

**CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
CAMPUS TIMÓTEO**

HENRIQUE FERREIRA CASTRO

**USO DE ALGORITMOS DE RECONHECIMENTO
DE PADRÕES APLICADOS À IDENTIFICAÇÃO DE
OBJETOS: UM ESTUDO DE CASO NO
RECONHECIMENTO DE FRUTAS**

Timóteo

**2019
HENRIQUE FERREIRA CASTRO**

USO DE ALGORITMOS DE RECONHECIMENTO DE PADRÕES APLICADOS À IDENTIFICAÇÃO DE OBJETOS: UM ESTUDO DE CASO NO RECONHECIMENTO DE FRUTAS

Monografia apresentada à Coordenação de
Engenharia de Computação do Campus
Timóteo do Centro Federal de Educação
Tecnológica de Minas Gerais para obtenção do
grau de Bacharel em Engenharia de Computação.

Orientador: Odilon Corrêa da Silva

Timóteo

2019

Henrique Ferreira Castro

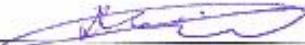
**Uso de Algoritmos de Reconhecimento de Padrões
Aplicados À Identificação de Objetos: Um Estudo de Caso
no Reconhecimento de Frutas**

Trabalho de Conclusão de Curso
apresentado ao Curso de Engenharia de
Computação do Centro Federal de Educação
Tecnológica de Minas Gerais, campus Timóteo,
como requisito parcial para obtenção do título de
Engenheiro de Computação.

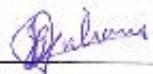
Trabalho aprovado. Timóteo, 13 de dezembro de 2019:



Prof. M. Odilon Corrêa da Silva
Orientador



Prof. M. Douglas Nunes da Oliveira
Professor Convidado



Prof. Dr. João Batista Queiroz Zuliani
Professor Convidado

Timóteo
2019

Agradecimentos

Primeiramente agradecer à meus pais, Valéria e Itamar, pelo suporte durante o tempo necessário para este trabalho, assim como no curso. Também agradeço a meu orientador, Odilon, pelos conselhos e pela paciência no andar do processo. Um muito obrigado também à minha amiga Laura por não me deixar desistir e me ajudar em momentos críticos. Agradeço também à todos os outros não citados, que direta ou indiretamente fizeram parte da minha formação.

Resumo

Apesar da grande aplicabilidade, fazer com que um computador seja capaz de reconhecer e classificar objetos em imagens pode não ser uma tarefa fácil na computação. Dito isso, este trabalho descreve um processo de reconhecimento e classificação de objetos em imagens utilizando 5 algoritmos classificadores diferentes e avaliando os resultados de cada um. Para tanto, foram aplicadas técnicas de processamento digital de imagens e reconhecimento de padrão. Para validar o funcionamento do processo proposto, desenvolveu-se um software utilizando recursos das bibliotecas OpenCV e scikit-learn para Python. O estudo de caso proposto foi o de identificação e classificação de frutas, e nos testes realizados foi possível identificar quais algoritmos se sobressaem em relação à outros de acordo com sua precisão. O estudo de caso ajudou a mostrar na prática o processo proposto, indicando ser viável o uso de algoritmos de classificação em problemas reais.

Palavras-chave: Processamento digital de imagens. Reconhecimento de Padrão. Algoritmos de Classificação.

Abstract

Despite its great applicability, making a computer capable of recognizing and classifying objects in images may not be an easy task in computing. That say, this essay describes a process of recognition and classification of objects in images using 5 different classifying algorithms and evaluating the results of each one. To this end, digital image processing and pattern recognition techniques were applied. To validate the operation of the proposed process, a software was developed using resources from the OpenCV and scikit-learn libraries for Python. The proposed case study was the identification and classification of fruits, and in the tests performed it was possible to identify which algorithms stand out in relation to others according to their precision. The case study helped to show the proposed process in practice, indicating that the use of classification algorithms in real problems is feasible.

Key-Words: Digital image processing. Pattern Recognition. Classification Algorithms.

Lista de Ilustrações

Figura 1 - Fluxograma do Processo Proposto por (Pinto, 2015)	16
Figura 2 – Diagrama geral do PDI.....	18
Figura 3 – Abordagens de segmentação. (A) - Imagem original em níveis de cinza. (B) Imagem segmentada através de binarização. (C) Imagem segmentada por detecção de bordas.....	20
Figura 4 - Exemplo de Decicion Tree	26
Figura 5 - Fluxograma do Processo.	27
Figura 6 – Exemplos das Frutas da Categoria 1. (A) Imagem de Fundo e Imagem com Banana. (B) Imagem de Fundo e Imagem com Maçã. (C) Imagem de Fundo e Imagem com Tomate....	32
Figura 7 - Imagens de Teste Para Identificação de Bananas na Categoria 1.	33
Figura 8– Exemplos das Frutas da Categoria 2. (A) Imagem de Fundo, Banana Madura e Banana Verde. (B) Imagem de Fundo, Maçã Madura e Maçã Verde. (C) Imagem de Fundo, Tomate Maduro e Tomate Verde.	34
Figura 9 - Imagens de Teste Para Identificação e Classificação de Maçãs na Categoria 2.....	35

Lista de Tabelas

Tabela 1 - Especificação das Categorias	30
Tabela 2 - Tabela de Resultados Com Configurações Padrão na Categoria 1.....	33
Tabela 3 - Tabela de Resultados Com Configurações Padrões na Categoria 2	35
Tabela 4 - Tabela de Resultados Com Configurações Ajustadas na Categoria 2.....	36
Tabela 5 - Tabela de Resultados Com Configurações Padrões na Categoria 3	36
Tabela 6 - Tabela de Resultados Com Configurações Ajustadas na Categoria 3.....	37

Listas de Abreviaturas e Siglas

RNA	Redes Neurais Artificiais
K-NN	<i>K-Nearest Neighbor</i>
PDI	Processamento Digital de Imagens
RGB	<i>Red, Green and Blue</i>
HSV	<i>Hue, Saturation and Value</i>
2D	Duas dimensões
HSL	Hue, Saturation and Lightness
SVM	Support Vector Machine
SVC	Support Vector Classifier
RF	Random Forest
DT	Decision Tree
GNB	Gaussian Naive-Bayes

Sumário

1 Introdução.....	10
1.1 O problema e sua importância.....	11
1.2 Justificativa.....	12
1.3 Objetivos	13
2 Materiais e Métodos.....	14
3 Estado da Arte	15
4 Fundamentação Teórica	18
4.1 Processamento digital de imagens	18
4.1.1 Aquisição de Imagens.....	18
4.1.2 Pré-Processamento de Imagens	19
4.1.3 Segmentação	19
4.1.4 Representação.....	21
4.2 Padrão de Cores	21
4.2.1 Padrão de Cores RGB.....	21
4.2.2 Padrão de Cores HSL	21
4.3 Reconhecimento de Padrões em Imagens	22
4.3.1 Aprendizagem Supervisionada	23
4.3.2 Aprendizagem Não Supervisionada	23
4.4 Algoritmos de Classificação.....	23
4.4.1 Algoritmo K-NN	23
4.4.2 Algoritmo Random Forest.....	24
4.4.3 Algoritmo Support Vector Machine (SVM).....	25
4.4.4 Algoritmo Decision Tree (Árvore de Decisão).....	25
4.4.5 Algoritmo Naive Bayes	26
5 Especificação do Processo	27
5.1 Aquisição de Imagem	28
5.2 Redimensionamento e Pré-Processamento.....	28
5.3 Segmentação e Representação	28
5.4 Treinar Algoritmo	29
5.5 Testar	29
6 Casos de Teste.....	30

6.1 Base de Dados	30
6.1.1 Composição da Base de Imagens de Treino.....	30
6.1.2 Composição da Base de Imagens de Teste	31
6.2.1 Configuração Padrão dos Algoritmos.....	31
6.2.2 Configuração Ajustada dos Algoritmos.....	31
6.2 Categoria 1	32
6.2.1 Resultados Configuração Padrão.....	33
6.3 Categoria 2.....	34
6.3.1 Resultados Configuração Padrão.....	35
6.3.2 Resultados Configuração Ajustada	36
6.4 Categoria 3.....	36
6.4.1 Resultados Configuração Padrão.....	36
6.4.2 Resultados Configuração Ajustada	37
6.5 Análise dos resultados.....	37
7 Conclusão.....	39
8 Referências Bibliográficas.....	41

1 Introdução

Entre os sentidos que o ser humano possui, o da visão pode ser considerado o mais eficiente para captar as informações originadas do ambiente ao seu redor. Esse sentido pode ser considerado uma tarefa de representação e processamento de informações, o que o torna adequado ao tratamento computacional. A visão é um processo que produz, através das imagens do mundo externo uma descrição útil para o observador, e ao mesmo tempo não há sobrecarga com informações desnecessárias (MARR, 1982). Humanos podem reconhecer objetos através de sua forma, cor, movimento, textura, entre outros. Com isso tem se estudado intensamente formas de garantir essa descrição através de métodos computacionais.

Algumas aplicações e alguns tipos de imagens requerem uma interação visual bastante intensa (SHIRAI, 1987). Com isso, viu-se necessário reproduzir o modo como o sistema visual humano funciona, já que algumas técnicas de adequação de imagem são importantes para explorar de forma eficiente os recursos dos sistemas de processamentos de imagens (FISCHLER; FIRSCHEIN, 1987).

Um sistema de visão artificial busca, com o auxílio de diversas áreas (biologia, medicina, eletrônica, matemática, etc), conseguir um conjunto de técnicas e métodos para dar suporte ao desenvolvimento de teorias e produtos que possam ser aplicáveis na prática, como por exemplo, automatizar processos de controle de qualidade, identificação de doenças e exploração de ambientes (FACON, 1996). O processamento e a análise de imagens é a ciência que permite analisar e manipular imagens digitais a partir de um computador. Uma estação de processamento de imagens deve oferecer os meios de digitalizá-las, visualizá-las e manipulá-las (MASCARENHAS; VELASCO, 1989).

Sendo assim, pode-se dizer que as possibilidades de avanço na indústria são inúmeras, porém esse desenvolvimento ainda é um desafio. O cérebro humano possui um alto poder de processamento que ainda não foi muito bem compreendido e consequentemente não replicado para as atuais tecnologias (PEDRINI; SCHWARTZ, 2008).

Algoritmos de Visão Computacional têm a capacidade de agilizar bastante o trabalho humano na indústria, sendo possível realizar tarefas de modo mais preciso e eficaz. Se bem avaliado e tendo tomado os cuidados adequados, a utilização de um bom algoritmo pode gerar um aumento significativo na produtividade e na garantia da qualidade dos produtos (FELICIANO; SOUZA; LETA, 2005).

Dentro da visão computacional existem alguns métodos utilizados para fazer as análises e classificações que se pretende no presente trabalho, uma delas, que trata classificação, descrição ou agrupamento de objetos é a de reconhecimento de padrões. Reconhecimento de padrões tem como objetivo principal classificar informações extraídas de conjuntos de dados (FRICK, 2008). O reconhecimento de padrões busca realizar uma descrição das imagens

baseado nas características próprias de cada uma através de meios computacionais, evitando assim a subjetividade causada por um operador humano. Formalmente, o reconhecimento de padrões é definido como um processo pelo qual um sinal é recebido e à ele é atribuído uma classe dentre um número já pré-determinado de classes (HAYKIN, 2001).

Com a evolução da computação, pesquisas foram desenvolvidas buscando criar recursos e ferramentas computacionais para otimizar e facilitar a utilização do reconhecimento de padrões na análise de imagens. Vale ressaltar que muitas dessas ferramentas também fornecem recursos que podem ser utilizados em outros tipos de análises de imagem, como Redes Neurais Artificiais e Métodos Exatos. Revisando a literatura, algumas ferramentas computacionais foram encontradas:

- BoofCV - Biblioteca open source Java para visão computacional em tempo real;
- MATLAB - Software de alta performance voltado para o cálculo numérico;
- OpenCV - Visão computacional open source;
- SimpleCV - Plataforma de visão computacional usando Python;
- VXL - Bibliotecas C++ para implementação e pesquisa de visão computacional.

1.1 O problema e sua importância

Frutas estão sujeitas a danos e perdas durante seu processo de produção e é na pós-colheita que se concentram os maiores prejuízos. Problemas como processamento, embalagem, manuseio e transporte inadequados, técnicas de conservação ruins, entre outros, podem vir a produzir alterações que degradam a qualidade do produto (DA SILVA; MORAIS, 2000).

A qualidade de frutas frescas é definida por várias características as quais podem ser mais ou menos atrativas para o consumidor, dentre essas pode-se destacar maturação, tamanho, peso, forma, cor, presença de manchas, etc. A maioria dessas características é tradicionalmente inspecionada por pessoas treinadas para a função. Contudo, o uso de um sistema computacional pode prover um substancial aumento de informação em tempo real, com precisão, velocidade e sem o uso de contato destrutivo, garantindo a manutenção do padrão de qualidade e redução dos custos (MELO, 2015).

Dentro do processo da pós-colheita é necessário uma avaliação precisa e rápida (MANICA et al., 1965), e ainda nos dias atuais essa avaliação representa uma das questões críticas no setor em virtude dos altos custos de tempo exigidos no processo. Pequenas e médias empresas geralmente contam somente com a capacidade humana para esse trabalho, o que ainda causa bastante impacto na produção.

A busca constante por maior produtividade e qualidade dos produtos do agronegócio é impulsionada pelo avanço tecnológico. A visão computacional atualmente pode ser aplicada na agricultura em diversas áreas, como por exemplo a agricultura de precisão (BRAMLEY, 2009),

veículos aéreos não-tripulados (HERWITZ et al., 2004), sistemas de informações geográficos (PEREIRA, 2013), análise da área foliar danificada (SANTOS, 2009), entre outros.

A multidisciplinaridade e os avanços da computação possibilitam novas práticas agrícolas e abrem a oportunidade para inserção de sistemas cada vez mais automatizados na agricultura. Nesse contexto, é possível utilizar um algoritmo de reconhecimento de padrões implementado com técnicas de visão computacional para auxiliar a prática de identificação e classificação de frutas? Qual dos algoritmos é o mais indicado para esse caso? É possível avaliar a precisão e desempenho computacional desses algoritmos?

A hipótese desse trabalho é que uma solução de baixo custo utilizando técnicas de visão computacional e reconhecimento de padrões pode ser utilizada no processo de reconhecimento e classificação de frutas.

1.2 Justificativa

Atualmente a triagem manual dos frutos e vegetais ainda possui algumas desvantagens como alta intensidade de trabalho, baixa produtividade e dificuldade de padronização na classificação (ARIVU; PRAKASH; SARMA, 2012). Com isso tem-se necessária uma melhora na classificação dos produtos, para que se alcance uma avaliação precisa e rápida. E dentre os métodos mais adequados para se trabalhar com visão computacional e análise de imagens temos o uso de métodos exatos, redes neurais artificiais (RNA) e reconhecimento de padrões.

No caso das RNA, foi descartado para este trabalho pois acredita-se não ser necessário um grande poder de adaptação e capacidade de representação não-linear que elas possuem, pois para esse poder de adaptação funcionar seria necessário desenvolver algoritmos de treinamento que permitam tratar de forma eficiente as mudanças (SILVA, 1998), gerando um custo computacional maior.

Foi descartado também a utilização de métodos exatos, pois apesar do seu baixo custo computacional, ele se baseia na realização de operações simples sobre imagens (PINTO, 2015) e considerando que algumas frutas podem ter formatos mais complexos, acredita-se que possa ser necessário um trabalho computacional maior para se obter bons resultados.

Embora diversos trabalhos de classificação de frutas tenham sido propostos, como pode ser visto em Carvalho et al (2014) e Simões e Costa (2003), particularmente a comparação entre algoritmos que podem ser utilizados na resolução dos casos permanece ainda como um problema em aberto. Desde modo, o presente trabalho propõe uma metodologia para utilizar diferentes algoritmos de reconhecimento de padrão no reconhecimento e identificação de frutas baseado em informações visuais, com ênfase na comparação entre esses algoritmos.

1.3 Objetivos

Este trabalho tem como objetivo geral utilizar e avaliar cinco algoritmos de reconhecimento de padrões através da identificação e classificação de frutas. Para isso, são propostos os seguintes objetivos específicos:

1. Descrever o funcionamento dos cinco algoritmos de reconhecimento de padrões propostos
 - a. Algoritmo K-NN;
 - b. Algoritmo Random Forest;
 - c. Algoritmo Support Vector Machine;
 - d. Algoritmo Decision Tree;
 - e. Algoritmo Naive Bayes.
2. Adaptar algoritmos de reconhecimento de padrões no reconhecimento de frutas;
3. Avaliar o desempenho dos algoritmos por meio de casos de teste;

2 Materiais e Métodos

As tarefas de codificação realizadas serão feitas com a linguagem de programação Python (versão 2.7.5) e a biblioteca OpenCV (versão 3.4.1). Será utilizado o ambiente de desenvolvimento Visual Studio Code.

A codificação, compilação e execução de todo software produzido será realizada em um computador com as seguintes configurações:

- Processador: Intel Core i7 7700HQ @ 2.80GHz;
- Memória RAM: 1x16 GBytes DDR4;
- Disco Rígido: 1TB (5400 RPM).

Atividades realizadas durante a elaboração do projeto:

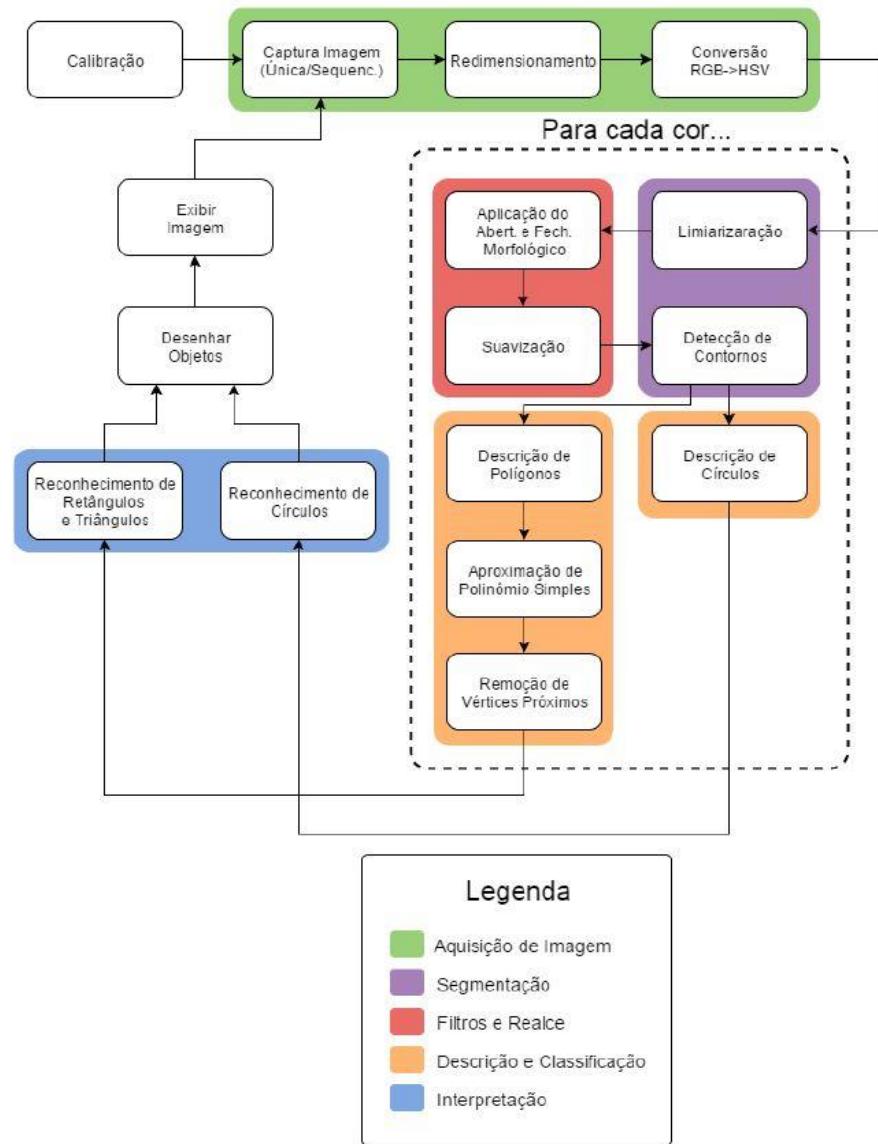
1. Estudo direcionado sobre visão computacional e reconhecimento de padrões;
2. Pesquisa e avaliação de trabalhos com base em visão computacional;
3. Pesquisa e avaliação das etapas de processamento de imagens em visão computacional;
4. Pesquisa e avaliação de trabalhos que utilizaram visão computacional e reconhecimento de padrões;
5. Pesquisa e avaliação de trabalhos que utilizaram diferentes algoritmos de classificação no reconhecimento de padrões;
6. Pesquisa e avaliação de problemas enfrentados no agronegócio que poderiam ser resolvidos com a ajuda da visão computacional;
7. Estudo direcionado sobre reconhecimento e classificação de diferentes frutas em relação a cor, maturação e tamanho;
8. Elaboração de um sistema de visão computacional para reconhecimento de frutas que permita a utilização de cinco algoritmos de classificação diferentes;
9. Elaboração e realização de testes do estudo de caso para a realização da análise comparativa dos cinco algoritmos de classificação implementados para o reconhecimento das frutas;
10. Documentação das atividades desenvolvidas;
11. Apresentação do trabalho para a banca examinadora;
12. Aplicação das correções do trabalho sugeridas pela banca examinadora.

3 Estado da Arte

Com o potencial que os sistemas computacionais passaram a oferecer e o avanço tecnológico, o número de trabalhos que buscam fazer o reconhecimento de objetos através da análise de imagens aumentou bastante nos últimos anos. Na literatura foram encontrados alguns casos que serão descritos a seguir.

O primeiro exemplo é de Pinto (2015), que propõe um processo de reconhecimento e rastreamento de objetos por cor e forma (figura 1), em imagens ou sequências de imagens capturadas em tempo real. Em seu trabalho os objetos utilizados eram de sua maioria compostos por figuras geométricas, como triângulos, quadrados, etc. Sendo assim, foram aplicadas técnicas de processamento digital de imagens e foi utilizado métodos exatos no processo de reconhecimento dos objetos. Em seus resultados, foram encontradas taxas de acerto satisfatórias com um tempo de resposta inferior à 33 milissegundos, evidenciando a simplicidade, eficiência e flexibilidade para o objetivo proposto. Como já explicado na seção de justificativa, a ideia de usar métodos exatos parte do pressuposto de que serão tratados objetos mais simples, sendo assim, ideal para o que fora proposto pelo autor. Entretanto, o mesmo não se aplica ao problema deste trabalho.

Figura 1 - Fluxograma do Processo Proposto por (Pinto, 2015)



Fonte: Pinto (2015)

Siqueira (2010) propôs um método desenvolvido para localização automática de cavidades em imagens ecocardiográficas fetais. Por se tratar de um problema relacionado à saúde e portanto que necessita de maior precisão em relação ao tempo necessário para se fazer a localização, o método desenvolvido no trabalho foi inspirado em uma metodologia proposta por Comaniciu, Ramesh, Meer (2003, p. 564) que trata o rastreamento de objetos em imagens utilizando redes neurais. Foi usado também para melhorar o desempenho computacional um método de busca hierárquica, que ajudou a diminuir o tempo gasto para a conclusão das tarefas.

Um outro trabalho que merece destaque, já voltado para o estudo de caso das frutas, é o de Simões e Costa (2003). Ele demonstra a classificação de laranjas baseada em padrões visuais. Esse trabalho utiliza padrões pré-estabelecidos para a classificação da fruta. Nele é

proposto uma classificação através de dois aspectos distintos: cor e qualidade. No caso da qualidade, a laranja é classificada de acordo com a ocorrência de defeitos de ordem mecânica, patológica, presença e intensidade de manchas e podridão. Já o parâmetro cor é descrito de acordo com 5 diferentes classes, que passam do verde escuro, ao amarelo e finalmente pelo laranja. Assim como no trabalho descrito anteriormente, neste também são utilizados alguns passos para se obter o resultado final. No primeiro é feita uma classificação das cores dos pixels da imagem, no segundo a segmentação da imagem, no terceiro é feita uma extração de padrões e no quarto os padrões encontrados são comparados com os padrões previamente conhecidos. Já para a automação da classificação foi proposta a seguinte abordagem:

- i) redes neurais artificiais do tipo perceptron multicamada para a classificação de pixels;
- ii) segmentação de imagens pelo algoritmo do crescimento da região;
- iii) contagem simples para extração de padrões;
- iv) agrupamento em células de Voronoi para a comparação dos padrões extraídos.

Entre os trabalhos que fazem comparações entre algoritmos, merece destaque o de Bezerra (2006), em que foi feito uma análise e comparação entre diferentes algoritmos de classificação utilizados no reconhecimento de imagens, como o K-NN, que será utilizado nesse trabalho, o POC-NN e o S-POC-KNN-RBF, que também utiliza redes neurais. Os resultados obtidos puderam mostrar que de um modo geral, o algoritmo que mescla os três classificadores se saiu melhor em relação à taxa de erro, critério de comparação utilizado entre os algoritmos, mas necessitou de um poder computacional muito maior que os demais, que apesar de possuírem taxas de erro maiores, não obtiveram resultados considerados ruins para a situação.

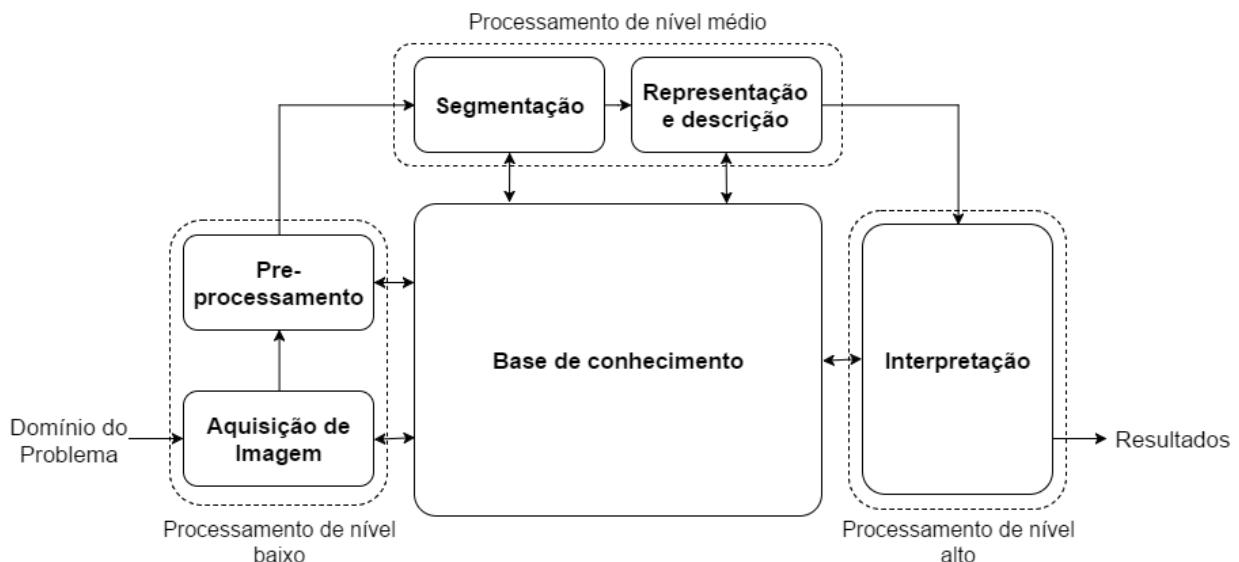
4 Fundamentação Teórica

4.1 Processamento digital de imagens

Segundo Gonzalez e Woods (2010), o processamento digital de imagens constitui a realização de processos digitais em imagens digitais por um computador digital. Basicamente, estes processos podem ser divididos em três categorias: processos de nível baixo, médio e alto (Figura 2).

Processos de nível baixo cobrem a aquisição da imagem e seu tratamento por meio do pré-processamento. Processos de nível médio envolvem a segmentação da imagem em áreas contendo os objetos de interesse, suas definições das formas de representação e por fim, a suas descrições. Processos de nível alto são os que tentam interpretar o conjunto de objetos descritos.

Figura 2 – Diagrama geral do PDI.



Fonte: Gonzales e Woods (2010).

4.1.1 Aquisição de Imagens

Em geral, a aquisição de imagens envolve algum tipo de transformação na imagem, como o redimensionamento ou a conversão do padrão de cores, e nesse caso pode ser tão simples quanto receber uma imagem que já esteja num formato digital ou então mais complexo como converter sinais analógicos transmitidos por um sensor convertendo informações ópticas

do mundo real em sinal elétrico, o que transforma uma imagem real em digital (GONZALEZ; WOODS, 2010).

4.1.2 Pré-Processamento de Imagens

Nessa etapa busca-se modificar e preparar os valores dos pixels da imagem com o objetivo de facilitar as operações subsequentes e obter melhores resultados (AWCOCK; THOMAS, 1996). Cada problema tem sua necessidade específica, portanto essa etapa depende de cada situação apresentada. Ou seja, algumas técnicas podem funcionar bem em determinado tipo de imagem e serem totalmente inadequadas em outros.

O pré-processamento pode ser dividido em duas partes: realce e restauração de imagens (AWCOCK; THOMAS, 1996). O realce é utilizado para melhorar a qualidade da imagem ou enfatizar algum aspecto particular da imagem. Já a restauração é um processo que procura reconstruir ou recuperar uma imagem que for degradada, usando algum conhecimento do que originou a degradação e aplicando o processo inverso para a recuperação (GONZALEZ; WOODS, 2010).

Uma técnica importante de realce é a equalização do histograma. O histograma de uma imagem em níveis de cinza representa as frequências relativas de ocorrência de vários níveis de cinza da imagem (JAIN, 1989). Em outras palavras, informa quantos pixels da imagem possui determinado nível de cinza, definido entre 0(preto) a 255(branco). Geralmente histogramas mal distribuídos apresentam imagens ruins (muito claras ou muito escuras). A equalização serve justamente para melhor uniformizar os valores dos níveis de cinza de tal forma que se obtenha uma melhor qualidade na imagem.

Como exemplo de técnicas de restauração podemos citar os filtros mais comuns na redução de ruídos, como os filtros da média e mediana (SONKA; HLAVAC; BOYLE, 1998), o filtro de crimmins e o alpha trimmed (BUSSE; CRIMMINS; FIENUP, 1995).

4.1.3 Segmentação

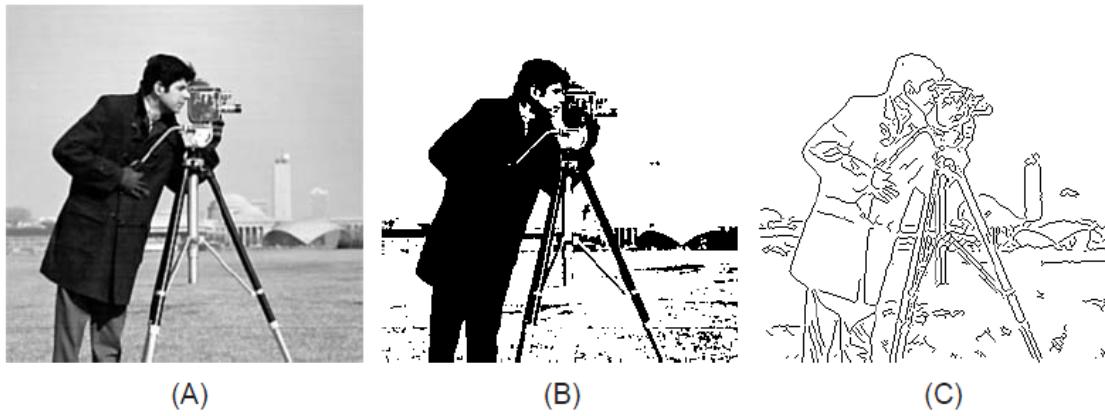
De acordo com Albuquerque (2003), segmentar uma imagem é nada mais do que separar a imagem como um todo nas partes que a constituem e que se diferenciam entre si. É comum denominar os grupos de pixels de interesse, ou pixels que fornecem alguma informação para o processamento da imagem, como “objetos”. Assim como é denominado “fundo” da imagem, o grupo de pixels que podem ser desprezados ou que não terão utilidade para o processamento.

Ainda de acordo com Albuquerque (2003), a segmentação é considerada a etapa mais crítica no tratamento da informação entre todas as etapas, pois é aqui que são definidas todas

as regiões de interesse que serão utilizadas no processamento e na análise posteriormente. Sendo assim, qualquer erro ou distorção presente nesta etapa vai refletir nas etapas seguintes, de forma a gerar no resultado final do processo resultados não desejados, contribuindo negativamente para a eficiência do projeto.

Vale lembrar que não existe um modelo formal para a segmentação, ela é um processo empírico e adaptativo, no caso, deve-se procurar sempre a adequação às características particulares de cada tipo de imagem e o objetivo que o processamento pretende alcançar.

Figura 3 – Abordagens de segmentação. (A) - Imagem original em níveis de cinza. (B) Imagem segmentada através de binarização. (C) Imagem segmentada por detecção de bordas



Fonte: Barelli (2018)

Geralmente são utilizadas dois tipos de abordagem para a segmentação, sendo a similaridade entre pixels ou a descontinuidade entre eles. No caso da primeira, a técnica mais utilizada é a binarização, que possui uma forma eficiente e simples do ponto de vista computacional. Geralmente é utilizada quando as amplitudes dos níveis de cinza são suficientes para caracterizar os objetos da imagem. Já nas técnicas baseadas em descontinuidade entre os pixels, procura-se determinar variações brutas do nível de luminância entre pixels vizinhos. Geralmente com isso procura-se delimitar os contornos ou bordas dos objetos na imagem. Como pode ser visto na imagem acima, a figura A tem a imagem original em níveis de cinza. Já na figura B é possível ver a imagem segmentada e retornando apenas níveis de preto e branco de acordo com os seus níveis de cinza. Por último, a figura C mostra a técnica de detecção de bordas, que é baseada em descontinuidade entre os pixels, sendo possível perceber claramente as bordas que destacam os objetos principais da imagem.

4.1.4 Representação

Na etapa anterior, foi mostrado que a segmentação define qual a região de interesse da imagem, que possui os agrupamentos de pixels denominados objetos. Com isso, a etapa de representação consiste em representar e descrever tais agrupamentos. De acordo com Figueirêdo (2007), esta etapa tem como resultado a extração de informação numérica dos objetos de interesse, e resulta em uma estrutura de dados denominada vetor de características.

Segundo Gonzalez e Woods (2010), os objetos numa região de interesse geralmente são representados ou por seus elementos externos (bordas) ou por seus elementos internos (pixels da região), sendo assim, uma representação é escolhida quando a atenção primária estiver voltada para a característica da forma, já a representação interna é usada caso a atenção esteja voltada para cor ou textura.

4.2 Padrão de Cores

4.2.1 Padrão de Cores RGB

Imagens no padrão RGB são matrizes de três dimensões que podem ser divididas e visualizadas em 3 matrizes 2D distintas, em que cada uma corresponde à um canal de cor, sendo R (Red - Vermelho), G (Green - Verde), B (Blue - Azul). Em imagens digitais cada pixel possui um espaço de armazenamento de 3 x 1 byte, sendo possível representar em cada byte o valor de intensidade de cada cor primária, indo de 0 (ausência da cor) até 255 (total intensidade da cor).

Este padrão é o mais usado no dia-a-dia, já que são geralmente as três cores primárias utilizadas para exibição de imagens em dispositivos como monitores, televisões, etc (SOLOMON, BRECKON, 2011).

4.2.2 Padrão de Cores HSL

O padrão de cores HSL é um modelo geralmente representado por um cone de três dimensões, também possuindo três parâmetros distintos, sendo H (Hue - Matiz ou tonalidade), S (Saturation - quantidade de branco misturado com a cor) e L (Lightness - luminosidade da cor).

Objetos representados neste padrão geralmente possuem uma forma mais consistente de serem percebidas e compreendida pelo ser humano do que nos canais RGB (SOLOMON; BRECKON, 2011). Por esse motivo, geralmente este é mais utilizado em sistemas de processamento digital de imagens baseado na percepção de cor pelo ser humano.

4.3 Reconhecimento de Padrões em Imagens

A capacidade que o ser humano possui de reconhecer padrões e classificar objetos sempre impressionou cientistas dos mais diversificados campos de estudo, principalmente àqueles que se dedicam a imitar a natureza humana (BARELLI, 2018). O campo da ciência denominado Reconhecimento de Padrões é o que estuda técnicas para descrever padrões de objetos procurando classificá-los.

No ramo da Visão Computacional, é essencial que as técnicas de reconhecimento sejam utilizadas, pois com elas se tornam possíveis a classificação automática de um objeto de interesse, ou seja, elas fazem com que os computadores se tornem capazes de enxergar o mundo, reconhecendo por exemplo, placas de carro, caracteres, faces humanas, entre outros. Vale ressaltar que tanto para nós, humanos, quanto para as máquinas, reconhecer objetos só é possível pois cada um tem suas características particulares, sendo necessárias três etapas para o reconhecimento ser possível. As duas primeiras são a Segmentação da imagem e a representação de características que foram abordadas em tópicos anteriores.

Na terceira etapa, a de reconhecimento, procura-se definir cada objeto de interesse em uma determinada classe. Contextualizando, pode-se definir uma classe como um conjunto de padrões que possuem características iguais. Para entender melhor, se considerarmos as figuras geométricas mais simples, pode-se dizer que cada uma possui uma quantidade de lados e que seus ângulos internos somados terão um determinado valor, no caso do triângulo, por exemplo, são três lados e a soma igual à 180° , ou seja, triângulo é uma classe com características padronizadas.

Os algoritmos que classificam o objeto a partir das características que possuem são conhecidos como classificadores. Existem diversos desses, sendo os mais populares os Bayesianos, o K-NN (K-Nearest Neighbor), a Lógica Nebulosa e as Redes Neurais Artificiais (BARELLI, 2018). Como existe uma variedade de classificadores, existe entre eles vantagens e desvantagens com relação à precisão e custo computacional, ficando a cargo do desenvolvedor do sistema decidir qual será implementado.

Além de definir o algoritmo a ser utilizado na classificação, a participação humana também é fundamental para definir as classes e as características consideradas para o processo e para que os classificadores consigam fazer seu trabalho, eles precisam aprender alguns critérios antes para executar suas tarefas. Entre esses critérios, existem dois métodos de aprendizagem bastante usados, que são a aprendizagem supervisionada e a não supervisionada.

4.3.1 Aprendizagem Supervisionada

Esse modelo de aprendizagem faz com que o classificador seja treinado para reconhecer padrões a partir de objetos já conhecidos, ou seja, deve-se apresentar um conjunto de características de objetos e suas respectivas classes anteriormente ao classificador. Com esses dados, o classificador é treinado para identificar automaticamente padrões que possuem essas informações, o que o torna capaz de classificar novos objetos.

Um algoritmo de classificação baseado em aprendizagem supervisionada e frequentemente usado em sistemas de visão computacional é o K-Nearest Neighbors (K-NN) (BARELLI, 2018).

4.3.2 Aprendizagem Não Supervisionada

Nesse tipo de abordagem, necessita-se apenas de uma lista de características de diversos objetos para ser feito o treinamento, sendo assim, a diferença para o método anterior, é que este não necessita das classificações dos objetos, ou seja, descarta-se a necessidade de um supervisor. Esses classificadores são capazes de identificar classes por conta própria, considerando as semelhanças das características de cada objeto.

4.4 Algoritmos de Classificação

Nos últimos anos várias aplicações com classificação de padrões foram surgindo, e a área de classificação de imagens foi uma das que obtiveram maior número de avanços, podendo citar como exemplo a classificação de dígitos, caracteres, faces, impressões digitais, entre outras.

No caso da classificação de padrões, o algoritmo responsável precisa primeiramente “aprender” a classificar os padrões do problema ao qual se deseja obter o reconhecimento, sendo assim, é necessário treinar o algoritmo de forma a torná-lo capaz de classificar um padrão desconhecido inicialmente em uma das classes existentes, já que a fase de treinamento possui um peso muito grande no desempenho do algoritmo (BEZERRA, 2006).

4.4.1 Algoritmo K-NN

Em 1966, Cover e Hart propuseram o algoritmo Nearest Neighbor (Vizinho Mais Próximo, em tradução livre), um método de estimativa de densidade simples e de fácil implementação

(WEBB, 2002). Com o passar dos anos foram surgindo variantes desse algoritmo e entre elas o K-NN, onde K representa uma constante.

O K-NN é um algoritmo que propõe uma modificação em relação ao algoritmo original. Essa modificação se dá na fase de teste/classificação especificamente, onde a principal diferença é que nesse caso o algoritmo faz uso dos K-vizinhos mais próximos, e não apenas o vizinho mais próximo, como é proposto no algoritmo original (WEBB, 2002). O parâmetro K faz com que o algoritmo consiga uma classificação mais refinada, mas é válido lembrar que o valor ótimo de K vai variar de um problema para outro, fazendo com que seja necessário para cada base de dados, um teste com diferentes valores de K de forma a descobrir o melhor valor para o problema proposto.

Para fazer a estimativa de um novo padrão que se quer reconhecer o algoritmo calcula os K-vizinhos mais próximos do padrão desconhecido e o classifica como pertencente a classe que aparece com mais frequência entre os vizinhos. Algumas vezes pode ocorrer o problema de haver um empate entre classes nos vizinhos mais próximos, fazendo com que seja necessário rodar o algoritmo de uma forma recursiva (usando K-1 e classificando de novo) até que uma das classes apareça com maior frequência (WEBB, 2002).

Apesar da evolução se comparado ao NN, o K-NN ainda é um algoritmo que possui deficiências, principalmente por continuar a armazenar os padrões de treinamento na memória e ainda exigir um grande esforço computacional (BEZERRA, 2006).

4.4.2 Algoritmo Random Forest

O algoritmo Random Forest é um método de classificação proposto por Breiman (2001). Nele, se consiste uma técnica de agregação de classificadores do tipo árvore, ou seja, são construídos de uma forma que a estrutura seja composta sempre de maneira aleatória (GHOSAL, 2009).

O Random Forest é um classificador baseado no método Bagging (BREIMAN, 1996), ou seja, para cada árvore, é utilizado um conjunto de treinamento diferente, que por sua vez é formado por N instâncias de treinamento escolhidas aleatoriamente (uma amostra, por exemplo). Para cada nó da árvore gerada, aleatoriamente são escolhidos um número de atributos que orientam o direcionamento do nó, que de acordo com uma métrica, será baseado na melhor discriminação de classes do treinamento.

Para determinar a classe de uma instância, o algoritmo combina o resultado de várias árvores de decisão, através de um mecanismo de votação. Cada árvore é responsável por dar uma classificação, ou um voto para uma classe. No final, a classificação é dada àquela classe que recebeu o maior número de votos entre todas as árvores da floresta (BREIMAN, 2001).

Em geral, a vantagem deste classificador é permitir base de dados com grande número de atributos, porém, isso o torna suscetível a um superajuste em algumas bases (DINIZ, MENDES NETO, FONTES, 2003).

4.4.3 Algoritmo Support Vector Machine (SVM)

O Algoritmo SVM faz basicamente uma máquina linear, em que a ideia principal é constituir um hiperplano como superfície de decisão de forma que a margem de separação entre exemplos positivos e negativos seja máxima (HAYKIN, 1999). Este método foi desenvolvido por Vladimir Vapnik junto de alguns colaboradores (BOSER et al. 1992; CORTES & VAPNIK, 1995; VAPNIK, 1995).

O treinamento da SVM envolve a resolução de um problema quadrático, dependente dos vetores de treinamento e de parâmetros especificados pelo usuário. Neste método, a partir de um espaço de entrada de padrões não-linearmente separáveis é formado um novo espaço de características, em dimensão outra onde os padrões serão linearmente separáveis. Assim, um hiperplano de separação ótimo entre os exemplos é construído (VAPNIK, 1995).

O algoritmo pode ser descrito da seguinte forma: dadas 'D' amostras de treinamento $\{x_i, y_i\}$, com $i=1, 2..., D$, onde $x_i \in RM$ é uma representação vetorial de um conjunto e $y_i \in \{-1, 1\}$ é sua classe associada. Nesse processo existe uma distribuição de probabilidade $P(x, y)$ desconhecida da qual os dados de treinamento serão retirados. Ou seja, o processo de treinamento consiste em treinar um classificador de forma que este aprenda um mapeamento $x \rightarrow y$ por meio de exemplos (classes) de treinamento $\{x_i, y_i\}$ de forma que a máquina seja capaz de classificar um exemplo (x, y) ainda não visto que siga a mesma distribuição de probabilidade (P) dos exemplos de treinamento.

4.4.4 Algoritmo Decision Tree (Árvore de Decisão)

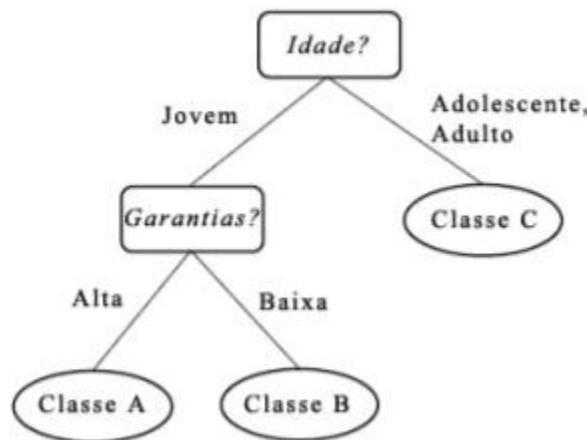
O método de classificação por Árvore de Decisão, funciona como um fluxograma em forma de árvore, onde cada nó (não folha) indica um teste feito sobre um valor (por exemplo, idade < 10). As ligações entre os nós representam os valores possíveis do teste do nó superior, e as folhas indicam a classe (categoria) a qual o registro pertence. Após a árvore de decisão montada, para classificarmos um novo registro, basta seguir o fluxo na árvore (mediante os testes nos nós não-folhas) começando no nó raiz até chegar a uma folha.

Pela estrutura que formam, as árvores de decisões podem ser convertidas em Regras de Classificação. O sucesso das árvores de decisão, deve-se ao fato de ser uma técnica extremamente simples, não necessita de parâmetros de configuração e geralmente tem um bom

grau de assertividade. Apesar de ser uma técnica extremamente poderosa, é necessário uma análise detalhada dos dados que serão usados para garantir bons resultados.

(QUINLAN, 1986) apresenta diversas técnicas para reduzir a complexidade das árvores de decisão geradas. Em um artigo recente (YANG et al, 2007) apresentam um algoritmo para extrair regras acionáveis, ou seja, regras que são realmente úteis para a tomada de decisões. Um exemplo de árvore de decisão pode ser visto na figura 4.

Figura 4 - Exemplo de Decicion Tree



Fonte: Han e Kamber (2006)

4.4.5 Algoritmo Naive Bayes

O classificador Naive Bayes é baseado no teorema de probabilidade de Bayes. No teorema de Bayes, a probabilidade condicional de um evento X pertencer a uma classe K pode ser calculada através das probabilidades condicionais em achar eventos particulares em cada classe e da probabilidade incondicional do evento em cada classe.

O cálculo da probabilidade de Bayes é um problema de classificação de padrões pois ele acha a probabilidade de que um dado elemento X pertence à uma classe K permitindo decidir uma classe ótimo escolhendo a classe com a maior probabilidade de ocorrências entre todas as classes possíveis, o que pode minimizar os erros de classificação.

Ele recebe o nome de “Naive” (ingênuo, em inglês) pois desconsidera a correlação entre variáveis. Por exemplo, se uma fruta é rotulada como “Limão”, caso ela também seja escrita como “Verde” ou “Redonda”, o algoritmo não leva em consideração a correlação entre os fatores, pois o trata como formas independentes.

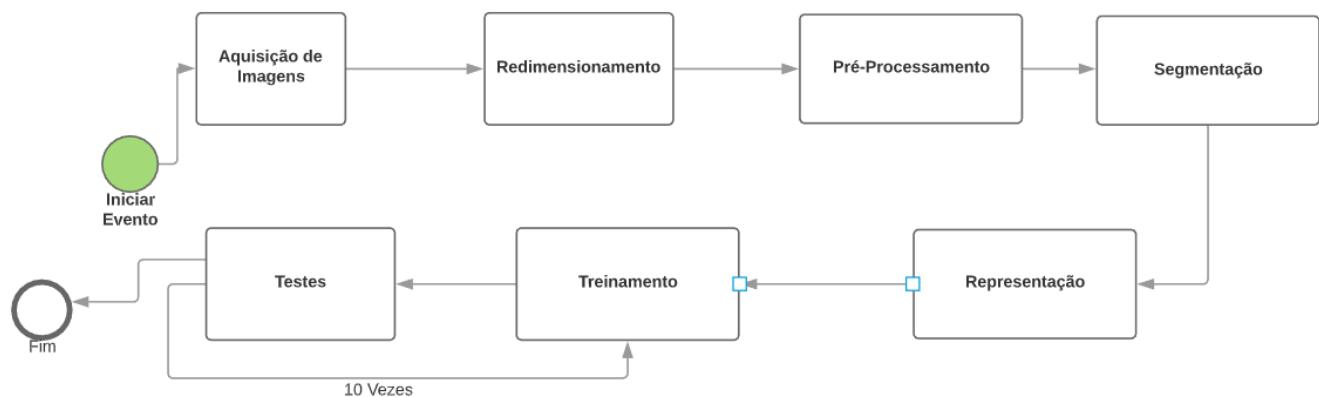
5 Especificação do Processo

Tomando como base todo o referencial teórico e os trabalhos já realizados na área, este trabalho faz uma captura e pré-processamento de imagens, treinamento e testes com os algoritmos classificadores. Para que os testes ofereçam valores consistentes são levantadas às seguintes restrições:

- Toda imagem deve ter iluminação adequada;
- As imagens devem ter ângulos diferentes para simular dificuldades e tentar refletir maior realidade;
- As cores do fundo devem se diferenciar das cores dos objetos a serem identificados;
- Os objetos a serem identificados devem estar totalmente contidos na imagem;

Um visão geral deste processo pode ser visualizada no fluxograma da Figura 5. Nas seções a seguir serão descritas todas as etapas deste fluxograma com o objetivo de esclarecer o funcionamento do processo proposto.

Figura 5 - Fluxograma do Processo.



Fonte: Elaborado pelo Autor

5.1 Aquisição de Imagem

Nesta etapa é feito a captura das imagens utilizadas para treinar os algoritmos. Por conveniência, espera-se imagens no padrão RGB. Nesta etapa serão capturadas diversas imagens.

Após a aquisição das imagens, são necessárias algumas transformações antes de avançar para as próximas etapas do processo.

5.2 Redimensionamento e Pré-Processamento

Esta etapa consiste em pegar as imagens capturadas na fase anterior e como o nome já diz, redimensioná-las para um tamanho menor e padrão. No trabalho em questão, as imagens são redimensionadas para 72x72 pixels. Esse processo serve para diminuir o custo computacional do treino dos algoritmos nos próximos passos, já que serão utilizados um menor número de pixels à serem classificados.

Junto à etapa anterior, também é feito um pré-processamento da imagem com a troca do sistema de cores RGB para uma representação de cores HSV. Como descrito no Referencial Teórico, essa parte do processo facilita o algoritmo a classificar as imagens.

5.3 Segmentação e Representação

Após a etapa de pré-processamento, vem as etapas de segmentação e representação na qual o principal objetivo é de simplificar e extrair as características de uma imagem para facilitar sua análise na próxima etapa do processo.

Enquanto a segmentação é usada para localizar os objetos e as formas na imagem a representação busca como o próprio nome já diz representar à essência da imagem, seja pela cor, pela forma dos objetos, etc. Assim o resultado dessas duas etapas pode ser usado diretamente nos algoritmos de classificação.

5.4 Treinar Algoritmo

Com as imagens pré-processadas e segmentadas, temos então um treinamento dos algoritmos. Nessa etapa as imagens são concatenadas com suas respectivas descrições, o que possibilita o algoritmo de ler os dados para treino e saber de qual classe a imagem se trata, sendo assim torna-se possível classificar uma nova imagem na fase de testes.

5.5 Testar

Com o algoritmo já treinado, são separadas novas imagens que passarão pelas etapas de redimensionamento, pré-processamento e segmentação assim como as imagens de treino. Após as etapas o algoritmo é capaz de ler os dados da imagem e assim classificar de acordo com a classe que os dados mais se assimilaram.

6 Casos de Teste

Para validar os algoritmos citados foram realizados testes utilizando imagens em padrão RGB providas pela câmera de um celular. Estas imagens são compostas por objetos de interesse ou não, dispostos em um mesmo plano. Foram definidas três categorias de testes, como pode ser visto na Tabela 1. Além disso, foram utilizadas duas configurações diferentes nos algoritmos, sendo uma configuração padrão com variáveis de entrada com valores já pré-estabelecidos pela biblioteca *scikit-learn* e outra com valores ajustados, na tentativa de melhorias os resultados obtidos.

Tabela 1 - Especificação das Categorias

Categoria	Categoria 1	Categoria 2	Categoria 3
Frutas diferentes	1	1	3
Frutas Verdes e Maduras	Não	Sim	Sim

Na intenção de se ter uma base consistente de resultados, cada categoria foi testada dez vezes em cada configuração, ou seja, em cada cena de teste foram feitos dez treinos e dez testes.

6.1 Base de Dados

Para um melhor entendimento de como foram realizados os testes, esta seção tem como finalidade deixar de forma explícita os detalhes das bases de dados de imagens utilizadas no processo, assim como os detalhes de alguns valores e constantes utilizados nos algoritmos.

6.1.1 Composição da Base de Imagens de Treino

Para compor uma base alguns critérios foram adotados, sendo assim, foram utilizadas 10 frutas de cada tipo e pra cada fruta 50 imagens diferentes. Ou seja, para compor a base de imagens de bananas maduras foram utilizadas 10 bananas e pra cada uma dessas bananas foram registradas 50 imagens em diferentes ângulos e rotações, totalizando uma quantidade de 500 imagens. Sendo assim, esse processo foi repetido para todos os outros tipos utilizados, sendo eles banana verde e madura, maçã verde e madura e tomate verde e maduro.

6.1.2 Composição da Base de Imagens de Teste

Como a intenção do trabalho era de identificar e classificar diferentes frutas, para compor as imagens da base de teste teve-se o cuidado de não repetir as frutas utilizadas na base de treino. Ou seja, para compor as imagens de bananas maduras que seriam testadas não foram utilizadas nenhuma das 10 bananas utilizadas para compor a base de treino, teve-se o cuidado de registrar imagens de outra banana. Esse processo também foi repetido para os outros tipos utilizados nas três categorias.

6.2.1 Configuração Padrão dos Algoritmos

Os algoritmos da biblioteca *scikit-learn* vêm com alguns valores padrões em suas variáveis de entrada que podem ser encontrados na documentação da biblioteca, sendo assim nessas seção e na seção 6.2.2 serão informados apenas os valores modificados nesse trabalho, ou seja, caso o valor de alguma variável não esteja entre os itens abaixo o valor utilizado foi o que a própria biblioteca já utiliza como padrão.

- K-NN
 - n_jobs= -1
 - weights= 'distance'
 - n_neighbors=11
- SVC
 - Kernel='linear'
- RF
 - Max_depth=2
- DT
 - Max_depth=2
- GNB
 - Apenas valores padrões da biblioteca.

6.2.2 Configuração Ajustada dos Algoritmos

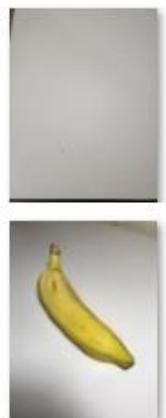
Com os valores obtidos na primeira parte de testes usando os valores padrões das bibliotecas viu-se que alguns algoritmos não obtiveram um resultado muito satisfatório, sendo assim alguns ajustes foram realizados na intenção de aumentar a precisão desses algoritmos em questão. Enquanto isso, com o intuito de confirmar o percentual de acertos dos três melhores algoritmos, foram mantidas as mesmas configurações.

- K-NN
 - Não houve mudança.
- SVC
 - Não houve mudança.
- RF
 - Max_depth=5
 - N_estimators=10
 - Max_features=1
- DT
 - Max_depth=5
- GNB
 - Não houve mudança.

6.2 Categoria 1

A primeira categoria escolhida para teste foi a de identificação da fruta. Nesse caso específico, foram testadas dez imagens em cada um dos algoritmos, sendo cinco imagens de fundo (ou seja, sem a fruta, nosso objeto de interesse) e cinco imagens com a fruta que se deseja identificar. Foram utilizadas três frutas diferentes, sendo elas banana, maçã e tomate, vide figuras 6 e 7 abaixo. Os resultados serão vistos abaixo na Tabela 2.

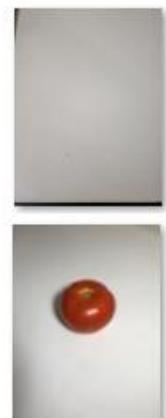
Figura 6 – Exemplos das Frutas da Categoria 1. (A) Imagem de Fundo e Imagem com Banana. (B) Imagem de Fundo e Imagem com Maçã. (C) Imagem de Fundo e Imagem com Tomate.



(A)



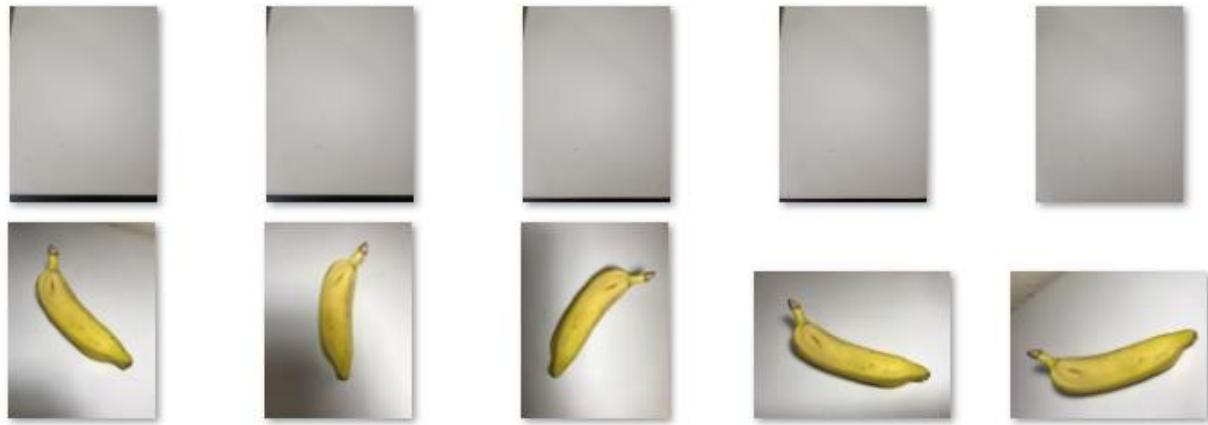
(B)



(C)

Fonte: Elaborado pelo Autor.

Figura 7 - Imagens de Teste Para Identificação de Bananas na Categoria 1.



Fonte: Elaborado pelo Autor.

6.2.1 Resultados Configuração Padrão

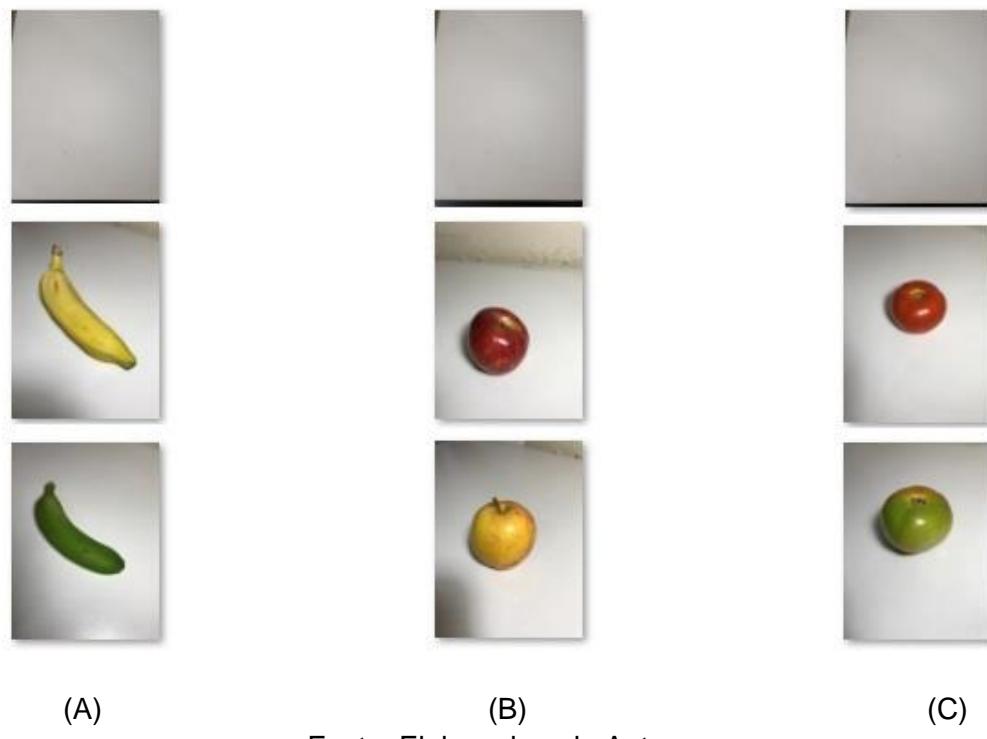
Tabela 2 - Tabela de Resultados Com Configurações Padrão na Categoria 1

Algoritmo	Banana	Maçã	Tomate
K-NN	100%	100%	100%
SVC	100%	100%	100%
RF	100%	100%	100%
DT	100%	100%	100%
GNB	100%	100%	100%

6.3 Categoría 2

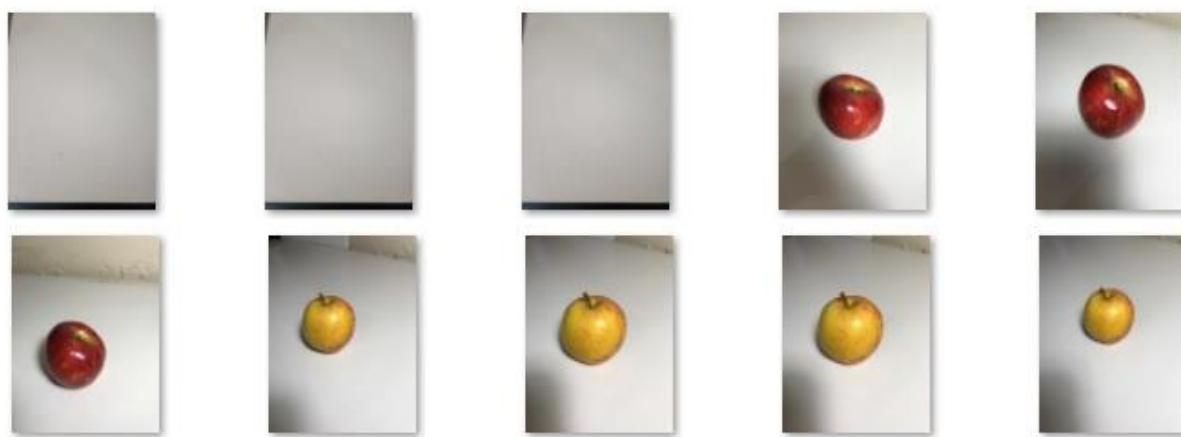
Para a segunda categoria aumentou-se a dificuldade com o intuito de testar a capacidade dos algoritmos classificarem as frutas em verde ou madura. Ou seja, nas dez imagens de testes foram adicionadas também frutas verdes e os algoritmos testados para saber se seriam capazes de identificar quando há fruta na imagem e em caso positivo, dizer se está madura ou verde. Nas figuras 8 e 9 podem ser vistos os exemplos de teste utilizados. Os resultados são vistos nas tabelas 3 e 4.

Figura 8 – Exemplos das Frutas da Categoria 2. (A) Imagem de Fundo, Banana Madura e Banana Verde. (B) Imagem de Fundo, Maçã Madura e Maçã Verde. (C) Imagem de Fundo, Tomate Maduro e Tomate Verde.



Fonte: Elaborado pelo Autor.

Figura 9 - Imagens de Teste Para Identificação e Classificação de Maçãs na Categoria 2.



Fonte: Elaborado pelo Autor.

6.3.1 Resultados Configuração Padrão

Tabela 3 - Tabela de Resultados Com Configurações Padrões na Categoria 2

Algoritmo	Banana	Maçã	Tomate
K-NN	90%	100%	100%
SVC	90%	100%	100%
RF	76%	99%	85%
DT	76%	99%	85%
GNB	80%	100%	100%

6.3.2 Resultados Configuração Ajustada

Tabela 4 - Tabela de Resultados Com Configurações Ajustadas na Categoria 2

Algoritmo	Banana	Maçã	Tomate
K-NN	90%	100%	100%
SVC	90%	100%	100%
RF	92%	100%	97%
DT	92%	100%	97%
GNB	80%	100%	100%

6.4 Categoria 3

A terceira categoria consiste numa junção das duas categorias acima com todos os três tipos de frutas. Sendo assim, dessa vez foram passadas 25 imagens para o teste com as frutas em seus estados maduro e verde e também imagens de fundo. Essa categoria foi planejada com a ideia de testar como os algoritmos se comportariam no caso de frutas diferentes mas com formato e coloração parecidos. Os resultados podem ser vistos na Tabela 5 e 6.

6.4.1 Resultados Configuração Padrão

Tabela 5 - Tabela de Resultados Com Configurações Padrões na Categoria 3

Algoritmo	Percentual de Acertos
K-NN	96%
SVC	92%
RF	55,2%
DT	54,8%
GNB	80%

6.4.2 Resultados Configuração Ajustada

Tabela 6 - Tabela de Resultados Com Configurações Ajustadas na Categoria 3

Algoritmo	Percentual de Acertos
K-NN	96%
SVC	92%
RF	83,6%
DT	83,6%
GNB	80%

6.5 Análise dos resultados

Com o final dos testes e todos os resultados obtidos e calculados, pode-se observar certas características e certos padrões peculiares entre os algoritmos que serão detalhados mais abaixo.

De um modo geral, na primeira categoria, todos obtiveram 100% de acerto em identificar quando havia fruta e quando não havia fruta na imagem nas duas configurações diferentes.

Já na segunda categoria, foi possível perceber que os algoritmos K-NN, SVC e GNB tiveram problemas com os testes quando a fruta era banana, alcançando um índice de 90%, 90% e 80% de acerto, respectivamente, enquanto os algoritmos RF e DT tiveram maiores dificuldades em todas as três frutas na primeira configuração de testes, já que não alcançaram 100% de acertos em nenhuma delas, mas obtiveram melhora ao se ajustar os dados de entrada dos algoritmos.

Na terceira e última categoria pode-se perceber que o algoritmos K-NN e SVC obtiveram os melhores resultados, 96% e 92%, respectivamente. Enquanto isso os algoritmos RF e DT tiveram as piores performances em ambas as configurações, tendo inicialmente obtido 55,2% e 54,8%, respectivamente, e 83,6% após os ajustes de entrada. Enquanto isso, o algoritmo GNB conseguiu um percentual de 80% de acertos, o que o fez sair da terceira para a ultima posição entre os melhores percentuais após os ajustes feitos em RF e DT.

Abaixo serão discutidas mais detalhadamente como cada algoritmo se comportou. Vale ressaltar que a Categoria 1 não será detalhada pois obteve resultados perfeitos em todos os testes.

Começando com o algoritmo K-NN, apesar de ter tido um alto índice de acertos no total das 3 categorias, pode-se reparar que ele obteve problemas quando a fruta era do tipo banana. Ao analisar quais as imagens o algoritmo mais errou (última banana da Figura 7), foi possível reparar juntamente com as imagens de treino, que a posição da fruta pode ser um fator diferencial para o algoritmo na hora de treinar e classificar as imagens e objetos de interesse.

Continuando a análise, o algoritmo SVC também obteve um bom percentual de acertos nas 3 categorias, porém foi possível reparar que assim como o algoritmo K-NN, dependendo do ângulo da fruta, o resultado foge do esperado, o que ocorreu em dois casos específicos (ao classificar uma banana e ao classificar um tomate). Apesar disso, com alguns ajustes, espera-se que a ferramenta se torne mais poderosa nesse tipo de atividade.

Com relação ao algoritmo RF e ao DT, os dois mostraram que para obter bons resultados precisa-se de um melhor estudo sobre o funcionamento do algoritmo, ou seja, quando se testou com as configurações padrões já vindas da biblioteca, os resultados deixaram a desejar principalmente na categoria 3, porém ao serem feitas algumas mudanças nas variáveis de entrada que os algoritmos utilizam, ambos obtiveram melhora considerável em todas as frutas, apesar de não atingirem 100% na categoria.

E por último o algoritmo GNB mostrou que nas duas primeiras categorias quando a quantidade de frutas era de apenas uma, sendo madura e verde, houve o mesmo problema dos algoritmos K-NN e SVC, pois percebeu-se maiores dificuldades ao classificar bananas enquanto as outras frutas tenham obtiveram 100% de sucesso. Mas ao aumentar a dificuldade como foi feito na categoria 3, o algoritmo passou a agir de forma mais aleatória, apesar de acertar 80% dos casos, foi possível perceber que todas as frutas tiveram erros em algum teste realizado.

7 Conclusão

Com base nas pesquisas e estudos realizados, conforme referencial bibliográfico, este trabalho propôs um processo capaz de utilizar e avaliar cinco algoritmos de reconhecimento de padrão no reconhecimento de imagens. Neste processo foi possível descrever e avaliar esses cinco algoritmos usando casos de teste para reconhecer e classificar frutas com formas e cores variadas em uma imagem.

Deste modo, o objetivo geral desta dissertação foi alcançado, assim como os seus objetivos específicos. A eficiência do método proposto pode ser demonstrada por meio de um estudo de caso com três categorias de testes, o que permitiu que o processo fosse avaliado em graus de dificuldades diferentes.

Com os resultados obtidos, foi possível perceber que o algoritmo K-NN foi o que obteve os melhores resultados, tendo alcançado 96% de acertos na ultima categoria de testes, porém como foi visto com os ajustes feitos nos algoritmos RF e DT, é esperado que com ajustes, todos os algoritmos possam chegar à melhores resultados.

O desenvolvimento desta pesquisa apresentou algumas dificuldades, sendo as principais delas relacionadas a parte de pré-processamento e extração de recursos nas imagens e ao aumento da base de dados, pois a preocupação em tratar a informação para melhorar o desempenho computacional e conseguir gerar resultados satisfatórios levou à investigação de diferentes técnicas. No entanto, essa busca por melhores soluções enriqueceu o desenvolvimento do trabalho.

De forma geral, a especificação de um processo de natureza simples, eficiente e flexível é relevante para auxiliar e incentivar trabalhos futuros e assim ampliar a quantidade de situações e problemas capazes de serem beneficiados e resolvidos através de recursos computacionais. Com isso, as principais contribuições deste trabalho são:

- Levantamento, avaliação e documentação das etapas do processamento digital de imagens;
- Especificação e implementação de um processo de reconhecimento e classificação de imagens;
- Implementação de protótipo de software de reconhecimento e classificação de frutas em imagens;
- Avaliação de precisão de cinco algoritmos diferentes no reconhecimento e classificação de frutas;

Entre os trabalhos futuros, considerados para a extensão deste projeto, estão:

- Avaliar os algoritmos em diferentes casos de testes, com maiores bases de dados;
- Propor implementação de protótipo para cenários reais;

- Melhorar o processo e implementá-lo em uma arquitetura de hardware;
- Avaliar o desempenho dos algoritmos (custo de processamento, tempo de treinamento e tempo de execução).

8 Referências Bibliográficas

ARIVU, C. V. G.; PRAKASH, G.; SARMA, A. S. S. Online image capturing and processing using vision box hardware: Apple grading. International Journal of Modern Engineering Research (IJMER), v. 2, p. 639–634, 2012.

AWCOCK, G. J.; THOMAS, R. Applied Image Processing. McGraw-Hill. 1996.

BARELLI, F. Introdução à Visão Computacional. Casa do Código. São Paulo - SP. 2018.

BEZERRA, M. E. R. Métodos Baseados na Regra do Vizinho Mais Próximo para Reconhecimento de Imagens. Dissertação (Dissertação em Engenharia da Computação) – Escola Politécnica de Pernambuco. Recife. 2006.

BOSER, B.; GUYON, I.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. Annual Workshop on Computational Learning Theory, Pittsburgh. Proceedings...San Mateo, CA: Morgan Kayfmann, p. 144-152. 1992.

BRAMLEY, R. G. V. Lessons from nearly 20 years of Precision agriculture research, development, and adoption as a guide to its appropriate application. Crop and Pasture Science, 60 (3), pp. 197-217, 2009.

BREIMAN, L. Bagging Predictors. Machine Learning, Vol. 24, páginas 123-140. Kluwer Academic Publishers, Holanda, 1996.

BREIMAN, L. Random Forest. In Journal of Machine Learning, Vol.45, páginas 5-32. Kluwer Academic, Holanda, 2001.

BUSSE, L. J.; CRIMMINS, T. R.; FIENUP, J. R. A model based Approach to Improve the Perfomance of the Geometric Filtering Speckle Reduction Algorithm. IEEE Ultrasonics Symposium 95CH35844, páginas 1353-6. 1995.

CARVALHO, J. N.; GURJÃO, E. C.; Mata, M. E. R. M. C.; Duarte, M. E. M. Classificação Automática de Frutas por Análise de Imagem - O Caso da Manga Tommy Atkins. WTA 2014 - VIII Workshop de Tecnologia Adaptativa. São Paulo. 2014.

COMANICIU, D.; RAMESH, V.; MEER, P. Kernel-based object tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, n. 5, páginas 564-577, maio 2003.

CORTES, C.; VAPNIK, V. Support vector networks. Machine Learning. v. 20, n. 3, p. 273-297, 1995.

DA SILVA, F.M.; MORAIS, A. M. M. B. Boas Práticas Pos-Colheita para Frutos Frescos, ESB/UCP/ORGAL, Porto, Portugal, 2000.

DINIZ, F. A; MENDES NETO, F. M; LIMA JÚNIOR, F. C; FONTES, L. M. de O. RedFace: um sistema de reconhecimento facial baseado em técnicas de análise de componentes principais e autofaces: comparação com diferentes classificadores. Revista Brasileira de Computação Aplicada (ISSN 2176-6649), Passo Fundo, v. 5, n. 1, p. 42-54, abr. 2013.

ESQUEF, I. A.; ALBUQUERQUE, Marcelo Portes de; ALBUQUERQUE, Márcio Portes de. Processamento Digital de Imagens 2003 (Artigo de Divulgação Científica).

FACON, J. Morfologia Matemática: Teoria e Exemplos. J. Facon Editor, Curitiba-PR, 1996.

FELICIANO, F.; DE SOUZA, L. I.; LETA, F. (2005). VISÃO COMPUTACIONAL APLICACADA À METROLOGIA DIMENSIONAL AUTOMATIZADA: CONSIDERAÇÕES SOBRE SUA EXATIDÃO. Engevista.

FIGUEIRÊDO, R. de C. Um estudo de algoritmos de processamento de imagem e reconhecimento de padrão em imagem digital capturada em ambiente aberto. Dissertação (Pós-Graduação em Mecatrônica). UFBA. Salvador. 2007.

FISCHLER, M.A.; FIRSCHEIN, O. Intelligence, the eye, the brain and the computer. Addison Wesley, 1987.

FRICK, M. A. D. Caracterização de minério de ferro por visão computacional. 2008. Dissertação (Mestrado em Engenharia de Produção) – Centro de Tecnologia, Universidade Federal de Santa Maria, Santa Maria, 2008.

GHOSAL V.; TIKMANI P.; GUPTA P. Face Classification Using Gabor Wavelets and Random Forest, In Proceedings of the Canadian Conference on Computer and Robot Vision (CRV '09), IEEE Computer Society, Washington, DC, USA, páginas 68-73, 2009.

GONZALEZ, R. C.; WOODS, R. E. Processamento Digital de Imagens. third. São Paulo: Pearson Prentice Hall, 2010.

HAN, J; KAMBER, M. Data Mining: Concepts and Techniques. Elsevier, 2006.

HAYKIN, S. Neural Networks: A Comprehensive Foundation. 2. ed. New Jersey, Prentice Hall, 1999.842 p.

HAYKIN, S. Redes Neurais: princípios e prática. 2.ed. Porto Alegre: Bookman, 2001.

HERWITZ, S. R.; JOHNSON, L. F.; DUNAGAN, S. E.; HIGGINS, R. G.; SULLIVAN, D. V.; ZHENG, J.; LOBITZ, B. M.; LEUNG, J. G.; GALLMEYER, B. A.; AOYAGI, M.; SLYE, R. E.;

BRASS, J. A. Imaging from an unmanned aerial vehicle: agricultural surveillance and decision support. *Computers and Electronics in Agriculture*. Volume 44, Issue 1, July 2004, Pages 49–61.

JAIN, A. K. *Fundamentals of Digital Image Processing*. 1.Ed. Prentice Hall. Red Star Teodoro Sampaio. São Paulo - SP. 1989.

MANICA, I.; ICUMA, I.M.; MALAVOLTA, E.; RAMOS, V.H.V.; OLVEIRA, M.E.; CUNHA, M.M.; JUNQUEIRA, N.T.V. *Tecnologia, produção, agroindústria e exportação da manga*. Ed. Cinco Continentes, Porto Alegre – RS, 2001.

MARR, D. *Vision: a Computational Investigation into the Human Representation and Processing of Visual Information*. New York: W. M. Freeman and Company, 1982.

MASCARENHAS, N.D.A.; VELASCO, F.D.R. *Processamento Digital de Imagens*. 2^a edição. IV Escola Brasileiro-Argentina de Informática, Termas de Rio Hondo, Argentina, 1989.

MELO, R. C. Automação do processo de classificação de manga (*Mangifera indica Linn*) cv. Palmer por meio dos descritores cor e calibre. *Dissertação (Dissertação de Mestrado em Engenharia Mecatrônica)*. UFRPE. Recife. 2015.

PEDRINI, H.; SCHWARTZ, W. R. *Analise de Imagens Digitais: Princípios, Algoritmos e Aplicações*. São Paulo: Thomson Learning, 2008.

PEREIRA, L. A. A. *Classificação automática de áreas cafeeiras em imagens de satélite, utilizando Redes Neurais Artificiais*. *Dissertação de Mestrado*. Lavras – MG. UFLA, 2013.

PEREIRA, R. C. *Técnica de rastreamento e perseguição de alvo utilizando o algoritmo Haar Cascade aplicada a robôs terrestres com restrições de movimento*. *Dissertação (Dissertação de Mestrado em Engenharia Mecatrônica)*. UFRN. Natal. 2017.

PINTO, P. R. *Utilizando Métodos Exatos no Reconhecimento e Rastreamento de Objetos*. *Dissertação (Dissertação em Engenharia de Computação) – CEFET-MG*. Timóteo. 2015.

SANTOS, A. P. O. *Desenvolvimento de Descritores de Imagens para Reconhecimento de Padrões de Plantas Invasoras (Folhas Largas e Folhas Estreitas)*. *Dissertação de mestrado, Programa de Pós Graduação em Ciência da Computação, Universidade Federal de São Carlos*, 2009.

SHIRAI, Y. *Three Dimensional Computer Vision*. Springer-Verlag, 1987.

SILVA, L. N. C. *Análise e síntese de estratégias de aprendizado para redes neurais artificiais*. 1998. 210 p. *Dissertação (Mestrado em Engenharia Elétrica) - Universidade Estadual de Campinas, Campinas*.

SIMÕES, A. S.; COSTA, A. H. R. Classificação de Laranjas Baseada em Padrões Visuais. 6º Simpósio Brasileiro de Automação Inteligente (SBAI). Sociedade Brasileira de Automática. Bauru, setembro 14-17, páginas 77-81.

SIQUEIRA, M. L. Reconhecimento Automático de Padrões em Imagens Ecocardiográficas. Dissertação (Tese parcial de Doutorado em Ciência da Computação). UFRS. Porto Alegre. 2010.

SOLOMON, C.; BRECKON, T. Fundamentos de Processamento Digital de Imagens. LTC. São Paulo. 2011.

SONKA, M.; HLAVAC, V.; BOYLE, R. Image Processing, Analysis and Machine Vision. Thomson-Engineering. 1998.

VAPNIK, V.N. The Nature of Statistical Learning Theory. New York, Springer-Verlag, 332 p. 1995.

WEBB, A. R. Statistical Pattern Recognition. 2. ed. Chichester: John Wiley & Sons, 2002.