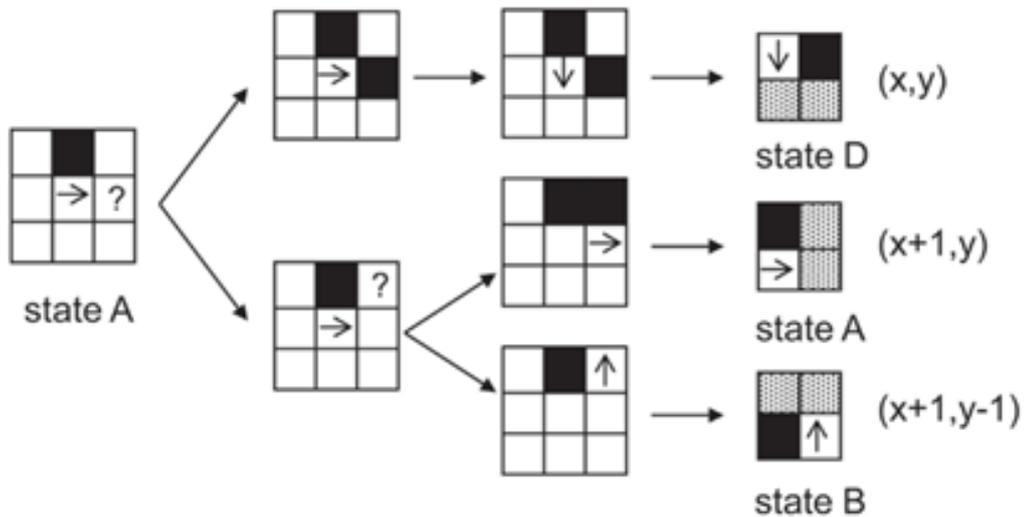
O próximo capítulo mostra a metodologia adotada nesta pesquisa, os passos que foram seguidos até a concepção de um algoritmo que possa ser utilizado para reconhecimento da *Ipomoea* sp. ou uma dicotiledônea, em uma plantação de cana-de-açúcar. Discutimos como os conceitos apresentados nesse capítulo e nos anteriores são utilizados para resolver o problema proposto.

Figura 23 - Diagrama de estados do Autômato Finito Determinístico.



Fonte: Liu e Gu (2014)

Figura 24 - Transições de estados a partir do estado A.



Fonte: Liu e Gu (2014)

6 METODOLOGIA

Investigar as maneiras de identificar uma planta foi o primeiro passo realizado na pesquisa. Várias áreas possuem interesse no desenvolvimento de um método automatizado para reconhecimento de uma determinada espécie. A agricultura está entre as principais áreas que se beneficiariam desse conhecimento.

Há diversas propostas para realizar o reconhecimento de plantas. Estudando a literatura percebe-se que há uma predominância de métodos que utilizam as folhas para realizar o reconhecimento, e diferenciação entre espécies de plantas, por ser a parte, geralmente, mais visível. Assim, foi assumido o pressuposto inicial de que as folhas da corda de viola e da cana-de-açúcar seriam, pelos menos inicialmente, utilizadas para obter os parâmetros necessários para realizar o processo de diferenciação.

As técnicas estudadas, que realizam o reconhecimento de plantas, utilizam diversas características das folhas, tais como: a forma, textura, venação, área, dentre outras. Há trabalhos que se propõem a diferenciar uma diversidade extensa de espécies e, por conta disso, utilizam várias características das folhas para melhorar a taxa de reconhecimento. Nesta pesquisa, a aplicação intrinsecamente restringe a identificação à somente duas espécies. Por conta disso, buscou-se trabalhar com a utilização de uma característica ou a menor quantidade possível delas para realizar a tarefa, tendo em vista a redução no tempo de reconhecimento.

Dentre as possíveis características que poderíamos utilizar para diferenciar as plantas, primeiramente, escolhemos a venação. Dado que a corda de viola é uma dicotiledônea e a cana-de-açúcar uma monocotiledônea, o padrão de venação difere consideravelmente. Assim, o próximo passo foi entender como os pesquisadores realizam a extração da estrutura de venação de uma folha.

6.1 Método proposto para extração da venação das folhas

Segundo a revisão bibliográfica realizada, geralmente a cor da venação não é relevantemente diferente do restante da folha. Por isso, é necessária uma etapa de

pré-processamento que consiste na aplicação de algumas técnicas de forma a otimizar o processo de extração das veias pertencentes as folhas. Perez-Sanz, Navarro e Egea-Cortines (2017) apontam a etapa de pré-processamento como fundamental para contornar o problema com a variação da luminosidade, sendo que a técnica a ser aplicada para correção dependerá da qualidade e do tipo da imagem.

Na etapa de pré-processamento, primeiramente, entendemos que a imagem original pode ser reduzida, com o objetivo de diminuir o tempo de processamento da imagem, requisito deste trabalho. Desta forma, foi analisado qual seria o menor tamanho de uma imagem contendo uma folha que ainda permita o reconhecimento da venação. Neste aspecto, como definimos a utilização da biblioteca do OpenCV para o processamento das imagens, foram analisadas as técnicas disponíveis nesta biblioteca que são utilizadas para reduzir o tamanho de uma imagem. O OpenCV dispõe das seguintes estratégias para redução de uma imagem: *nearest-neighbor*, *bi-linear*, *pixel area relation*, *bicubic* e *Lanczos* (ver o Apêndice A para detalhes).

Após a redução da imagem, há a transformação da imagem colorida em escala de cinza. A literatura estudada afirma que a conversão utilizando a fórmula padrão e o espaço de cores RGB não retorna bom resultado. Um dos espaços mais adequado para trabalhar com folhas é o HSI.

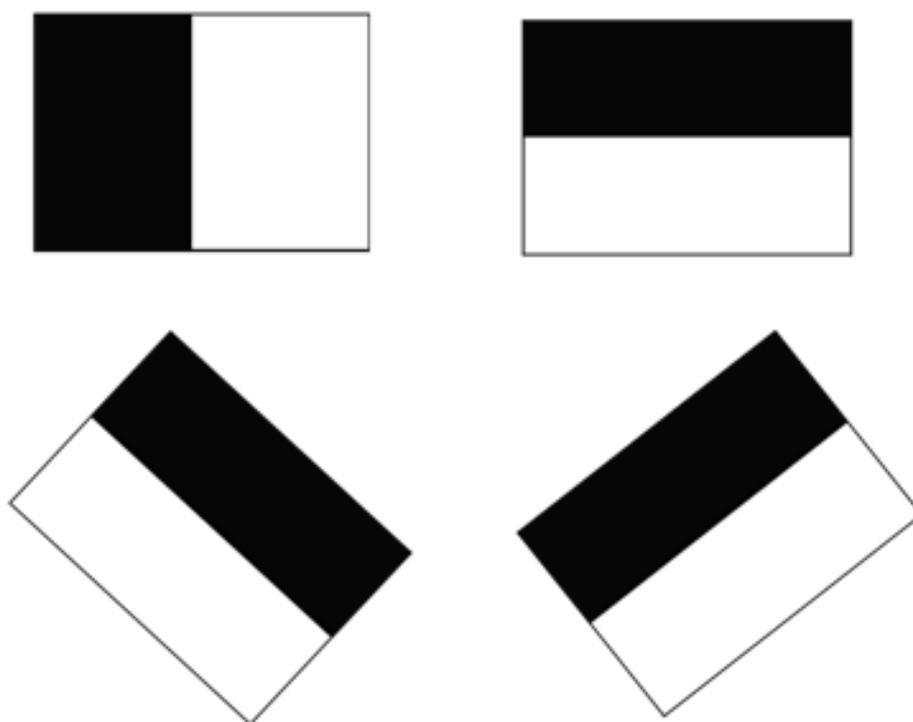
A última etapa no estágio do pré-processamento é a utilização de um algoritmo para realizar a binarização da imagem, separando os pixels pertencentes a veia e ao restante da folha. O algoritmo de Canny é amplamente utilizado nessa etapa. Porém, é necessário analisar quais os valores de dois parâmetros que são requeridos para sua aplicação, os valores mínimo e máximo, que determinam o limiar a ser considerado para dizer o que é veia ou não. Esse valor foi investigado através de experimentos, aumentando ou diminuindo os valores e analisando a imagem de saída.

A imagem binarizada, resultado da aplicação do algoritmo de Canny, deve ser utilizada para aplicação de uma técnica que identifique o padrão da veia apresentada na imagem. Para analisar essa estrutura, identificamos que alguns pesquisadores utilizaram a transformada de Hough. Essa técnica foi utilizada em uma base de dados para analisar o seu desempenho neste contexto. Como o tempo de processamento foi maior do que o esperado, outra técnica foi analisada, o Haar-like *features*.

A aplicação do Haar-like *features* é amplamente utilizada para reconhecimento de faces com baixo tempo de resposta. Como as folhas apresentam as venações com um padrão reconhecível e o tempo de processamento é importante na nossa aplicação, decidimos investigar o desempenho desta técnica para descobrir o padrão de venação das folhas. Para isso, primeiramente definimos os *templates* que serão utilizados para a extração das características das folhas. A Figura 25 mostra os *templates* propostos para análise das veias de uma folha. A ideia é que, independente da rotação da folha, caso ela seja monocotiledônea, um dos *templates* terá uma ocorrência maior do que nas folhas dicotiledôneas.

O diferencial desta técnica são os *templates* propostos para diferenciar a corda de viola ou as dicotiledôneas em geral, da cana-de-açúcar.

Figura 25 - *Templates* propostos para reconhecimentos do padrão de venação.



Fonte: Elaboração Própria.

Uma das vantagens da utilização do Haar-like *features* é a possibilidade de aplicar o conceito de imagem integral, de forma que a verificação de ocorrência de *templates* na imagem seja ainda mais rápida. A aplicação dos *templates* gera quatro matrizes de ocorrência de cada *template* na imagem.

Um aspecto observado nas imagens retiradas das plantas em ambientes reais é que há partes do céu ou do solo que a compõem e que não precisariam ser analisadas. Para isso, definimos uma estratégia de dividir a imagens em quadrados menores e analisamos o percentual de verde na área, bem como realizar testes para definir qual seria o tamanho do quadrado mais apropriado, de forma a reduzir o espaço de buscas.

Decidimos utilizar um sistema fuzzy para auxiliar na definição do que seria um quadrado com pequena, média ou grande quantidade de pixels verdes, para o tamanho do *template* e classificação da área como pertencente a uma folha dicotiledônea ou monocotiledônea.

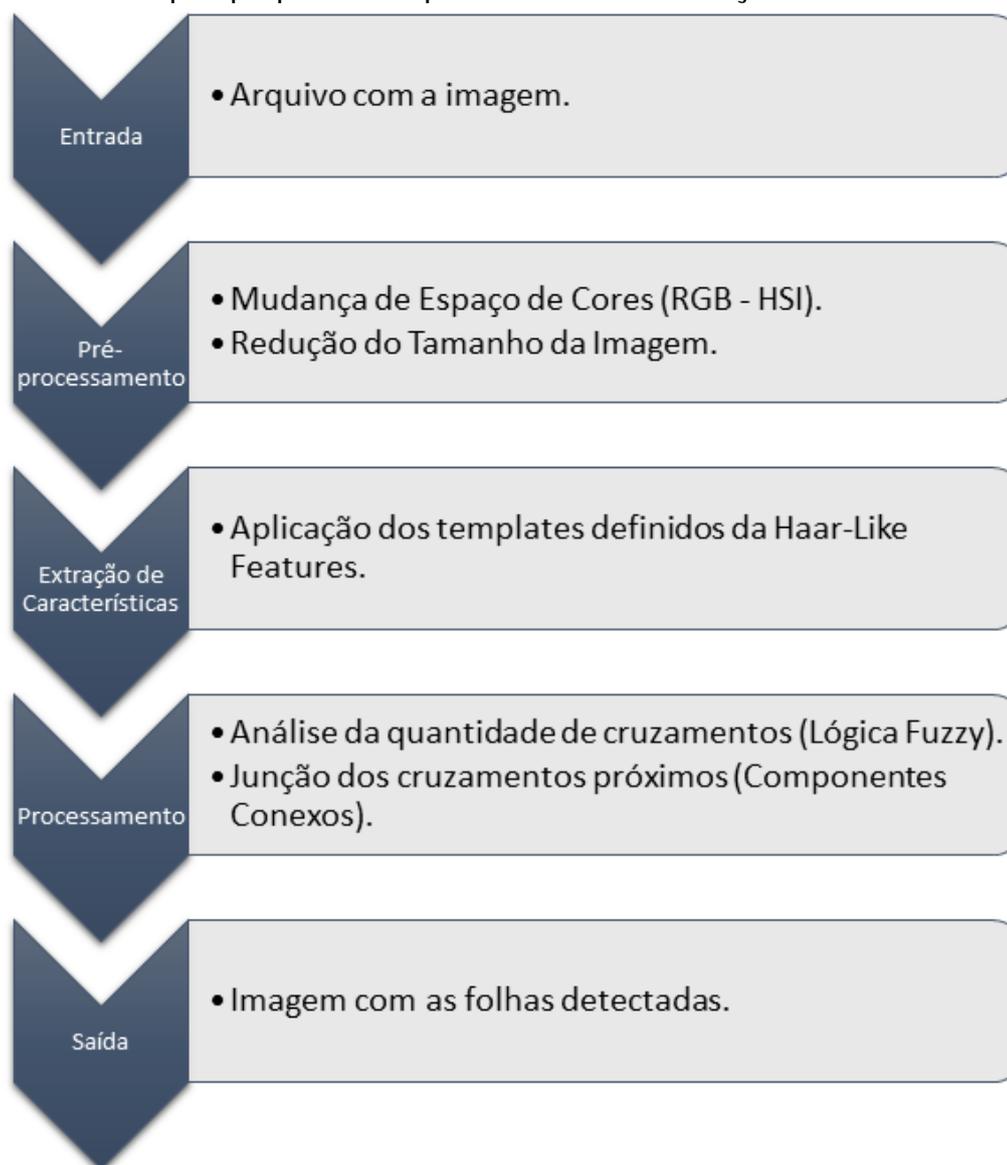
Definimos três variáveis a serem utilizadas no sistema de inferência fuzzy: percentual de pixels verde na área, tamanho do *template* e total de cruzamentos entre templates encontrados. No próximo Capítulo, descrevemos os testes realizados para identificar a função de pertinência relativa a cada variável linguística mencionada acima.

Os passos propostos anteriormente podem ser resumidos conforme a Figura 26. Primeiro será realizada a aquisição da imagem, seja por um arquivo ou por uma câmera acoplada em algum artefato. Depois vem a etapa do pré-processamento no qual a imagem é reduzida, convertida para o HSI e o algoritmo de Canny é aplicado. O terceiro bloco é realizado pelo Haar-like *features*. Ainda na etapa de processamento, a imagem é dividida em quadrados, o percentual de área verde é estimado, o tamanho do *template* é definido, depois a quantidade de intersecções é avaliada e a parte classificada. Utilizamos também o conceito de componentes conexos para analisar os quadrados próximos e estimar os que pertencem a mesma folha.

Essa primeira etapa da pesquisa realizar-se-á com imagens controladas de uma determinada base de dados. A próxima etapa será a investigação de quais ajustes devem ser feitos para aplicação do método elaborado em ambientes reais.

Nesse caso, optamos por montar uma base de dados com imagens reais, aplicar a técnica e analisar a saída. A resposta da técnica em relação a luminosidade, oclusão das folhas e seus tamanhos variados também foi avaliada. Outro aspecto, é a investigação da resolução das imagens, bem como a distância em relação a planta para que ocorra a identificação. Além disso, como os *templates* também possuem um tamanho

Figura 26 - Etapas propostas no processo de identificação das ervas daninhas.



Fonte: Fonte: Elaboração própria.

que deve ser definido, buscamos uma forma de relacionar a resolução da imagem e o tamanho dos *templates* a serem aplicados.

É importante realizar testes para identificar o comportamento da técnica em relação a algumas condições existentes nos ambientes reais. As folhas aparecem em diversas posições, assim, a técnica deve ser invariante a rotação. Além disso, é recorrente a ocorrência de folhas oclusas e diversas condições de luminosidade das imagens. O próximo Capítulo, descreve os testes realizados a fim de verificar a resposta a essas condições.

7 RESULTADOS

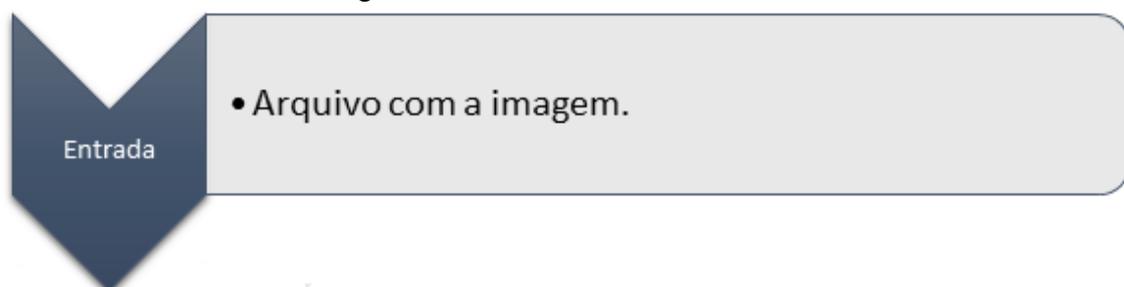
Abaixo descrevemos os procedimentos realizados a fim de investigar os procedimentos metodológicos necessários para resolver o problema proposto nesta pesquisa.

7.1 Entrada

Para o processamento de imagens escolhemos o OpenCV (*Open Source Computer Vision Library*), por ser uma biblioteca amplamente utilizada para este fim. Além disso, sua licença é BSD, livre para uso acadêmico e comercial. Também é compatível com diversos sistemas operacionais, tais como: Windows, Linux, Mac OS, iOS e Android.

A base de dados utilizada para teste do algoritmo em ambiente controlado foi o Flavia Dataset (WU et al., 2007). Esta base de dados contém folhas de 33 espécies, algumas monocotiledôneas e a maioria dicotiledônea. No total, a base dispõe de 1908 imagens, cuja resolução é 1600 x 1200 pixels. Foram selecionadas 131 imagens, 59 monocotiledôneas e 72 dicotiledôneas. As imagens foram escolhidas devido à similaridade com as folhas da cana-de-açúcar e da corda de viola. Definidas as imagens de entrada, usadas no primeiro estágio conforme a etapa definida na Figura 27, a próxima etapa consiste no pré-processamento destas. Na próxima seção será descrito o processo de pré-processamento.

Figura 27 - Processo de entrada.

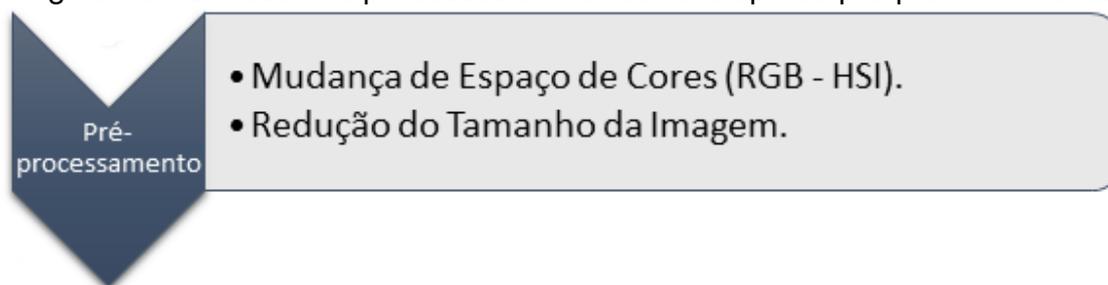


Fonte: Fonte: Elaboração própria.

7.2 Pré-processamento

Nessa etapa, serão realizadas as etapas conforme descritas na Figura 28. As imagens de entrada passam pelo estágio de pré-processamento, no qual é realizado a redução da imagem, a mudança no espaço de cores e depois a segmentação.

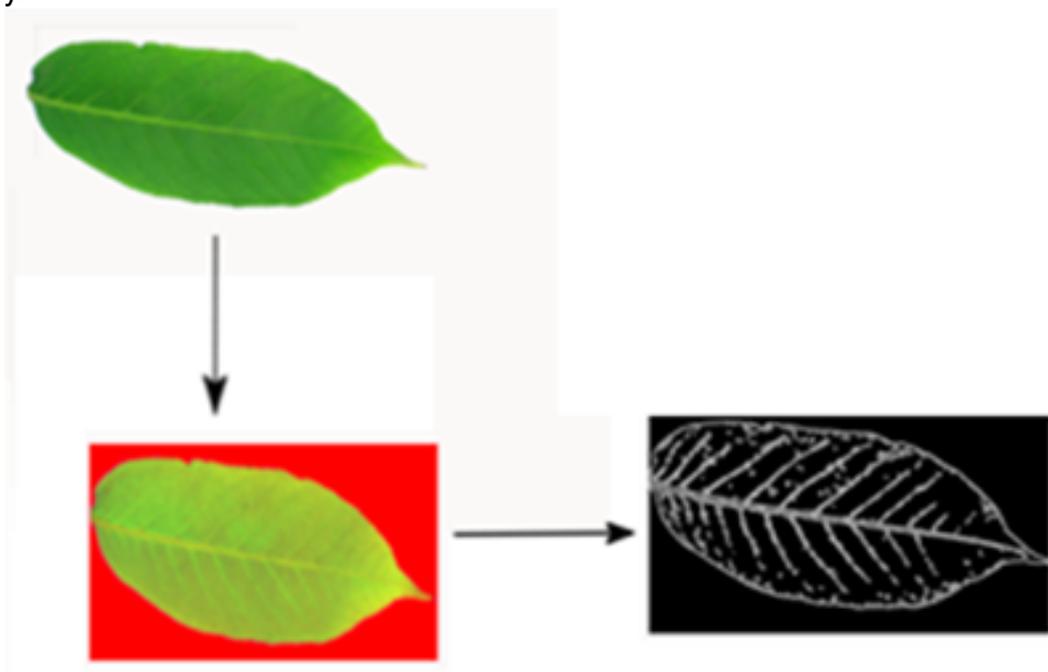
Figura 28 - Síntese dos processos utilizados na etapa de pré-processamento.



Fonte: Fonte: Elaboração própria.

Na Figura 29 é mostrado um resumo dos passos realizados nesta fase. A imagem de entrada sofre uma mudança no espaço de cores, do RGB para o HSI. Depois a imagem é binarizada utilizando o método de detecção de bordas de Canny.

Figura 29 - Resumo das etapas de pré-processamento. Primeiro a imagem de entrada, depois a mudança para o HSI e por fim o resultado da aplicação do algoritmo de Canny.



Fonte: Elaboração própria.

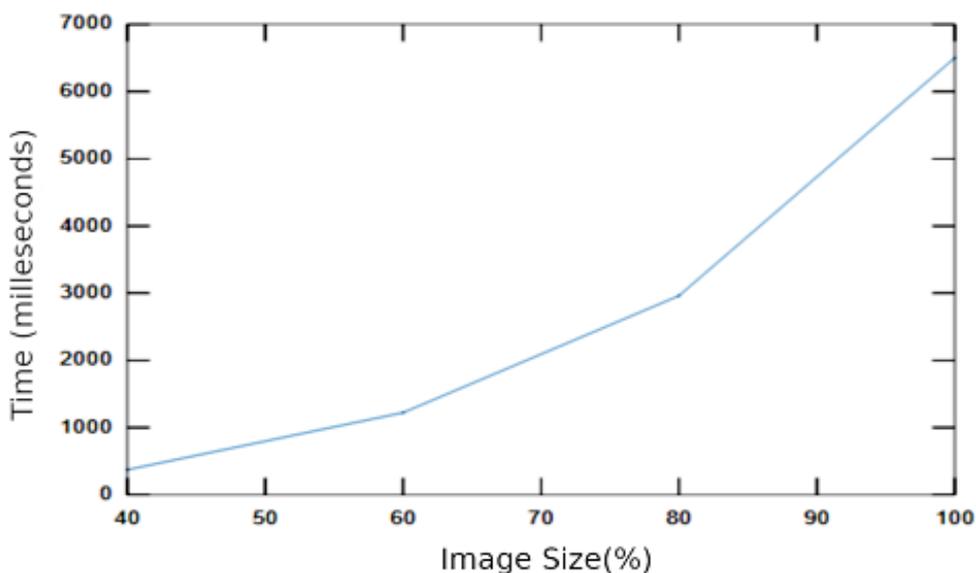
Após a etapa de pré-processamento, são realizados testes com técnicas para ob-

termos os padrões encontrados nas folhas. Primeiramente, utilizamos a transformada de Hough, conforme descrito na próxima seção.

7.3 Utilização da Transformada de Hough

Primeiramente, utilizamos a transformada de Hough para reconhecimento da veinação das folhas. Nosso objetivo era encontrar uma técnica que analisasse uma imagem em cerca de 100 milissegundos, dado que a mesma contém somente uma folha e foi retirada em um ambiente controlado sem incidência de sombras ou com possibilidade de oclusão. Conforme pode ser visto no gráfico na Figura 30, mesmo com a redução da imagem, o tempo de processamento estaria longe de ser alcançado. Além disso, no caso desta técnica, percebe-se que a redução da imagem ocasiona dificuldade em reconhecer as linhas pertencentes a folha.

Figura 30 - Relação entre o tempo de processamento e o tamanho da imagem utilizando a transformada de Hough.



Fonte: Elaboração própria.

Por conta do desempenho utilizando a transformada de Hough, decidimos testar uma outra técnica: o Haar-Like *features*, cujos procedimentos são descritos na seção abaixo.

7.4 Utilizando o Haar-like *features*

Trabalhar com uma imagem de alta resolução permite encontrar maiores detalhes, o que podem resultar em melhor acurácia. Porém, o tempo de processamento possui relação direta com a resolução da imagem, por isso testes foram executados no sentido de obter os parâmetros necessários para uma ponderação entre uma boa acurácia e um adequado tempo de resposta, relacionados com o menor tamanho possível a imagem.

Várias simulações foram feitas para ajustar alguns parâmetros. Primeiramente, como os templates podem possuir tamanhos variados, investigamos qual seria o seu tamanho ideal, bem como qual o fator de redução da imagem original que ainda permitisse uma boa acurácia.

7.5 Tamanho dos templates

A Tabela 2 mostra a acurácia em relação ao tamanho dos *templates* utilizados nas imagens do Flávia Dataset, ou seja, em ambiente controlado. Utilizamos um *threshold* que corresponde a quantidade mínima de pontos a serem encontrados em uma determinada imagem, em pelo menos 2 ou 3 *templates* diferentes, para que seja considerada dicotiledônea.

O tamanho do *template* refere-se á quantidade de pixels na linha e coluna que é variável a depender de sua orientação. O *template* de tamanho 3 e 2 apresentam taxas de acerto semelhantes, porém quanto maior o *template*, por estarmos utilizando o conceito de imagem integral, menor será o tempo para analisar a imagem. Assim, escolhemos o *template* 3, que tem tamanho 3 x 2, para os *templates* na vertical e 2 x 3 na horizontal. Nos outros dois *templates*, que procuram padrões nas diagonais, utilizou-se uma matriz 4 x 4 devido aos cálculos da imagem integral.

Ressaltando que todos os tamanhos de *templates* foram testados em imagens de folhas em ambiente controlado, contendo uma folha por imagem e com o fundo branco. Em ambiente real, como há uma diversidade de tamanho de folhas, decidimos testar uma técnica que utilizasse um ajuste de acordo com a área a ser percorrida.

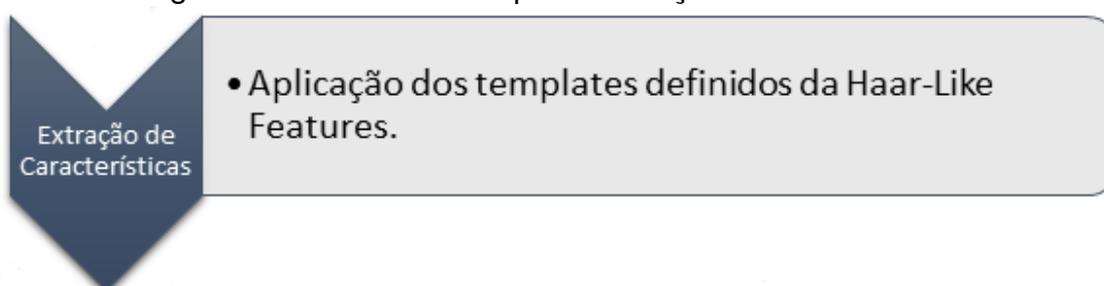
Tabela 2 - Acurácia relacionada com o tamanho do template.

Tamanho do Template	Acurácia
Tamanho 4	90%
Tamanho 3	95%
Tamanho 2	95%

Fonte: Elaboração própria.

Esta etapa, utilizando os templates Haar-Like Features, está compreendida na fase de extração de características conforme Figura 31

Figura 31 - Síntese da etapa de extração de características.



Fonte: Fonte: Elaboração própria.

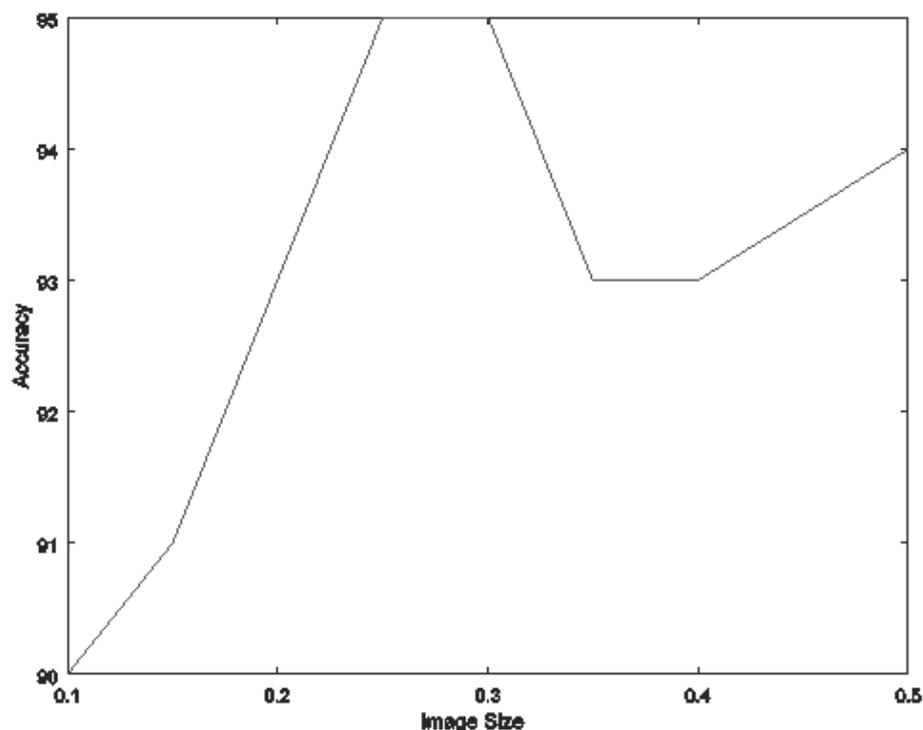
7.6 Tamanho da Imagem

A Figura 32 mostra o gráfico obtido nos testes realizados. Neste, é possível observar que é possível obter uma acurácia de 95% utilizando 30% ou 25% do tamanho relativo a imagem original. Como a quantidade de pixels na imagem influencia no tempo final do processamento dos *templates*, escolhemos trabalhar com 25% do tamanho da imagem original.

7.7 Taxa de acerto com as imagens rotacionadas

Após definido o tamanho da imagem a ser utilizado, foi verificado como a técnica proposta se comporta nas imagens com diferentes rotações, pois no ambiente real as folhas se encontram em várias disposições. Para isso, cada imagem de monocotiledônea foi rotacionada em 6° de diferença em relação a outra e as dicotiledôneas em 5°. Então, ao aplicar novamente o algoritmo nas imagens rotacionadas, a taxa de acerto obtida foi a mesma para a imagem reduzida a 25% do tamanho original, que é 95%.

Figura 32 - Relação entre o tamanho da imagem utilizada e a acurácia.



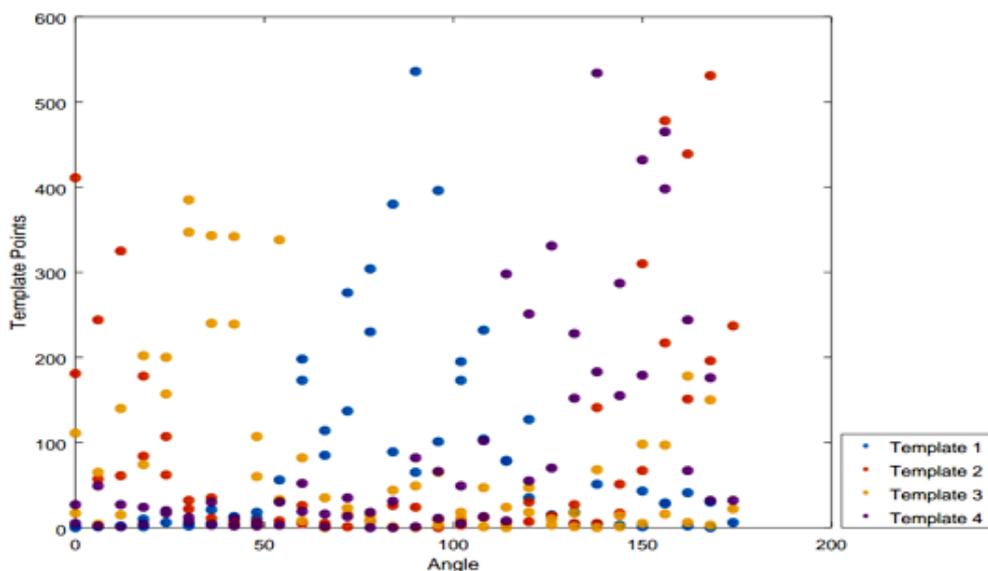
Fonte: Elaboração própria.

Desta forma, a técnica utilizada independe da rotação da folha, o que é desejável para aplicação em um ambiente real. Liu et al. (2018), utilizando uma rede neural convolucional para classificação de plantas e também utilizando o Flavia Dataset, ressalta que a técnica por eles utilizada não é invariante a rotação, sendo que os autores realizam a rotação das imagens inserindo novas imagens ao conjunto de dados de forma a cobrir um maior espaço de rotações possível.

Os gráficos das Figuras 33 e 34 mostram a quantidade de pontos encontrados para cada um dos 4 *templates*, em cada imagem, cuja venação central está posicionada em um determinado ângulo. Percebe-se que no gráfico da Figura 33, que se refere as imagens de monocotiledôneas, um determinado *template* se destaca em um ângulo. Já para dicotiledôneas, mostrado Figura 34, não há a prevalência de um determinado *template*. O resultado está de acordo com o esperado, pois como na monocotiledônea a venação é paralela, há predominância de um *template*. Com a mudança na rotação o *template* que corresponde ao padrão da venação muda, mas continua a predominância de um deles. Por exemplo, no intervalo que corresponde

as imagens rotacionadas com o ângulo entre 50 e 100 graus há a predominância do *template 1* (azul), que procura as linhas verticais na folha. Como nesse intervalo a veias da folha ficam próximas de 90° , é natural que o *template* que encontre as linhas verticais tenha mais ocorrência. Assim acontece com os outros *templates* na monocotiledônea, onde as correspondências aumentam de acordo com a proximidade entre a forma do *template* e o ângulo da folha.

Figura 33 - Quantidade de correspondências encontradas de cada *template* em ângulos diferentes nas folhas monocotiledôneas.



Fonte: Elaboração própria.

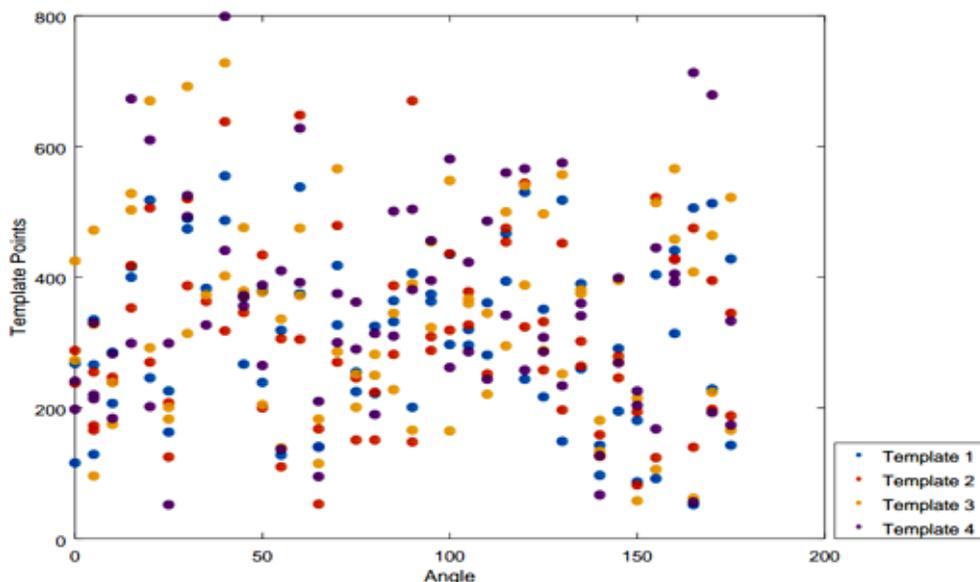
Com base nisso, para identificar uma dicotiledônea é necessário encontrar uma quantidade de ocorrência em uma região de mais de um *template*.

7.8 Desempenho da metodologia

O tempo médio que essa técnica leva para processar cada imagem contida no Flavia Dataset, utilizando o tamanho de *template* e o tamanho da imagem que adotamos como padrão, é de 75 milissegundos em média. Para essas simulações foi utilizado um computador com o processador Intel Core i5-2400 3.10GHz e com 6 GB de memória RAM.

A acurácia da técnica desenvolvida é comparada com outras para reconhecimento de uma determinada espécie utilizando a folha e que utilizaram o Flavia Dataset para

Figura 34 - Quantidade de correspondências encontradas de cada template em ângulos diferentes nas folhas dicotiledôneas.



Fonte: Elaboração própria.

treinamento e validação do modelo, conforme a Tabela 3. O resultado da Tabela 3 foi obtido utilizando como entrada imagens sem edição retiradas da Flavia Dataset.

Após os testes utilizando imagens com uma única folha e retiradas em um ambiente controlado, passamos para os testes com imagens retiradas em um ambiente não controlado, e com várias folhas na mesma imagem, o que implica em aumento na complexidade do algoritmo. Nas próximas seções serão apresentados os passos realizados para identificação das plantas em um ambiente não controlado. O objetivo é identificar cada folha de dicotiledônea, a erva daninha alvo, existente na imagem.

Tabela 3 - Comparação entre técnicas para reconhecimento de folhas, em ambiente controlado.

Técnica	Acurácia (%)
Wang et al. (2018)	98,53
Raju e Kumar (2016)	94,37
Muthevi e Uppu (2017)	84,78
Este trabalho	95,0

Fonte: Elaboração própria

7.9 Dataset de imagens em ambiente não controlado

Foi realizada visita a um campo de cultivo de cana-de-açúcar orgânica pertencente a Usina São Francisco, localizada em Sertãozinho, São Paulo, latitude -21.174470982176377, longitude -48.118241520597955, retirando imagens de plantas que estavam infestadas pela erva daninha em questão. As imagens foram retiradas em áreas de primeira, segunda e terceira soca, que são os ciclos de plantio da cana-de-açúcar. Quando é plantada a primeira vez em um área temos a primeira soca, depois da colheita há o replantio o que seria a segunda soca e depois ocorre a terceira. Fomos em duas épocas do ano, no mês de março e novembro. As imagens foram retiradas em diferentes horários no período diurno, desde das 8 horas ao meio dia, e também das 15 horas as 17. Destas imagens, foram selecionadas 146 imagens para testar e aprimorar o algoritmo que desenvolvemos. Essas imagens possuem tamanho de 2448 x 3264, com 300 dpi e foram retiradas com uma câmera Nikon Coopix de 16 Mega Pixels. O objetivo é testar o reconhecimento do máximo de folhas nas imagens obtidas. Outras etapas foram adicionadas de forma a tratar as novas demandas advindas das características de um ambiente não controlado.

O primeiro passo foi definir uma estratégia para varrer a imagem e analisar as áreas para definir se contém uma folha de uma dicotiledônea ou não. Na próxima seção será descrito o processo de varredura utilizado para analisar as imagens.

7.10 Varredura da imagem

Vale ressaltar que nossa pesquisa busca a definição de uma técnica que possa ser executada em tempo real. Nesse sentido, entendemos que reduzir o espaço de busca é uma estratégia que pode contribuir para alcançar esse objetivo. Como pode ser visto na Figura 35, há áreas na imagem que são compostas, em sua maioria, pelo solo ou pelo céu. Dado que usamos mais de 4 *templates* para varrer a imagem, um conjunto de pixels será acessado pelo menos 4 vezes e operações aritméticas serão feitas baseadas nesses valores, há uma diminuição no processamento se definirmos áreas que não precisariam ser analisadas por não conter pixels verdes na mesma.

A estratégia aplicada foi dividir a imagem em quadrados e analisar o percentual de

Figura 35 - Erva Daninha e cana-de-açúcar em ambiente real.



Fonte: Elaboração própria.

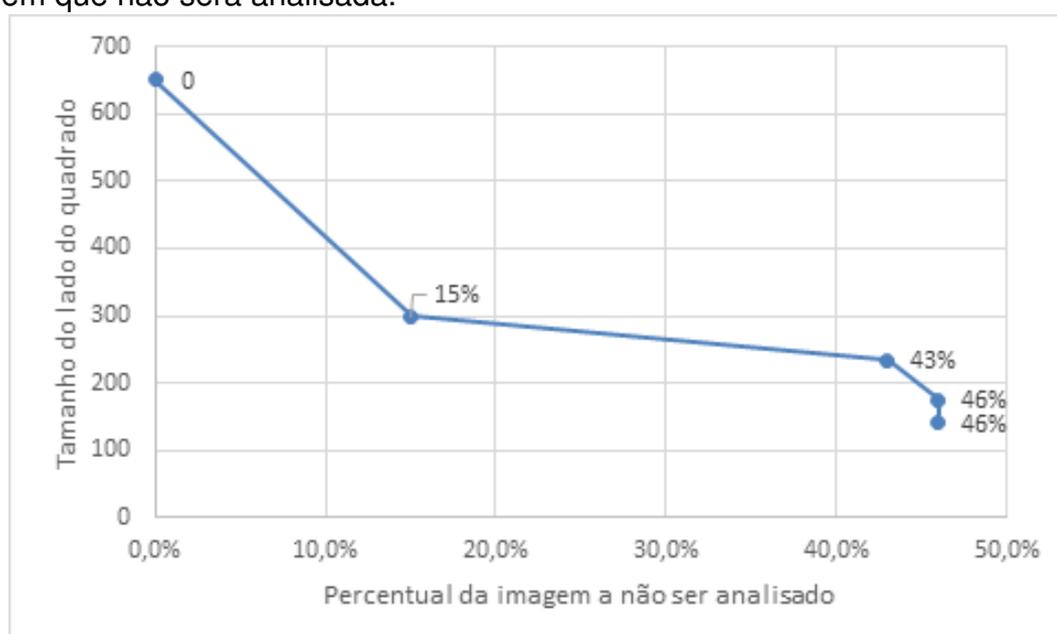
área verde em cada um deles. Para isso, usando o espaço de cores RGB, os valores dos três componentes em cada pixel são verificados e se o valor do componente verde for maior do que o componente azul e também maior pelo menos do que 90% do valor do componente vermelho, o pixel é definido como pertencente a área de interesse a ser analisada. Esses valores foram definidos analisando os valores dos pixels nas imagens.

Fizemos testes para definir a melhor largura do quadrado necessário para dividir a imagem. No gráfico apresentado na Figura 36 são mostrados os resultados obtidos. Em uma imagem de 1400 x 3200 pixels, utilizamos a largura como referência. Primeiramente, testamos se a imagem fosse dividida em blocos de 650 x 650 pixels, qual o percentual da imagem não precisaria ser analisado. Nesse primeiro teste, levamos em considerações as áreas que não possuem nenhum pixel verde. Em seguida, os

valores dos quadrados foram reduzidos. Ao utilizarmos os quadrados com a largura de 175 e 140 obtivemos o mesmo resultado que foi uma redução em 46%, no melhor caso, no espaço a ser verificado na imagem. Como quadrados maiores permitem a utilização de *templates* maiores para analisar a área e, quanto maior o *template*, graças a utilização do conceito de imagem integral, o tempo de processamento também é reduzido, optamos por utilizar a largura dos quadrados em 175 pixels. A Figura 36 apresenta o gráfico dessa relação entre o tamanho do quadrado definido e o percentual da área que não será analisada. Na Figura 37 temos o exemplo de uma imagem dividida em quadrados e o percentual de pixels verdes em cada subdivisão.

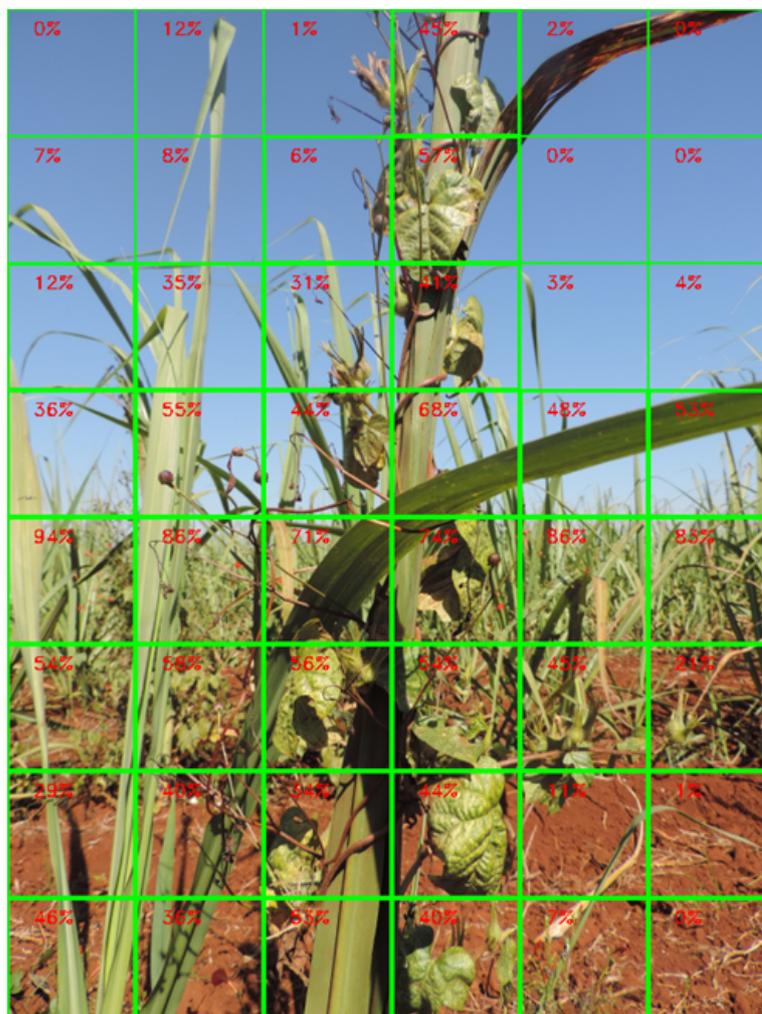
Definido o tamanho da largura do quadrado a ser utilizado, buscamos definir o tamanho de cada *template* a ser utilizado em cada área, a depender do percentual de área verde no mesmo. Como mencionado anteriormente, quanto maior o tamanho do *template* a utilização do conceito de imagem integral auxilia na redução no tempo de processamento. É esperado que quanto maior a área verde, maior a venação existente nessa área. Conseqüentemente, podemos utilizar um tamanho de *template* maior. Outro aspecto importante a ressaltar é que a utilização de um *template* maior em uma área cuja venação possua um tamanho menor ao referido *template* pode implicar em uma diminuição na acurácia, pela falta de correspondências.

Figura 36 - Gráfico que relaciona a área em pixels do quadrado e o percentual da imagem que não será analisada.



Fonte: Elaboração própria.

Figura 37 - Imagem dividida em quadrados mostrando o percentual de verde em cada um deles.

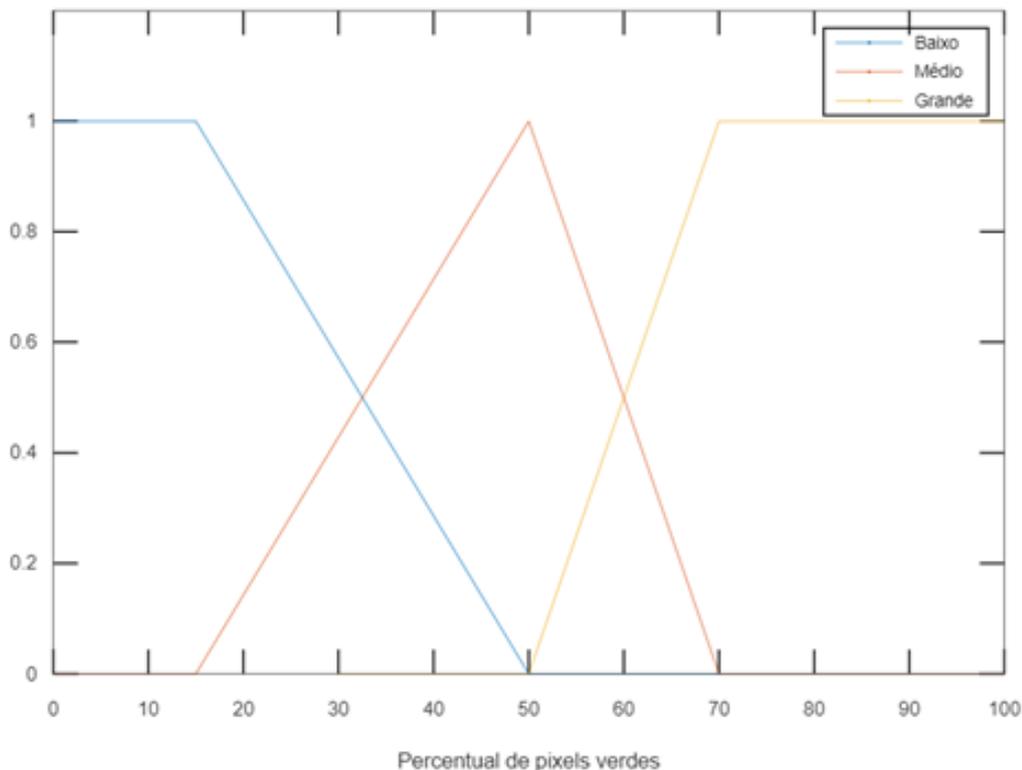


Fonte: Elaboração própria.

Após definida a forma de varredura da imagem, iniciamos a definição dos componentes de um sistema com lógica Fuzzy. Primeiramente, fizemos testes a fim de definir a função de pertencimento entre cada variável escolhida. As Figuras 38, 39 e 40 mostram a função triangular das variáveis escolhidas.

No gráfico da Figura 38 é exibida a função triangular para a classificação de uma determinada área em relação a porcentagem de pixels verdes existentes na mesma. A área pode ter um percentual de pixels considerado baixo, médio ou grande, sendo estes os valores linguísticos possíveis. Por exemplo, se uma área possuir 10% de pixels verdes, observando o gráfico notamos que $f_V(10) = 1$. Caso a área possua 20% de pixels verdes a $f_V(20) = 0.1$ para médio, $f_V(20) = 0.9$ para baixo e $f_V(20) = 0$ para grande.

Figura 38 - Gráfico da função de pertinência referente ao percentual de pixels verdes existentes em uma determinada área da folha.



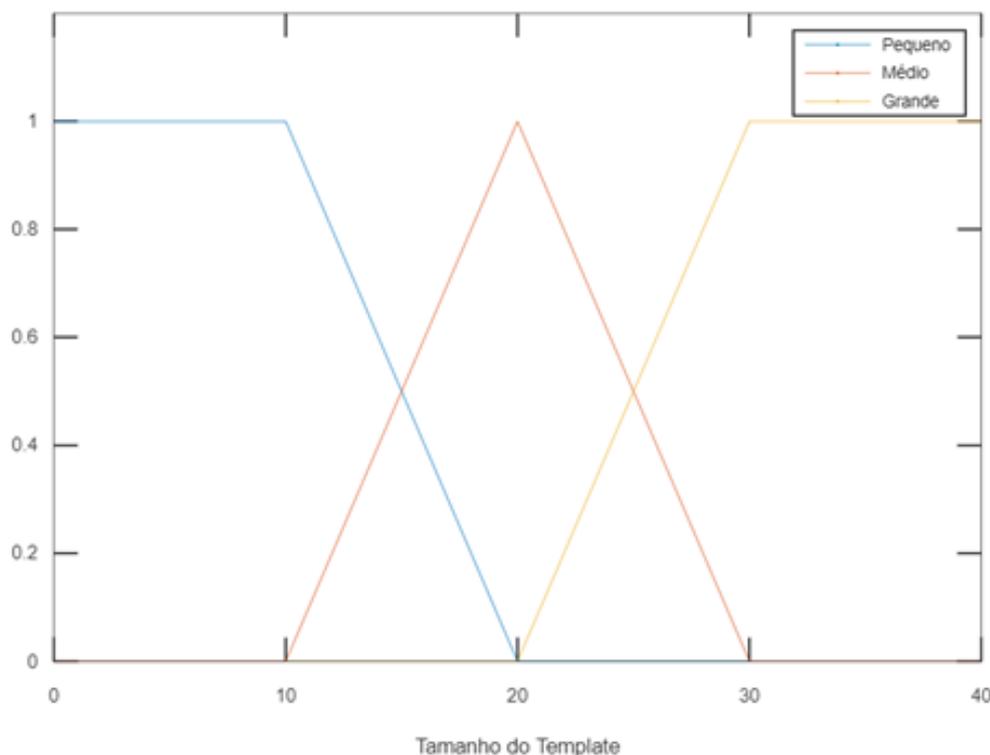
Fonte: Elaboração própria.

Já no gráfico da Figura 39 é mostrado a função triangular referente a variável linguística tamanho do *template*. Para essa variável, há três valores linguísticos possíveis: pequeno, médio e grande. Por exemplo, caso um *template* tenha o tamanho 20 temos que $f_T(25) = 0,5$ para grande, $f_T(25) = 0,5$ para médio e $f_T(25) = 0$ para pequeno.

No gráfico da Figura 40, temos a variável ocorrência de cruzamentos entre *templates*. Como a venação das monocotiledôneas é paralela, em uma imagem de uma folha em um ambiente controlado haveria somente a ocorrência de um *template*, sem cruzamentos. No ambiente real porém, como há sobreposição de folhas, pode aparecer cruzamentos mesmo em ocorrência de folhas monocotiledôneas. Por este motivo o valor da função de pertinência quando o total de cruzamentos é 1 resulta em 1 na classificação de monocotiledônea e $f_C(1) = 0$ para dicotiledônea.

Na fase de inferência, primeiramente obtemos a quantidade de pixels verde de uma determinada área. A depender do valor da função de pertinência é utilizado o respectivo tamanho de *template*. Por exemplo, se o valor de pixels verde for 60%, a

Figura 39 - Gráfico da função de pertinência referente ao tamanho do template a ser utilizado.



Fonte: Elaboração própria.

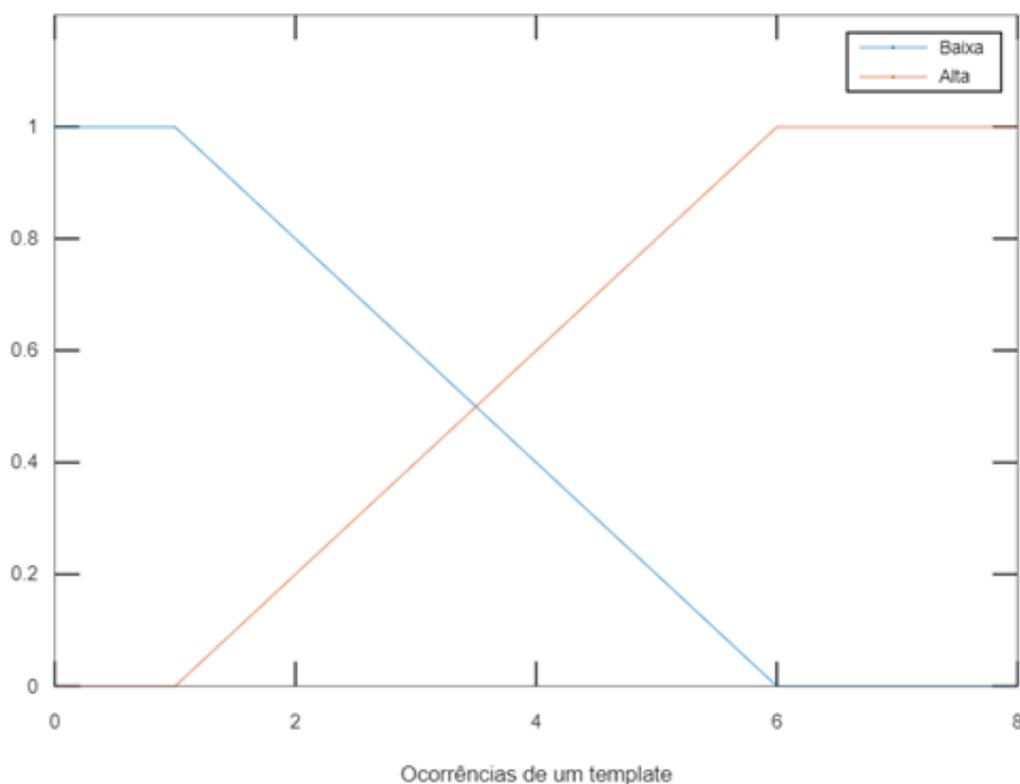
função de pertinência é igual a $f_V(60) = 0,5$ para médio, $f_V(60) = 0,5$ para grande e $f_V(60) = 0$ para pequeno. Dessa forma, utilizaremos o tamanho de *template* 25 que possui os seguintes valores da função de pertinência: $f_T(25) = 0,5$ para grande, $f_T(25) = 0,5$ para médio e $f_T(25) = 0$ para pequeno.

O último gráfico referente a lógica fuzzy adotada nesta pesquisa, está na Figura 40, e leva em consideração a variável linguística ocorrências de um *template*. Os possíveis valores da variável linguística são: baixa ou alta. Em relação a função de pertinência, caso haja a correspondência de um mesmo *template* 5 vezes em uma determinada área, a função de pertinência assume os seguintes valores: $f_C(5) = 0,4$ correspondente á baixa ocorrência e $f_C(5) = 0,6$ para alta ocorrência.

A regra de inferência lógica definida é que se forem encontradas ocorrências altas de pelo menos dois *templates* diferentes em uma área, então assumimos que essa área pertence a uma folha monocotiledônea, caso contrário a uma folha dicotiledônea.

Foi realizado um teste para avaliar o comportamento da técnica proposta frente á

Figura 40 - Gráfico da função de pertinência referente ao tipo da folha a qual pertence uma determina área.



Fonte: Elaboração própria.

oclusão foliar. No ambiente real é comum que a área de uma folha não seja capturada em sua totalidade, seja por estar oclusa por outras folhas ou por outro material como a argila, solo ou poeira, ou pelo ângulo em que é capturada em relação a sua curvatura. Nesse sentido, escolhemos uma folha que fosse totalmente capturada e editamos a imagem com certas porcentagens dela, de acordo com a imagem. Na imagem a folha foi segmentada com 10%, 25%, 50% e 100% do tamanho original, conforme mostrado na Figura 41.

Foi realizada a identificação dos elementos conectados que possuem intersecções com dois gabaritos. O gráfico mostrado na Figura 43 mostra o maior número de intersecções obtidas em cada porcentagem da folha da imagem apresentada na Figura 41. Analisando o gráfico, percebe-se que com 50% da folha são encontrados mais de 20 cruzamentos. Com este número de cruzamentos é possível afirmar o reconhecimento de uma folha dicotiledônea com alta precisão. Isso é importante para mostrar a viabilidade de utilização da técnica em ambientes reais, independente da ocorrência

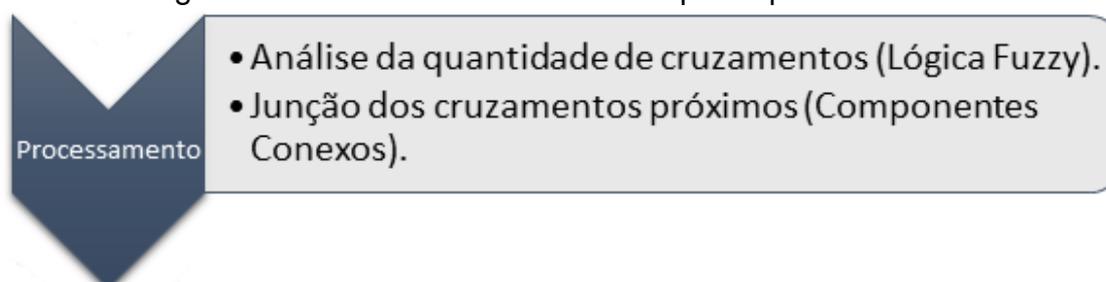
Figura 41 - Folha segmentada artificialmente para simular a oclusão.



Fonte: Elaboração própria.

de folhas oclusas. É importante destacar que, dependendo da porcentagem de folha exposta, é mais provável identificá-la corretamente. Essa etapa finaliza a etapa de processamento, Figura 42.

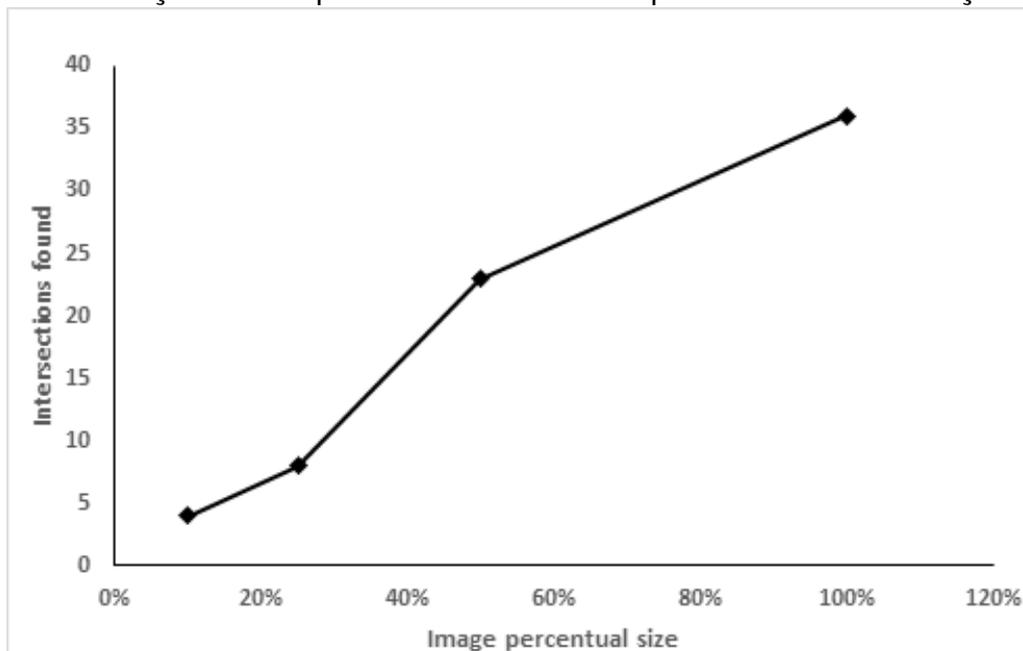
Figura 42 - Técnicas utilizadas na etapa de processamento.



Fonte: Fonte: Elaboração própria.

A Figura 44b mostra a saída final depois da utilização de componentes conexos de forma a agrupar as ocorrências de mais de um *template* que estão próximas. Na imagem, quanto mais ocorrências de *templates* ocorrem próximas, mais vermelho é o quadrado que circunda a região. Isto torna possível a seleção da folha mais provável de ser uma dicotiledônea na imagem.

Figura 43 - Relação entre o percentual da folha e a quantidade de intersecções obtidas.



Fonte: Elaboração própria.

Em condições como apresentadas na Figura 44a, com percentual baixo de folhas sombreadas, a taxa de acerto é cerca de 90,0%. Porém, em outras situações, com percentual alto de sombreamento a taxa de acerto cai para 55,0%. Por fim, de 146 imagens com e sem dicotiledôneas, o algoritmo proposto reconheceu 119, possuindo uma taxa de acerto média de 81,6%. A Tabela 4 mostra o comparativo com outras técnicas utilizadas para reconhecimento de plantas em tempo real.

Para processar cada imagem foi utilizado um computador com processador core i7-8565U com 1,8 GHz de frequência de clock, com 16 GB de memória RAM DDR4. O sistema operacional utilizado foi o Linux Ubuntu 20.04. O tempo médio de processamento de cada imagem foi de 0,146 segundo ou 146 milissegundos. O tempo gasto no pré-processamento são aproximadamente 0,108 segundo ou 108 milissegundos. Para o processo de binarização e demais técnicas para detectar o padrão da venação são utilizados em média 0,038 segundos ou 38 milissegundos.

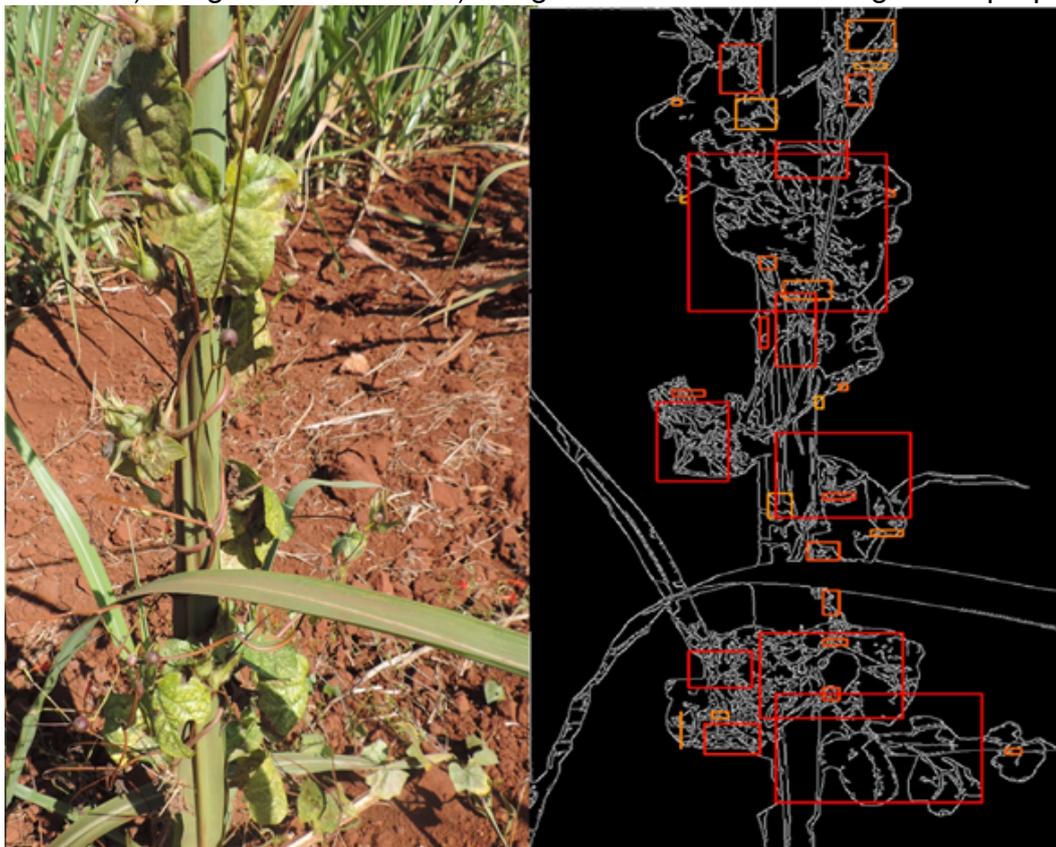
No próximo capítulo, discorreremos acerca da influência do sombreamento no percentual de reconhecimento da técnica proposta neste trabalho.

Em relação a complexidade do algoritmo proposto, levando em consideração que o número de pixels em uma imagem corresponde a n , a proposta calcula a imagem integral que está diretamente relacionada ao número de pixels na imagem, o que

Tabela 4 - Comparação entre técnicas para reconhecimento de folhas em ambiente real.

Técnica	Acurácia (%)	Tempo de Processamento	Hardware
Chavan e Nandedkar (2018)	93,64	970ms	Processador Intel Xeon E52695 v4 (64 GB RAM) e NVIDIA TITAN Xp GPU (12 GB RAM)
Dyrmann, Christiansen e Midtiby (2017)	96,7	143ms	Lenovo Thinkpad X200, equipado com um Processador 2.4GHz Intel Core 2 Duo
Lottes et al. (2018)	85.5	99,1	200 a 800 ms
Ashraf e Khan (2020)	86,0	30,965ms	GeForce GTX 980 GPU Computador com 4 GB Random Access Memory (RAM), Processador 1.70 GHz with core i5-3317 CPU
Sujaritha et al. (2017)	92,9	20ms	Raspberry Pi 2 (Model B+)
Este trabalho	81,6	146ms	Computador com core i7-8565U, 1,8 GHz de frequência de clock, 16 GB of RAM DDR4

Figura 44 - a) Imagem de entrada b) Imagem de saída final do algoritmo proposto.



Fonte: Elaboração própria.

resultaria em uma complexidade $O(n)$. Em seguida, são realizadas as operações relacionadas aos 4 *templates*. Essas operações dependem do tamanho do *template* utilizado, mas como este utiliza regiões e cada operação considera duas regiões no *template*, são realizadas 8 operações. Para cada *template* aplicado a toda a imagem teríamos $O(8.(n)/t)$, e para os quatro modelos teríamos $O(4.8.(n/t))$, o que é igual a $O(32(n/t))$, onde t é o tamanho do modelo. No processo de contar pixels verdes para dividir por regiões, o número de operações depende do número de pixels, então a complexidade seria $O(n)$. Teríamos a complexidade total como $O(n + 32(n/t) + n)$, ou simplesmente $O(n)$.

Como Sujaritha et al. 2017 é considerado o mais comparável a este trabalho que pudemos encontrar, analisamos a complexidade do algoritmo. Os autores usam quatro *features* e um classificador fuzzy no processo de detecção de ervas daninhas. As *features* utilizadas são: matriz de co-ocorrência de nível de cinza de segunda ordem normalizada, recursos de textura de Laws, Gabor *wavelet* e rotação- método robusto de decomposição *wavelet*. O Gabor *wavelet* tem uma complexidade de $O(n.\log(n))$ (QIU;

ZHOU; CRANDALL, 1999) que é maior que a complexidade do algoritmo proposto nesta pesquisa. Da mesma forma, Redes Neurais Convolucionais, a tendência atual em reconhecimento de imagem, tem uma complexidade de $O(n \cdot \log(n))$ para inferência, considerando a implementação de menor complexidade que pudemos encontrar (PARK; LEE; SIM, 2020). A complexidade da fase de treinamento varia de acordo com o algoritmo utilizado, mas não deve ser considerada em uma comparação justa.

8 DISCUSSÕES

Nesse capítulo, apresentaremos alguns aspectos que entendemos ser importante ressaltarmos com base os testes que realizados e todo o trabalho desenvolvido.

Um aspecto que inicialmente destacamos é a importância de aplicarmos técnicas de pré-processamento as imagens a serem analisadas para que o padrão de venação seja ressaltado, e também o espaço de busca diminuído utilizando um algoritmo de redução da imagem. Como alguns artigos descritos nesse documento apresentam a transformada de Hough para obtenção do padrão de venação em folhas, decidimos utilizá-la. O tempo de processamento da técnica se mostrou elevado por conta da proposta da técnica de usar, resumidamente, senoides em diversas direções e verificar a que cruza mais pontos. O Haar-like *features* faz uma análise menos extensiva para capturar os padrões, dependendo do tipo de *template* que será usado.

O tamanho e o aspecto do *template* são fatores que influenciam o desempenho da técnica. O formato do *template* deve ser capaz de capturar o padrão pré estabelecido. O tamanho também pode ocasionar dificuldades no reconhecimento. Um *template* com tamanho muito pequeno pode até capturar mais padrões, mas com impacto no tempo de processamento. Por isso, utilizamos a lógica fuzzy de forma a ajustar o tamanho do *template* com a imagem a ser aplicada.

É importante salientar que é necessário utilizar um valor abaixo de 255 como parâmetro (Equação 6) ao avaliar a diferença entre as médias das regiões com pixels pretos e brancos. É de se esperar, sobretudo em ambientes não controlados, que a ocorrência das veias não sejam lineares, por isso a verificação de correspondência utilizando um valor menor auxilia na melhoria da eficácia. Testando empiricamente, encontramos o valor de 160 para *templates* com tamanho grande e 200 para tamanhos pequenos. Tais valores são guias para problemas similares, mas devem ser ajustados caso a aplicação mude.

Além de capturar o padrão da folha é importante notar o desempenho da técnica considerando determinadas condições inerentes ao ambiente real. A verificação se a técnica consegue capturar os padrões da venação de forma invariante ao angulo em

que a folha se encontra é de suma importância. A técnica apresentada nesse trabalho se mostrou eficiente nesse aspecto.

Outro aspecto importante é a aplicação da técnica tendo em vista a possibilidade de oclusão. Como a folha da *Ipomea* sp. possui simetria bilateral, pressupondo que não houve algum dano a folha, se a maioria da folha estiver visível, a técnica aqui descrita é capaz de realizar o reconhecimento.

Analisando as imagens é possível ver o efeito do sombreadamento na etapa de limiarização para obtenção da venação. Nota-se na Figura 45 que no local em que há incidências de sombras os padrões de venação das folhas praticamente desaparecem. Na mesma figura o sombreadamento não é completo e o algoritmo ainda encontra algumas correspondências de *templates*, porém a taxa de acerto fica em 55,5% das folhas em condições como esta, em que a folha está em sua maioria sombreada.

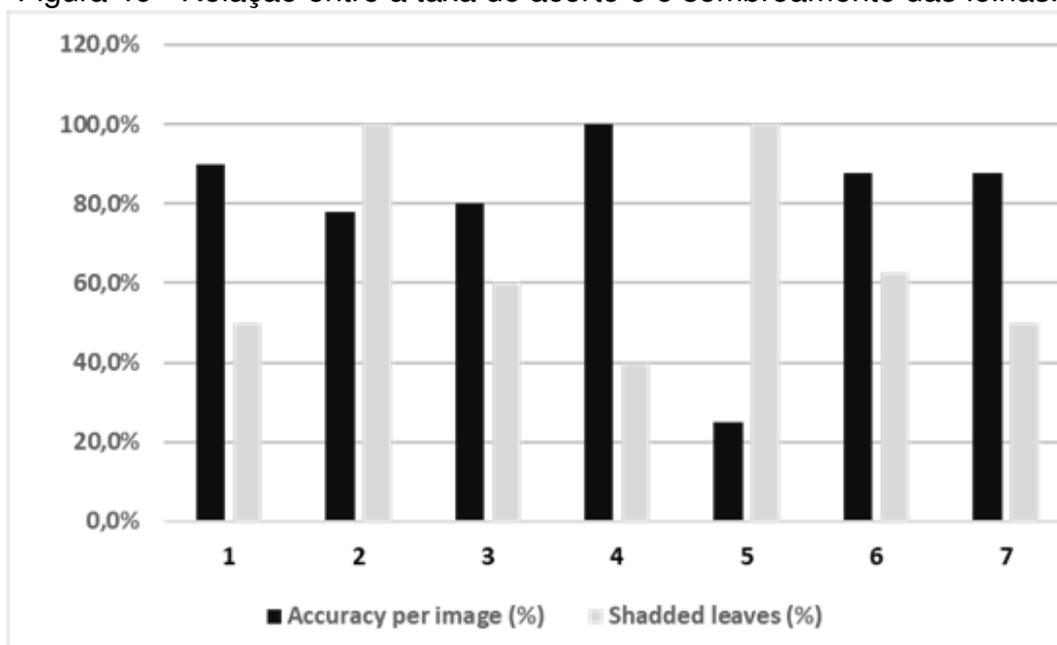
Figura 45 - a) Imagem de entrada com maioria da área das folhas sombreadas. b) Imagem de saída após a aplicação da técnica.



Fonte: Elaboração própria.

A Figura 46 mostra a relação a quantidade de folhas sombreadas e a taxa de acerto. A barra referente a imagem 2 e a 5, que são as imagens com maior número de folhas sombreadas total ou parcialmente, apresenta o menor valor em relação a taxa de acerto. Por outro lado, a barra referente a imagem 4 possui o maior valor de acurácia e apresenta o menor número de folhas sombreadas. Assim, quanto menor for a incidência de luz na folha menor será a taxa de reconhecimento obtida.

Figura 46 - Relação entre a taxa de acerto e o sombreamento das folhas.



Fonte: Elaboração própria.

Pudemos verificar o desempenho do algoritmo sobre algumas folhas que foram danificadas pelo frio, pois na noite anterior à coleta das imagens ocorreu uma geada. As folhas danificadas geralmente apresentam partes mais amareladas ou pouco mais enrugadas, mas não apresentaram dificuldade no processo de reconhecimento como as folhas sombreadas.

Para reduzir os efeitos relativos ao sombreamento, testamos uma solução usando o conceito de Superpixels, onde possíveis áreas de uma imagem agrupam pixels com características semelhantes através da luminosidade de cada Superpixels, que é obtida através da mudança da imagem para o espaço de cores YCrCb, onde o Y representa a luminância, tentando ajustar os Superpixels com menor luminância com os Superpixels mais próximos. Porém, esta solução obviamente ocasiona um aumento no tempo de processamento de uma imagem. Além disso, após a aplicação da técnica a venação aparece mais uniformizada com o restante da folha, o que dificulta a limiarização. Assim, optamos por posteriormente avaliar uma solução que utilize uma forma externa de iluminação ou utilização de *flash* para captura das imagens.

Com relação ao tempo de processamento, é interessante notar como os passos na fase de pré-processamento consomem mais tempo do que o conjunto de técnicas utilizadas para identificação do padrão de venação e classificação como erva daninha

ou não. Isso ressalta o bom desempenho que o padrão discriminativo possui para realização da tarefa.

Um aspecto importante a ser discutido é a utilização da técnica proposta para mapeamento de ocorrência espacial das ervas daninhas trepadeiras referidas no trabalho. Papadopoulos et al. (2018) ressaltam que a ocorrência das ervas daninhas em uma plantação não ocorre de forma homogênea, mas há a concentração da mesma em determinadas áreas. Shiratsuchi (2001) acrescenta que essas áreas possuem diversas densidades e que são conhecidas como reboleiras. Tanto Papadopoulos et al. (2018) como Shiratsuchi (2001) reconhecem que é o importante o mapeamento dessas áreas para o melhor controle dessas ervas, seja de forma manual ou utilizando um produto químico. Assim, nossa proposta pode ser utilizada em um artefato para identificar essas áreas de concentração da erva daninha, possibilitando a melhor utilização de recursos para combate dessas plantas.

No geral, entendemos que os testes mostraram que a técnica tem bom potencial de aplicação em tempo real para reconhecimento de ervas daninhas em plantações de cana-de-açúcar, apesar da dificuldade apresentada em situações com a maioria das folhas sombreadas.

9 CONCLUSÃO

O reconhecimento em ambiente real da *Ipomoea* sp. levando em consideração determinadas adversidades como sobreposição, sombreamento e diferentes tamanhos das folhas é uma tarefa complexa, sobretudo quando se deseja realizar o processamento de imagens em tempo real. Como a cana de açúcar é monocotiledônea e a erva daninha dicotiledônea, a utilização do padrão de venação para distinção entre as plantas, dada que uma é paralela e a outra reticulada, se mostrou adequado para distinção entre as mesmas.

Para obtenção do padrão de venação das folhas, a utilização de padrões discriminativos como o Haar-like *features* apresenta resultados que mostram a viabilidade do uso da técnica proposta neste trabalho para o reconhecimento de uma erva dicotiledônea na plantação de cana de açúcar. A utilização do conceito de Imagem Integral auxilia na redução do tempo de processamento e aplicação para reconhecimento em tempo real. Outro aspecto importante, é que não é necessária uma base de dados extensa para o processo de treino e teste de um padrão discriminativo, resumindo-se a ajuste de parâmetros que pouco variam para a aplicações similares.

A técnica proposta se mostra invariante a rotação, o que é desejado em ambientes reais, e também invariante ao tamanho da folha, que também é outro desafio para reconhecimento já que as folhas podem ter tamanhos variados. Outra vantagem também é o reconhecimento, segundo os testes realizados, de folhas com partes oclusas ou danificadas, sendo possível o reconhecimento de folhas com até 50% de sua área aparente na imagem. Um aspecto importante encontrado é a influência da intensidade da luz incidente nas folhas, pois se houver uma parte considerável da folha sombreada (maior que 50%), a taxa de reconhecimento é diretamente afetada.

Outro aspecto importante é a utilização da lógica fuzzy de forma a definir termos para classificar se uma determinada área possui uma taxa de pixels verde baixa, média ou alta. A mesma lógica também pode ser usada para definir o tamanho de um *template* de acordo com a área de pixels verdes e depois avaliar, utilizando a quantidade de correspondências obtidas, onde há folhas de dicotiledôneas e monocotiledô-

neas na imagem.

A solução proposta ainda foi testada com imagens que apresentam duas espécies de plantas: dicotiledôneas e monocotiledôneas. Há soluções da literatura para reconhecimento de plantas em ambientes reais com desempenho similar a este trabalho, porém assumindo somente a ocorrência de uma espécie de planta por imagem. A utilização do conceito de componentes conexos é importante para identificação de diversas folhas de dicotiledônea em uma mesma imagem.

9.1 Contribuições do trabalho

Utilização de uma técnica que utiliza padrões discriminativos para identificação da venação de folhas. Definição dos *templates* necessários para encontrar o padrão das folhas.

Utilização da lógica fuzzy para definição, de maneira automática, do tamanho adequado do *template* e classificação da folha. Utilização do conceito de componentes conexos para agrupar áreas que pertençam a uma mesma folha.

Integração dessas três técnicas para reconhecimento de uma erva daninha em ambiente real.

9.2 Trabalhos futuros

Realização de testes utilizando um artefato ou um robô para avaliar de forma mais ampla o desempenho da técnica proposta.

Investigar a utilização de alguma artefato físico para diminuir a influência do sombreamento no reconhecimento das veias.

Investigar uma forma de realizar a remoção da planta, dado que foi realizado a identificação da ocorrência da erva daninha. As técnicas atuais da literatura envolvem deposição de químicos pontual, corte mecânico ou choque elétrico com o intuito de danificar o caule.

Por último, avaliar a execução do algoritmo desenvolvido utilizando GPU, computação paralela ou ambas.

REFERÊNCIAS

- AHDEROM, S.; APOPEI, B. A novel method for detecting morphologically similar crops and weeds based on the combination of contour vi nguyen thanh le. *Oxford University Press*, p. 1–16, 2020.
- AMEND, S. et al. Weed Management of the Future. *KI - Künstliche Intelligenz*, Springer Berlin Heidelberg, v. 33, n. 4, p. 411–415, 2019. ISSN 1610-1987. Disponível em: <<https://doi.org/10.1007/s13218-019-00617-x>>.
- AMLEKAR, M. M. et al. Leaf shape extraction for plant classification. In: . [S.l.: s.n.], 2015. p. 1–4.
- ANDRES, A. et al. Predictions for weed resistance to herbicides in brazil: A botanical approach. *Herbicide Resistance in Weeds and Crops*, 2017.
- ASHRAF, T.; KHAN, Y. N. Weed density classification in rice crop using computer vision. *Computers and Electronics in Agriculture*, Elsevier, v. 175, n. February, p. 105590, 2020. ISSN 0168-1699. Disponível em: <<https://doi.org/10.1016/j.compag.2020.105590>>.
- BAKHSIPOUR, A. et al. Weed segmentation using texture features extracted from wavelet sub-images. *Biosystems Engineering*, Elsevier Ltd, v. 157, p. 1–12, 2017. ISSN 15375110. Disponível em: <<http://dx.doi.org/10.1016/j.biosystemseng.2017.02.002>>.
- BATAINEH, B. et al. Expert Systems with Applications A novel statistical feature extraction method for textual images : Optical font recognition. *Expert Systems With Applications*, Elsevier Ltd, v. 39, n. 5, p. 5470–5477, 2012. ISSN 0957-4174. Disponível em: <<http://dx.doi.org/10.1016/j.eswa.2011.11.078>>.
- BAWDEN, O. et al. Robot for weed species plant-specific management. n. March 2016, p. 1179–1199, 2017.
- BHULLAR, M. S. et al. Control of Morningglories (Ipomoea spp.) in Sugarcane (Saccharum spp.) . *Weed Technology*, v. 26, n. 1, p. 77–82, 2012.
- CAO, Y. et al. Leaf Vein Extraction and Angle Measurement Using Hue Information and Line Detection. In: *2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*. [S.l.: s.n.], 2017. v. 1, p. 223–225.
- CHAMP, J. et al. Instance segmentation for the fine detection of crop and weed plants by precision agricultural robots. *Applications in Plant Sciences*, v. 8, n. 7, p. 1–10, 2020. ISSN 15372197.
- CHAVAN, T. R.; NANDEDKAR, A. V. AgroAVNET for crops and weeds classification : A step forward in automatic farming. *Computers and Electronics in Agriculture*, Elsevier, v. 154, n. April, p. 361–372, 2018. ISSN 0168-1699. Disponível em: <<https://doi.org/10.1016/j.compag.2018.09.021>>.

CHENG, G. et al. Learning Rotation-Invariant and Fisher Discriminative Convolutional Neural Networks for Object Detection. *IEEE Transactions on Image Processing*, IEEE, v. 28, n. 1, p. 265–278, 2019.

CODIZAR, A. L.; SOLANO, G. Plant Leaf Recognition by Venation and Shape Using Artificial Neural Networks. 2016.

DUDA, R. O.; HART, P. E. Use of the hough transformation to detect lines and curves in pictures. v. 15, n. 1, 1972.

DYRMANN, M.; CHRISTIANSEN, P.; MIDTIBY, H. S. Estimation of plant species by classifying plants and leaves in combination. n. March, p. 202–212, 2017.

ELSTONE, L. et al. High Speed Crop and Weed Identification in Lettuce Fields for Precision Weeding. 2020.

FATICHAH, C. et al. A hybrid fuzzy morphology and connected components labeling methods for vehicle detection and counting system. *International Journal on Smart Sensing and Intelligent Systems*, v. 9, p. 765–779, 2016. ISSN 11785608.

FUZITA, B. A. O C ONTROLE DE Merremia aegyptia NA C ULTURA DA C ANA - DE -A ÇÚCAR Efficacy of Herbicides Applied During the Dry and Wet Seasons for the Control of Merremia aegyptia in Sugarcane. p. 631–642, 2009.

GAO, Y. et al. A survey on recent object detection techniques useful for monocular vision-based planetary terrain classification. *Robotics and Autonomous Systems*, Elsevier B.V., v. 62, n. 2, p. 151–167, 2014. ISSN 09218890. Disponível em: <<http://dx.doi.org/10.1016/j.robot.2013.11.003>>.

GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. USA: Prentice Hall, 2008.

GUAN, P. P.; YAN, H. BLOOD CELL IMAGE SEGMENTATION BASED ON THE HOUGH TRANSFORM AND FUZZY CURVE TRACING. *2011 International Conference on Machine Learning and Cybernetics*, IEEE, v. 4, p. 1696–1701, 2011.

HE, L.; CHAO, Y. A very fast algorithm for simultaneously performing connected-component labeling and euler number computing. *IEEE Transactions on Image Processing*, IEEE, v. 24, n. 9, p. 2725–2735, 2015. ISSN 10577149.

HU, Y. et al. Designing , Prototyping , and Testing a Flexible Suturing Robot for Transanal Endoscopic Microsurgery. *IEEE Robotics and Automation Letters*, IEEE, v. 4, n. 2, p. 1669–1675, 2019.

HUNG, C. et al. Autonomous intelligent system for fruit yield estimation. p. 545–550, 2016.

INÁCIO, D.; RIEDER, R. Computer vision and arti fi cial intelligence in precision agriculture for grain crops : A systematic review. v. 153, n. April, p. 69–81, 2018.

IQBAL, Z. et al. An automated detection and classification of citrus plant diseases using image processing techniques: A review. *Computers and Electronics in Agriculture*, Elsevier, v. 153, n. July, p. 12–32, 2018. ISSN 01681699. Disponível em: <<https://doi.org/10.1016/j.compag.2018.07.032>>.

- JIN, L.; LIANG, H.; FILTER, A. M. Deep Learning for Underwater Image Recognition in Small Sample Size Situations. n. 61379007, p. 3–6, 2017.
- JUNEJA, K.; RANA, C. Multi-featured and fuzzy-filtered machine learning model for face expression classification. *Wireless Personal Communications*, Springer US, 2020. ISSN 1572834X. Disponível em: <<https://doi.org/10.1007/s11277-020-07620-8>>.
- KAMATH, R.; BALACHANDRA, M.; PRABHU, S. Paddy Crop and Weed Discrimination : A Multiple Classifier System Approach. v. 2020, 2020.
- KAUR, R.; SINGH, A. FUZZY LOGIC : AN OVERVIEW OF DIFFERENT APPLICATION AREAS. v. 18, n. 8, p. 677–689, 2019.
- KIM, Y. J. et al. Vision-Based Direction Determination of a Mobile Robot in Indoor Environment. *5th IEEE International Conference on Digital Ecosystems and Technologies (IEEE DEST 2011)*, IEEE, v. 5, n. June, p. 153–157, 2011.
- KSHIRSAGAR, K. G. IMPACT OF ORGANIC SUGARCANE FARMING ON ECONOMICS. p. 172–184, 2005.
- LEVRATTI, A. et al. Full length Article TIREBOT : A collaborative robot for the tire workshop. v. 57, n. November 2018, p. 129–137, 2019.
- LI, Y. et al. Local Haar-like features in edge maps for pedestrian detection. *Proceedings - 4th International Congress on Image and Signal Processing, CISP 2011*, IEEE, v. 3, p. 1424–1427, 2011. ISSN 978-1-4244-9306-7.
- LI, Z. et al. Automatic image annotation using fuzzy association rules and decision tree. *Multimedia Systems*, Springer Berlin Heidelberg, v. 23, p. 679–690, 2017. ISSN 09424962.
- LIU, J. et al. Plant Leaf Classification Based on Deep Learning. IEEE, p. 3165–3169, 2018.
- LIU, Q. et al. Microscopic image enhancement of Chinese Herbal Medicine based on fuzzy set. *2017 2nd International Conference on Image, Vision and Computing, ICIVC 2017*, p. 299–302, 2017.
- LIU, Y.; GU, G. Efficient algorithm for obtaining connected components in bi-level images. *IEICE Electronics Express*, v. 11, n. 3, p. 1–6, 2014.
- LOTTE, P. et al. Fully Convolutional Networks With Sequential Information for Robust Crop and Weed Detection in Precision Farming. IEEE, v. 3, n. 4, p. 2870–2877, 2018.
- LYTRIDIS, C. et al. An overview of cooperative robotics in agriculture. *Agronomy*, v. 11, n. 9, 2021. ISSN 2073-4395. Disponível em: <<https://www.mdpi.com/2073-4395/11/9/1818>>.
- MEGAHED, M.; MOHAMMED, A. Modeling adaptive E-Learning environment using facial expressions and fuzzy logic. *Expert Systems with Applications*, Elsevier Ltd, v. 157, p. 113460, 2020. ISSN 09574174. Disponível em: <<https://doi.org/10.1016/j.eswa.2020.113460>>.

- MILIOTO, A.; MAR, C. V. Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in cnns. 2018.
- MONQUERO, P. A. et al. Monitoramento do banco de sementes de plantas daninhas em áreas com cana-de-açúcar colhida mecanicamente. *Planta Daninha*, v. 29, n. 1, p. 107–119, 2011.
- MUTHEVI, A.; UPPU, R. B. Leaf classification using completed local binary pattern of textures. *2017 IEEE 7th International Advance Computing Conference (IACC)*, p. 870–874, 2017. Disponível em: <<https://api.semanticscholar.org/CorpusID:23414287>>.
- OLSEN, A. et al. DeepWeeds : A Multiclass Weed Species Image Dataset for Deep Learning. *Scientific Reports*, Springer US, n. December 2018, p. 1–12, 2019. ISSN 2045-2322. Disponível em: <<http://dx.doi.org/10.1038/s41598-018-38343-3>>.
- OREN, M. et al. Pedestrian detection using wavelet templates. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 193–199, 1997. ISSN 10636919.
- PALANIAPPAN, S.; NATARAJAN, S. Parallel realization of single pass connected component analysis on a multi-core architecture. *International Conference on Communication and Signal Processing, ICCSP 2014 - Proceedings*, IEEE, p. 582–586, 2014.
- PAPADOPOULOS, A. V. et al. Weed mapping in cotton using ground-based sensors and GIS. *Environmental Monitoring and Assessment*, Environmental Monitoring and Assessment, v. 190, n. 10, 2018. ISSN 15732959.
- PAPAGEORGIU, C. P.; OREN, M.; POGGIO, T. General framework for object detection. *Proceedings of the IEEE International Conference on Computer Vision*, IEEE, p. 555–562, 1998.
- PARK, J.; LEE, J.; SIM, D. Low-complexity CNN with 1D and 2D filters for super-resolution. *Journal of Real-Time Image Processing*, Springer Berlin Heidelberg, v. 17, n. 6, p. 2065–2076, 2020. ISSN 18618219. Disponível em: <<https://doi.org/10.1007/s11554-020-01019-1>>.
- PEREZ-SANZ, F.; NAVARRO, P. J.; EGEA-CORTINES, M. Plant phenomics: An overview of image acquisition technologies and image data analysis algorithms. *GigaScience*, Oxford University Press, v. 6, n. 11, p. gix092, 2017.
- PETRÔNIO, L. C. de F. et al. The organic sugar business in brazil: main producing regions and international market insertion. *Iheringia - Serie Botanica*, Fundacao Zoobotanica do Rio Grande do Sul, v. 77, 2022. ISSN 24468231.
- PEUWNUAN, K.; WORARATPANYA, K.; PASUPA, K. Modified adaptive thresholding using integral image. *2016 13th International Joint Conference on Computer Science and Software Engineering, JCSSE 2016*, 2016.
- POLYMENI, S. et al. The impact of 6g-iot technologies on the development of agriculture 5.0: A review. *Electronics*, MDPI AG, v. 12, p. 2651, 6 2023. ISSN 20799292.

- POSSEBON, I. F.; POLLI, H. Q. Cultivo Orgânico Da Cana De Açúcar. *Revista Interface Tecnológica*, v. 17, n. 1, p. 517–529, 2020. ISSN 1807-3980.
- QIU, S.; ZHOU, F.; CRANDALL, P. E. Discrete Gabor transforms with complexity $O(N\log N)$. *Signal Processing*, v. 77, n. 2, p. 159–170, 1999. ISSN 01651684.
- RAJA, R. et al. Rtd-seps: Real-time detection of stem emerging points and classification of crop-weed for robotic weed control in producing tomato. *Biosystems Engineering*, Elsevier Ltd, v. 195, p. 152–171, 2020. ISSN 15375110. Disponível em: <<https://doi.org/10.1016/j.biosystemseng.2020.05.004>>.
- RAJU, N.; KUMAR, D. J. N. Leaf classification based on Shape and Edge feature with k -NN Classifier. p. 548–552, 2016.
- RIGHETTO, A. J. et al. Predicting weed invasion in a sugarcane cultivar using multispectral image. v. 4763, 2019. Disponível em: <<https://doi.org/10.1080/02664763.2018.1450362>>.
- ROBERTO, V.; JUNIOR, A. Cultura orgânica da cana-de-açúcar, manejo ecológico e biodiversidade faunística associada. 2015.
- ROSHANIANFARD, A. ScienceDirect Performance evaluation of harvesting Performance evaluation of harvesting robot for heavy-weight crops Performance evaluation. *IFAC-PapersOnLine*, Elsevier B.V., v. 51, n. 17, p. 332–338, 2018. ISSN 2405-8963. Disponível em: <<https://doi.org/10.1016/j.ifacol.2018.08.200>>.
- SEBASTIAN, B.; BEN-TZVI, P. Support vector machine based real-time terrain estimation for tracked robots. *Mechatronics*, Elsevier Ltd, v. 62, 10 2019. ISSN 09574158.
- SHIRATSUCHI, L. S. Mapeamento da variabilidade espacial das plantas daninhas com a utilização de ferramentas da agricultura de precisão. 2001.
- SOLOMON, S. K. S. S.; SHARMA, L.; PATHAK, V. P. J. A. D. Green Technologies for Improving Cane Sugar Productivity and Sustaining Soil Fertility in Sugarcane-Based Cropping System. *Sugar Tech*, Springer India, v. 21, n. 2, p. 186–196, 2019. ISSN 0974-0740. Disponível em: <<https://doi.org/10.1007/s12355-019-00706-z>>.
- SORI, H. et al. Effect for a paddy weeding robot in wet rice culture. *Journal of Robotics and Mechatronics*, v. 30, n. 2, p. 198–205, 2018. ISSN 18838049.
- SUJARITHA, M. et al. Weed detecting robot in sugarcane fields using fuzzy real time classifier. *Computers and Electronics in Agriculture*, Elsevier B.V., v. 134, p. 160–171, 2017. ISSN 0168-1699. Disponível em: <<http://dx.doi.org/10.1016/j.compag.2017.01.008>>.
- SVYRYDOV, A.; KUCHUK, H.; TSIAPA, O. Improving efficiency of image recognition process : Approach and case study. *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, IEEE, p. 593–597, 2018.
- TALPUR, N.; SALLEH, M. N. M.; HUSSAIN, K. An investigation of membership functions on performance of ANFIS for solving classification problems. *IOP Conference Series: Materials Science and Engineering*, v. 226, n. 1, 2017. ISSN 1757899X.

- TODERO, I. et al. Scientia Horticulturae Formulation of a bioherbicide with metabolites from *Phoma* sp. *Scientia Horticulturae*, Elsevier, v. 241, n. July, p. 285–292, 2018. Disponível em: <<https://doi.org/10.1016/j.scienta.2018.07.009>>.
- UMAMAHESWARI, S.; ARJUN, R.; MEGANATHAN, D. Weed Detection in Farm Crops using Parallel Image Processing. IEEE, 2018.
- UTSTUMO, T. et al. Robotic in-row weed control in vegetables. *Computers and Electronics in Agriculture*, Elsevier, v. 154, n. 7034, p. 36–45, 2018. ISSN 01681699. Disponível em: <<https://doi.org/10.1016/j.compag.2018.08.043>>.
- VARATHARAJAN, R.; MANOGARAN, G.; PRIYAN, M. K. A big data classification approach using LDA with an enhanced SVM method for ECG signals in cloud computing. *Multimedia Tools and Applications*, Multimedia Tools and Applications, v. 77, n. 8, p. 10195–10215, 2018. ISSN 15737721.
- VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. [S.l.: s.n.], 2001. v. 1, p. I–I. ISSN 1063-6919.
- WÄLDCHEN, J.; MÄDER, P. *Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review*. [S.l.]: Springer Netherlands, 2018. v. 25. 507–543 p. ISSN 18861784. ISBN 0123456789.
- WANG, A.; ZHANG, W.; WEI, X. A review on weed detection using ground-based machine vision and image processing techniques. *Computers and Electronics in Agriculture*, Elsevier, v. 158, n. November 2018, p. 226–240, 2019. Disponível em: <<https://doi.org/10.1016/j.compag.2019.02.005>>.
- WANG, Z. et al. Leaf Recognition Based on DPCNN and BOW. *Neural Processing Letters*, Springer US, v. 47, n. 1, p. 99–115, 2018. ISSN 1573-773X.
- WENDEL, A.; UNDERWOOD, J. Self-Supervised Weed Detection in Vegetable Crops Using Ground Based Hyperspectral Imaging. p. 5128–5135, 2016.
- WILSON, G. et al. Robot-enabled support of daily activities in smart home environments. *Cognitive Systems Research*, Elsevier B.V., v. 54, p. 258–272, 2019. ISSN 1389-0417. Disponível em: <<https://doi.org/10.1016/j.cogsys.2018.10.032>>.
- WU, S. G. et al. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network. p. 1–6, 2007.
- WU, X. Design and Implementation of Computer Vision based In-Row Weeding System. p. 4218–4224, 2019.
- YOGHEEDHA, K. Automatic Vehicle License Plate Recognition System based on Image Processing and Template Matching Approach. *2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, IEEE, p. 1–8, 2018.
- ZADEH, L. A. Fuzzy sets. *INFOR ATIO AND CONTROL*, v. 8, p. 338–353, 1965.

ZHENG, X.; WANG, X. Leaf vein extraction using a combined operation of mathematical morphology. p. 0–3, 2010.

ZHOU, S.; SUN, B.; SHI, J. An SPC Monitoring System for Cycle-Based Waveform Signals Using Haar Transform. v. 3, n. 1, p. 60–72, 2006.

ZHU, X.; ZHU, M.; REN, H. Method of plant leaf recognition based on improved deep convolutional neural network. *Cognitive Systems Research*, Elsevier B.V., v. 52, p. 223–233, 2018. ISSN 13890417. Disponível em: <<https://doi.org/10.1016/j.cogsys.2018.06.008>>.

APÊNDICE A - CONSIDERAÇÕES SOBRE O PROCESSAMENTO DE IMAGENS DESTE TRABALHO

A.1 Pré-processamento

A.1.1 Tamanho da Imagem - redução, transformação em escala de cinza

Na representação de uma imagem colorida, um determinado sistema de cores é adotado. O objetivo desse sistema é uma padronização que facilita a especificação de cores (GONZALEZ; WOODS, 2008). Há muitos sistemas de cores utilizados atualmente, cada um orientado a aplicação na qual a imagem é utilizada. Para fins deste trabalho, abordaremos dois sistemas de cores o RGB e o HSI.

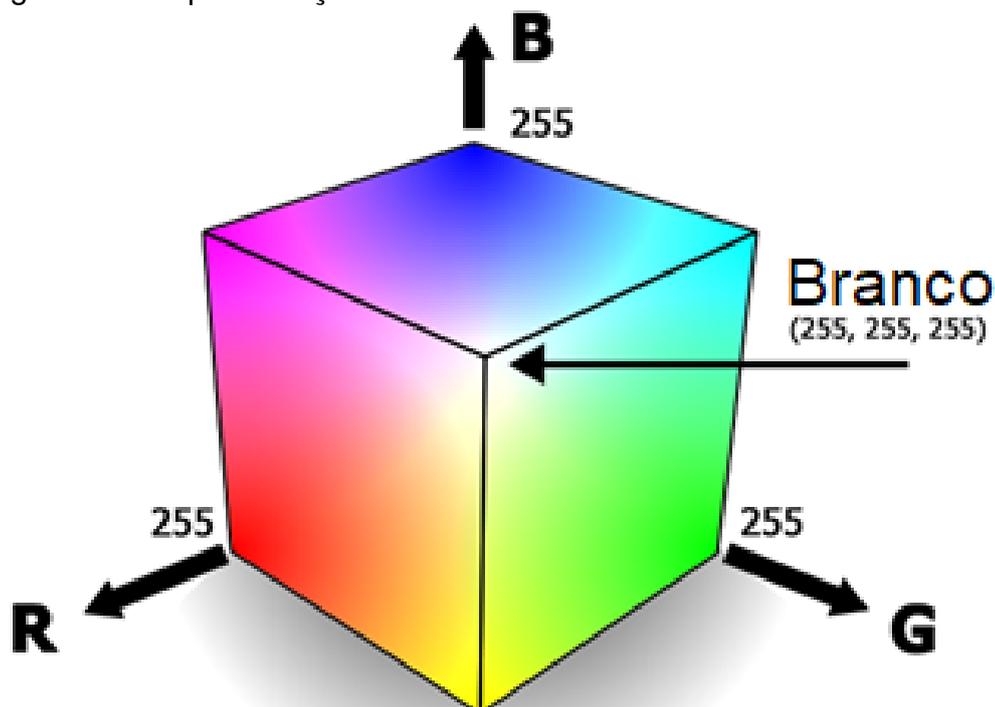
A.2 Sistema RGB

O modelo RGB utiliza três componentes das cores primárias: vermelho (Red), verde (Green) e azul (Blue). Baseia-se no sistema de coordenadas cartesianas, conforme é mostrado na Figura 47.

O sistema RGB utiliza três valores para representar cada imagem, sendo que cada valor correspondente a uma das três cores primárias. O número de bits utilizados para representar cada pixel no espaço RGB é chamado de profundidade de pixel (GONZALEZ; WOODS, 2008).

Cada pixel tem a profundidade de 24 bits, se cada um dos três componentes é representado por 8 bits. A imagem RGB com 24-bit é conhecida como *full-color*. O total de cores que podem ser representadas com 24 bits é $(2^8)^3 = 16.777.216$ possíveis cores. Uma imagem colorida, utilizando o RGB, é adquirida usando três filtros

Figura 47 - Representação do Sistema RGB em coordenadas cartesianas.



Fonte: Adaptada de Cunha (2016).

sensíveis ao vermelho, verde e azul. Cada filtro produz três imagens monocromáticas que são os componentes da imagem em RGB.

A.3 Sistema HSI

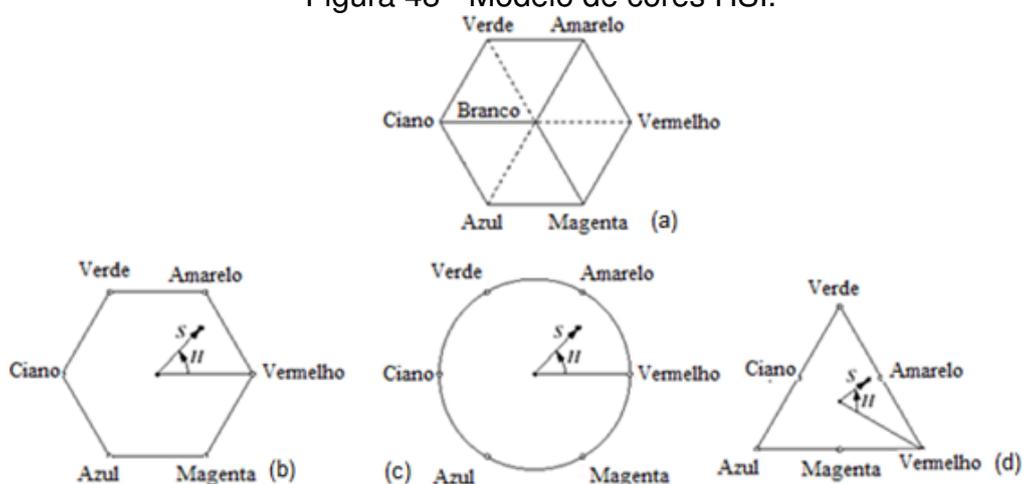
Gonzalez e Woods (2008) afirmam que o RGB e outros sistemas de cores são idealmente aplicados para implementação de hardware. No entanto, tais sistemas não se ajustam bem em aplicações que necessitam de interpretação humana. Por exemplo, ninguém descreve a cor de um automóvel informando o percentual de cada uma das cores primárias que a compõem.

O modelo HSI (Hue Saturation Intensity) pretende se ajustar melhor a forma como os humanos descrevem um objeto. Ao descrever um objeto nós o fazemos baseado na matiz (hue), saturação e no brilho. A matiz é um atributo da cor que descreve uma cor pura, já a saturação mede o grau em que a cor pura é diluída na luz branca. O brilho é difícil de mensurar, por ser subjetivo. Por isso, utiliza-se a intensidade, que incorpora o brilho, para descrever a percepção da cor, sendo esta mensurável.

A Figura 48 mostra a representação do HSI. Na Figura 48a podem ser visualizadas

as cores primárias separadas por 120° e as cores secundárias separadas em 60° em relação as cores primárias, sendo que também separadas por 120° entre si. A 48b mostra um hexágono com um determinado ponto, que representa uma cor, e a matiz é calculada a partir de um ângulo relacionado ao ponto. A saturação é o tamanho do vetor em relação ao ponto e a origem. O eixo de intensidade vertical, 48d, é o tamanho do vetor até um ponto de uma cor, e o ângulo que este vetor faz com o eixo do vermelho (GONZALEZ; WOODS, 2008).

Figura 48 - Modelo de cores HSI.



Fonte: Adaptada de Gonzalez e Woods (2008).

Gonzalez e Woods (2008) ressaltam que o sistema RGB é ideal no processo de aquisição da imagem, mas oferece uma descrição limitada das cores. Por isso, na maioria dos sistemas temos a aquisição das imagens em RGB e depois é realizada a conversão desta para o HSI ou outro sistema de cores.

A.3.1 A.1.3.1 Conversão de RGB para HSI

Como é mais natural a aquisição da imagem no sistema RGB, a seguir é mostrado como calcular, conforme Gonzalez e Woods (2008), cada componente do HSI a partir do RGB. O componente H é obtido usando a seguinte equação:

$$H = \begin{cases} \theta, & B \leq 0 \\ 360 - \theta & B > 0 \end{cases} \quad (\text{A.1})$$

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{[(R-G)^2+(R-B)(G-B)]^{\frac{1}{2}}}} \right\} \quad (\text{A.2})$$

O componente Saturação é dado pela equação:

$$S = 1 - \frac{3}{R + G + B}[\min(R, G, B)] \quad (\text{A.3})$$

O componente de Intensidade é dado pela equação:

$$I = \frac{1}{3}(R + G + B) \quad (\text{A.4})$$

A.4 Interpolação

A interpolação é uma técnica muito utilizada em processamento de imagens, em operações como: ampliação, redução, rotação e correções geométricas (GONZALEZ; WOODS, 2008). Conceitualmente, pode ser entendida como o processo de usar um dado conhecido para estimar valores em locais desconhecidos.

Existem diferentes técnicas para realizar a interpolação de uma imagem. A técnica do vizinho mais próximo é reconhecida como a de implementação mais simples. Nesta técnica, por exemplo, no caso de aumento de uma imagem, o valor da intensidade do novo pixel receberá o mesmo valor do pixel mais próximo a ele. Gonzalez e Woods (2008) apontam que esta técnica possui a tendência de produzir distorção severa de bordas retas e artefatos indesejáveis. Por isso, outras técnicas são propostas como a interpolação bi-linear, que utiliza os quatro vizinhos mais próximos para estimar a intensidade de uma determinada localização. A equação para calcular o valor de um determinado ponto, denotado por $v(x, y)$ é expressa como:

$$v(x, y) = ax + by + cxy + d \quad (\text{A.5})$$

onde os quatro coeficientes são determinados através dos valores dos quatro vizinhos mais próximos do ponto. A interpolação bi-linear possui resultado melhor do que a técnica do vizinho mais próximo, porém apresenta um aumento modesto na carga computacional (GONZALEZ; WOODS, 2008).

Outra técnica utilizada é a interpolação bi-cúbica. Esta técnica, utiliza os dezesseis vizinhos mais próximos para estimar o valor de um determinado ponto. A técnica utiliza

a seguinte equação para o cálculo:

$$v(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j \quad (\text{A.6})$$

Em linhas gerais, a interpolação bi cúbica preserva mais detalhes em uma imagem do que a bi-linear. A Figura 49 mostra a imagem original comparada com outras imagens que mostram a aplicação das 3 técnicas descritas acima. Na Figura 49b, que utiliza o vizinho mais próximo, é possível notar uma espécie de degrau, sobretudo na transição de uma região entre cores diferente, como por exemplo no chapéu. Na Figura 49c não se nota os degraus da técnica anterior, porém a figura tem um aspecto borrado. A Figura 49d apresenta o resultado visual melhor, porém é a técnica que exige maior tempo de processamento.

A.5 Segmentação

Na segmentação a imagem é subdividida em objetos e regiões que os compõe. Esse processo é fundamental para abstrair o conhecimento necessário para o reconhecimento de determinado padrão. A segmentação é uma das tarefas mais difíceis no âmbito de processamento de imagem (GONZALEZ; WOODS, 2008). A próxima seção descreve o algoritmo de Canny que é amplamente utilizado no processo de segmentação de imagens.

A.5.1 Detector de Bordas Canny

Segundo Gonzalez e Woods (2008), a abordagem do Canny consiste em três objetivos básicos:

1. Taxa baixa de erro. Todas as bordas devem ser encontradas e as encontradas devem ser próximas as reais.
2. Os pontos da borda devem ser bem localizados. A localização dos pontos encontrados deve ser próxima aos pontos reais.
3. Único ponto de resposta. O detector não deve identificar vários pixels de borda onde existe somente um.

Figura 49 - Comparação de técnicas de interpolação da imagem. (a) Imagem Original (b). Vizinho mais próximo (c) Bi-linear (d) Bi-cúbica.



Fonte: Parsania e Virparia (2018).

Expressar esses objetivos de forma matemática é o que constitui o trabalho do detector Canny. Os seguintes passos são realizados pelo algoritmo de Canny:

1. Leitura da imagem (I) para ser processada.
2. Criação de uma máscara Gaussiana unidimensional para convoluir com I . O desvio padrão (s) da Gaussiana é um parâmetro para o detector de bordas.
3. Criação de uma máscara unidimensional para a primeira derivada da Gaussiana nas direções x e y ; chamá-los de G_x e G_y . O mesmo valor é usado no passo 2.
4. Convoluir a imagem I com o G ao longo das linhas para obter o componente x e a imagem I_x , e aplicar nas colunas para obter o componente y da imagem I_y .

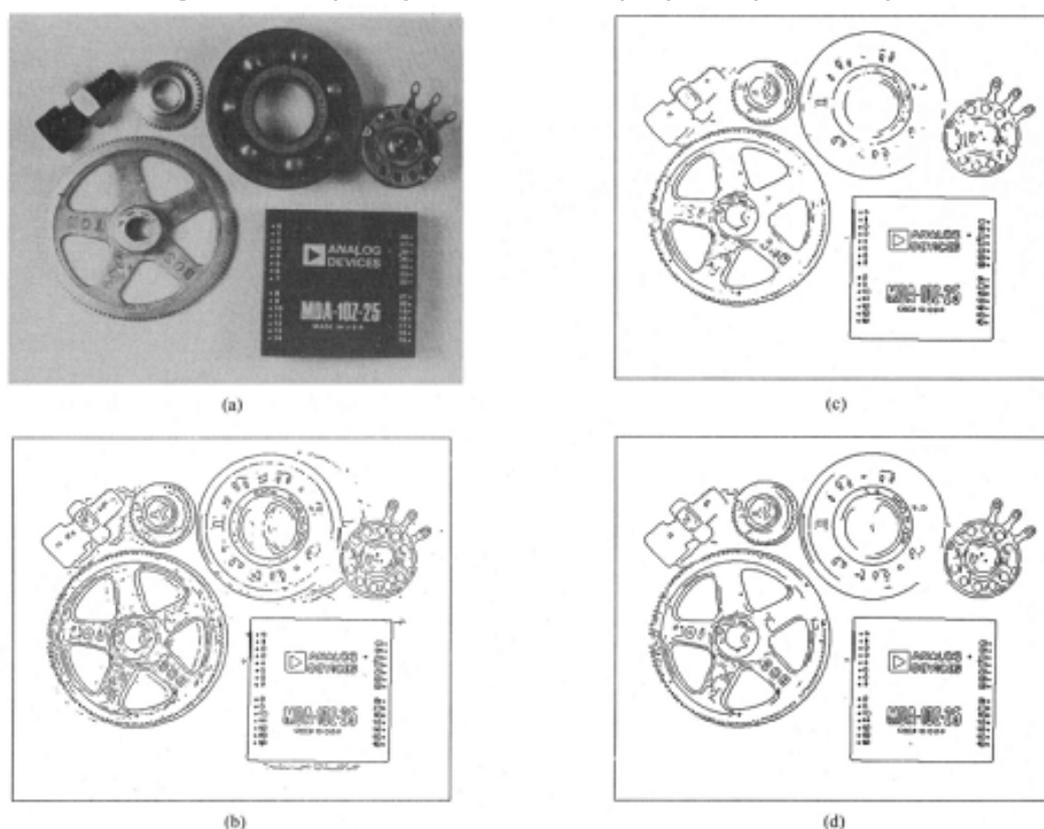
5. Convoluir I_x com G_x para obter I'_x , o componente x de I convoluido com a derivada da Gaussiana, e convoluir I_y com G_y para obter I'_y .
6. Calcular a magnitude da resposta da borda combinando os componentes x e y .

A magnitude do resultado pode ser calculada em cada pixel (x, y) como:

$$M(x, y) = \sqrt{I'_x(x, y)^2 + I'_y(x, y)^2} \quad (\text{A.7})$$

O passo final é a supressão dos valores não máximos, onde os pixels que não possuem a máxima local são removidos. A Figura 50 mostra o exemplo retirado do artigo original de aplicação da técnica para a identificação de bordas. Na Figura 50a temos a imagem original de entrada e nas demais a aplicação de um valor de *threshold* diferente. Nota-se que as bordas resultantes são influenciadas pelo valor de *threshold*.

Figura 50 - Aplicação do método proposto por Canny.



Fonte: Canny (1986).

Gonzalez e Woods (2008) ressaltam que, na prática, os pixels que pertencem às bordas raramente são completamente encontrados seja por conta de ruídos, espaços nas bordas devido a iluminação não uniforme, entre outros efeitos que introduzem va-

lores de intensidade espúrios e descontínuos. Algumas técnicas são utilizadas para tentar unir esses espaços com o restante das bordas. Dentre elas, temos a transformada de Hough.

A.5.2 Transformada de Hough

Conforme Duda e Hart (1972), uma linha reta pode ser especificada por um ângulo θ e sua distância algébrica do ponto de origem ρ . A equação da reta correspondente a esta geometria é:

$$X \cos \theta + Y \sin \theta = \rho \quad (\text{A.8})$$

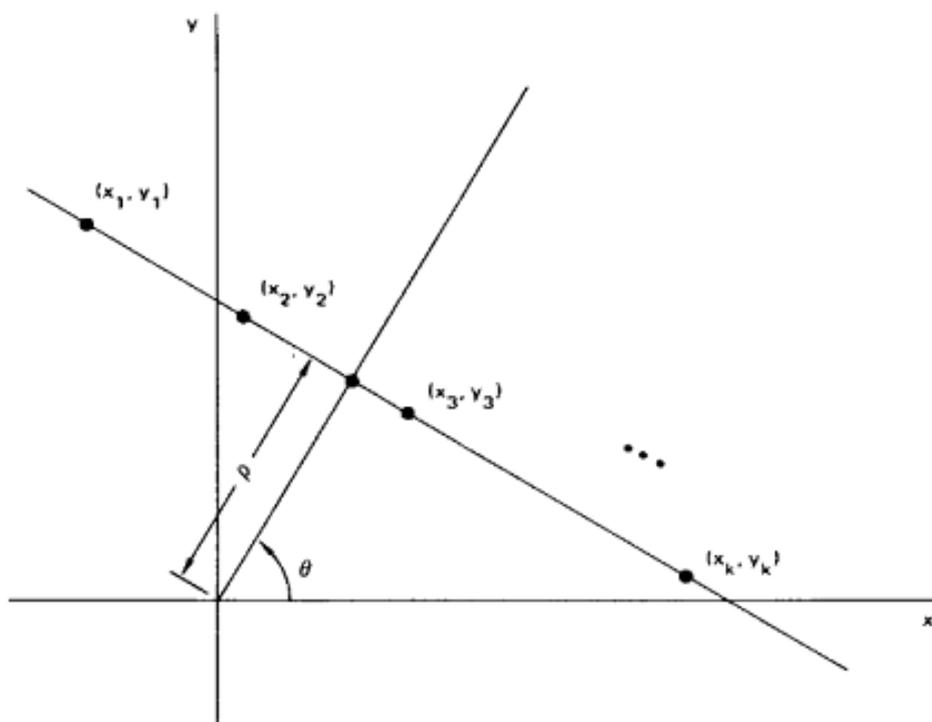
Restringindo o valor de θ para o intervalo $[0, \phi]$, então os parâmetros normais para uma linha são únicos. Assim, cada linha no plano $x - y$ corresponde a um único ponto no plano $\theta - \rho$. Para encontrar uma linha que passe por um conjunto de pontos, esses pontos (x_i, y_i) são transformados em curvas sinusoidais no plano $\theta - \rho$ definido pela equação:

$$\rho = X_i \cos \theta + Y_i \sin \theta \quad (\text{A.9})$$

Os pontos colineares possuirão curvas com ponto de intersecção comum, conforme mostra a Figura 51. Duda e Hart (1972) resumem as propriedades de sua técnica da seguinte forma:

1. Um ponto no plano da imagem corresponde a uma curva sinusoidal no plano dos parâmetros.
2. Um ponto no plano dos parâmetros corresponde a uma linha reta no plano da imagem.
3. Os pontos que pertencem a mesma linha no plano da imagem correspondem a curvas com um ponto comum no plano dos parâmetros.
4. Os pontos pertencentes as mesmas curvas no plano dos parâmetros correspondem as linhas no mesmo ponto no plano da imagem.

Figura 51 - Identificação de pontos colineares.



Fonte: Duda e Hart (1972).

Guan e Yan (2011) utilizam a transformada de Hough para segmentação automática de células sanguíneas, o que é importante para os diagnósticos e para as análises clínicas. Kim et al. (2011) utilizam a técnica para determinar rotas e realizar a movimentação de um robô em um ambiente fechado.