



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS SOBRAL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO (PPGEEC)
MESTRADO ACADÊMICO EM ENGENHARIA DE COMPUTAÇÃO

PATRÍCIA DIAS ARAÚJO

ANÁLISE E CLASSIFICAÇÃO DA QUALIDADE DO *LINTER* E DO ÓLEO DE
ALGODÃO UTILIZANDO TÉCNICAS DE VISÃO COMPUTACIONAL

SOBRAL

2018

PATRÍCIA DIAS ARAÚJO

ANÁLISE E CLASSIFICAÇÃO DA QUALIDADE DO *LINTER* E DO ÓLEO DE ALGODÃO
UTILIZANDO TÉCNICAS DE VISÃO COMPUTACIONAL

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PPGEEC) do Campus Sobral da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica e de Computação. Área de Concentração: Sistemas de Informação. Linha de Pesquisa: Algoritmos e Computação Distribuída.

Orientador: Prof. Dr. Jarbas Joaci de Mesquita Sá Junior

SOBRAL

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

A69a Araújo, Patrícia.

ANÁLISE E CLASSIFICAÇÃO DA QUALIDADE DO LÍINTER E DO ÓLEO DE ALGODÃO
UTILIZANDO TÉCNICAS DE VISÃO COMPUTACIONAL / Patrícia Araújo. – 2019.
87 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Campus de Sobral, Programa de Pós-Graduação
em Engenharia Elétrica e de Computação, Sobral, 2019.

Orientação: Prof. Dr. Jarbas Joaci de Mesquita Sá Junior.

1. Reconhecimento de padrões, Análise de textura, Óleo de algodão, línter.. I. Título.

CDD 621.3

PATRÍCIA DIAS ARAÚJO

ANÁLISE E CLASSIFICAÇÃO DA QUALIDADE DO *LINTER* E DO ÓLEO DE ALGODÃO
UTILIZANDO TÉCNICAS DE VISÃO COMPUTACIONAL

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PPGEEC) do Campus Sobral da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica e de Computação. Área de Concentração: Sistemas de Informação. Linha de Pesquisa: Algoritmos e Computação Distribuída.

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. Jarbas Joaci de Mesquita Sá
Junior (Orientador)
Universidade Federal do Ceará (UFC / Campus de
Sobral)

Prof. (a). Dra. Fátima Nelsizeuma Sombra de
Medeiros
Universidade Federal do Ceará (PPGETI / UFC)

Prof. Dr. Iális Cavalcante de Paula Júnior
Universidade Federal do Ceará (UFC / Campus de
Sobral)

Aos meus pais Josué Monteiro Araújo e Maria Francisca Dias Araújo por todo apoio e amor incondicional e aos meus irmãos John Lennon Dias Araújo e Jorlene Dias Araújo Trablusi por serem meus amigos e cúmplices.

AGRADECIMENTOS

Nada nessa vida conquistamos sozinhos. Às vezes é preciso que alguém nos motive, que acredite em você mais do que você mesmo, uma palavra amiga que nos faz acreditar que no final tudo vai dar certo. Tenho muito o que agradecer:

À Deus por sua presença constante que me fez superar todos os momentos difíceis.

A minha família por se fazer presente mesmo estando distante, em especial meus pais por suas ligações diárias, demonstração de carinho, preocupação e amor. Não tenho dúvidas que vocês são os melhores que Deus poderia me dar.

Ao meu amigo Jonas por ter me incentivado a fazer o mestrado.

Ao Prof. Dr. Jarbas Joaci de Mesquita Sá Junior por me orientar em minha dissertação de mestrado, compartilhar seu conhecimento e por toda sua paciência. Muito obrigada!

Ao Prof. Dr. Iális Cavalcante de Paula Júnior, por ser um bom professor e por seu auxílio durante o período do mestrado, tenho você como exemplo.

À Universidade Federal do Ceará, seus professores e funcionários que foram alicerces para a realização deste trabalho.

Aos colegas do mestrado, em especial Maria Victoria, Medeiros Neto, Livia Sousa, Lucivando Ribeiro, Miquéias Araújo e Alânio Ferreira por todas as palavras de apoio e incentivo.

Ao meu namorado Artur Almeida por seu apoio e compreensão, por todos os momentos em que acreditou em mim e me fez acreditar que eu era capaz.

Aos irmãos Lira, Raquel, Bruno e Victor pelo carinho e por me receberem sempre que precisei de estadia.

À Pâmela, minha amiga e irmã, a sua amizade fez os dias difíceis serem mais fáceis.

Ao Senhor Mena e Dona Antônia por todas as palavras de apoio e incentivo.

Ao meu amigo Elton, por todo apoio, incentivo e orações.

À Indústria Rações Golfinho pelo auxílio na pesquisa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento da pesquisa de mestrado via bolsa de estudos.

A todos que contribuíram de forma direta e indireta para a realização deste trabalho, o meu muito obrigada.

“Em seu coração o homem planeja o seu caminho,
mas o Senhor determina os seus passos.”

(Provérbios 16:9)

RESUMO

Os métodos de classificação utilizados em processos industriais estão em constante aprimoramento e um dos fatores está no alinhamento tecnológico em planta fabril. Porém, muitas tecnologias não possuem custos viáveis para implantação em chão de fábrica. Um exemplo são os métodos utilizados para análises de plumas e fibras curtas de algodão (*linter*), os quais utilizam de tecnologia aplicada e estão disponíveis no Brasil apenas em alguns laboratórios de análise que atendem às normas brasileiras de qualidade, o que torna a produção mais cara e demanda maior tempo para disponibilização do produto ao mercado consumidor. O *linter* é um subproduto do algodão, matéria-prima de larga aplicabilidade, utilizado na fabricação do papel-moeda, algodão hidrofílico, placas de circuito impresso e tecidos cirúrgicos. A importância de sua classificação vem da valorização da qualidade do produto no mercado interno e externo, os quais apresentam diferentes padrões, levando a uma produção da matéria-prima com características de qualidade distintas. Outro subproduto do algodão que necessita de análises laboratoriais durante seu processo de fabricação é o óleo de algodão, que tem como característica coloração intensa em seu estado bruto e que, quando usado para fins alimentícios, passa por procedimentos de clareamento e redução de acidez, os quais podem ser realizados pelo processo de refino ou semirrefino. Neste trabalho foi estudado o processo de semirrefino, no qual são analisadas concentrações de sabões, qualidade e estabilidade do óleo, verificando ao fim do processo a ausência de carotenoides e determinando se está apropriado para consumo. Considerando a importância da classificação do *linter* e a técnica do semirrefino do óleo de algodão, este trabalho propõe desenvolver metodologias de análise semiautomática para a classificação de ambos os casos, proporcionando velocidade no envio dos produtos para o mercado consumidor, obtenção de dados confiáveis e alinhamento entre os setores de produção e qualidade. Assim, são propostas duas metodologias utilizando processamento digital de imagens e reconhecimento de padrões. Na classificação do *linter* para obtenção de características foram aplicados os descritores de textura e comparados diferentes classificadores. Na classificação das etapas do semirrefino foram utilizados diferentes modelos de cores para extração de características, que foram agrupadas pelos classificadores. Com este estudo, resultados promissores foram obtidos tanto para a classificação do processo de semirrefino do óleo de algodão como do *linter*.

Palavras-chave: Reconhecimento de padrões, Análise de textura, Óleo de algodão, *linter*.

ABSTRACT

The classification methods applied to industrial processes are in constant enhancement and one of its factors is the technological alignment in industrial plant. However, many technologies are not economically practicable for factory floor implantation. As an example the methods utilized on feathers and cotton short fibers (linter) analysis which use applied technology and are only available in a few Brazilian analysis laboratories that comply with national quality standards, leading to higher production costs and longer periods to make the product available on market. Linter is a cotton's byproduct, raw material of wide applicability, used in the manufacture of currency paper, hydrophilic cotton, printed circuit boards and surgical fabrics. Its classification importance comes from the appreciation of product quality in the internal and external market, which present different standards, leading to a raw material production with different quality characteristics. Another byproduct that requires laboratory examination during its manufacture is the cotton oil, characterized by intense coloring in its raw state, and, when it is used for alimentary purposes, goes through a lightening procedure and acidity reduction, which can be carried out by the refining or semi-refining processes. This study analyzes the semi-refining process, in which are analyzed the soaps concentrations, oil's quality and stability, verifying at the end of the process the carotenoids absence and determining if it is suitable for consumption. Considering the significance of the linter categorization and the cotton oil semi-refining technique, this study proposes the semi-automatic analysis methodologies development for both classification methods, providing readiness in the delivery of products to the consumer market, obtaining reliable data and alignment between production and quality sectors. Thus, two methodologies are proposed using digital image processing and pattern recognition. In the linter classification, texture descriptors were applied to obtain characteristics and different classifiers were compared. In the classification of the semi-refining stages, different color models were used to extract characteristics, which were grouped by classifiers. With this study, promising results were obtained both for the classification of cotton oil semi-refining process and linter.

Keywords: Pattern recognition, Texture analysis, Cotton oil, linter.

LISTA DE TABELAS

Tabela 1 – Níveis de classificação do <i>linter</i> conforme sua área de contaminação.	28
Tabela 2 – Esquema de um neurônio.	58
Tabela 3 – Classificação do <i>linter</i> em duas classes (boa qualidade e qualidade ruim), classificador KNN e validação <i>leave-one out</i>	69
Tabela 4 – Classificação do <i>linter</i> em duas classes (boa qualidade e qualidade ruim), classificador QDA e validação <i>leave-one out</i>	70
Tabela 5 – Classificação do <i>linter</i> em duas classes (boa qualidade e qualidade ruim), classificador LDA e validação <i>leave-one out</i>	70
Tabela 6 – Classificação do <i>linter</i> em cinco classes (alta qualidade Japão, média quali- dade Japão, padrão China, baixa qualidade e péssima qualidade), classificador KNN e validação <i>leave-one out</i>	71
Tabela 7 – Classificação do <i>linter</i> em cinco (alta qualidade Japão, média qualidade Japão, padrão China, baixa qualidade e péssima qualidade), classificador QDA e validação <i>leave-one out</i>	71
Tabela 8 – Classificação do <i>linter</i> em cinco classes (alta qualidade Japão, média quali- dade Japão, padrão China, baixa qualidade e péssima qualidade), classificador LDA e validação <i>leave-one out</i>	71
Tabela 9 – Matriz de confusão, para KNN = 4, método <i>Leave-one-out</i> e modelo de cor RGB.	72
Tabela 10 – Matriz de confusão, para KNN = 4, método <i>Leave-one-out</i> e modelo de cor YIQ.	72
Tabela 11 – Matriz de confusão, com classificador QDA, método <i>Leave-one-out</i>	73
Tabela 12 – Matriz de confusão utilizando ELM, método K-fold e os modelos de cores RGB e YIQ.	73
Tabela 13 – Acurácia média da classificação dos estados de semirrefino do óleo de algo- dão, conforme classificador e modelo de cor aplicados.	74

LISTA DE ABREVIATURAS E SIGLAS

<i>KNN</i>	<i>K Vizinhos Mais Próximos, K Nearest Neighbors</i>
<i>LBP</i>	<i>Padrão Binário Local, Local Binary Pattern</i>
<i>LDA</i>	<i>Análise por Discriminante Linear, Linear Discrimination Analysis</i>
<i>MLP</i>	<i>Perceptron Multicamadas, MultiLayer Perceptron</i>
<i>QDA</i>	<i>Análise por Discriminante Quadrático, Quadratic Discrimination Analysis</i>
<i>RGB</i>	<i>Vermelho, Verde, Azul, Red, Green, Blue</i>
ABRAPA	Associação Brasileira dos Produtores de Algodão
AGOPA	Associação Goiana dos Produtores de Algodão
AMPA	Associação Maranhense dos Produtores de Algodão
AMPASUL	Associação Sul-Mato-Grossense dos Produtores de Algodão
CBRA	Centro Brasileiro de Referência em Análise de Algodão
CIDAO	Companhia Industrial de Algodão e Óleos
DF	Dimensão Fractal
DFT	Transformada Discreta de Fourier
ELM	Máquina de Aprendizagem Extrema, <i>Extreme Learning Machine</i>
FIALGO	Fundo de Incentivo à Cultura do Algodão
HVI	Instrumento de Alto Volume, <i>High Volume Instrument</i>
NIR	Espectroscopia no Infravermelho Próximo, <i>Near Infrared Spectroscopy</i>
PIB	Produto Interno Bruto
PLS-DA	<i>Análise Discriminante pelos Mínimos Quadrados Parciais, Partial Least Square Discriminant Analysis</i>
RM	Ressonância Magnética
RNA	Rede Neural Artificial
Sampled LFP	<i>Sampled Local Fuzzy Pattern</i>
SIMCA	<i>Modelagem Independente Flexível por Analogia de Classe, Soft Independent Modeling by Class Analogy</i>
TAPPI	<i>Associação Técnica da Indústria de Celulose e Papel, Technical Association of the Pulp and Paper Industry</i>
USDA	Departamento de Agricultura dos Estados Unidos
UV-vis	Ultravioleta-visível

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Objetivos do Trabalho	16
1.2	Organização da Dissertação	16
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	Óleo Vegetal de Algodão	18
2.1.1	<i>Processo de Semirrefino do Óleo de Algodão</i>	18
2.1.1.1	<i>Estado Bruto</i>	19
2.1.1.2	<i>Primeiro Refino</i>	19
2.1.1.3	<i>Segundo Refino</i>	19
2.1.1.4	<i>Lavagem</i>	19
2.1.1.5	<i>Secagem do Óleo</i>	20
2.1.1.6	<i>Branqueamento</i>	20
2.1.2	<i>Métodologias Utilizadas para Classificação do Óleo de Algodão</i>	20
2.1.2.1	<i>Métodos de Análise</i>	20
2.1.2.2	<i>Near Infrared Spectroscopy - NIR</i>	21
2.1.2.3	<i>Ultravioleta Visível</i>	22
2.1.2.4	<i>Modelos de Classificação</i>	23
2.2	<i>Linter</i>	24
2.2.1	<i>Métodos Utilizados Para Análise de Qualidade do linter</i>	26
2.2.1.1	<i>Classificação Manual / Visual</i>	26
2.2.1.2	<i>Classificação Tecnológica por HVI</i>	27
2.2.1.3	<i>Classificação do Linter pelo princípio de TAPPI T213</i>	28
2.3	Visão Computacional	29
2.3.1	<i>Aquisição da Imagem e Digitalização</i>	30
2.3.2	<i>Pré-processamento</i>	31
2.3.2.1	<i>Modelos de Cores</i>	31
2.3.2.2	<i>Modelo de Cores RGB</i>	32
2.3.2.3	<i>Modelo de Cores YIQ</i>	33
2.3.3	<i>Segmentação</i>	34
2.3.4	<i>Extração de Características</i>	34

2.3.5	<i>Classificação e Reconhecimento de Padrões</i>	35
2.3.6	<i>K-Nearest Neighbors (K-NN)</i>	35
2.3.7	<i>Linear Discriminant Analysis (LDA)</i>	36
2.3.8	<i>Quadratic Discriminant Analysis - QDA</i>	37
2.3.9	<i>Análise de Textura</i>	37
2.3.10	<i>Descritores de Textura</i>	38
2.3.10.1	<i>Matriz de Coocorrência</i>	38
2.3.10.2	<i>Descritores de Haralick</i>	39
2.3.10.3	<i>Dimensão Fractal</i>	42
2.3.10.4	<i>Aplicação da dimensão fractal em análise de textura</i>	44
2.3.10.5	<i>Lacunaridade</i>	45
2.3.10.6	<i>Local Binary Pattern - (LBP)</i>	46
2.3.10.7	<i>Local Fuzzy Pattern - (LFP)</i>	47
2.3.10.8	<i>Filtros de Gabor</i>	51
2.3.10.9	<i>Descritores de Fourier</i>	52
2.3.11	<i>Método de Validação</i>	54
2.3.11.1	<i>Rede Neural</i>	54
2.3.11.2	<i>Extreme Learning Machine - (ELM)</i>	55
2.3.11.3	<i>K-fold e Leave-one-out</i>	56
2.3.11.4	<i>Métricas para Avaliar os Classificadores</i>	57
2.3.12	<i>Aplicação de Análise de Textura em Inspeção Industrial</i>	59
3	METODOLOGIA	60
3.1	CLASSIFICAÇÃO DO LINTER	60
3.1.1	<i>Amostras</i>	61
3.1.2	<i>Aquisições de imagens e pré-processamento</i>	61
3.1.3	<i>Extração das características</i>	62
3.1.3.1	<i>Parâmetros dos descritores de textura</i>	62
3.1.4	<i>Classificadores</i>	65
3.2	CLASSIFICAÇÃO DO ÓLEO VEGETAL	65
3.2.1	<i>Aquisições de imagens e pré-processamento</i>	66
3.2.2	<i>Extração das características</i>	67
3.2.3	<i>Classificação</i>	68

3.3	Considerações finais	68
4	RESULTADOS	69
4.1	Classificação do <i>Linter</i>	69
4.2	Classificação do Óleo de Algodão	72
5	CONCLUSÕES	75
5.1	Trabalhos futuros:	76
5.2	Trabalhos publicados	76
	REFERÊNCIAS	77

1 INTRODUÇÃO

O cultivo de algodão representa uma das atividades tradicionais do Brasil desde o século XIX, onde empresas como a Embrapa Algodão e a Companhia Industrial de Algodão e Óleos (CIDAÓ) deram início a estudos e pesquisas visando obter melhorias na qualidade do produto e na sua aplicação. Desse modo, mediante geração e transferência de tecnologias, obteve-se o desenvolvimento da cotonicultura (VAINSENCHE, 2009) e (VAINSENCHE, 2012).

O desenvolvimento tecnológico junto com o aumento da produtividade agrícola proporcionaram a transformação do Brasil de principal importador mundial de algodão para o terceiro maior exportador do produto em 12 anos (DIAS, 2016). Atualmente está entre os cinco maiores produtores de algodão, atrás da Índia, da China, dos EUA e do Paquistão (TEIXEIRA; NETO, 2017). Chegando ao Produto Interno Bruto (PIB) do setor algodoeiro para o ano agrícola 2016/2017 em um valor estimado de 74,11 bilhões de dólares (UNIVERSOAGRO, 2017).

Através da colheita do algodão obtêm-se as plumas de algodão, as amêndoas e o *linter*, que são fibras curtas aderidas às amêndoas. As plumas são separadas para comercialização na indústria têxtil e as sementes com as fibras curtas são destinadas às indústrias produtoras de torta de algodão (ração animal), óleo vegetal e demais subprodutos (ALGOTEC, 2010).

Um processo importante para o setor algodoeiro é o de beneficiamento, o qual é responsável pela separação das plumas das sementes do algodão e armazenamento de forma a reduzir as contaminações provenientes do campo, mantendo a qualidade do produto colhido. Mesmo com o beneficiamento ainda ocorre da matéria-prima chegar à indústria apresentando impurezas, o que ocasiona perdas na produção e, conseqüentemente, deságio no preço do produto final (SILVA *et al.*, 2010). A qualidade das fibras de algodão é determinada conforme o seu grau de contaminação envolvendo: sisal, restos de cultura, como folhas, bráctea e galhos, sementes de plantas daninhas, pedras, entre outros (ALGOTEC, 2010). Tais impurezas, além de elevar o custo de produção, já que torna necessário o acréscimo do procedimento de limpeza em chão fabril, dificulta a classificação da qualidade do produto. Um exemplo é na fabricação do *linter*, em que boa parte da produção é destinada a processos que exigem determinados níveis de pureza, o que faz necessária a especificação do percentual de celulose e de impurezas presentes no produto.

Dois subprodutos do caroço de algodão que são responsáveis por mais de cinquenta por cento do faturamento das esmagadoras de algodão do Brasil são: óleo vegetal e *linter* (NEVES; PINTO, 2012). Estes em seu processo de fabricação passam por etapa de classificação

e em seguida por análise laboratorial de qualidade, o que demanda maior investimento e tempo para disponibilização do produto ao mercado consumidor. Na literatura existem metodologias analíticas para a classificação de qualidade de óleo vegetal. Por exemplo, Veras et al. (2010) utilizaram a metodologia de Espectroscopia no Infravermelho Próximo (*Near Infrared Spectroscopy* - NIR) e a metodologia do Ultravioleta-visível (UV-vis) para realizar a classificação do biodiesel a partir da análise de qualidade do óleo utilizado como matéria-prima, aplicando como técnica de reconhecimento de padrões o método de Modelagem Independente Flexível por Analogia de Classe (*Soft Independent Modeling by Class Analogy* - SIMCA). Balabin e Safieva (2011) também desenvolveram métodos de classificação de óleo utilizando espectroscopia NIR e a técnica de classificação multivariada de Análise Discriminante pelos Mínimos Quadrados Parciais (*Partial Least Square Discriminant Analysis* - PLS-DA).

O *linter* possui um mercado consumidor com suas próprias exigências, o qual proporcionou o desenvolvimento de um produto com melhor qualidade, características diferentes, mas aparências semelhantes. Os diferentes níveis de qualidade do *linter* por vezes ocasionam dificuldades na identificação de amostras em chão fabril, a qual é realizada por inspeção visual, sendo assim um método subjetivo e sujeito a erros. A qualidade do *linter* é determinada conforme a contaminação de suas fibras, qual ocorre com a presença de matéria estranha como fragmentos de corpo vegetal (folhas, caules, casca e caroço), material plástico, açúcares (pegajosidade), minerais (areia, poeira), químicas (óleo, graxa) ou material metálico (ALGOTEC, 2010).

Outro método disponível para a classificação das fibras é utilizando o Instrumento de Alto Volume (*High Volume Instrument* - HVI) em que um grande volume de amostras pode ser analisado em um curto período de tempo, sendo analisados parâmetros físicos como: cor, resistência, espessura e contaminação (FONSECA; SANTANA, 2002). Contudo, para o uso do HVI é preciso um alto investimento com equipamentos, infraestrutura laboratorial adequada e a presença de um analista treinado. Um exemplo é o laboratório de classificação visual e por HVI construído pela Associação Goiana dos Produtores de Algodão (AGOPA), com o apoio do Governo do Estado de Goiás, cujo valor investido ultrapassou os 4 milhões de reais (CASCELLI *et al.*, 2016).

No Brasil o Centro Brasileiro de Referência em Análise de Algodão (CBRA) é o responsável pela verificação e padronização dos processos classificatórios do algodão brasileiro, garantindo a qualidade e a credibilidade dos resultados aferidos nos diversos laboratórios instalados no território nacional. Atualmente o CBRA é composto por 14 unidades que juntas totalizam

64 máquinas de HVI. Tendo programas de qualidade como o *Standard Brasil HVI – SBHVI* com investimento de cerca de 50 milhões de reais, porém, ainda possui uma margem de análise aleatória da safra nacional de 1% (MAGALHÃES, 2014).

Tanto as técnicas desenvolvidas para a classificação do óleo como do *linter* possuem custos elevados em instrumentos e de manutenção, sendo notável a importância do desenvolvimento de metodologias alternativas de fácil implementação em chão fabril, de baixo custo e que forneçam dados confiáveis. O uso de técnicas de visão computacional é uma alternativa atrativa, tendo como fatores positivos: o uso da imagem digital, que pode substituir o sistema visual humano, eliminando o caráter subjetivo da análise; apresentar menos custos que as técnicas laboratoriais já utilizadas; e redução de tempo para classificação do produto. Com o intuito de melhorar o desempenho da análise de qualidade em indústrias que produzem subprodutos do caroço de algodão, este trabalho apresenta a proposta de desenvolvimento de metodologias para análise e classificação do óleo de algodão e do *linter*.

1.1 Objetivos do Trabalho

O objetivo geral deste trabalho consiste em apresentar metodologias de análises semiautomáticas para aplicação industrial na classificação do *linter* e do processo de semirrefino do óleo de algodão, através do reconhecimento de padrões em imagens digitais do produto em processo de fabricação, utilizando técnicas de visão computacional. Objetivos específicos:

- Realizar o pré-processamento das imagens obtidas tanto do óleo como do *linter* para formar as bases de imagens;
- Aplicar técnicas de extração de características nas imagens para obter informações relevantes que permitam a classificação dos produtos em análise;
- Realizar comparações entre técnicas de reconhecimento de padrões para a classificação do óleo de algodão, tendo como parâmetro a cor;
- Classificar a qualidade do *linter* aplicando análise de textura, realizando a comparação entre distintos descritores de textura presentes na literatura.

1.2 Organização da Dissertação

Neste primeiro capítulo são introduzidas a motivação para o estudo, análise e classificação do *linter* e etapas do processo de semirrefino do óleo vegetal de algodão. É realizada uma

revisão bibliográfica acerca do tema e são ressaltados os objetivos da dissertação. No Capítulo 2 são apresentadas algumas metodologias de extração de características por análise de textura e cor. Em extração de textura são apresentados os Descritores de Haralick, Dimensão Fractal, Lacunaridade, *Local Binary Pattern* - (*LBP*), *Local Fuzzy Pattern* - (*LFP*) e Filtros de Gabor. Já na extração de características de cor são vistos diferentes modelos de cores para obtenção de atributos. São apresentados também alguns métodos de classificação e validação. No Capítulo 3 é detalhada a metodologia proposta, que está dividida em: aquisição de imagens, modelos de cores adotados, extração das características de textura e cor, segmentação e classificação. O Capítulo 4 trata da análise dos resultados obtidos para a classificação do *linter* e processo de semirrefino do óleo vegetal de algodão. As conclusões obtidas e propostas para trabalhos futuros são apresentadas no Capítulo 5.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão apresentadas as bases teóricas para o desenvolvimento da pesquisa. As seções 2.1 e 2.2 apresentam as características do óleo de algodão e do *linter*, bem como alguns modelos literários utilizados para a classificação destes. Nas seções 2.3 a 2.6 é apresentada uma fundamentação teórica relacionada às ferramentas e técnicas de visão computacional, bem como os métodos de reconhecimento de padrões, extração de características por análise de textura e estratégias de validação.

2.1 Óleo Vegetal de Algodão

O óleo da semente de algodão possui uma larga aplicabilidade no Brasil, sendo o mais antigo óleo vegetal produzido industrialmente (CARDELLO *et al.*, 1995). É frequentemente utilizado como medida padrão de qualidade para outros óleos, possuindo um sabor leve de castanha. Normalmente é reluzente de cor dourado claro ao amarelo-avermelhado. Assim como os demais óleos, seu nível de cor depende do grau de refinamento (CAMPESTRE, 2017).

No processo de refino do óleo de algodão ocorre a eliminação da umidade de substâncias coloidais, coloridas, voláteis, inorgânicas e dos ácidos graxos. Este processo, além de melhorar a aparência do óleo, deixa-o em condições adequadas para consumo alimentício (MANDARINO *et al.*, 2015) (MANDARINO; ROESSING, 2001).

O óleo de algodão é utilizado em diversos segmentos e a partir do óleo bruto são obtidos o biodiesel, a borra e o óleo refinado (CAMPESTRE, 2017). No Brasil o óleo bruto de algodão está entre as três principais matérias-primas vegetais mais utilizadas na fabricação do biodiesel, ficando atrás apenas do óleo de soja e da gordura bovina (SCHUTTE *et al.*, 2017). O óleo refinado possui diversas aplicações, tais como na indústria alimentícia, na fabricação de margarinas, biscoitos e chocolates. É também utilizado no setor farmacêutico, domissanitários, na fabricação de lubrificantes e sabões (CAMPESTRE, 2017).

2.1.1 Processo de Semirrefino do Óleo de Algodão

O processo de semirrefino apresentado a seguir é o processo realizado na indústria em que foram obtidas as amostras para análise. Para obter o óleo semirrefinado, o mesmo passa por seis etapas: estado bruto; primeiro refino; segundo refino; lavagem; secagem do óleo e braqueamento. Os subitens a seguir apresentam breves descrições sobre cada etapa.

2.1.1.1 *Estado Bruto*

Inicialmente tem-se o óleo bruto, que, após ser extraído por processo de prensagem do caroço de algodão, é destinado para centrífugas onde ocorre a separação de óleo e água e sólidos por meio da força da centrifugação, dando assim início ao processo de semirrefino.

2.1.1.2 *Primeiro Refino*

Nesta etapa é dosada uma solução de soda cáustica conforme a quantidade e qualidade do óleo. A aplicação da solução é realizada para que ocorra a neutralização alcalina do óleo vegetal, por meio de reação dos ácidos graxos livres com a solução de soda. Estes ácidos graxos irão se tornar em sabões que serão retirados do óleo neutro por processo físico.

Após a dosagem de soda, o óleo tem sua temperatura elevada em torno de 65°C, mantendo-se nesse estado por todo o processo de retenção. Durante esse processo ocorre o primeiro clareamento do óleo.

2.1.1.3 *Segundo Refino*

Na etapa de segundo refino é realizada uma pré-lavagem, em que após a neutralização do óleo é removida parte da quantidade de sabões gerados no processo de primeiro refino (400 a 700 ppm de sabões). Durante o processo de pré-lavagem a temperatura das centrífugas é elevada a 70°C, sem apresentar necessidade de retenção.

Visto que o óleo de algodão apresenta alta acidez e que a etapa de primeira lavagem não é suficiente para a retirada de todos os resíduos de sabões, torna-se necessária a realização de outro estágio de lavagem.

2.1.1.4 *Lavagem*

No segundo processo de lavagem é acrescentado cerca de 20% de água aquecida em relação ao volume de óleo e a temperatura das centrífugas é elevada a 80 °C. O processo é finalizado com a retirada de quase todo o volume de água contendo o material residual de sabões.

2.1.1.5 *Secagem do Óleo*

A secagem do óleo é realizada por um secador a vácuo, retirando assim o excesso de água que possa permanecer do processo anterior.

2.1.1.6 *Branqueamento*

No processo de branqueamento ocorre a adição de um auxiliar filtrante ao óleo, sendo retido posteriormente em um clarificador a 110 °C e só então filtrado. Este processo tem a finalidade de reduzir a quantidade de impurezas e substâncias que atuam como agentes catalíticos, as quais podem prejudicar a qualidade de cor e estabilidade do óleo. Além disso, a clarificação pode corrigir possíveis falhas que venham a ocorrer durante os processos de neutralização e lavagem, além de facilitar a desodorização (OLIVEIRA, 2001).

2.1.2 *Métodologias Utilizadas para Classificação do Óleo de Algodão*

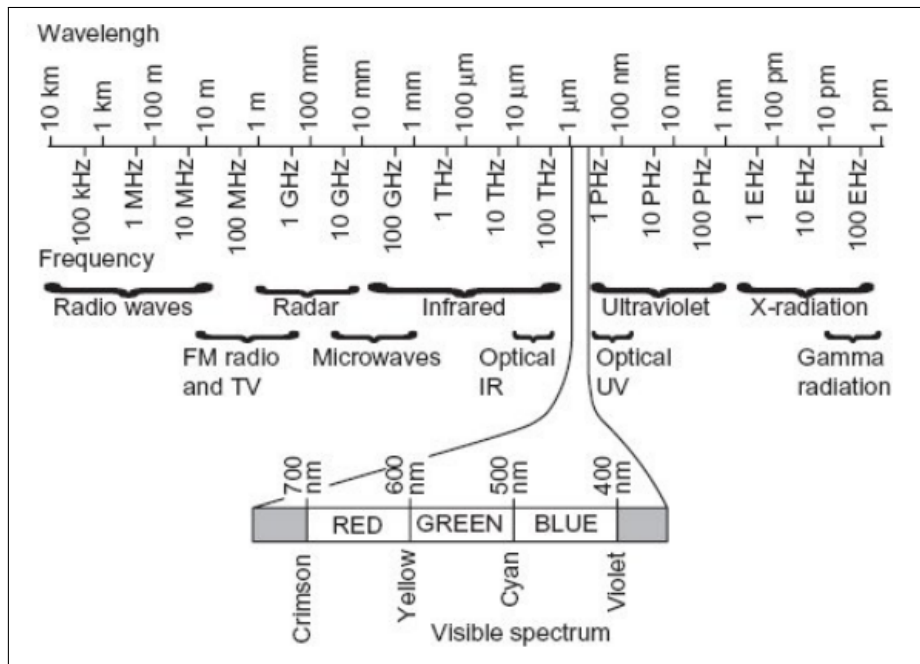
Para cada aplicação do óleo de algodão é necessário que este esteja atendendo a determinadas características físicas e químicas, estando ele em seu estado bruto ou refinado. Na produção do biodiesel é importante a observação dos dados de índice de acidez e de teor de umidade, os quais devem apresentar valores baixos, em que a presença de ácidos monocarboxílicos e água são desfavoráveis ao bom desempenho do processo produtivo de biodiesel (BARROS *et al.*, 2013).

No uso alimentício o óleo tem que passar pelo processo de branqueamento, no qual um fator utilizado para a determinação de sucesso do processo é a baixa presença de clorofila, que é um composto prejudicial tanto no nível oxidativo do óleo como na sua aparência visual, ou seja, na sua cor. Uma técnica utilizada para realizar a medição dos níveis de clorofila é a espectrofotoscopia UV-vis (PECEGO, 2014).

2.1.2.1 *Métodos de Análise*

A seguir são apresentadas duas metodologias utilizadas para análise de óleos brutos e refinados por meio de espectroscopia. As técnicas espectroscópicas têm por base o índice de radiação emitida ou absorvida por moléculas ou espécies químicas de interesse, as quais podem ser classificadas em relação à região do espectro eletromagnético utilizado na medição (SKOOG *et al.*, 2007b). A Figura 1 apresenta várias bandas do espectro eletromagnético, inclusive nas

Figura 1 – Região do espectro magnético.



Fonte – (MIGNANI *et al.*, 2012).

faixas que serão abordadas, sendo elas a do UV-vis (400-780 nm) e do infravermelho próximo (780-2500 nm) (MIGNANI *et al.*, 2012).

2.1.2.2 Near Infrared Spectroscopy - NIR

O método da Espectroscopia no Infravermelho Próximo - NIR utiliza uma técnica não destrutiva, a qual fornece informações químicas e físicas da amostra analisada. A quantificação é realizada utilizando modelos matemáticos e por análise de dados multivariados. A análise por espectroscopia no infravermelho próximo foi realizada pela primeira vez pelo Departamento de Agricultura dos Estados Unidos (USDA) em meados da década de 1960 para o estudo de maturação de frutas. Desde então esse método tem sido utilizado em indústrias alimentícias, têxteis, farmacêuticas, entre outras (METROHM, 2013).

Na classificação de óleo o método NIR é utilizado na detecção de adulterações ou impurezas no produto (HOURANT *et al.*, 2000). Essa classificação é realizada com o uso da radiação eletromagnética da luz, em que por meio da transferência da energia entre a luz e a matéria são obtidos dados classificatórios (METROHM, 2013).

A espectrometria NIR é uma ferramenta eficiente em determinações analíticas. No entanto, as bandas espectrais de baixa intensidade e muito sobrepostas, geradas por matrizes complexas, por vezes dificultam a obtenção de uma boa correlação entre grupos de átomos

presentes numa molécula e a sua composição nas amostras. Para melhor classificação, as informações dos espectros gerados precisam passar por um tratamento, em que é fundamental a aplicação de ferramentas quimiométricas, as quais utilizam métodos ou técnicas estatísticas para interpretar melhor os dados obtidos (NUNES, 2008)(PEREIR, 2007).

2.1.2.3 Ultravioleta Visível

A espectroscopia UV-visível a partir do comprimento de onda específico para o composto em estudo é capaz de determinar tanto a concentração como a composição da amostra. Assim, ao realizar a classificação de óleos, é possível identificar a presença de substâncias químicas (PALMER *et al.*, 2000).

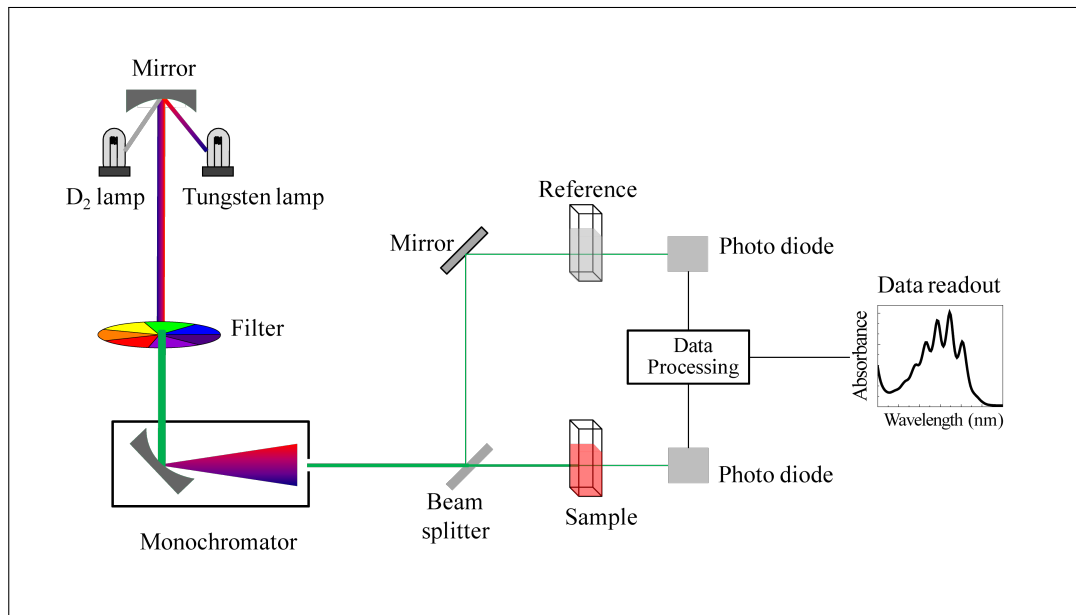
Os instrumentos que utilizam UV-visível são denominados de espectrofotômetros, que podem ser de feixe simples ou duplo. O que diferencia um do outro é que no simples a luz passa pela amostra, já no duplo a luz é emitida por um divisor de feixe, o qual de forma alternada conduz o feixe de luz para a amostra ou para uma célula de referência, diversas vezes por segundo (SKOOG *et al.*, 2007a).

Para realizar a análise de uma amostra por espectroscopia no ultravioleta visível, esta deve ser inserida no caminho óptico do espectrofotômetro. Após a emissão da luz é analisado quanto de luz que foi absorvida pela amostra, logo a transmitância da amostra será a relação da luz emitida pela quantidade que foi absorvida, ou seja, o espectro UV-visível é essencialmente um gráfico que relaciona o valor da absorvância e o comprimento de onda na faixa do ultravioleta visível (SKOOG *et al.*, 2007a). Assim, quanto maior o valor de absorvância, maior é a presença de substâncias indesejáveis. A Figura 2 mostra o esquema de um Espectrômetro UV-visível de feixe duplo.

Os dados obtidos pelo espectrômetro UV-Vis por vezes possuem respostas com sobreposição de bandas associadas a duas ou mais substâncias presentes em uma amostra. A obtenção de respostas com essas características sucedem da classe de espectro emitido pelo equipamento e de sua simples correlação com a estrutura molecular em análise. Dessa forma, substâncias que apresentam arranjos moleculares diferentes, porém, mesma capacidade de absorção de luz, podem apresentar espectros UV-Vis com perfis semelhantes e bandas localizadas nas mesmas regiões de comprimentos de onda (SKOOG *et al.*, 2007a).

Respostas obtidas pelo espectrômetro UV-Vis com sobreposições e classificação de perfis distintos, porém, valores de banda bem próximos, podem ocasionar dificuldades em sua

Figura 2 – Esquema de um espectrômetro UV-visível de feixe duplo.



Fonte – (SOBARWIK, 2013).

interpretação. Uma vez que são obtidas matrizes complexas é fundamental o uso de métodos quimiométricos para facilitar na elucidação dos perfis espectrais (VÉRAS *et al.*, 2012).

2.1.2.4 Modelos de Classificação

Para a classificação de dados existem diversos métodos, sendo eles estatísticos supervisionados ou não. A seguir são apresentados dois modelos de classificação utilizados em estudos de classificação de óleos.

Modelagem Independente Flexível por Analogia de Classe

O método SIMCA utiliza a comparação de semelhanças para o reconhecimento de padrões (BEEBE *et al.*, 1998). Para a realização da classificação dos padrões são consideradas as informações de distribuição da população e o grau de confiança da classificação estimado. Sendo assim, é possível prever as novas amostras como pertencentes a uma ou mais classes ou a nenhuma das classes determinadas (FERREIRA *et al.*, 1999).

A modelagem independente flexível por analogia de classe tem o propósito de criar um espaço limitado para cada classe. Assim, é realizada a distinção das classes pela distância e presença de resíduos em cada classe (CALAS *et al.*, 1997). Na classificação de uma amostra são vistas a projeção da mesma no espaço dos escores e sua dinâmica das fronteiras da classe em

questão, sendo repetido esse processo para todas as classes (SRIVASTAVA; CARTER, 1983). Um erro que pode ocorrer nesse processo de classificação é quando os modelos utilizados não possuem características suficientes para a classificação da amostra em análise, podendo caracterizar esta como pertencente a uma classe incorreta (FERREIRA, 2002).

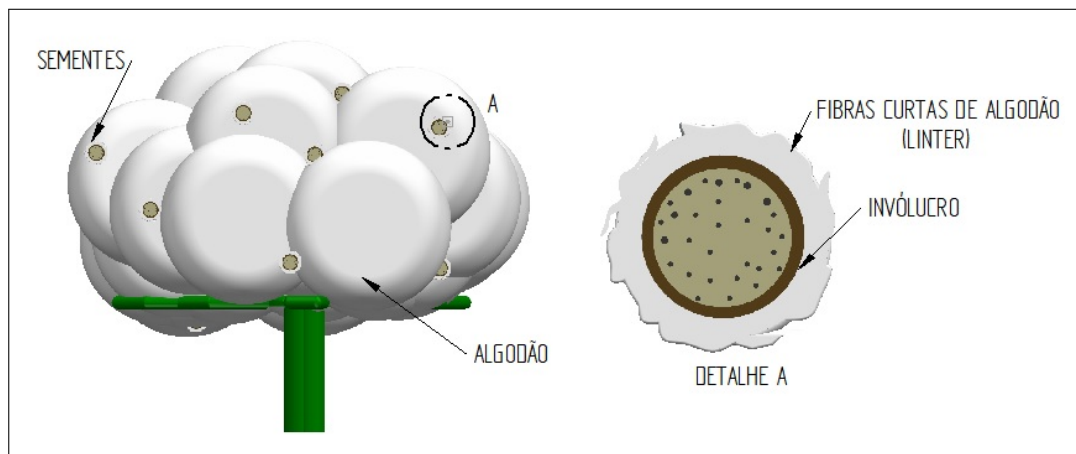
Análise Discriminante pelos Mínimos Quadrados Parciais

A técnica de Análise Discriminante por Quadrados Mínimos Parciais (PLS-DA) é uma variação do algoritmo de regressão por Quadrados Mínimos Parciais (PLS), método de calibração inversa que tem por objetivo correlacionar a resposta instrumental (matriz x) com um vetor y de concentrações. No PLS-DA a diferença consiste no fato de que o vetor y conterá valores da classe a qual a amostra pertence (MASSART *et al.*, 1998)(BARKER; RAYENS, 2003). Para reduzir os erros no processo de determinação de classes é utilizado um limiar no ponto em que o número de falsos positivos e falsos negativos sejam minimizados (BALLABIO; CONSONNI, 2013), reduzindo assim o efeito da colinearidade dos dados do modelo e maximizando a correlação entre as medidas de absorvância predictoras e a variável resposta (MEVIK; CEDERKVIST, 2005).

2.2 Linter

O *linter* é formado por uma porção de fibras curtas aderidas ao invólucro da semente de algodão. Estas são compostas principalmente por celulose de larga aplicabilidade na fabricação de inúmeros produtos, como algodão cirúrgico, papel e papelões especiais, diversos tipos de embalagem, fibras de celulose sintéticas, papel-moeda, acetatos e nitratos de celulose, dentre outros (BELTRÃO; AZEVEDO, 2008). A Figura 3 apresenta a região do *linter* no invólucro das sementes de algodão.

Figura 3 – Demonstrativo da pluma de algodão e região do *linter* no invólucro das sementes de algodão.



Fonte – Própria autora.

Para a remoção do *linter* das sementes de algodão podem ser utilizados meios mecânicos, químicos ou por flambagem, sendo que os dois últimos processos destroem o *linter*, enquanto o primeiro o preserva (BELTRÃO; AZEVEDO, 2008). No processo de deslinteramento mecânico são utilizadas máquinas denominadas deslinteradeiras, compostas por bicas de ferro que dosam a quantidade de sementes. Ao dosar pequenas porções de sementes, estas são direcionadas a rolos de serras com pentes, os quais são responsáveis por retirar o *linter* do invólucro do grão. O grão só é retirado da máquina quando cerca de 96% da fibra é retirada (MANDARINO; ROESSING, 2001).

Ao finalizar o processo de retirada do *linter* este pode ser classificado conforme suas características de qualidade, sendo normalmente analisados os parâmetros de cor, resistência, espessura e contaminação (DOLENSI, 2013). A contaminação que ocorre nas fibras do algodão é proveniente de partículas estranhas, podendo ser fragmentos de corpo vegetal (folhas, caules, casca e caroço), açúcar (pegajosidade), material plástico, minerais (areia, poeira), metálicas ou químicas (óleo, graxa) (ALGOTEC, 2010). A Figura 4 apresenta o *linter* com contaminação ao nível aceitável para comercialização.

Figura 4 – Amostra de *linter* com pequeno índice de contaminação.



Fonte – Própria autora.

2.2.1 Métodos Utilizados Para Análise de Qualidade do *linter*

Existem três métodos para análise do *linter*:

- Classificação manual/visual realizada por um profissional qualificador, especialista em qualidade de fibras;
- Classificação tecnológica com o uso do Instrumento de Alto Volume (*High Volume Instrument - HVI*);
- Classificação da qualidade do *linter* por base de gráficos de estimativas de impurezas disponibilizados pela Associação Técnica da Indústria de Celulose e Papel (*Technical Association of the Pulp and Paper Industry - TAPPI*) (DOLENSI, 2013).

2.2.1.1 Classificação Manual / Visual

O profissional responsável por determinar a qualidade das fibras pelo padrão visual deve ser habilitado conforme o sistema americano, tendo que considerar a cor, percentual de matéria estranha e o estado da amostra (LANA *et al.*, 2014). As informações obtidas são comparadas com um padrão que é organizado em caixas com amostras utilizadas atualmente por mais de dezesseis países (REZENDE, 2010).

O padrão americano classifica as amostras (folhas) de *linter* em grau visual. Sendo este classificado por três dígitos: O primeiro dígito refere-se principalmente ao brilho e à

homogeneidade de aspecto. Quanto mais brilho e menos rugosidade, menor o número, que pode ser classificável de 1 a 8. O segundo dígito refere-se à tonalidade. Quanto mais alto o número, mais amarelo é o algodão, que é classificável de 1 a 5, sendo: 1 - Branco; 2 - Ligeiramente Creme; 3 - Creme; 4 - Avermelhado; e 5 - Amarelado Estanhado. O terceiro dígito representa a folha, que varia de 1 a 7. A folha recebe a numeração conforme seu índice de impurezas, quanto maior o nível de partículas entranhas na amostra, maior será a numeração utilizada. A matéria estranha quando detectada pelo classificador deve ser informada por uma menção específica chamada de “*call*”. No relatório de classificação, o classificador deve mencionar o tipo de matéria encontrada e o nível de contaminação. Conforme define a IN-MAPA 063/2002, a combinação dos dois primeiros dígitos referentes a tipo e tonalidade representa o grau de cor do algodão (LANA *et al.*, 2014).

A classificação visual apresenta sensibilidade à qualidade da luz ambiente. Exigindo assim, a sua realização em ambientes normalizados conforme a NBR 12276/1991, a qual define e orienta várias especificações, tais como cor de piso e paredes, mesas e disposição de amostras, entre outros (LANA *et al.*, 2014).

2.2.1.2 Classificação Tecnológica por HVI

O sistema de classificação do HVI foi desenvolvido para realizar medições em larga escala de amostras de algodão e do *linter* em um espaço mínimo de tempo. A análise por HVI é realizada em laboratórios com estrutura adequada e composto por profissionais capacitados a manusear equipamentos integrados de alta produção. Na classificação de fibras por HVI, além de avaliar a quantidade de impurezas, são analisadas de forma objetiva suas características como: resistência (STR), comprimento (UHML), micronaire (Finura da Fibra), SFI (Índice de Fibras Curtas), Alongamento (ELG), Maturidade (MR), RD (Refletância/Brilho) e +B (Grau de amarelamento) (FONSECA; SANTANA, 2002).

As amostras para análise do *linter* consistem em duas porções. Cada porção é tirada de lados opostos do fardo e organizada manualmente em faixas. Cada porção deve possuir entre 150 a 300 mm de comprimento e cerca 150 mm de largura, devendo ter aproximadamente a mesma massa. A amostra total submetida a teste deve pesar em torno de 225 gramas e ser identificada por etiqueta, indicando o número do fardo do qual foi retirada (USDA, 2005).

Tabela 1 – Níveis de classificação do *linter* conforme sua área de contaminação.

Nível de Contaminação	Áreas Contaminadas
1	< 30
2	30 - 75
3	75 - 120
4	> 120

Fonte – (DOLENSI, 2013).

2.2.1.3 Classificação do *Linter* pelo princípio de TAPPI T213

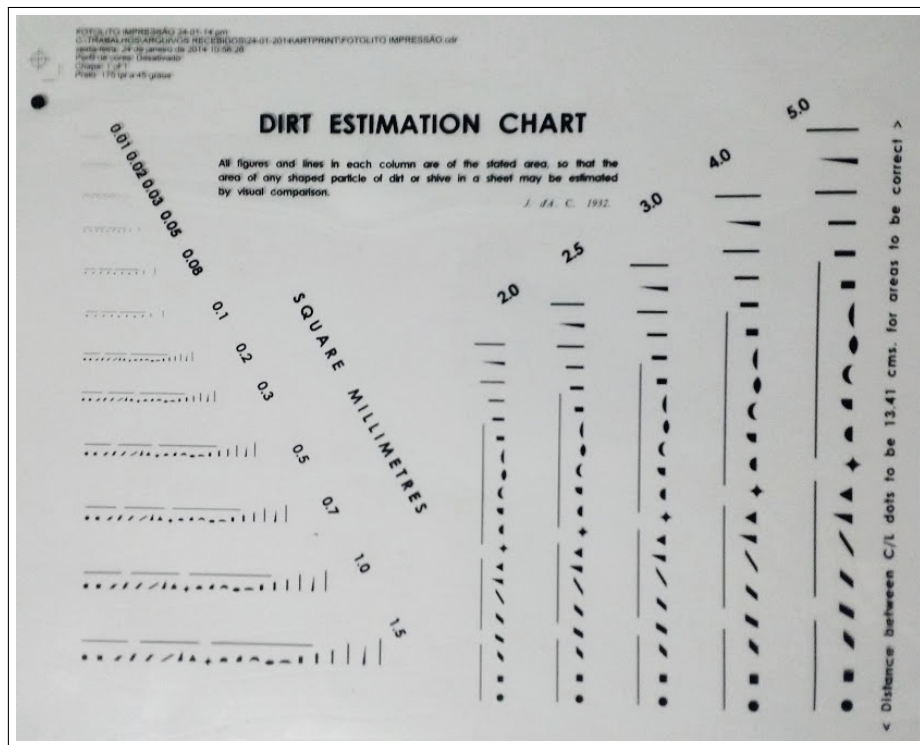
O modelo TAPPI T213 utiliza como ferramenta a transparência *dirt estimation chart* a qual é desenvolvida com base no gráfico estimativo de sujeiras da TAPPI. Essa transparência é usada para medir o tamanho de pontos, defeitos ou inclusões no intervalo de 0,02 a 5 milímetros quadrados em papel e outros materiais industriais como os têxteis e de plásticos (TAPPI, 2000).

Para aplicação desse procedimento é realizada a preparação de uma amostra de *linter* em forma de folha, a qual será analisada com o auxílio de uma luz refletora, permitindo assim a visualização de ambos os lados da folha. Na análise será determinada a área total de sujeiras que sejam iguais ou maiores do que 0,1 milímetro quadrado (DOLENSI, 2013). O nível de sujeira já está tabelado e conforme a área de sujeira encontrada em 20 gramas de *linter* por milímetro quadrado é realizada a classificação do *linter* (DOLENSI, 2013). A Tabela 2 apresenta os níveis classificáveis do *linter* conforme área de sujeira apresentada.

A execução da análise segue os seguintes passos:

- 1º Colocar a folha de *linter* sobre o iluminador;
- 2º Com o uso de uma *dirt estimation chart*, tipo James d’A Clark 1932, a qual é apresentada na Figura 5, determinar o tamanho das áreas de sujeira;
- 3º No decorrer da análise, localizar uma área de sujeira por vez e determinar o tamanho da área. Se o tamanho for igual ou maior que 0,1 milímetro quadrado é anotado em uma planilha de controle o valor da área. Dar continuidade ao processo até que cada área de sujeira seja localizada e os seus valores sejam anotados;
- 4º Tendo todas as áreas visíveis de sujeira localizadas de um lado da folha, realizar o mesmo procedimento do outro lado da folha;
- 5º Com todas as áreas de sujeiras localizadas, é realizada a soma das que possuem tamanho igual ou maior do que 0,1 milímetro quadrado. O resultado é expresso

Figura 5 – *Dirt estimation chart*, tipo James d'A Clark 1932.



Fonte – Própria autora.

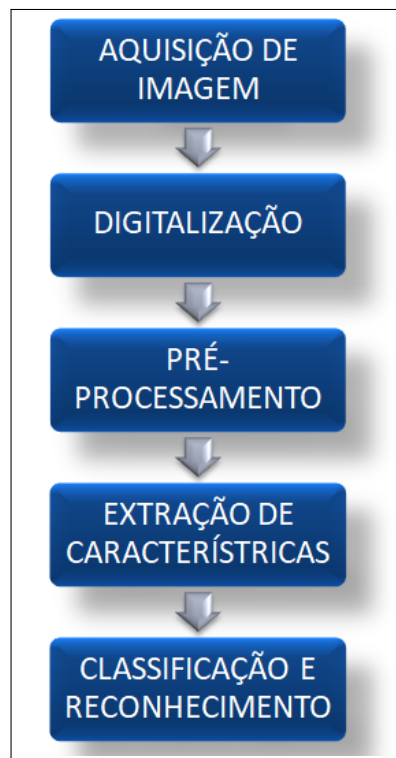
em milímetro quadrado de sujeira por 20 gramas de amostra (DOLENSI, 2013).

2.3 Visão Computacional

A visão computacional é a área da ciência que estuda e atua no desenvolvimento de teorias e metodologias voltadas à obtenção automática de informações pertinentes contidas em imagens (CROWLEY; CHRISTENSEN, 1995). Estas informações proporcionam o reconhecimento, manipulação e classificação de objetos que compõem uma imagem (BALLARD; BROWN, 1982), contribuindo para o avanço no desenvolvimento de máquinas inteligentes (JAHNE, 2005).

A visão computacional pode ser considerada uma área recente, sendo os primeiros trabalhos realizados por volta dos anos 70 (LUCIANO; HONORATO, 2010). Os trabalhos desenvolvidos com visão computacional objetivam auxiliar na resolução de problemas de alta complexidade, visando imitar a cognição humana e a capacidade do ser humano em tomar decisões de acordo com as informações presentes em imagens (RODRIGUES, 2015). Atualmente as tecnologias da área de visão computacional estão sendo utilizadas em distintas áreas como: aeronáutica, astronomia, medicina, multimídia, robótica, industrial, sistemas produtivos, entre

Figura 6 – Etapas de um sistema de processamento de imagem.



Fonte – Própria autora.

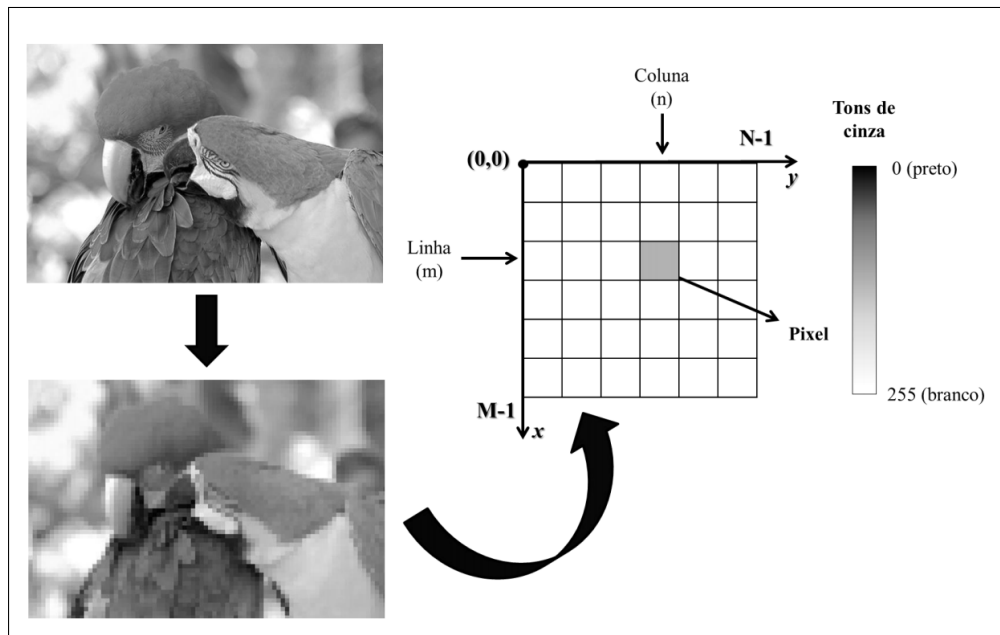
muitas outras (RUDEK *et al.*, 2011).

Quando se trata do sistema de visão computacional, dois níveis de abstração são normalmente determinados: processamento de imagens (baixo nível) e análise de imagens (alto nível) (GONZALEZ; WOODS, 2008) (PEDRINI; SCHWARTZ, 1998). O processamento digital de imagens consiste em um conjunto de técnicas, tais como: aquisição da imagem, digitalização, pré-processamento, segmentação, pós-processamento, extração de características, classificação e reconhecimento, como ilustra a Figura 6. A seguir, serão apresentadas breves descrições de cada uma destas etapas.

2.3.1 Aquisição da Imagem e Digitalização

No processo de aquisição de uma imagem digital são realizadas três etapas: rastreamento, amostragem e quantização. No rastreamento é determinado o endereço de cada *pixel* da imagem. Na amostragem, por sensoriamento é medida a intensidade de luz de cada *pixel*, sendo obtido, assim, o nível de cinza do *pixel*. Na quantização, o nível de cinza é representado por um valor inteiro. Desta forma, é possível gerar uma matriz (CASTLEMAN, 1996) a qual terá tamanho $N \times M$ *pixels* (N : linhas e M : colunas) de valores de *pixel* $p(x, y)$ inteiros positivos, que

Figura 7 – Imagem monocromática e a sua representação digital.



Fonte – (QUEIROZ; GOMES, 2014) adaptado por (GONÇALVES, 2015).

indicam a intensidade de cor em cada posição (x, y) da imagem (GONZALEZ; WOODS, 2008). A Figura 7 apresenta uma imagem monocromática e a sua representação digital.

2.3.2 Pré-processamento

O pré-processamento tem o propósito de melhorar a qualidade da imagem. Nessa etapa ocorre a exclusão de características do objeto a ser reconhecido, por meio de métodos específicos que facilitam a sua identificação como, por exemplo, destaques de contornos, bordas, destaque de figuras geométricas etc. (RODRIGUES, 2015). A aplicação desses métodos exclui informações não úteis e que possam vir a atrapalhar o trabalho de reconhecimento (MOREIRA, 2002).

O pré-processamento é dividido em duas fases distintas: a segmentação de imagem e a extração de características (CASTLEMAN, 1996).

2.3.2.1 Modelos de Cores

Em trabalhos que utilizam a cor como uma característica classificatória é comum o uso de diferentes modelos de cores para análises comparativas e obtenção de melhores dados descritivos. Segundo Gonzalez e Woods (2008) um modelo de cor é a denominação dada a uma coordenada tridimensional onde cada ponto representa uma cor (MILANEZ, 2013).

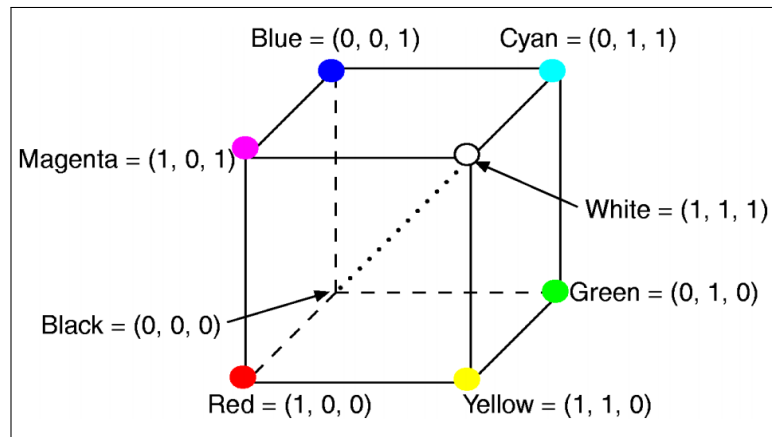
Os modelos de cores mais utilizados são RGB (*red, green, blue*), CMY (*cyan, magenta, yellow*), CMYK (*cyan, magenta, yellow and black (key)*), HSI (*hue, saturation e intensity*) YIQ e HSV (*hue, saturation e value*) (PLATANIOTIS; VENETSANOPOULOS, 2000)(MILANEZ, 2013).

Conforme as características do modelo de cor, este é destinado à distintas aplicações. O RGB é orientado para hardware, utilizado para monitores e câmeras de vídeo a cores; os modelos CMY e CMYK são modelos para impressão colorida; o modelo HSI corresponde estritamente à forma como os seres humanos descrevem e interpretam as cores, separando o componente intensidade das informações de cores (matiz e saturação) em uma imagem colorida (GONZALEZ; WOODS, 2008) e o modelo YIQ baseia-se na segmentação dos sinais de cor RGB em um sinal de luminosidade ou luminância (Y) e dois sinais de cromaticidade ou diferença de cor (IQ) (SOUZA *et al.*, 2009) .

2.3.2.2 *Modelo de Cores RGB*

O modelo RGB é o mais comum dos espaços, pois é diretamente suportado nos ecrãs a cores, tendo em vista também que qualquer cor é expressa sobre a junção das cores Vermelho (R), Verde (G) e Azul (B). Esse modelo baseia-se no mecanismo de formação de cores do olho humano, o qual relaciona a combinação de radiações monocromáticas das três cores primárias (vermelho, verde e azul), tornando-se então a formação de cores como um processo aditivo. Assim, as cores podem ser representadas por um cubo definido sobre os eixos R, G e B, que assumem valores de (0-255), ou seja, 256 níveis de cores. Desse modo, cada canal de cor é formado por um conjunto de 8 bits, resultando em um imagem de 24 bits ou 16,7 milhões de cores (GONZALEZ; WOODS, 2008) (SOLOMON; BRECKON, 2011). A Figura 8 apresenta o cubo que representa o modelo de cores RGB.

Figura 8 – Cubo representativo do modelo de cores RGB.



Fonte – (CARVALHO, 2008).

2.3.2.3 Modelo de Cores YIQ

O sistema NTSC (*National Television Standards Committee*) utilizou o modelo de cor YIQ, o qual foi criado no período em que as TVs passaram a projetar em cores, sendo assim, responsável por possibilitar que as emissões dos sistemas de televisão em cores fossem conciliáveis com os receptores em preto e branco. A conversão do sistema RGB para o sistema YIQ pode ser obtida de duas formas, a primeira é dada pelas equações 2.1, 2.2 e 2.3 (SOUZA *et al.*, 2009).

$$Y = 0.299R + 0.587G + 0.114B, \quad (2.1)$$

$$I = 0.74(R - Y) - 0.27(B - Y), \quad (2.2)$$

$$Q = 0.48(R - Y) + 0.41(B - Y). \quad (2.3)$$

A segunda metodologia de transformação do sistema RGB para o sistema YIQ é realizada pela multiplicação de matrizes presentes na Figura 9.

Figura 9 – Cubo representativo do modelo de cores RGB.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & -0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Fonte – (SOUZA *et al.*, 2009) (PLATANIOTIS; VENETSANOPOULOS, 2000).

2.3.3 Segmentação

A segmentação é um processo complexo, uma vez que, além de tentar traduzir para o computador um processo cognitivo realizado pela visão humana, possui metodologia essencialmente empírica que deve ser ajustada conforme o que se deseja analisar. Assim existem diferentes estruturas para a classificação de uma imagem, podendo ser, geometria, cor, textura, topologia, entre outras (SOARES, 2008).

Na segmentação de uma imagem ocorre o processo de geração de regiões representadas por suas características espaciais e espectrais, em que à proporção que cada região é obtida, ela é nomeada e os seus atributos estatísticos são extraídos (ERTHAL *et al.*, 1991). Geralmente tem por base duas características dos tons de cinza de uma imagem: a descontinuidade e a similaridade (RODRIGUES, 2015). O método da similaridade baseia-se no agrupamento de *pixels* conforme sua semelhança com os *pixels* vizinhos (SCHINDEWOLF *et al.*, 1994). Já o método de descontinuidade analisa as características que sobressaem em uma imagem, seja por haver tons de cinza diferentes na região na qual estão inseridas (caso de pontos e linhas) ou por demarcarem variações repentinas de tons de cinza entre regiões (caso de bordas e linhas) (SALDANHA; FREITAS, 2010).

A segmentação é classificada como supervisionada quando existe alguma informação a priori dos tipos de textura presentes na imagem e como não supervisionada caso as regiões tenham que ser agrupadas com base em medidas de similaridade (BRAVO, 2016).

2.3.4 Extração de Características

A extração de características é o processo que define a forma de representação e descrição dos dados, visando obter características úteis que possam ajudar nos processos de distinção entre objetos. A representação constitui-se em determinar se os objetos da cena serão representados por suas fronteiras ou por uma região completa. A representação por fronteira é utilizada quando se tem interesse nas dimensões, forma, cantos ou pontos de inflexão. Quando se pretende analisar as características de propriedade interna, como cor e textura, utiliza-se a representação por regiões (GONZALEZ; WOODS, 2008).

As informações selecionadas das imagens e consideradas importantes para o que se deseja analisar serão adicionadas a um vetor de características. Um vetor de características é uma interpretação numérica resumida de uma imagem ou parte dela, representando suas

particularidades mais significativas. O vetor de características é um vetor n -dimensional contendo esses valores. Assim, os dados representativos de uma imagem podem ser armazenados para sua utilização em distintos processos de classificação (SOARES, 2008).

2.3.5 *Classificação e Reconhecimento de Padrões*

O trabalho de reconhecimento de padrões tem o propósito de realizar a identificação de objetos (padrões) em um número de classes ou categorias. Este pode ser realizado a partir de duas etapas: o pré-processamento e o reconhecimento propriamente dito (THEODORIDIS; KOUTROUMBAS, 2008).

Na etapa de reconhecimento e interpretação os dados obtidos na imagem são avaliados, onde cada objeto é rotulado, resultando assim, na caracterização do conjunto de dados (SOUSA, 2015). O objeto só pode ser reconhecido se de forma prévia for conhecida sua geometria. Não há como reconhecer algo desconhecido, ou seja, que não tenha sido previamente armazenado em um banco de dados (TRUCCO; VERRI, 1998).

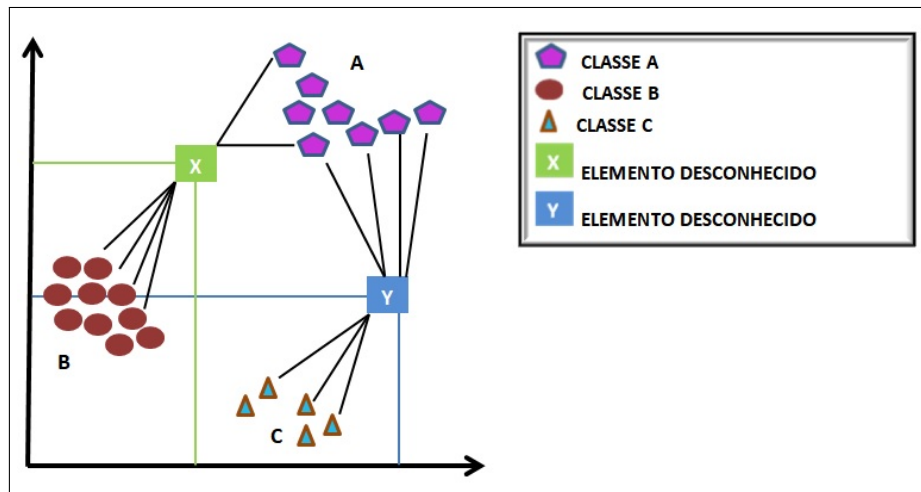
Nas próximas seções serão abordados os métodos de reconhecimento de padrões utilizados para o desenvolvimento deste trabalho.

2.3.6 *K-Nearest Neighbors (K-NN)*

A metodologia K-NN, foi introduzida pela primeira vez por Hodges em seu estudo de análise discriminante (EVELYN; HODGES, 1989). É uma variação do algoritmo NN (Nearest Neighbor), que usa os K-vizinhos mais próximos e não apenas o vizinho mais próximo (WEBB, 2002). Apresenta modelo de classificação simples e objetiva, separando os dados em análise com base em seus similares assumidos entre várias classes. Assim, dado um padrão X (desconhecido), o K-NN calcula a distância entre X e seus K-vizinhos mais próximos, depois X é atribuído à classe que apresenta maior frequência dentre os seus K-vizinhos (WEBB, 2002). Possui aprendizagem *fuzzy* não-paramétrica, ou seja, que não faz suposições sobre a distribuição de dados. Além disso, baseia-se em instâncias que podem ser representadas como pontos em um espaço euclidiano (EVELYN; HODGES, 1989).

Na Figura 10 observa-se o classificador K-NN num espaço bidimensional, onde existem três classes com suas amostras e dois elementos desconhecidos (X e Y). Pretende-se classificar esses dois elementos utilizando sete vizinhos mais próximos ($K = 7$) para cada ponto desconhecido, os quais são apresentados pelas linhas que os conectam. O elemento desconhecido

Figura 10 – Classificador K-vizinhos mais próximos.



Fonte – Própria autora, tendo como base (MARTINS *et al.*, 2005).

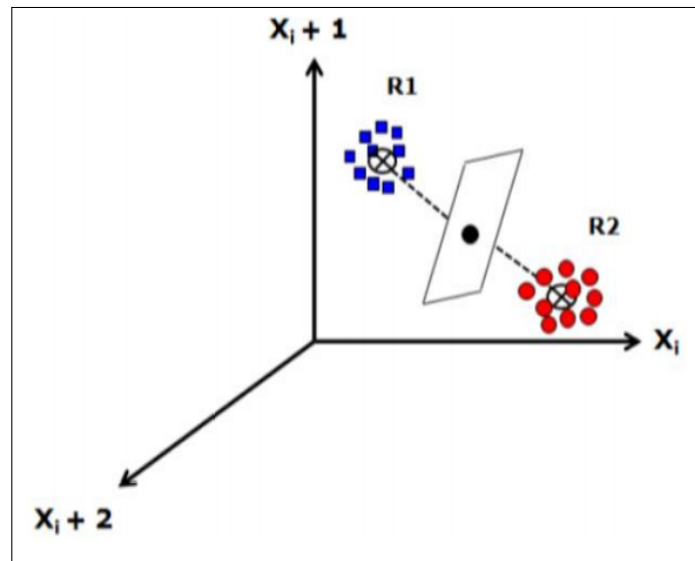
X pertence à classe B devido ao fato de o elemento X apresentar maior número de amostras na classe B, enquanto que o elemento desconhecido Y pertence à classe A (MARTINS *et al.*, 2005).

2.3.7 Linear Discriminant Analysis (LDA)

A Análise Discriminante Linear (LDA) é um dos métodos de reconhecimento de padrões mais utilizados e constitui-se na determinação de funções discriminantes, as quais maximizam a distância entre as classes e minimizam a distância entre os objetos ou amostras dentro de cada classe (KUMAR *et al.*, 2014) (PONTES, 2009). Essas novas variáveis são obtidas por meio de combinações lineares das variáveis originais, de modo a obter a melhor discriminação das classes de amostras (PONTES, 2009).

A Figura 11 mostra um esquema de obtenção de uma análise discriminante linear e como ocorre a classificação de um dado elemento. Os centros de gravidade (os asteriscos) dos grupos compostos por círculos e quadrados são inicialmente estabelecidos. Em seguida traça-se um plano de decisão paralelo e perpendicular à linha que integra os centros de gravidade dos grupos, formando as regiões R1 e R2. Se um ponto qualquer estiver presente na região R1, o mesmo é classificado como pertencente ao grupo 1. Caso contrário, obrigatoriamente pertencerá à categoria 2 (NETO *et al.*, 2009).

Figura 11 – Representação esquemática da Análise Discriminante Linear.



Fonte – (NETO *et al.*, 2009) com alterações de (COSTA, 2015).

2.3.8 Quadratic Discriminant Analysis - QDA

A análise discriminante quadrática (QDA) é uma metodologia estatística simples e flexível que considera que os dados analisados possuem distribuição normal e que tem por finalidade separar os objetos de uma população em duas ou mais classes. Para análise as variáveis dependentes representam as classes e as variáveis independentes representam as características dos objetos que definem cada classe (FRIEDMAN, 1989)(JÚNIOR; FARIA, 2007)(KHATTREE; NAIK, 2000).

O método QDA considera para cada classe uma matriz de covariância e vetor médio próprios, resultando em funções de decisões quadráticas. Por meio da matriz de covariância são obtidas informações referentes ao formato da distribuição dos *pixels* das classes, no espaço multiespectral. Assim, classes que possuem o mesmo vetor médio podem ser distinguidos com razoável grau de precisão, desde que as respectivas matrizes de covariância sejam suficientemente distintas (FUKUNAGA; HAYES, 1989).

2.3.9 Análise de Textura

A textura é uma expressão referente as propriedades que retratam um objeto ou imagem (SKLANSKY, 1978), e podem ser divididas em termos de distribuição espacial e variação de luminosidade contidas na imagem. A textura também descreve o arranjo estrutural das superfícies e relações entre regiões vizinhas, que na maioria das vezes são constituídas por

padrões repetitivos, cuja disposição pode ser periódica ou aleatória (PEDRINI; SCHWARTZ, 1998). Gonzalez e Woods (2008) representam a textura por medidas que mensuram suas propriedades de rugosidade, suavidade e regularidades.

O estudo da textura já se estende por mais de cinco décadas, em que boa parte das pesquisas são motivadas pela grande quantidade de texturas presentes em imagens e pela eficiência do sistema visual humano em detectá-las e diferenciá-las. Embora o sistema visual humano apresente capacidade de realizar o reconhecimento e descrição de texturas, é extremamente difícil padronizar sua definição. Tal dificuldade implica em uma vasta quantidade de definições de textura na literatura (JULESZ, 1962).

Através da análise de texturas é possível discriminar regiões que apresentam as mesmas características de reflectância, e portanto, mesmas cores em determinada combinação de bandas. Isso torna a textura um eficiente descritor regional capaz de ajudar na melhoria da exatidão do processo de reconhecimento de imagens (MANJUNATH *et al.*, 2002). Para Sklansky (1978) uma região em uma imagem apresenta uma textura única se um conjunto de características estatísticas ou outras propriedades locais da imagem forem constantes, com pouca variação ou proximidades periódicas.

2.3.10 *Descritores de Textura*

A textura é um descritor que fornece medidas de propriedades de um objeto ou imagem, tais como regularidade, suavidade e rugosidade. Além disso, ela também é adotada como característica de grande importância na classificação e reconhecimento de padrões (GONZALEZ; WOODS, 2008). Foram analisados diferentes algoritmos para gerar vetores de características contendo informações de textura do *linter*. A seguir, são apresentadas as descrições desses métodos.

2.3.10.1 *Matriz de Coocorrência*

A matriz de coocorrência pode ter dependência espacial dos níveis de cinza computadas em vários ângulos (0° , 45° , 90° e 135°) e distâncias. Esta se baseia no relacionamento espacial entre os *pixels* que compõem uma textura. Cada elemento que compõe a matriz de coocorrência simboliza a frequência com que um *pixel* de nível de cinza X e outro de nível de cinza Y ocorrem na imagem, separados por uma distância x de linhas e y de colunas. Essa abordagem fundamenta-se na dedução de que a referência de textura na imagem está contida

no comportamento da distribuição espacial média ou global dos níveis de cinza na imagem (GONZALEZ; WOODS, 2008). Assim, ao obter os dados da matriz de coocorrência tem-se um conjunto de quatro valores para cada uma das 14 características introduzidas por Haralick. A média e o alcance de cada uma dessas 14 medidas, calculadas sobre os quatro valores, compõem um conjunto de 56 características que podem ser usadas como entradas para o classificador. Nesse conjunto de 56 características, algumas estão fortemente correlacionadas entre si. Para determinar as principais características de classificação, um subconjunto ou combinações lineares podem ser selecionados por um procedimento de seleção de recursos (HARALICK *et al.*, 1973).

2.3.10.2 *Descritores de Haralick*

A textura foi caracterizada como uma concepção bidimensional, onde uma dimensão possui as propriedades primitivas da tonalidade e a outra corresponde às relações espaciais entre essas propriedades. Haralick *et al.* (1973) assinalaram que a concepção de tonalidade e textura não são independentes, de modo que em algumas imagens a tonalidade é dominante, e em outras imagens, é a textura. As abordagens de estatística, estrutura e espectro são as três principais técnicas utilizadas na classificação de imagens para a descrição de texturas (HARALICK *et al.*, 1973).

Haralick *et al.* (1973) introduziram um conjunto de 14 características para definir textura, entre elas, uniformidade, densidade, rugosidade, regularidade e intensidade. As características propostas são medidas estatísticas de segunda ordem e são extraídas a partir da matriz de coocorrência. A descrição do que cada medida representa é apresentada a seguir:

Entropia

É uma medida que representa a desordem dos pixels em uma imagem, tendo por dados as informações contidas em $p(i, j)$ de sua matriz de coocorrência. Assim, a entropia apresenta um valor alto quando os valores da matriz de coocorrência são iguais e baixo quando a concentração de valores na diagonal é alta ou quando as entradas da matriz de coocorrência possuem valores baixos.

$$Entropia = -\sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j) \log(p(i, j)). \quad (2.4)$$

Segundo Momento Angular

É a medida da homogeneidade da imagem. Uma imagem é considerada homogênea quando esta apresenta poucas variações de níveis de cinza.

$$\text{Segundo Momento Angular} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (p(i, j))^2. \quad (2.5)$$

Contraste

É a variação entre os *pixels* de determinado ponto, ou seja, mede a presença de transição abrupta de níveis de cinza (bordas) em uma imagem.

$$\text{Contraste} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j)(i - j)^2. \quad (2.6)$$

Momento da Diferença Inversa

Denota a homogeneidade da matriz de coocorrência. O momento da diferença é máximo quando a concentração dos valores na diagonal da matriz de coocorrência for máxima.

$$\text{Momento da Diferença Inversa} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{1}{1 + (i + j)^2} p(i, j). \quad (2.7)$$

Correlação

Corresponde a uma ideia de linearidade de dependências de tons de cinza em uma imagem. O valor da correlação é alto em uma imagem onde existe uma certa ordenação local dos níveis de cinza. A equação que representa essa medida é apresentada a seguir:

$$\text{Correlação} = \frac{\sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}, \quad (2.8)$$

em que μ_x , μ_y e σ_x , σ_y são as médias e os desvios padrão de p_x e p_y .

Média da Soma

É um indicador da média de tonalidades entre pares de *pixels*.

$$\text{Média da Soma} = \sum_{i=2}^{2Ng} i \cdot P_{x,y}(i). \quad (2.9)$$

Entropia da Soma

Indica quantidade de desorganização entre os *pixels*.

$$\text{Entropia da Soma} = -\sum_{i=2}^{2Ng} P_{x,y}(i) \log(P_{x,y}(i)). \quad (2.10)$$

Variância da Soma

Indica a variação da tonalidade do fundo da imagem.

$$\text{Variância da Soma} = \sum_{i=2}^{2Ng} (i - \text{Entropia da Soma})^2 P_x + y(i). \quad (2.11)$$

Variância

Mede a dispersão em relação à média, ou seja, indica a variação de tonalidades na imagem.

$$\text{Variância} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (i - \mu_x \mu_y)^2 p(i, j). \quad (2.12)$$

Variância da Diferença

Similar à variância da soma, mas com a média centrada em zero.

$$\text{Variância da Diferença} = \text{Variância de } p_{x-y}. \quad (2.13)$$

Entropia da Diferença

Semelhante à entropia da soma, é um indicador de quantidade de desordem entre os pixels.

$$\text{Entropia da Diferença} = -\sum_{i=0}^{Ng-1} p_{x-y}(i) \log(P_{x-y}(i)). \quad (2.14)$$

Informação de Correlação 1

É um apontador da correlação fundamentada na medida da entropia para elementos independentes e pares de elementos na imagem. A equação 2.15 representa essa característica estatística:

$$\text{Informação de Correlação 1} = \frac{HXY - HXY1}{\max[HX, XY]}, \quad (2.15)$$

em que

$$HXY = -\sum_i \sum_j p(i, j) \log(p(i, j)), \quad (2.16)$$

$$HXY1 = -\sum_i \sum_j p(i, j) \log(p_x(i) p_y(j)), \quad (2.17)$$

$$HX = \text{entropia de } p_x, \quad (2.18)$$

$$HY = \text{entropia de } p_y. \quad (2.19)$$

Informação de Correlação 2

É a medida da correlação baseada na semelhança entre entropias de elementos independentes da imagem e pares de elementos da imagem. A equação 2.20 representa essa característica estatística:

$$\text{Informação de Correlação 2} = (1 - \exp[-2.0(HXY2 - HXY)]^1/2), \quad (2.20)$$

em que

$$HXY2 = -\sum_i \sum_j p_x(i) p_y(j) \log(p_x(i) p_y(j)). \quad (2.21)$$

2.3.10.3 *Dimensão Fractal*

A dimensão fractal é uma medida estatística que indica o nível de complexidade e ocupação do espaço euclidiano por um objeto fractal ou uma quantidade que é auto-similar em alguma região do espaço ou intervalo de tempo (NAKAMURA *et al.*, 1985) (PETERS, 1994). Quando aplicada a texturas, a dimensão fractal atua como medida dos *pixels* que compõem essa textura, onde o nível de complexidade, por sua vez, está diretamente alusivo a sua aparência e à homogeneidade da textura. Assim, é possível quantificar a textura analisada em termos de homogeneidade, a fim de possibilitar sua comparação com outras texturas (CHAUDHURI; SARKAR, 1995) (SMITH *et al.*, 1996).

Existem diferentes métodos para mensurar a dimensão fractal e, entre eles, os principais métodos são: *Bouligand-Minkowski*, *Massa-Raio* e *Box-counting* (BACKES, 2010).

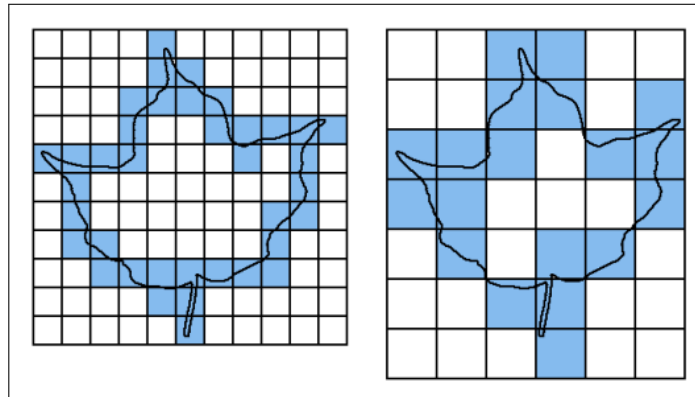
Bouligand-Minkowski

O método *Bouligand-Minkowski* apresenta uma grande sensibilidade em detectar as diferentes mudanças estruturais da forma sob análise. A estimativa da dimensão fractal utilizando o método de *Bouligand-Minkowski* é dada por meio do estudo da área de influência de um objeto. Essa área é criada a partir da dilatação desse objeto por um disco com raio de valor determinado (TRICOT, 1995)

Massa-Raio

O método *Massa-Raio* tem por base o estudo da distribuição dos pontos de uma forma em análise dentro de uma vizinhança r . Assim, círculos de raio r são sobrepostos sobre

Figura 12 – Divisão de uma imagem pelo método contagem de caixas para diferentes tamanhos de caixas.



Fonte – (BACKES, 2010).

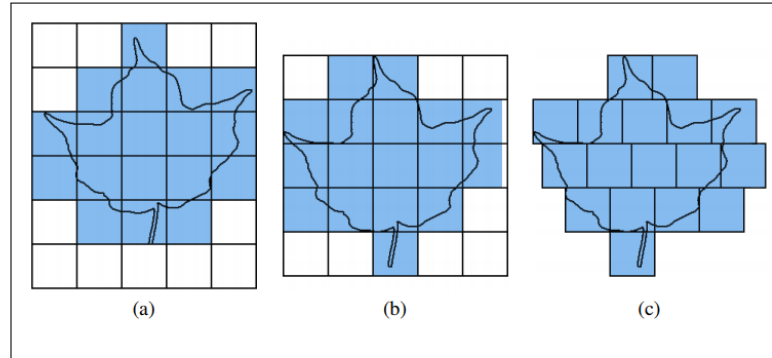
a forma, de modo a estimar a quantidade de pontos da forma presente no interior do círculo. Conforme o raio r do círculo utilizado aumenta, o número de pontos contidos no interior do círculo e a massa da forma aumentam também (FERNÁNDEZ; JELINEK., 2001)(CASERTA *et al.*, 1995).

Contagem de caixas

O método por contagem de caixas é um dos métodos mais conhecidos e utilizados para medir a dimensão fractal de um objeto. Tendo como fator suas características de fácil execução e simplicidade dos cálculos utilizados. Seu cálculo fundamenta-se na sobreposição de uma malha de quadrados sobre uma imagem e na correspondente contagem do número de quadrados necessários para cobrir o objeto contido na imagem. A Figura 12 apresenta a sobreposição de malhas distintas sobre uma mesma imagem, onde são aplicados diferentes tamanhos de caixas (COELHO; COSTA, 1995).

Ao utilizar a metodologia de contagem de caixas é possível realizar diversos refinamentos na sua técnica. Além do posicionamento da malha de quadrados sobre a imagem e da variação do número de caixas que interceptam o objeto, existe ainda a viabilidade de inserir os quadrados de forma independente, reduzindo ao máximo a contagem de caixas (JIAN *et al.*, 2006). A Figura 13 apresenta diferentes maneiras de sobrepor as caixas sobre a imagem.

Figura 13 – Diferentes maneiras de sobrepor as caixas sobre a imagem. a) sobreposição da malha de quadrados; b) sobreposição alinhada da malha de quadrados; c) Caixas ajustadas de forma independente.



Fonte – (BACKES, 2010).

2.3.10.4 Aplicação da dimensão fractal em análise de textura

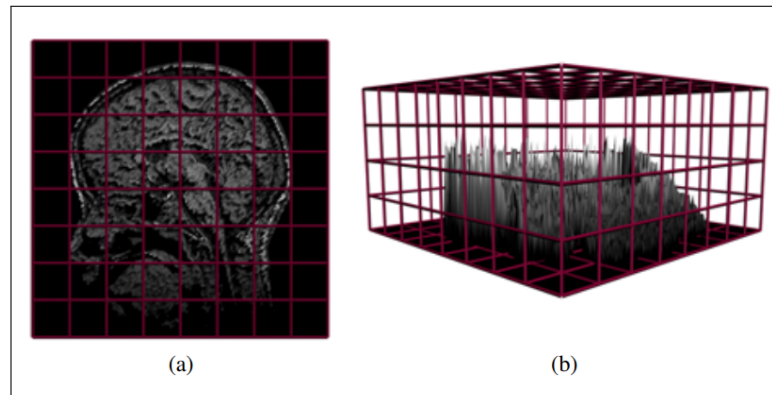
A dimensão fractal ao ser aplicada em análise de textura passa a atuar como uma medida de complexidade da organização dos *pixels* que constituem as imagens. O nível de complexidade obtido está diretamente relacionado ao aspecto visual, ou seja, homogeneidade da textura, possibilitando a sua quantificação bem como a comparação com outras texturas (CHAUDHURI; SARKAR, 1995)(SMITH *et al.*, 1996).

A metodologia por contagem de caixas quando aplicada na classificação de imagens em tons de cinza, como é o caso na análise de textura, considera a intensidade dos *pixels* como a altura daquele ponto da imagem. Ou seja, sua equação representativa deixa de realizar a contagem de quadrados para realizar a contagem de cubos de aresta r como apresentado na Figura 14 (BACKES; BRUNO, 2006).

As alterações que ocorrem na análise de textura produzem um novo ($N_A(r)$), em que ($N_A(r)$) é agora o número de cubos que interceptam a imagem A , sem que isso modifique as relações que estimam o valor da dimensão fractal. Logo a equação representativa da Dimensão fractal pelo método contagem de caixas é dada pela equação 2.22.

$$DF_{BoxCounting} = \lim_{x \rightarrow 0} \frac{\log(N_A(r))}{\log(r)}. \quad (2.22)$$

Figura 14 – Divisão de uma imagem pelo método *BoxCounting* aplicada em imagem em tom de cinza.



Fonte – (BACKES, 2010).

2.3.10.5 Lacunaridade

O conceito de lacunaridade foi desenvolvido para identificação de padrões de texturas de objetos fractais (PLOTNICK *et al.*, 1993). Fractais, por sua vez, são formas geométricas incapazes de serem classificadas nos padrões da geometria euclidiana, uma vez que apresentam três características essenciais que os descrevem e os diferenciam de outras formas: dimensão fractal, auto-semelhança em diferentes níveis de escala e sua complexidade infinita (GULICK, 1992).

A lacunaridade estabelece uma medida fundamentada no grau de invariância à translação que um fractal apresenta. Ela descreve o modo de organização e distribuição dos *pixels* em uma determinada região da imagem, ou seja, ela avalia o preenchimento do espaço ao medir a distribuição espacial dos “*gaps*” ou buracos existentes na imagem, diferentemente da dimensão fractal, que mede quão preenchido está o espaço. Assim é factível mensurar a homogeneidade de uma imagem ou de parte dela, de forma a torná-la comparável com outras imagens (PLOTNICK *et al.*, 1996).

Vários métodos para estimar a lacunaridade estão sendo desenvolvidos, sendo os mais populares baseados no uso do algoritmo *gliding-box*. Esse algoritmo faz uso dos momentos de probabilidade de primeira e segunda ordem da imagem com o propósito de estimar a sua lacunaridade. Tendo comportamento semelhante ao algoritmo do *Box-counting* utilizado para calcular a Dimensão Fractal. Nele, uma caixa de lado r é sobreposta ao canto superior esquerdo da imagem e então é realizada a contagem do número de pontos da imagem. Durante sua execução o mesmo processo é aplicado para todas as linhas e colunas da imagem, gerando uma

distribuição de frequência da massa da imagem. O número de caixas de lado r contendo uma massa S da imagem é designado por $n(S, r)$ e o total de caixas contadas por $N(r)$. Essa distribuição de frequência é então convertida para uma distribuição de probabilidade $Q(S, r)$, onde

$$Q(S, r) = n(S, r)/N(r). \quad (2.23)$$

O primeiro e o segundo momentos dessa distribuição são determinados como:

$$Z^1 = \sum S Q(S, r), \quad (2.24)$$

$$Z^2 = \sum S^2 Q(S, r). \quad (2.25)$$

A lacunaridade para uma caixa de tamanho r é então definida como:

$$\lambda(r) = Z^2 / (Z^1)^2. \quad (2.26)$$

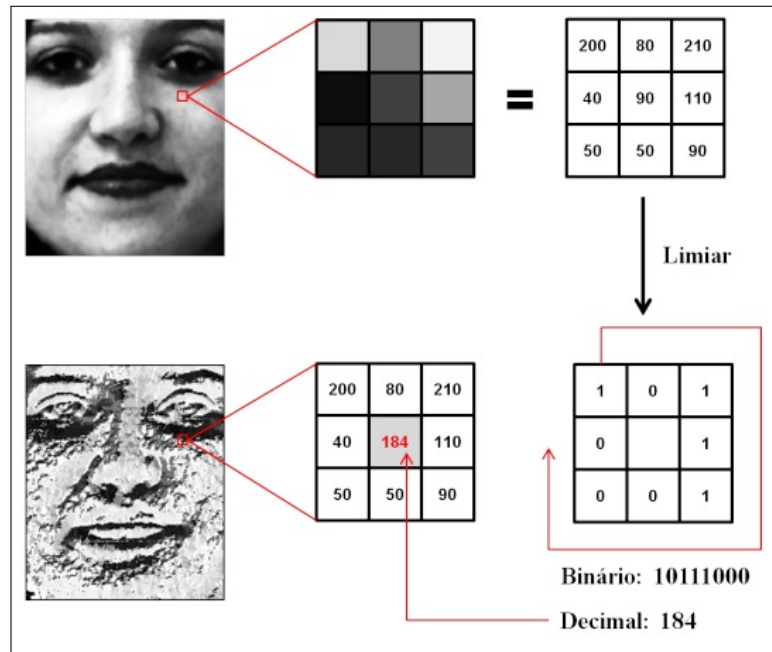
Outras características correspondentes à lacunaridade podem ser adquiridas alterando o tamanho da caixa utilizada no *gliding-box* (PLOTNICK *et al.*, 1996).

2.3.10.6 Local Binary Pattern - (LBP)

O LBP é um descritor estatístico desenvolvido para obter uma assinatura de textura analisando as intensidades relativas entre o *pixel* central de um círculo e os *pixels* a sua volta. O *pixel* central é utilizado como limiar para os *pixels* da circunferência, assim se um *pixel* da vizinhança tiver valor maior que o *pixel* central este é definido como um, caso contrário, será definido como zero. Um vetor binário é formado, concatenando todos os códigos binários obtidos, como no exemplo mostrado na Figura 15. Os números binários obtidos são classificados como padrões binários locais ou códigos LBP (HUANG *et al.*, 2011) (OJALA *et al.*, 2002).

A conversão do valor binário para um número decimal gera uma contagem para o valor do histograma. O algoritmo funciona como um laço onde a cada iteração um *pixel* é utilizado como *pixel* central e portanto um número é formado. Ao final é formado um histograma com valores decimais, o qual é utilizado como vetor de característica da imagem. Na aplicação do LBP pode-se escolher o tamanho do círculo a ser utilizado para composição do número binário, e os valores máximos de vizinhos e raio. Se uma janela de 3×3 for escolhida, haverá 8 *pixels* a ser utilizados para formar o número binário. O histograma e, conseqüentemente, o vetor de características, terá $2^8 = 256$ posições. Assim, o histograma de rótulos LBP, calculados sobre uma região, pode ser explorado como um descritor de textura (OJALA *et al.*, 2002).

Figura 15 – Aplicação básica do LBP a uma imagem em tons de cinza. Valor binário resultante 10111000 e decimal 184.



Fonte – (AMARAL; THOMAZ, 2012).

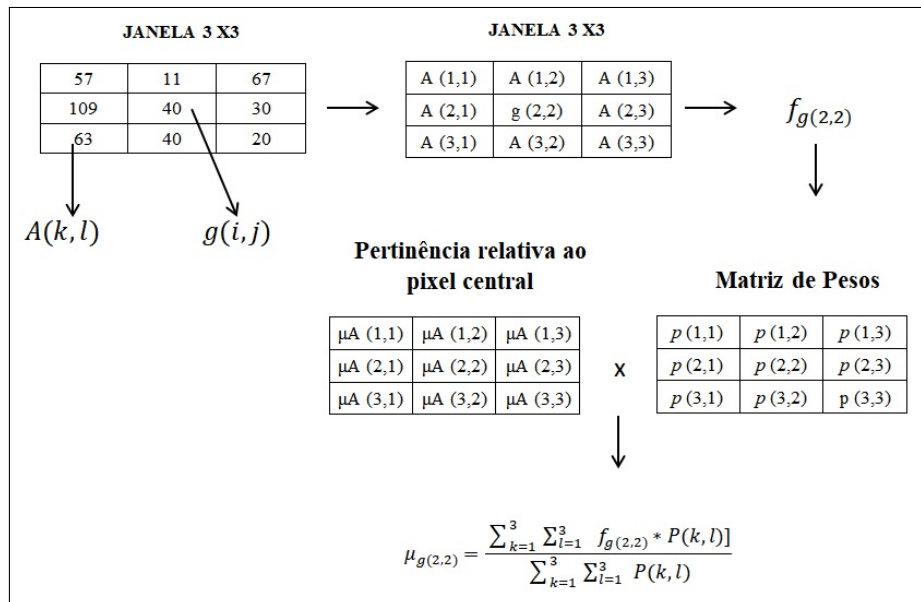
2.3.10.7 Local Fuzzy Pattern - (LFP)

Vieira et al. (2012) propuseram o *Local Fuzzy Pattern - (LFP)* como uma metodologia para a classificação de texturas fundamentada em números *fuzzy*, capaz de lidar com problemas que envolvem graus de ambiguidade e incerteza. Essa metodologia teve por base o novo método de detecção de bordas apresentado por Boaventura e Gonzaga (2007), o *Fuzzy Number Edge Detector (FUNED)*, que foi desenvolvido para minimizar os defeitos apresentados pela falta de robustez do LBP contra ruídos.

O LFP é um descritor de textura local que estabelece o grau de pertinência (*membership degree*) de cada *pixel* da região de interesse, considerando a intensidade dos tons de cinza dos *pixels* como componentes de um conjunto *fuzzy* (VIEIRA *et al.*, 2012), ou seja, ao analisar a vizinhança de *pixels*, utiliza números *fuzzy* para determinar uma relação com propriedades descritivas (RIBEIRO *et al.*, 2014).

Analisando uma janela de $W \times W$ *pixels*, a metodologia assume que cada distribuição dos níveis de cinza dos *pixels* vizinhos ao *pixel* central é um conjunto *fuzzy*, devido ao grau de incerteza gerado pelo processo de aquisição da imagem, pré-processamento e ao ruído produzido por diversas fontes. Cada *pixel* central possui um identificador que indica a relação de seu valor

Figura 16 – Geração de um código LFP.



Fonte – (TRAVAINI, 2015).

de cinza para com os valores da vizinhança. O grau de pertinência do *pixel* central $g(i, j)$ para a vizinhança $W \times W$ é obtido pela média ponderada dos graus de pertinência dos pixels vizinhos, conforme a equação 2.27, na qual $F_{g(i,j)}$ é a função de pertinência aplicada para cada pixel vizinho e $P(k, l)$ é uma matriz de pesos de mesmo tamanho que a vizinhança $W \times W$. A Figura 16 exemplifica a geração de um código LFP.

$$\mu(g(i, j)) = \frac{\sum_{k=1}^w \sum_{l=1}^w F_{g(i,j)}(g, (k, l)) \times P(k, l)}{\sum_{k=1}^w \sum_{l=1}^w P(k, l)} \quad (2.27)$$

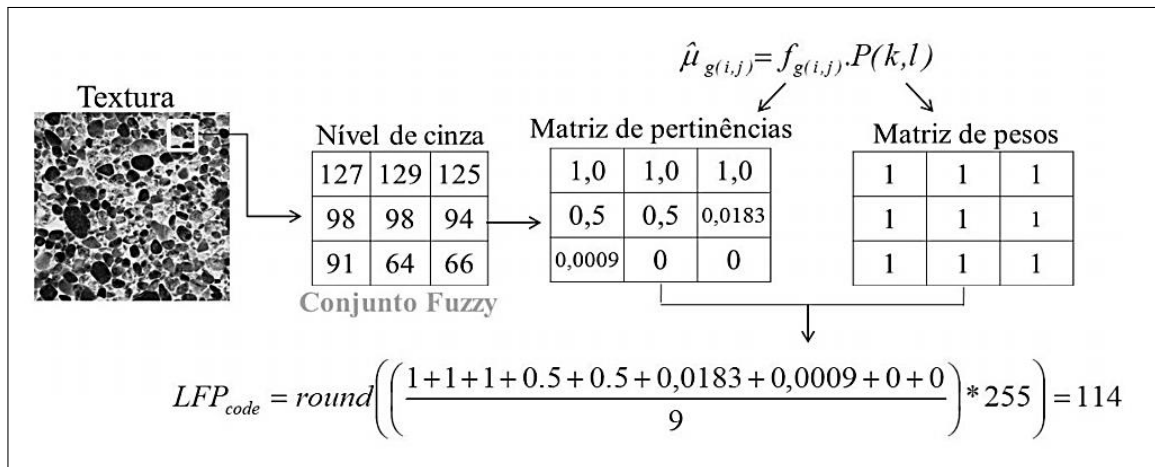
A função de pertinência pode ser definida pelo usuário durante a aplicação do método, por isso que sua função de pertinência é heurística e não única (VIEIRA *et al.*, 2012).

O LBP é paramétrico e passível de otimização por utilizar uma representação por números *fuzzy*. O método foi sugerido com três principais funções de pertinência que geram, por sua vez, as metodologias a seguir:

LFP-sigmóide

A função de pertinência que descreve o LFP-s é a função sigmóide, a qual é dada

Figura 17 – exemplo do cálculo do LFP-s para uma amostra de 3×3 .



Fonte – (VIEIRA *et al.*, 2012).

pela equação 2.28

$$f_{g(i,j)(A(k,l))} = \frac{1}{1 + e^{-\frac{A(k,l) - g(i,j)}{\beta}}}, \quad (2.28)$$

na qual β é a inclinação da curva sigmóide, $A(k,l)$ os *pixels* da vizinhança e $g(i,j)$ é o *pixel* central. A função LFP-s considera que os *pixels* de vizinhança apresentarão pesos iguais na constituição do grau de pertinência do *pixel* central, assim para uma vizinhança de 3×3 *pixels*, a matriz de pesos $P(k,l)$ é dada por:

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Tendo realizado a multiplicação dos graus de pertinência dos *pixels* da vizinhança pelos seus respectivos pesos, uma média ponderada de seus valores é estabelecida. Este valor, que pode variar entre 0 e 1, é convertido para uma escala de cinza (0-255). A Figura 17 apresenta um exemplo do cálculo do LFP-s para uma amostra de 3×3 *pixels*.

A Figura 18 representa a função de pertinência sigmóide.

Figura 18 – Função de pertinência sigmóide.



Fonte – (CHIERICI, 2015).

A partir dos códigos LFP-s já convertidos para escala de cinza, um histograma é montado, representando a identidade da textura analisada.

LFP - triangular

A LFP triangular é obtida através de uma função de pertinência triangular e simétrica, descrita pela equação 2.29

$$f_{g(i,j)} = \max\left(0, 1 - \frac{|g(i,j) - A(k,l)|}{\delta}\right), \quad (2.29)$$

em que $A(k,l)$ são os *pixels* na vizinhança $W \times W$, δ é o espalhamento de número *fuzzy* e a matriz de pesos é da forma:

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

O valor zero do *pixel* central desta matriz de pesos deve-se ao fato de que o *pixel* central da região analisada é excluído do cálculo de sua pertinência (BOAVENTURA, 2010).

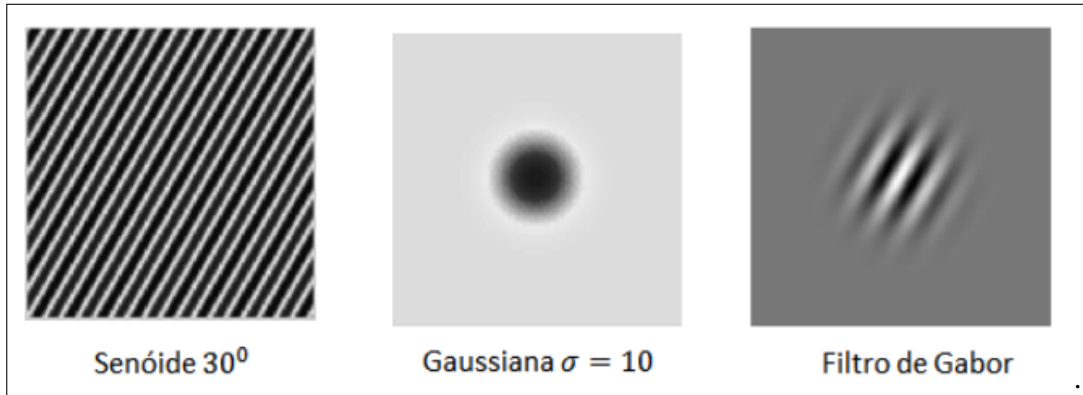
LFP - gaussiana

Outra função de pertinência é a função gaussiana, dada pela equação 2.30

$$f_{g(i,j)} = \exp\left(-\frac{A(k,l) - g(i,j)^2}{\sigma}\right), \quad (2.30)$$

em que σ é o desvio padrão.

Figura 19 – Um filtro de Gabor é formado por uma senóide convoluída com uma gaussiana no domínio do espaço.



Fonte – (NEIVA, 2016)

2.3.10.8 Filtros de Gabor

Os filtros de Gabor representam o tempo e a frequência de um sinal. A representação no domínio do tempo define a amplitude do sinal em cada momento, enquanto que a do domínio da frequência dispõe de senóides infinitamente longas (no domínio do tempo), definidas pela frequência, amplitude e fase (GABOR, 1946), assim expressando, respectivamente, as características de localização e textura, possibilitando a manipulação de vários parâmetros, como orientação, simetria e excentricidade. Os bancos de filtros de Gabor são formados usando combinações desses parâmetros (YIMING *et al.*, 2004).

A formação dos filtros de Gabor é dada por um sinal senoidal modulado por uma gaussiana, apresentada na equação 2.31

$$h(t) = e^{-\gamma^{-2}(t-t_0)^2} \cdot (\cos[2\pi f(t-t_0) + \delta] + j\sin[2\pi f(t-t_0) + \delta]), \quad (2.31)$$

em que γ é a constante da gaussiana, inversamente proporcional à largura da função, t_0 representa o centro da gaussiana, δ e f são, respectivamente, a fase da senoide complexa e sua frequência. A Figura 19 demonstra a formação de um filtro de Gabor (GABOR, 1946).

A aplicação de um filtro de Gabor em uma imagem pode ser feita de três formas: usando somente a parte real; usando somente a parte imaginária; ou usando a magnitude (CHAVES, 2015). A extração de características consiste em cada filtro $g_{\sigma\theta}(x,y)$ ser convoluído com a imagem, gerando assim uma saída $c_{\sigma\theta}(x,y)$ (NEIVA, 2016). O vetor de características obtido $v = [E_{11}, E_{12}, \dots, E_{\sigma_{max}\theta_{max}}]$ pode ser gerado calculando a energia de cada uma das imagens

c de acordo com a equação 2.32

$$E_{\sigma,\theta} = \Sigma_{x,y}[c_{\sigma,\theta}(x,y)]^2. \quad (2.32)$$

Os parâmetros determinados para a extração de características de uma imagem utilizando filtros de Gabor são a quantidade de variações de escala e orientações. As imagens obtidas após cada filtro ser convoluído com a imagem original são concatenadas para a geração do vetor de características (MANJUNATH; MA, 1996). Geralmente, utilizam-se cinco escalas e oito orientações, totalizando 40 filtros (DENG *et al.*, 2005) (OWUSU *et al.*, 2014).

2.3.10.9 Descritores de Fourier

Os descritores de Fourier são utilizados em análise de texturas, sendo capazes de codificar informações sobre a forma de um objeto em um conjunto de dados. A obtenção dos descritores pode ser realizada de forma simples ao utilizar os valores obtidos através de seu espectro de potência $P(u, v)$. Ao considerar que as frequências variam partindo do centro da imagem os descritores podem ser considerados com a soma dos valores do espectro localizados a uma certa distância do centro da imagem (WESZKA *et al.*, 1976), como demonstrado na equação 2.33

$$D(i) = \Sigma P(u, v), \forall d(u, v) = i, \quad (2.33)$$

em que $P(u, v) = |F(u, v)|^2$, i é um valor inteiro de distância e $d(u, v)$ é a distância da coordenada (u, v) ao centro da imagem transformada.

Através da transformada de Fourier, é possível analisar uma imagem como um conjunto de sinusóides espaciais em várias direções, tendo cada senoide uma frequência precisa (ACHARYA; RAY, 2005). Além disso, também é possível realizar a remoção de ruídos e desfoques (HSU, 2001).

Em processamento de imagens ou processamento de sinais, a série discreta de Fourier é chamada de Transformada Discreta de Fourier (DFT). A DFT é uma transformação de coordenadas, que resulta em componentes pertencentes aos números complexos. Cada coeficiente é obtido pela combinação linear dos elementos da entrada com o núcleo da transformada. O resultado da transformada de Fourier representa exatamente um período (HAYKIN; VEEN, 2001). Linearidade, periodicidade e inversão são algumas propriedades básicas da DFT. A

propriedade de inversão proporciona a definição do inverso da DFT e remoção da assimetria entre a sequência original de comprimento N e a sequência transformada de comprimento infinito. Uma consequência desta propriedade é que não existem duas sequências distintas que podem ter a mesma DFT (PUPIN, 2011)

Para utilizar a transformada de Fourier em imagens deve-se usar sua forma bidimensional, que nada mais é que a extensão do estado unidimensional. As propriedades da transformada de Fourier bidimensional são translação, periodicidade, convolução e correlação (GONÇALVES, 2004).

Segundo (SOLOMON; BRECKON, 2011) a função da transformada de Fourier bidimensional é dada pela equação 2.34

$$F(K_x, K_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{i(K_x x + K_y y)} dx dy. \quad (2.34)$$

Para imagens a transformada pode ser calculada de acordo com a função abaixo (em que $M \times N$ é referente ao tamanho da imagem):

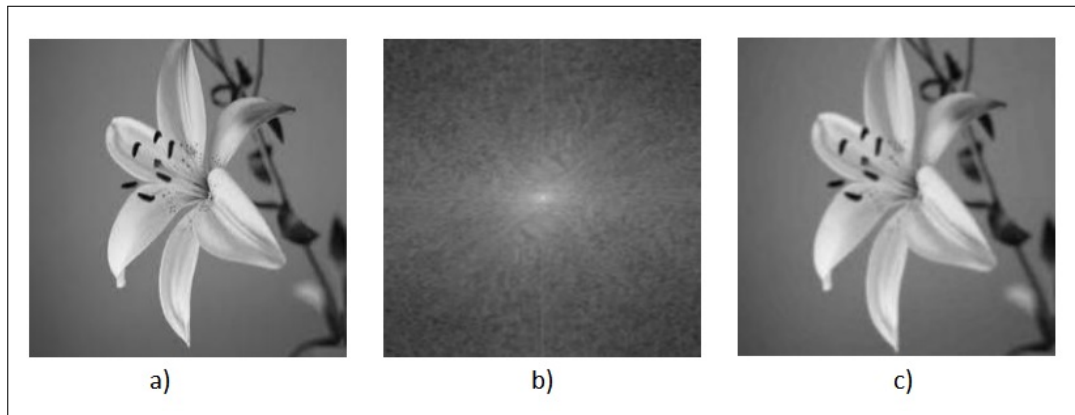
$$F(K_x, K_y) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2j\pi(K_x x/M + K_y y/N)}. \quad (2.35)$$

Pode-se mostrar que Transformada de Fourier em duas dimensões é invertível e sua inversa é dada pela equação 2.36

$$F(u - u_0, v - v_0) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(K_x, K_y) e^{i(K_x x + K_y y)} dK_x dK_y. \quad (2.36)$$

A Figura 20 apresenta uma imagem original, a aplicação da transformada de Fourier (b) e da inversa de Fourier (c).

Figura 20 – a) Representação da imagem original, b) Representação da transformada de Fourier, c) Representação da transformada inversa de Fourier.



Fonte – (PUPIN, 2011).

2.3.11 Método de Validação

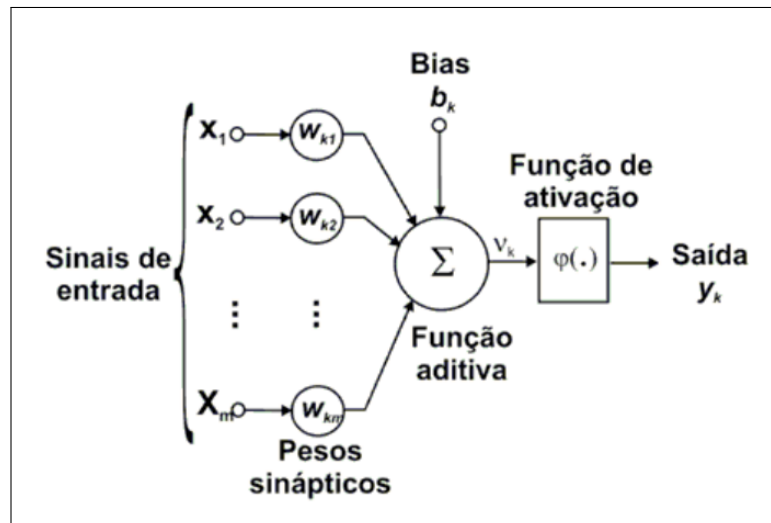
2.3.11.1 Rede Neural

A rede neural humana tem a predisposição natural de reter conhecimento experimental e torná-lo disponível para o uso (HAYKIN, 2001). Assim também é o comportamento de uma rede neural artificial (RNA), a qual pode ser definida como uma estrutura de processamento (rede), suscetível a implementação em dispositivos eletrônicos, formada por um número de unidades interconectadas (neurônios artificiais), em que cada unidade apresenta um comportamento particular de entrada/saída. O comportamento de cada unidade é diretamente relacionado aos parâmetros adotados, como: função de transferência, tamanho da vizinhança, interações com outras unidades, além das possíveis entradas externas (FERNANDES, 2005).

No modelo neural podem ser identificados três elementos básicos: o conjunto de sinapses, um somatório e uma função de ativação (HAYKIN, 2001). Em sua aplicação o conhecimento é obtido através de etapas de aprendizagem e pesos sinápticos, os quais são usados para armazenar o conhecimento. A Figura 21 apresenta o esquema de um neurônio.

Uma sinapse é o nome dado à conexão existente entre neurônios e a função de ativação simboliza o efeito que a entrada interna e o estado atual de ativação exercem na definição do próximo estado de ativação da unidade. Geralmente o estado de ativação é classificado como uma função algébrica (CASTRO, 1998).

Figura 21 – Esquema de um neurônio.



Fonte – (HAYKIN, 2001).

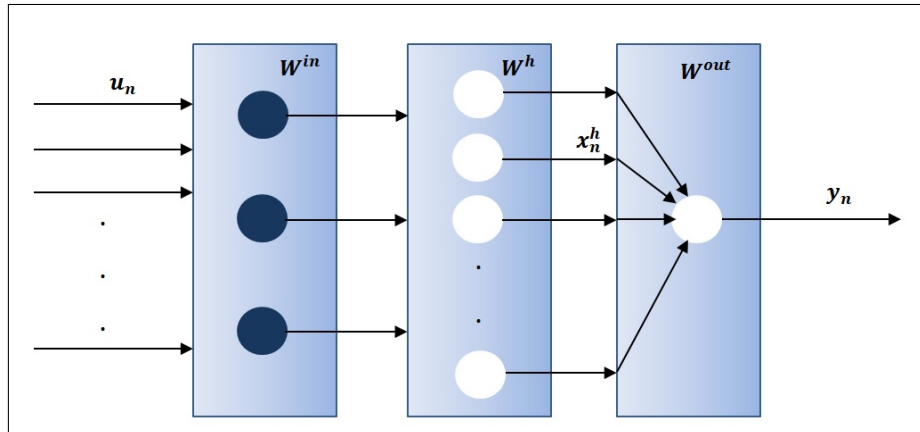
Diferentes tipos (arquiteturas) de redes neurais são formadas realizando a variação de alguns parâmetros, como: o tipo de neurônio utilizado, a função de ativação aplicada, quantidade de camadas existentes, o modo de conexão entre os diversos neurônios, entre outros (NIGRIN, 1993) (HAYKIN, 1998) (BISHOP, 1996).

2.3.11.2 Extreme Learning Machine - (ELM)

A ELM é uma rede neural de aprendizado extremo, tendo por características o fato de que sua camada oculta da rede não precisa ser treinada (YANG *et al.*, 2016). Proposta por Huang *et al.* (2004), com diferencial dos demais algoritmos clássicos de aprendizado que são baseados em gradientes para atingir um erro mínimo de treinamento, mas não consideram a magnitude dos pesos, a ELM tende a alcançar não apenas o menor erro de treinamento, mas também a menor norma de pesos (HUANG *et al.*, 2012). O que a leva a ter o melhor desempenho de generalização para redes neurais *feedforward*, intervenção mínima e os benefícios de escalabilidade, menos complexidade computacional e solução unificada para diferentes aplicações práticas (por exemplo, regressão, classificação binária e multiclasse) (HUANG *et al.*, 2004)(HUANG *et al.*, 2006) .

Os pesos nas camadas escondidas das ELMs são gerados de forma aleatória, ou seja, apenas aqueles da camada de saída passam por um processo de ajuste. Essas redes possuem a capacidade de aproximar qualquer mapeamento não linear, contínuo, limitado e diferenciável com precisão arbitrária, fato demonstrado por meio de uma abordagem construtiva em Huang

Figura 22 – Rede Neural ELM.



Fonte – Própria autora, tendo base (DELATORRE *et al.*, 2015).

et al. (2006). A precisão está aliada ao fato de que a inserção de novos neurônios na camada escondida eleva o potencial de aproximação da rede (HUANG *et al.*, 2004)(HUANG *et al.*, 2006). A Figura 22 apresenta esta arquitetura.

Os neurônios da camada intermediária possuem ativações dadas pela expressão a seguir:

$$X_n^h = f^h(W^h u_n + b), \quad (2.37)$$

na qual $u_n = [u_n, u_{n-1}, \dots, u_{n-k+1}]^T$, é o vetor que contém o sinal de entrada, W^h representa os coeficientes (pesos) da camada intermediária, b significa o bias de cada unidade da camada escondida e $f^h(\cdot)$ especifica as ativações dos neurônios da camada escondida (DELATORRE *et al.*, 2015). As saídas são geradas pela expressão:

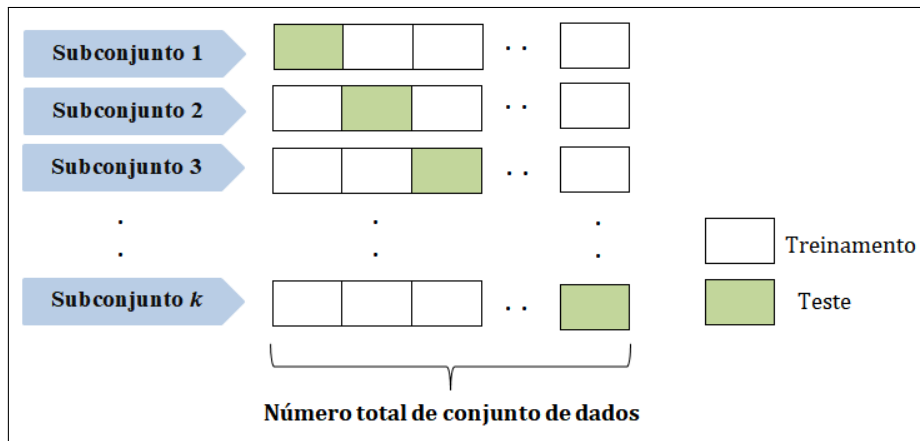
$$y = W^{out} X_n^h, \quad (2.38)$$

em que W^{out} representa os pesos da camada de saída, os quais podem ser obtidos pelo método dos mínimos quadrados (BROTTO *et al.*, 2017).

2.3.11.3 K-fold e Leave-one-out

A validação cruzada *K-fold* é uma técnica que utiliza todas as amostras disponíveis como amostras de treinamento e teste, por isso é conhecida como uma técnica computacional intensiva (DUCHESNE; BRUNO, 2005). Neste método de validação o conjunto de dados é dividido em k subconjuntos, contendo exemplos de todas as classes. A cada iteração, é determinado um subconjunto distinto para ser utilizado como teste, e os outros $k-1$ para treino.

Figura 23 – Visualização do funcionamento da validação cruzada *k-fold*.



Fonte – Própria autora, tendo base (MOURA, 2018).

Os processos de treinamento e teste são repetidos k vezes. Logo, ao considerar uma base de dados hipotética que contenha 50 registros e definindo $k = 5$, a base de dados será dividida em 5 subconjuntos. Concluída a divisão, um subconjunto será empregado na validação do modelo e os conjuntos restantes são utilizados como treinamento. O processo é então repetido k vezes, de forma que cada um dos k subconjuntos sejam usados para teste na validação do modelo. Por fim é calculada a média dos erros das k iterações. Na Figura 23 é mostrada a representação do funcionamento da validação cruzada *k-fold*.

O leave-one-out é um método simplificado da validação cruzada e extrema da validação *k-fold*. Nele os N padrões são divididos em dois conjuntos: um com 1 elemento e o outro com $N-1$ elementos. O método realiza a modificação dos conjuntos da mesma forma que ocorre no método da validação cruzada. Tendo, por exemplo, uma base de n dados com 100 padrões, primeiramente o treino é efetuado com 99 padrões e o teste é realizado com a amostra restante. Depois o treinamento é feito com outros 99 padrões e o teste é feito com a amostra remanescente. Ao término todas as amostras passaram pelo processo de treinamento, gerando 100 valores de erros. Então é calculada a média desses erros para gerar o erro final. Esse método possui a mesma vantagem de confiabilidade da validação cruzada, em que todos os padrões passam pelas fases de treinamento e de teste (DIAS, 2004)(RIBEIRO, 2003)

2.3.11.4 Métricas para Avaliar os Classificadores

Para avaliar os resultados de um classificador, utiliza-se a matriz de confusão, também denominada tabela de contingência, a qual mostra as predições corretas, bem como das confusões

geradas. Na matriz de confusão $K \times K$, os elementos da diagonal principal dão o número dos acertos do classificador. Os elementos das colunas i da matriz representam o número de amostras que o classificador identificou como sendo das classes 1 a k . Na Tabela 2 observa-se uma matriz de confusão de duas classes (MATHER; KOCH, 2011).

Tabela 2 – Esquema de um neurônio.

		Classe Predita	
		P	N
Classe Correta	P	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	N	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Fonte – (BRAGEIRO, 2015).

A partir desta matriz, quatro indicadores são obtidos: VP - número de amostras classificadas corretamente como positivas, FP - número de amostras classificadas erroneamente como positivas, FN - número de amostras classificadas erroneamente como negativas e VN - número de amostras classificadas corretamente como negativas. Em seguida, são calculadas as métricas (BRAGEIRO, 2015):

Acurácia (Acc): calcula a proporção de resultados verdadeiros (VP e VN) entre o total de classes.

$$ACC = \frac{VP + VN}{VP + FP + VN + FN}. \quad (2.39)$$

Precisão: calcula a precisão das predições de uma classe.

$$precisão = \frac{VP}{VP + FP}. \quad (2.40)$$

Sensibilidade (S): proporção de amostras positivas que foram corretamente identificadas como positivas.

$$S = \frac{VP}{VP + FN}. \quad (2.41)$$

Especificidade: proporção de amostras negativas que foram corretamente identificadas como negativas.

$$Especificidade = \frac{VN}{VN + FP}. \quad (2.42)$$

2.3.12 *Aplicação de Análise de Textura em Inspeção Industrial*

Em vários segmentos industriais, uma das propriedades visuais mais importantes que devem ser sujeitas ao controle de qualidade é a textura de matérias-primas e produtos. Exemplo desta aplicação está na detecção de defeitos em têxteis e inspeção automatizada do desgaste de tapetes e pinturas de automóveis por meio de análise de imagens. Siewe et al. (1988) propuseram um método para a avaliação do desgaste de tapetes, utilizando características de texturas simples que são computadas a partir da diferença entre as estatísticas de segunda e de primeira ordem dos níveis de cinza. Eles demonstraram que as características numéricas de texturas obtidas a partir dessas técnicas podem ser usadas com sucesso para caracterizar um tapete.

Muitas abordagens para classificação de textura se encontram hoje na literatura. Connors et al. (1983) utilizaram métodos de análise de texturas para detectar automaticamente defeitos em madeira serrada. A detecção de defeitos é realizada dividindo a imagem em sub-janelas e classificando cada sub-janela em uma das categorias de defeitos. Os recursos utilizados para executar esta classificação são baseados em características tonais tais como variância, média, assimetria, curtose de níveis de cinza, juntamente com características texturais calculadas a partir de matrizes de coocorrência de níveis de cinza, analisadas por meio das imagens da madeira.

Jain et al. (1990) também usaram as características de texturas, calculadas a partir de um banco de filtros de Gabor para classificar automaticamente a qualidade de superfícies metálicas pintadas. Diversas abordagens para inspeção por texturas também foram propostas por (FARROKHINIA, 1990) , (RAO; SCHUNCK, 1991) e (CHEN; JAIN, 1988).

Este capítulo abordou os fundamentos necessários para o desenvolvimento da metodologia proposta. O próximo capítulo apresenta os seus detalhes da implementação, que compreende as etapas de pré-processamento, extração de características e classificação do óleo vegetal de algodão e *linter*.

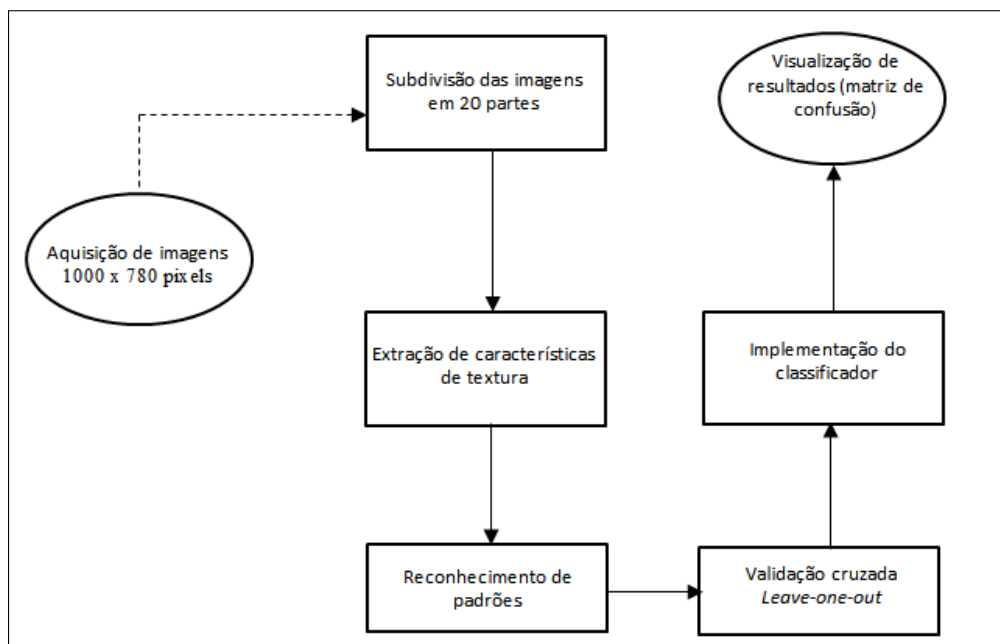
3 METODOLOGIA

Este capítulo detalha a metodologia proposta para a classificação do *linter* e óleo do algodão. Por se tratar de objetivos diferentes para a análise de cada tipo de material, este capítulo está dividido em duas partes. Na primeira parte será tratado o processo de classificação do *linter*, com explanação de cada um dos passos. A segunda parte corresponde à análise de imagens do óleo vegetal de algodão.

3.1 CLASSIFICAÇÃO DO LINTER

A Figura 24 apresenta de forma resumida a metodologia adotada para análise e classificação do *linter*. Inicia-se com a captura das imagens, após esta etapa, as imagens são subdivididas para obter uma base de dados maior. Posteriormente são aplicadas as metodologias de extração de características, as quais serão a base de entrada para os métodos de reconhecimento de padrões adotados: KNN, LDA e QDA. A forma de validação cruzada dos classificadores foi o método *leave-one-out*. Para facilitar a visualização dos desempenhos dos métodos utilizados, os valores de acurácia obtidos foram inseridos em matrizes de confusão.

Figura 24 – Metodologia proposta para classificação do *linter*.



Fonte – Própria autora.

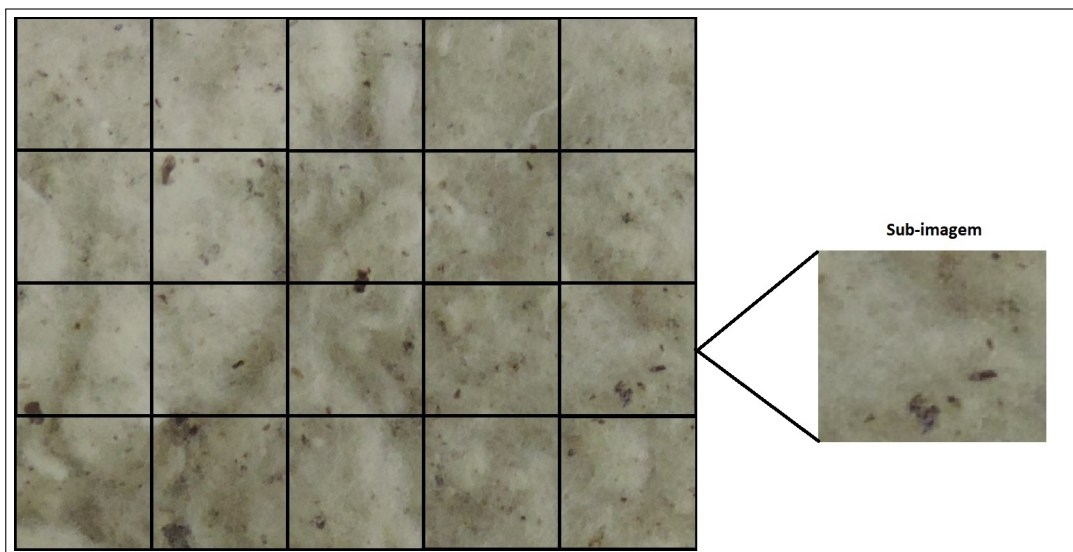
3.1.1 Amostras

As amostras de *linter* estudadas neste trabalho foram disponibilizadas pela Indústria e Comércio de Rações Golfinho Ltda, localizada na cidade de Sobral-Ce, Brasil.

3.1.2 Aquisições de imagens e pré-processamento

As imagens utilizadas para análise foram obtidas com uma câmera semiprofissional, no formato JPG. No total foram adquiridas 124 imagens com dimensões de 1000 x 780 *pixels*, contendo cinco classes comerciais distintas. Com o objetivo de aumentar a base de imagens e diminuir o tempo de processamento, as imagens são subdivididas em 20 recortes cada uma e organizadas conforme o padrão de qualidade do *linter*, tendo assim, 2470 imagens de 200 x 195 *pixels*. A Figura 25 apresenta como as imagens foram divididas em sub-imagens.

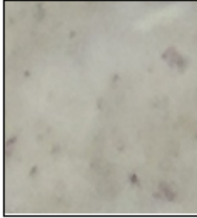




Figura 25 – Divisão da amostra médio Japão.



Fonte – Própria autora.

Para avaliar o comportamento do classificador conforme a complexidade de classificação, ou seja, tendo em consideração a quantidade de classes a serem determinadas, o grupo de sub-imagens foi analisado de duas formas: A primeira possuindo cinco classes, padrão china (510 sub-imagens), qualidade média Japão (490 sub-imagens), alta qualidade Japão (510 sub-imagens), baixa qualidade (500 sub-imagens) e péssima qualidade (460 sub-imagens). A segunda forma consistiu em agrupar as cinco classes de *linter* em dois grupos, boa qualidade (padrão China, alto Japão e médio Japão) e qualidade ruim (baixa qualidade e péssima qualidade).

Figura 26 – Classes comerciais do *linter*.

Classes do <i>Linter</i>				
Boa qualidade			Baixa qualidade	
Alta qualidade Japão	Média qualidade Japão	Padrão China	Baixa qualidade	Péssima qualidade
				

Fonte – Própria autora.

As amostras do *linter* foram adquiridas de lotes diferentes de cada tipo comercial da indústria. A Figura 26 apresenta as cinco categorias diferentes de classes do *linter* presentes no banco de dados.

Como apresentado na Figura 26, as classes de *linter* analisadas apresentam semelhanças, o que torna difícil identificá-las por uma inspeção visual. Apesar da similaridade visual do *linter* entre distintas classes, este subproduto do algodão também apresenta características diferentes, o que impacta no seu valor comercial.

3.1.3 Extração das características

A extração de características é a etapa em que as informações quantitativas da imagem são extraídas para descrever diferentes informações como cor, forma e textura. O método utilizado para classificação do *linter* foi por análise de textura.

Tanto os extratores de características como os classificadores aplicados neste trabalho foram desenvolvidos no software Matlab-R2016a. Inicialmente as imagens que compõem a base de dados foram convertidas em escala de cinza para extração dos dados e formação dos vetores de características.

3.1.3.1 Parâmetros dos descritores de textura

- Descritores de Haralick

De cada uma das sub-imagens resultantes foram extraídas as características de textura apresentadas por Haralick, as quais se baseiam na matriz de coocorrência. A matriz de

coocorrência relaciona cada um dos *pixels* com seus vizinhos e para cobrir toda a vizinhança é preciso gerar quatro matrizes simétricas, cada uma delas obtidas por um ângulo diferente (0° , 45° , 90° e 135°) (HARALICK *et al.*, 1973). Outro parâmetro definido é a distância (d), e neste trabalho foi adotado $d = 1$.

Para definir quais características apresentadas por Haralick proporcionam a melhor acurácia na classificação do *linter*, foi feita uma combinação com as 14 características apresentadas e os valores de acurácia obtidos foram comparados. Assim, foram definidos para obtenção do vetor de características os dados de contraste, entropia, correlação, média da soma, informação de correlação 2, momento da diferença inversa e energia. Para cada sub-imagem foram extraídas 7 características de textura e para cada medida quatro valores, logo é obtido um vetor com 28 características de cada imagem analisada.

- Dimensão fractal

O método utilizado foi o de contagem de caixas, em que o tamanho das caixas a serem utilizadas foi obtido por busca exaustiva. A melhor acurácia foi obtida com os tamanhos de caixa $d = [5, 10]$. O vetor de características para o método é formado por 1 atributo normalizado.

- Lacunaridade

O método de lacunaridade quantifica como o espaço em uma imagem está preenchido. Para análise de textura é utilizada uma binarização simples dentro da caixa determinada. Esta caixa é utilizada para medir a quantidade de *pixels* na imagem que possuem intensidade acima do valor médio de intensidade apresentado na caixa. O tamanho dos lados (l) do quadrado foi adotado por busca exaustiva, resultando em $l = 8$. O vetor de características resultante é de 1 atributo normalizado.

- LBP

Para cada sub-imagem é aplicado o descritor de textura. A seguir é construído um histograma registrando a distribuição das probabilidades dos códigos LBP gerados. O código LBP é calculado usando n pontos de amostragem em uma tabela de mapeamento. O ponto de amostragem é de 8 vizinhos tendo como mapeamento uma rotação invariável em quatro graus. Assim, esse procedimento gera um vetor de características contendo 32 atributos.

- LFP

Diferentemente do LBP, que tem uma formação direta do histograma de características, o LFP precisa que o valor de seu código seja quantizado, de modo a permitir a construção do histograma. Neste trabalho, o método LFP foi aplicado utilizando a função sigmóide, por meio de uma variante do LFP-s, *Sampled Local Fuzzy Pattern* (Sampled LFP). A escolha da resolução empregada na descrição de textura é feita de maneira heurística, visto que determinada característica pode ser mais evidente em determinada resolução.

O Sampled LFP é um método paramétrico que compõe a vizinhança de um *pixel* central a partir de uma vizinhança de simetria circular de raio R . Ou seja, a configuração da análise multirresolução é dada pelo número de vizinhos, ou amostras, e pelo valor de raio da vizinhança. A determinação do melhor valor de β , utilizado na definição do valor da pertinência de cada *pixel* da vizinhança, é chamada sintonização e foi feita por treinamento, repetindo o processo de classificação até obter a sensibilidade máxima.

Para o presente trabalho, o número de vizinhos adotado foi de 8, raio igual a 1 e o valor para β foi definido em 1,20. Essas faixas de valores foram definidas empiricamente. O vetor obtido possui 239 atributos.

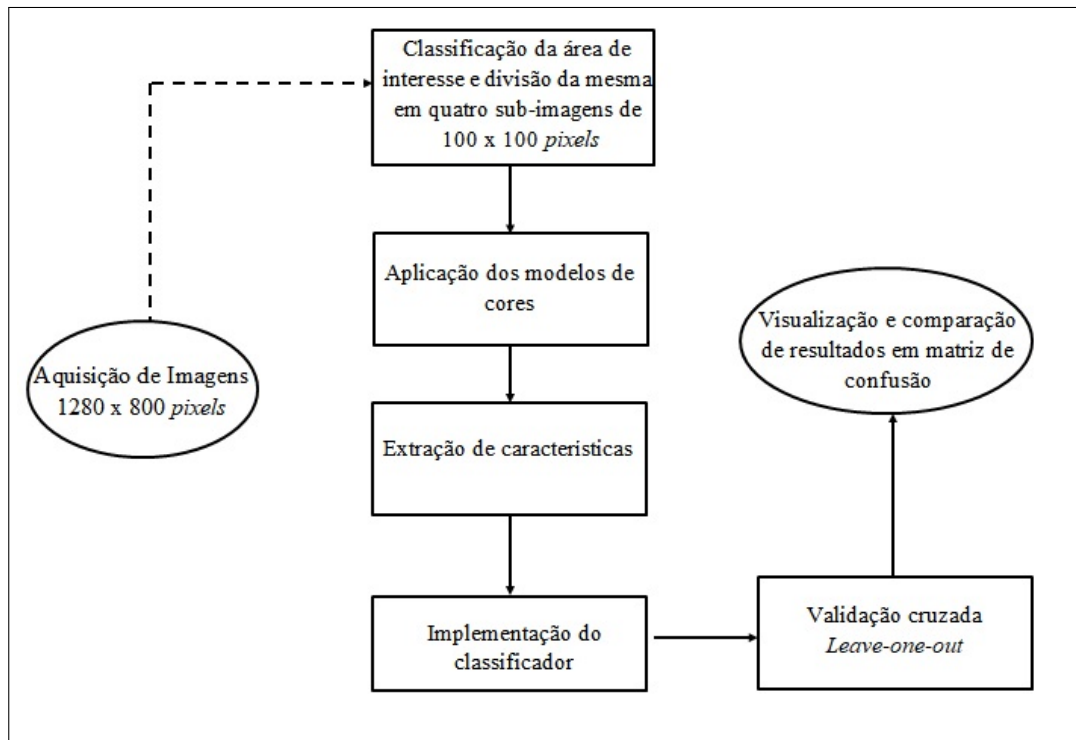
- Filtros de Gabor

Os parâmetros de um filtro de Gabor são a frequência W , o parâmetro de orientação θ e as escalas σ_x e σ_y da função gaussiana. As diferentes orientações σ e frequências W são, dessa forma, parâmetros-chave em uma aplicação de processamento de textura. Assim as imagens analisadas passam por uma convolução ao serem submetidas aos filtros de Gabor, apresentando em sua composição várias escalas e orientações de uma determinada configuração original (MANJUNATH; MA, 1996). Para a análise realizada neste trabalho, foi considerada uma família de 24 filtros (4 filtros de rotação e 6 filtros de escala), com frequências inferior e superior de 0,05 e 0,4, respectivamente. O vetor obtido possui 48 atributos.

- Descritores de Fourier

O vetor utilizado foi formado pelas características extraídas pelo espectro de potência, em que cada descritor de característica corresponde à soma dos valores do espectro de potência localizados a uma certa distância radial do centro da imagem. O vetor com os descritores de Fourier possui 96 atributos.

Figura 27 – Metodologia proposta para classificação do óleo vegetal.



Fonte – Própria autora.

3.1.4 Classificadores

Para classificação foram utilizados os classificadores: LDA, QDA e KNN com $K = 1$ e distância euclidiana. Todos os classificadores foram repetidos 20 vezes para avaliar a veracidade dos dados de acurácia. O método de validação cruzada adotado foi o *leave-one-out*.

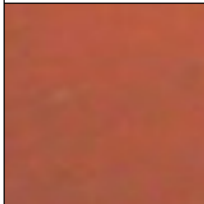
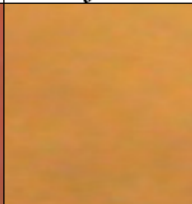


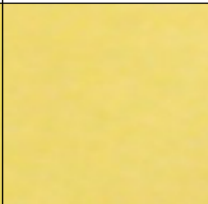
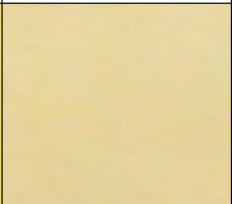
3.2 CLASSIFICAÇÃO DO ÓLEO VEGETAL

As amostras de óleo vegetal de algodão são compostas por seis níveis diferentes do seu processo de semirrefino. Todas as amostras foram disponibilizadas pela Indústria e Comércio de Rações Golfinho Ltda. A Figura 27 apresenta a metodologia proposta para análise e classificação do processo de semirrefino do óleo de algodão. O processo de classificação se inicia com a captura das imagens, posteriormente é determinada qual área das imagens é de interesse realizar análise. Com a área desejada classificada, esta é dividida em 4 sub-imagens. O próximo passo é a aplicação de diferentes modelos de cores, levando a obter diferentes vetores de características. Os dados obtidos serão utilizados para a descrição de padrões e validados por dois métodos de validação. Os resultados obtidos são apresentados por matriz de confusão.

3.2.1 Aquisições de imagens e pré-processamento

Durante o processo de semirrefino o óleo passa por diversas etapas e alterações químicas, que não são de fácil percepção visual. A Figura 28 apresenta o estado do óleo em cada fase do processo, sendo possível notar a semelhança entre as etapas de lavagem, secagem e branqueamento, o que pode prejudicar na hora de verificar se o processo de semirrefino foi concluído e em qual grau de eficiência.

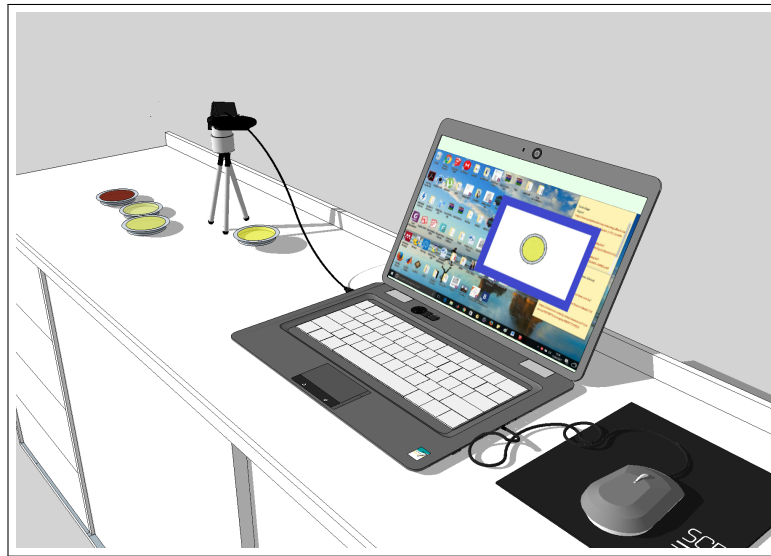
Figura 28 – Classes comerciais do óleo.

Estado do óleo em cada fase do processo de semirrefino					
<i>Óleo bruto</i>	<i>Primeiro refino</i>	<i>Segundo refino</i>	<i>Lavagem</i>	<i>Secagem do óleo</i>	<i>Branqueamento</i>
					

Fonte – Própria autora.

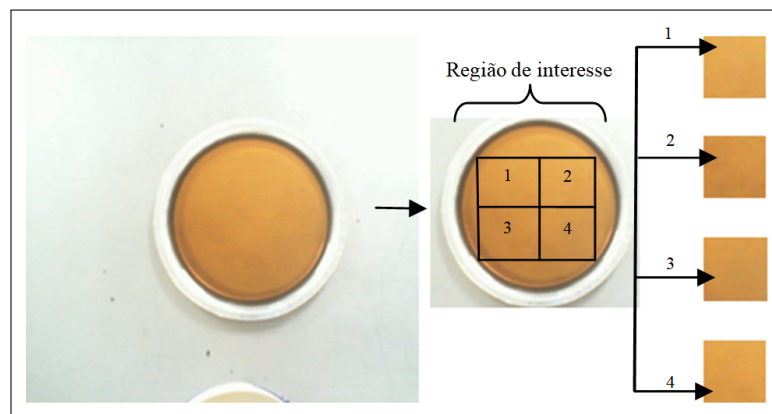
Com o intuito de que as imagens ao serem capturadas fossem obtidas nas mesmas circunstâncias, um sistema simples de captura foi montado. Neste sistema foi usado uma webcam Microsoft LifeCam HD-3000 conectada a um notebook e fixada a uma altura aproximada de 28,5 centímetros das amostras. Foram definidos padrões de volume, posição e características do recipiente a serem utilizados, sendo 5 ml de óleo para cada amostra e o recipientes em aço inoxidável para amenizar interferências na cor do óleo. A Figura 29 apresenta o sistema utilizado na aquisição das imagens.

Figura 29 – Sistema para captura de imagens.



Fonte – Própria autora.

Figura 30 – Subimagens do processo de Segundo refino.



Fonte – Própria autora.

As imagens obtidas possuem a dimensão de 1280 x 800 *pixels*, porém, além de boa parte da área capturada apresentar o que se pretende analisar, esta também apresenta elementos alheios ao óleo, como as bordas do recipiente utilizado para amostragem e a própria mesa. Com isso foi utilizado um algoritmo que identifica a área desejada e a partir desta área geram-se quatro subimagens de 100 x 100 *pixels*, totalizando 96 subimagens como base de dados. A Figura 30 apresenta um exemplo da imagem original obtida pelo sistema de aquisição e suas subimagens.

3.2.2 Extração das características

Inicialmente as imagens foram convertidas em diferentes modelos de cores. A seguir, foram analisados quais modelos representavam melhor as características de classificação do óleo

de algodão. Com isso foi adotado o uso dos modelos de cor RGB e YIQ para formação do vetor de características, uma vez que os dois tratam de cores importantes para a classificação do óleo vegetal de algodão. Na análise laboratorial é analisado um conjunto de parâmetros, entre eles: turbidez e tonalidade. Quando a amostra apresenta tons avermelhados, indica a presença de caratenóides. Se a taxa de cor vermelha for elevada, isso indica que o produto não é próprio para consumo. Assim são obtidos dois vetores de características, um para cada modelo de cor, abrangendo três atributos cada, contendo a informação da média dos valores de cada faixa de cor.

3.2.3 *Classificação*

A classificação foi realizada por três métodos distintos: QDA, KNN com $K = 4$ e ELM com duas camadas (13 neurônios ocultos e 4 neurônios de saída). As validações cruzadas adotadas foram *k-fold* com $k = 4$ e *leave-one-out*.

3.3 **Considerações finais**

Neste capítulo foram apresentadas as etapas para extração de características e classificação da metodologia proposta. Os detalhes e restrições da implementação são descritos em seguida. Os resultados obtidos nos vários métodos de classificação são mostrados e analisados no capítulo seguinte.

4 RESULTADOS

Este capítulo apresenta os resultados obtidos com o método de extração de características proposto neste trabalho, comparando as técnicas de classificação utilizadas. Na Seção 4.1 são apresentados os resultados das análises de imagens do *linter* com diferentes descritores de textura e na Seção 4.2 são apresentados os resultados da classificação do óleo vegetal aplicando dois modelos de cores e classificadores diferentes.

4.1 Classificação do *Linter*

Foram propostas duas formas de classificação da base de dados do *linter*. Na primeira a base é classificada em dois tipos de qualidade, já na segunda são considerados cinco tipos de classificação adotados pela indústria em que foram obtidas as amostras. O processo de obtenção de características foi realizado por alguns métodos descritivos de textura e posteriormente foram aplicados diferentes classificadores para conhecer a taxa de acerto em cada caso.

As Tabelas 3, 4 e 5 apresentam os resultados de classificação do *linter* em duas classes utilizando diferentes descritores de textura e os classificadores KNN, QDA e LDA respectivamente. Na aplicação dos três classificadores os descritores de Haralick apresentaram melhores respostas, uma vez que estes descritores têm como propriedade a distinção de texturas semelhantes. A melhor resposta obtida para a classificação do *linter* em duas classes foi com o uso do classificador LDA, o qual apresenta bons resultados quando utilizado na distinção de duas classes, além de apresentar baixo custo computacional, tendo 96,4% de acurácia, como apresentado na Tabela 5.

Tabela 3 – Classificação do *linter* em duas classes (boa qualidade e qualidade ruim), classificador KNN e validação *leave-one out*.

Método	Imagens corretamente classificadas	Taxa de acertos (%)
Descritor de Haralick	2,210	89,50
Dimensão fractal	1,515	61,13
Lacunaridade	1,778	71,99
LBP	2,142	86,76
LFP	1,668	67,53
Filtros de Gabor	1,454	58,80
Descritores de Fourier	1,690	68,46

Fonte – Elaborado pela autora.

Tabela 4 – Classificação do *linter* em duas classes (boa qualidade e qualidade ruim), classificador QDA e validação *leave-one out*.

Método	Imagens corretamente classificadas	Taxa de acertos (%)
Descritor de Haralick	2,172	87,93
Dimensão fractal	960	38,86
Lacunaridade	1,980	80,16
LBP	2,142	86,76
LFP	1,802	72,95
Filtros de Gabor	1,664	67,40
Descritores de Fourier	1,691	68,46

Fonte – Elaborado pela autora.

Tabela 5 – Classificação do *linter* em duas classes (boa qualidade e qualidade ruim), classificador LDA e validação *leave-one out*.

Método	Imagens corretamente classificadas	Taxa de acertos (%)
Descritor de Haralick	2,381	96,40
Dimensão fractal	960	38,86
Lacunaridade	1,981	80,20
LBP	1,340	54,29
LFP	1,827	74,00
Filtros de Gabor	1,750	70,85
Descritores de Fourier	2,142	86,72

Fonte – Elaborado pela autora.

As Tabelas 6, 7 e 8 apresentam os resultados de classificação do *linter* em cinco classes utilizando diferentes descritores de textura e os classificadores KNN, QDA e LDA respectivamente. A melhor resposta para a classificação do *linter* em cinco classes também foi obtida utilizando os descritores de Haralick, porém com o classificador QDA, o qual apresenta melhor comportamento em classificações de maior complexidade, no caso a distinção das cinco classes do *linter*. A melhor acurácia obtida foi de 64%, como apresentado na Tabela 7.

Tabela 6 – Classificação do *linter* em cinco classes (alta qualidade Japão, média qualidade Japão, padrão China, baixa qualidade e péssima qualidade), classificador KNN e validação *leave-one out*.

Método	Imagens corretamente classificadas	Taxa de acertos (%)
Descritor de Haralick	1,254	50,76
Dimensão fractal	460	18,62
Lacunaridade	797	32,36
LBP	1,326	53,68
LFP	746	30,20
Filtros de Gabor	1,358	55,00
Descritores de Fourier	923	37,36

Fonte – Elaborado pela autora.

Tabela 7 – Classificação do *linter* em cinco (alta qualidade Japão, média qualidade Japão, padrão China, baixa qualidade e péssima qualidade), classificador QDA e validação *leave-one out*.

Método	Imagens corretamente classificadas	Taxa de acertos (%)
Descritor de Haralick	1,580	64,00
Dimensão fractal	460	18,62
Lacunaridade	958	38,80
LBP	1,534	62,10
LFP	698	28,25
Filtros de Gabor	1,364	55,22
Descritores de Fourier	1,176	47,61

Fonte – Elaborado pela autora.

Tabela 8 – Classificação do *linter* em cinco classes (alta qualidade Japão, média qualidade Japão, padrão China, baixa qualidade e péssima qualidade), classificador LDA e validação *leave-one out*.

Método	Imagens corretamente classificadas	Taxa de acertos (%)
Descritor de Haralick	1,464	59,30
Dimensão fractal	460	18,62
Lacunaridade	963	39,00
LBP	1,448	58,62
LFP	780	32,38
Filtros de Gabor	1,457	58,98
Descritores de Fourier	1,467	59,39

Fonte – Elaborado pela autora.

4.2 Classificação do Óleo de Algodão

Com os melhores resultados para a classificação de etapas do processo de semirrefino do óleo, foram geradas matrizes de confusão para cada caso. Nelas, suas entradas diagonais apresentam a quantidade de amostras classificadas corretamente. Assim, é possível analisar os pontos do processo em que acontecem maior confusão e a taxa de acerto para cada classe. As matrizes de confusão para a classificação KNN nos modelos de cores RGB e YIQ são apresentadas nas Tabelas 9 e 10. Na Tabela 9, as classes de óleo bruto e secagem foram classificadas com 100% de acurácia. Já as classes de primeiro refino, segundo refino e lavagem do óleo foram as que apresentaram o menor percentual de classificação, tendo como fator a semelhança de características que estas etapas apresentam entre si. Utilizando o modelo de cor YIQ foram obtidos os resultados apresentados na Tabela 10, em que as classes de óleo bruto, lavagem e secagem apresentaram acurácia de 100%, tendo um resultado mais significativo tendo em vista as características do modelo de cor, como a consideração da intensidade.

Tabela 9 – Matriz de confusão, para KNN = 4, método *Leave-one-out* e modelo de cor RGB.

Óleo bruto	1º Ref.	2º Ref.	Lavagem	Secagem	Branqueamento	Acurácia (%)	Classe real
16	0	0	0	0	0	100	Óleo bruto
0	15	1	0	0	0	93,75	1º Ref.
0	1	11	4	0	0	68,75	2º Ref.
0	0	1	14	0	1	87,5	Lavagem
0	0	0	0	16	0	100	Secagem
0	0	0	1	0	15	93,75	Branqueamento

Fonte – Elaborado pela autora.

Tabela 10 – Matriz de confusão, para KNN = 4, método *Leave-one-out* e modelo de cor YIQ.

Óleo bruto	1º Ref.	2º Ref.	Lavagem	Secagem	Branqueamento	Acurácia (%)	Classe real
16	0	0	0	0	0	100	Óleo bruto
0	15	1	0	0	0	93,75	1º Ref.
0	1	11	4	0	0	68,75	2º Ref.
0	0	0	16	0	0	100	Lavagem
0	0	0	0	16	0	100	Secagem
0	0	0	1	0	15	93,75	Branqueamento

Fonte – Elaborado pela autora.

Para a classificação QDA, a matriz de confusão apresentou os mesmos resultados tanto para o modelo de cores RGB quanto para o modelo YIQ, que são apresentados na Tabela 11. As acurácias obtidas foram de 100% para as etapas de óleo bruto, lavagem e secagem do óleo,

e de 81,25% na classificação do processo de segundo refino. Ao utilizar o classificador QDA, foram obtidos acurácias melhores em todas as etapas do processo de semirrefino em relação ao classificador KNN.

Tabela 11 – Matriz de confusão, com classificador QDA, método *Leave-one-out*.

Óleo bruto	1º Ref.	2º Ref.	Lavagem	Secagem	Branqueamento	Acurácia (%)	Classe real
16	0	0	0	0	0	100	Óleo bruto
0	15	1	0	0	0	93,75	1º Ref.
0	0	13	2	0	1	81,25	2º Ref.
0	0	0	16	0	0	100	Lavagem
0	0	0	0	16	0	100	Secagem
0	0	0	1	0	15	93,75	Branqueamento

Fonte – Elaborado pela autora.

A Tabela 12 apresenta a matriz de confusão para a rede neural ELM, na qual os parâmetros utilizados para a classificação das seis classes de óleo foram: 13 neurônios e *4-fold*. A matriz demonstra a eficácia da ELM na classificação das etapas do processo de semirrefino do óleo de algodão, tendo margem de erro em apenas uma das classes, que é na etapa de segundo refino.

Tabela 12 – Matriz de confusão utilizando ELM, método K-fold e os modelos de cores RGB e YIQ.

Óleo bruto	1º Ref.	2º Ref.	Lavagem	Secagem	Branqueamento	Acurácia (%)	Classe real
16	0	0	0	0	0	100	Óleo bruto
0	16	0	0	0	0	100	1º Ref.
0	0	12	4	0	0	75	2º Ref.
0	0	0	16	0	0	100	Lavagem
0	0	0	0	16	0	100	Secagem
0	0	0	0	0	16	100	Branqueamento

Fonte – Elaborado pela autora.

A Tabela 13 apresenta a comparação entres as respostas obtidas com os distintos classificadores utilizados na classificação do processo de semirrefino do óleo. Com aplicação da ELM teve-se o melhor resultado de classificação e mesmo comportamento para os distintos modelos de cores utilizados.

Tabela 13 – Acurácia média da classificação dos estados de semirrefino do óleo de algodão, conforme classificador e modelo de cor aplicados.

Modelo de cor	Acurácia KNN(%)	Acurácia QDA (%)	Acurácia ELM (%)
RGB	90,6	94,7	95,8
YIQ	92,7	94,7	95,8

Fonte – Elaborado pela autora.

Por fim, os resultados obtidos indicam que é possível realizar a classificação do *linter* por análise de textura e do processo de semirrefino do óleo de algodão pelas características de cor, utilizando técnicas de visão computacional, as quais são viáveis para implementação em chão fabril. Para a classificação do processo de semirrefino do óleo de algodão foi obtida uma proposta eficaz e de baixo custo, quando comparada com as demais metodologias presentes na literatura. Já para classificação do *linter* foi obtida uma proposta útil na distinção de duas classes, porém, na indústria em que foi realizado o estudo é necessária a classificação de cinco classes, sendo proposto estudo para melhorar a acurácia obtida, viabilizando sua aplicação na área industrial.

5 CONCLUSÕES

Este trabalho teve por objetivo proporcionar metodologias viáveis e de fácil implantação em chão fabril para a classificação do *linter* e óleo vegetal de algodão, que venham apresentar potencial que substitua o atual processo manual de classificação por um método semiautomático, mais eficiente e confiável, impulsionando o estudo para o desenvolvimento de técnicas que possam ser utilizadas em chão fabril na classificação de óleos vegetais e/ou subprodutos do algodão.

Para o desenvolvimento das metodologias apresentadas foram obtidas bases de dados tanto para a classificação do processo de semirrefino do óleo de algodão como do *linter*, sendo utilizados os descritores de cor e textura como parâmetros de características. Através de técnicas de processamento de imagens e reconhecimento de padrões foram alcançadas boas respostas para a classificação dos dois subprodutos do algodão. A utilização dos descritores de textura de Haralick para a classificação do *linter* apresentou resultados promissores para as duas condições analisadas, uma considerando duas classes e utilizando o classificador LDA e a outra cinco classes e usando classificador QDA, obtendo acurácias de 96,40% e 64,00%, respectivamente. Logo, para a classificação do *linter* tem aplicação útil considerando a distinção de duas classes, porém, sendo necessário efetuar melhorias na classificação de cinco classes para aplicação industrial, onde o ideal seria acurácia na faixa de 90,00 a 95,00%.

Na classificação das etapas de semirrefino do óleo de algodão, os resultados mostraram que as técnicas de processamento digital de imagens são capazes de extrair atributos e informações necessárias para classificar em qual estágio do processo as amostras de óleo se encontram. Assim, obteve-se um método de classificação mais eficiente utilizando a rede neural ELM, modelo de cor RGB e método de validação *k-fold*, apresentando acurácia de 95,80%.

É importante ressaltar que este estudo demonstra a possibilidade e importância de executar os procedimentos de classificação analisados de forma semiautomática, tendo como diferencial dos métodos existentes:

- Redução do tempo de análise e de envio do produto ao mercado consumidor;
- Melhoria ergonômica da atividade de classificação;
- Confiabilidade na classificação dos produtos;
- Arquivamento de dados, evitando assim a intervenção humana nos resultados;
- Disponibilidade dos dados obtidos por produto em relação à qualidade, auxiliando no planejamento de melhoria contínua.

5.1 Trabalhos futuros:

A seguir, como continuação do estudo iniciado nesta dissertação, são apresentados os seguintes projetos:

- Para a classificação do *linter*, utilizar o algoritmo aplicando base de dados com descritores de características de texturas coloridas;
- Na classificação das etapas do processo do óleo vegetal, avaliar o desempenho do algoritmo proposto, com uma base de imagens maior.

5.2 Trabalhos publicados

Durante a pesquisa para a presente dissertação, os seguintes trabalhos foram publicados em eventos da área científica:

ARAÚJO, Patrícia D.; MERA MOYA, Victoria; DE PAULA JUNIOR, Iális C. **Classificação do óleo de algodão em processo de semirrefino usando técnicas de processamento de imagem** In: XXXV SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS - SBrT. Sociedade Brasileira de Telecomunicação, 2017, São Pedro-SP. p.137 - 141

ARAÚJO, Patrícia D.; MERA MOYA, Victoria; DE PAULA JUNIOR, Iális C. **Classification of cotton oil in the semi-refining process using image processing techniques: Image processing for industrial applications**; In: Seventh International Conference on Innovative Computing Technology (INTECH), 2017, Luton. p.21 - DOI: 10.1109/INTECH.2017.8102417

REFERÊNCIAS

- ACHARYA, T.; RAY, A. K. **Image Processing: Principles and Applications**. 1. ed. New Jersey: Wiley Interscience, 2005. ISBN 978-0-471-71998-4.
- ALGOTEC. **Manual de Orientações Técnicas para Algodoeiras de Mato Grosso do Sul**. AMPASUL and UNIFIBRA, 2010. Disponível em: <<http://rhes.ruralhorizon.org/uploads/documents/algotec2010.compressed.pdf>>. Acesso em: 26 Mar. 2018.
- AMARAL, V. do; THOMAZ, C. E. Extração e comparação de características locais e globais para o reconhecimento automático de imagens de faces. VIII Workshop de Visão Computacional, 2012. Disponível em: <https://fei.edu.br/~cet/VagnerAmaral_WVC2012.pdf>.
- BACKES, A. R. **Estudos de Métodos de Análise de Complexidade de Imagens**. Dissertação (Mestrado) — Universidade de São Paulo - USP, São Carlos – Brasil, 2010.
- BACKES, A. R.; BRUNO, O. M. Segmentação de texturas por análise de complexidade. Universidade de São Paulo, 2006. Disponível em: <https://www.researchgate.net/publication/228987626_Segmentacao_de_texturas_por_analise_de_complexidade>.
- BALLABIO, D.; CONSONNI, V. Classification tools in chemistry. part 1: linear models. pls-da. Analytical Methods, 2013.
- BALLARD, D. H.; BROWN, C. M. **Computer Vision**. 1st. ed. Englewood Cliffs: Prentice Hall, 1982.
- BARKER, M.; RAYENS, W. Partial least squares for discrimination. Journal of Chemometrics, 2003. Disponível em: <<https://doi.org/10.1002/cem.785>>.
- BARROS, E.; MEDEIROS, J. F.; PEREIRA, N. C. Análise das propriedades físico-química do óleo de soja degomado visando a produção de biodiesel. VIII Encontro Internacional de Produção Científica - EPCC, 2013. Disponível em: <http://www.cesumar.br/prppge/pesquisa/epcc2013/oit_mostra/Janaina_Fernandes_Medeiros_02.pdf>.
- BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M. B. **Chemometrics: A Practical Guide**. 1. ed. New York: Wiley Interscience, 1998. ISBN 9780471124511.
- BELTRÃO, N. E. M.; AZEVEDO, D. M. P. **O agronegócio do algodão no Brasil**. 2. ed. Brasília - DF: Embrapa Agroindústria Tropical; Embrapa Algodão, 2008. ISBN 978-85-7383-424-6.
- BISHOP, C. M. **Neural Networks for Pattern Recognition**. 1. ed. New York: Clarendon Press, 1996. ISBN 0198538642.
- BOAVENTURA, I. A. G. **Números fuzzy em processamento de imagens digitais e suas aplicações na detecção de bordas**. Tese (Doutorado) — Escola de Engenharia de São Carlos - USP, São Carlos, 2010. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/18/18152/tde-06052010-154227/publico/Ines.pdf>>.
- BRAGEIRO, L. D. **Classificação de Padrões Radiológicos por blocos em Imagens não segmentadas de Tomografia Computadorizada**. Dissertação (Mestrado) — Universidade Federal do Paraná, Curitiba, 2015. Disponível em: <<https://acervodigital.ufpr.br/bitstream/handle/1884/40319/R%20-%20-%20LUIZA%20DRI%20BAGESTEIRO.pdf?sequence=2&isAllowed=y>>.

BRAVO, M. J. A. **Análise do Descritor de Padrões Mapeados Localmente em Multiescala para Classificação de Textura em Imagens Digitais**. Dissertação (Mestrado) — Universidade Estadual Paulista, São José do Rio Preto – Brasil, 2016.

BROTTO, R. D. B.; BOCCATO, L.; ROMANO, J. M. T. Aplicação de máquinas de aprendizado extremo ao problema de identificação de sistemas não-lineares. XXXV Simposio Brasileiro de Telecomunicações e Processamento de Sinais- SBrT, 2017. Disponível em: <<http://www.sbrt.org.br/sbrt2017/anais/1570359796.pdf>>.

CALAS, M.; CORDINA, G.; BOMPART, J.; BEN, B. M.; JEI, T.; ANCELIN, M.; VIAL, H. Journal of medicinal chemistry, 1997. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/9357523>>.

CAMPESTRE. **Óleo de algodão**. [s.n.], 2017. Disponível em: <<http://www.campestre.com.br/oleos-vegetais/oleo-de-algodao/>>. Acesso em: 27 Fev. 2018.

CARDELLO; BORGHI, A. B. M. P.; VILA, M. M. D. C.; GONÇALVES, L. A. G. Clarificação convencional de óleo de sementes de algodão (*Gossypium hirsutum*). Alimentos e Nutrição Araraquara, n. 1, p. 77–87, 1995. Disponível em: <<http://serv-bib.fcfar.unesp.br/seer/index.php/alimentos/article/view/737/626>>.

CARVALHO, B. M. de. Representações de cores. Universidade Federal do Rio Grande do Norte - UFRN-35T56, 2008. Acesso em 01/07/2018.

CASCELLI, C. A.; SACRAMENTO, L.; MEDEIROS, R.; MENDES, A.; MENESES, E.; LAMARQUES, R.; SANTOS, B. J. dos. Relatório de gestão - biênio 2015-2016. Color, Brasília, 2016. Associação Brasileira dos Produtores de Algodão.

CASERTA, F.; ELDRED, W. D.; FERNANDEZ, E.; HAUSMAN, R.; STANFORD, L. R.; BULDEREV, S. V.; S., S.; STANLEY, H. Determination of fractal dimension of physiologically characterized neurons in two and three dimensions. Journal Neurosci Methods, 1995.

CASTLEMAN, K. R. **Digital image processing**. NJ - USA: Prentice Hall Press Upper Saddle River, NJ, USA, 1996. ISBN 0-13-211467-4.

CASTRO, L. **Análise e síntese de estratégias de aprendizado para redes neurais artificiais**. Dissertação (Mestrado) — Universidade Estadual de Campinas - Faculdade de Engenharia Elétrica e de Computação, Campinas – SP – Brasil, 1998. Disponível em: <ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/theses/lnunes_mest/indice.pdf>.

CHAUDHURI, B. B.; SARKAR, N. Texture segmentation using fractal dimension. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995.

CHAVES, F. E. **Análise Comparativa das Características de Filtros de Gabor, LBP e DCt para Reconhecimento Automático de Expressões Faciais**. Dissertação (Mestrado) — Universidade Estadual do Ceará, Centro de Ciências e Tecnologia, Fortaleza - Ceará, 2015. Programa de Pós-Graduação em Ciência da Computação.

CHEN, J.; JAIN, A. K. A structural approach to identify defects in textured images. IEEE International Conference on Systems, Man, and Cybernetics, 1988.

CHIERICI, C. E. de O. **Classificação de texturas com diferentes orientações baseada em descritores locais**. Dissertação (Mestrado) — Universidade de São Paulo - Escola de Engenharia de São Carlos, São Carlos, 2015. Departamento de Engenharia Elétrica e de Computação. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/18/18152/tde-27102015-103555/>>.

COELHO, R. C.; COSTA, L. d. F. The boxcounting fractal dimension: Does it provide an accurate subsidy for experimental shape characterization? if so, how to use it. *Anais do Sibgrapi*, 1995. P. 183 - 191, Ed. 95.

COSTA, G. B. da. **Classificação Multivariada de Biodiesel B100 e B5 Usando Imagens Digitais**. Dissertação (Mestrado) — Universidade Estadual da Paraíba, Campina Grande - PB – Brasil, 2015.

CROWLEY, J. L.; CHRISTENSEN, H. I. **Vision as Process Basic Research on Computer Vision Systems**. [S.l.]: Springer-Verlag Berlin e Heidelberg GmbH e Co. K, 1995. ISBN 9783642081972.

DELATORRE, A.; OKIDA, S.; TROJAN, F.; PONTES, J. C. A. de. Avaliação preliminar do desempenho de redes neurais feedforward para a previsão de preços de derivados da cana-de-açúcar. X Congresso Brasileiro de Agroinformática, 2015. Disponível em: <http://eventos.uepg.br/sbiagro/2015/anais/SBIAgro2015/pdf_resumos/4/4_sergio_luiz_stevan_jr_106.pdf>.

DENG, H.-B.; JIN, L.-W.; ZHEN, L.-X.; HUANG, J.-C. A new facial expression recognition method based on local gabor filter bank and pca plus lda. *International Journal of Information Technology*, Huntsville, Alabama, 2005. Disponível em: <<https://pdfs.semanticscholar.org/9eea/da49fc2c8a846b4dad1012ba8a7ee78a8bb7.pdf>>.

DIAS, D. N. **Identificação dos sintomas de ferrugem em áreas cultivadas com cana-de-açúcar**. Dissertação (Mestrado) — Universidade de São paulo, São Carlos, 2004. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/55/55134/tde-19052004-134621/publico/TDE-19052004-134621.pdf>>.

DIAS, F. O avanço da tecnologia e o aumento da produtividade. Setembro 2016. Disponível em: <<http://www.amapa-ma.com.br/2016/09/29/o-avanco-da-tecnologia-e-o-aumento-da-produtividade/>>. Acesso em: 19 Março. 2018.

DOLENSI, A. F. C. Procedimento operacional padrão - conteúdo de sujeira em linter de algodão. Sobral- CE, 2013. Relatório de procedimento operacional da industria em estudo de caso. Acesso em: 06 de 2016.

DUCHESNE, P.; BRUNO, R. **Statistical modeling and analysis for complex data problems**. 1. ed. [S.l.]: pringer Science Business Media, 2005.

ERTHAL, G. J.; FONSECA, L. M. G.; BINS, L. S.; VELASCO, F. R. D.; MONTEIRO, A. M. V. Um sistema de segmentação e classificação de imagens de satélite. IV Sipiósio Brasileiro de Computação Gráfica e Processamento de Imagens, 1991. Aplicações em Metereologia e Geoprocessamento.

EVELYN, F.; HODGES, J. L. J. Discriminatory analysis. nonparametric discrimination: Consistency properties. *International Statistical Review / Revue Internationale de Statistique*, 1989. Disponível em: <<https://www.jstor.org/stable/1403797>>.

- FARROKHINIA, F. Multi-channel filtering techniques for texture segmentation and surface quality inspection. *scinapse*, 1990. P. 573 - 583.
- FERNANDES, A. M. da R. **Inteligência artificial: noções gerais**. Brasil: Visual Books, 2005.
- FERNÁNDEZ, E.; JELINEK., H. Use of fractal theory in neuroscience: methods, advantages, and potential problems. US National Library of Medicine National Institutes of Health, 2001.
- FERREIRA, M. M. C. Multivariate qsar. *Journal of the Brazilian Chemical Society*, 2002. ISSN 0103-5053. Disponível em: <<http://dx.doi.org/10.1590/S0103-50532002000600004>>.
- FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O. Chemometrics i: multivariate calibration, a tutorial. *Química Nova - Scielo*, 1999.
- FONSECA, R. G. d.; SANTANA, J. C. F. d. **Resultados de ensaio HVI e suas interpretações (ASTM D-4605)**. 66. ed. Repositório de informação tecnológica da Embrapa, 16-Out-2007, 2002. ISSN 0100-6460. Disponível em: <<http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/275874>>. Acesso em: 02 Mar. 2018.
- FONSECA, R. G. da; SANTANA, J. C. F. de. **Resultados de Ensaio HVI e Suas Interpretações (ASTM D-4605)**. Campina Grande, PB: EMBRAPA, 2002. ISSN 0100-6460.
- FRIEDMAN, J. Regularized discriminant analysis. *journal of the american statistical association*. *Journal of the American Statistical Association*, 1989.
- FUKUNAGA, K.; HAYES, R. Effects of sample size in classifier design. *IEEE Transaction Pattern Analysis and Machine Intelligence*, 1989.
- GABOR, D. Theory of communication. *Journal of the Institution of Electrical Engineers - Part I: General*, Huntsville, Alabama, 1946.
- GONÇALVES, L. A. **Um estudo a transformada Rápida de Fourier e seu uso em processamento de imagens**. Dissertação (Mestrado) — Universidade Federal do Rio grande do Sul, Porto Alegre, 2004. Dissertação (Mestrado). Disponível em: <<https://www.lume.ufrgs.br/bitstream/handle/10183/6748/000446124.pdf?sequence=1>>.
- GONÇALVES, M. I. S. **Aplicação de Imagens Digitais e Análise Multivariada para classificação e Determinação de Parâmetros de Qualidade em Plumas de Algodão**. Dissertação (Mestrado) — Universidade Federal da Paraíba, João Pessoa, 2015.
- GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. Pearson Education, Inc, 2008. ISBN ISBN 0-13-168728-x. Disponível em: <http://web.ipac.caltech.edu/staff/fmasci/home/astro_refs/Digital_Image_Processing_3rdEd_truncated.pdf>.
- GULICK, D. **“Encounters with Chaos”, McGraw-Hill International Editions - Mathematics and Statistics Series**. 1. ed. [S.l.]: McGraw-Hill College, 1992. ISBN 978-0070252035.
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 1973. Disponível em: <<https://pdfs.semanticscholar.org/5140/f1dc83e562de0eb409385480b799e9549d54.pdf>>.
- HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. 2. ed. [S.l.]: Prentice Hall, 1998.

- HAYKIN, S. **Redes Neurais - Princípios e Práticas**. 2. ed. Brasil: Bookman, 2001.
- HAYKIN, S.; VEEN, B. V. **Sinais e Sistemas**. 2. ed. Porto Alegre, RS, Brasil: Bookman, 2001. ISBN 8573077417.
- HOURANT, P.; BAETEN, V.; MORALES, M. T.; MEURENS, M.; APARICIO, R. Oil and fat classification by selected bands of near-infrared spectroscopy. *Appl. Spectrosc*, 2000. Disponível em: <<https://www.osapublishing.org/as/abstract.cfm?URI=as-54-8-1168>>.
- HSU, H. P. **Análise de Fourier**. Rio de Janeiro: Coleção técnica, Prentice Hall, 2001.
- HUANG, D.; SHAN, C.; ARDEBILIAN, M.; WANG, Y.; CHEN, L. Local binary patterns and its application to facial image analysis: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2011. Disponível em: <<https://iris.cnrs.fr/Documents/Liris-5004.pdf>>.
- HUANG, G.-B.; ZHOU, H.; DING, X.; ZHANG, R. Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, man, and Cybernetics- Part B: Cybernetics*, 2012.
- HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: A new learning scheme of feedforward neural networks. *International Joint Conference of Neural Networks (IJCNN2004)*, 2004.
- HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing*, 2006.
- JAHNE, B. **Digital Image Processing**. Berlin: Springer-Verlag Berlin Heidelberg, 2005. ISBN 978-3-540-27563-3.
- JIAN, L.; CAIXIN, S.; QIAN, D. A new box-counting method for estimation of image fractal dimension. *International Conference on Image Processing/ IEEE*, 2006.
- JULESZ, B. Visual pattern discrimination. *IRE Transactions on Information Theory*, 1962.
- JÚNIOR, I. J. D. N.; FARIA, S. R. de. Proposição de modelo de análise discriminante para avaliação de desempenho de empresas de sucesso e fracasso do setor de indústrias manufatureiras. Universidade Católica de Brasília - UCB, 2007. Disponível em: <https://www.aedb.br/seget/arquivos/artigos07/1268_Proposicao\%20de\%20Modelo.pdf>.
- KHATTREE, R.; NAIK, N. Multivariate data reduction and discrimination with sas software. Cary, NC, Usa: SAS Institute INC, 2000.
- KUMAR, N.; BANSAL, A.; SARMA, G.; RAWAL, R. Chemometrics tools used in analytical chemistry: An overview. *Talanta*, 2014. Volume 127, Pages 186-199. Disponível em: <<https://doi.org/10.1016/j.talanta.2014.02.003>>.
- LANA, V.; RIBAS, P. V.; CHANSELME, J.-L. **Manual de Beneficiamento**. Cuiabá - MT: AMPA-IMamt, 2014. A classificação do algodão.
- LUCIANO, D. de M.; HONORATO, B. Visão computacional. Universidade Estadual de Campinas and, 2010. Em: <http://docplayer.com.br/3058305-Visao-computacional-danilo-de-milano-luciano-barrozohonorato-unicamp-universidade-estadual-de-campinas-ft-faculdade-de-tecnologia.html>.

MAGALHÃES, L. E. **Qualidade – Centro de Análise de Fibras de Algodão**. [S.l.: s.n.], 2014.

MANDARINO, J. M. G.; HIRAKURI, M. H.; ROESSING, A. C. **Tecnologia para produção do óleo de soja: descrição das etapas, equipamentos, produtos e subprodutos**. 2. ed. Londrina-PR: Publicado pela Embrapa Soja, 2015.

MANDARINO, J. M. G.; ROESSING, A. C. **Tecnologia para produção do óleo de soja: descrição das etapas, equipamentos, produtos e subprodutos**. 1. ed. Londrina-PR: Publicado pela Embrapa Soja, 2001.

MANJUNATH, B. S.; MA, W. Texture features for browsing and retrieval of image data. IEEE Transactions on Pattern Analysis and Machine Intelligence, August 1996.

MANJUNATH, B. S.; SALEMBIER, P.; SIKORA, T. **Introduction to MPEG-7: Multimedia Content Description Interface**. Italy: Wiley, 2002. ISBN 978-0-471-48678-7.

MARTINS, E. R. da S.; MARQUES, P. M. de A.; OLIVEIRA, L. F. de; TRAD, C. S. Caracterização de lesões intersticiais de pulmão em radiograma de tórax utilizando análise local de textura. Radiologia Brasileira 38(6), 2005.

MASSART, D.; M.VANDEGINSTE, B.; BUYDENS, L. C.; D.JONGS, S.; J.LEWI, P.; SMEYERS-VERBEKE, J. **Handbook of Chemometrics and Qualimetrics**. 1st. ed. Elsevier Science, 1998. ISBN 9780080887036. Disponível em: <<https://www.elsevier.com/books/handbook-of-chemometrics-and-qualimetrics/massart/978-0-444-82853-8>>.

MATHER, P. M.; KOCH, M. **Computer Processing of remotely-sensed images: an introduction**. [S.l.]: John Wiley & Sons, 2011.

METROHM. Nir spectroscopy-a guide to near-infrared spectroscopic analysis of industrial manufacturing processes. 2013. Disponível em: <http://www.mep.net.au/wpmep/wp-content/uploads/2013/05/MEP_Monograph_NIRS_81085026EN.pdf>. Acesso em: 20 Março, 2018.

MEVIK, B.; CEDERKVIST, H. R. Mean squared error of prediction (mse) estimates for principal component regression (pcr) and partial least squares regression (pls). Journal of Chemometrics banner, 2005. Disponível em: <<https://doi.org/10.1002/cem.887>>.

MIGNANI, A. G.; CIACCHERI, L.; MENCAGLIA, A. A.; CIMATO, A. Optical absorption spectroscopy for quality assessment of extra virgin olive oil. INTECH, 2012. Disponível em: <<https://mts.intechopen.com/books/olive-oil-constituents-quality-health-properties-and-bioconversions/optical-absorption-spectroscopy-for-quality-assessment-of-extra-virgin-olive-oil>>"title="OpticalAbsorptionSpectroscopyforQualityAssessmentofExtraVirginOliveOil">OpticalAbsorptionSpectroscopyforQualityAssessmentofExtraVirginOliveOil</>.

MILANEZ, K. **Classificação de óleos vegetais comestíveis usando imagens digitais e técnica de reconhecimento de padrões**. Dissertação (Mestrado) — Universidade Federal da Paraíba, João pessoa- PB, 2013. Dissertação de Mestrado. Disponível em: <<http://tede.biblioteca.ufpb.br:8080/handle/tede/7157>>.

MOREIRA, F. C. **Reconhecimento e classificação de padrões de imagens de núcleos de linfócitos do sangue periféricos humano com a utilização de redes neurais artificiais**.

Dissertação (Mestrado) — Universidade Federal de Santa Catarina, Florianópolis - SC, 2002. Dissertação de Mestrado.

MOURA, I. S. de. **Combinação de modelos de previsão climática**. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Norte, Natal, 2018.

NAKAMURA, Y.; YAMAMOTO, Y.; MURAOKA, I. Autonomic control of heart rate during physical exercise and fractal dimension of heart rate variability. US National Library of Medicine National Institutes of Health; Search database, 1985.

NEIVA, M. B. **Metodos de pré-processamento de texturas para otimizar o reconhecimento de padrões**. Dissertação (Mestrado) — Universidade de São Paulo -, São Carlos, 2016.

NETO, F. F. G.; MARINO, G.; ARAÚJO, M. C. U. de; GALVÃO, R. K. H.; PONTES, M. J. C.; MEDEIROS, E. P.; LIMA, R. S. classification of edible vegetable oils square wave voltammetry with multivariate data analysis. *Talanta*, 2009.

NEVES, M.; PINTO, M. **A Cadeia do Algodão Brasileiro: Desafios e Estratégias**. Brasil: Publicado pela Associação Brasileira dos Produtores de Algodão (ABRAPA), Redação Fundace, 2012.

NIGRIN, A. **Neural Networks for Pattern Recognition**. MA - USA: A Bradford Book - MIT-Press, 1993. ISBN 9780262140546.

NUNES, P. G. A. **Uma nova técnica para seleção de variáveis em calibração multivariada aplicada às espectroscopias UV-VIS e NIR**. Tese (Doutorado) — Universidade Federal da Paraíba, João Pessoa, 2008.

OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Gray scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Finland, 2002.

OLIVEIRA, C. G. **Proposta de modelagem para a clarificação de óleos vegetais – experimentos cinéticos e simulação do processo indústria**. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, Florianópolis - Santa Catarina – Brasil, 2001.

OWUSU, E.; ZHAN, Y.; MAO, Q. R. A neural-adaboost based facial expression recognition system. *Expert Systems with Applications*, Pergamon, Huntsville, Alabama, 2014. Disponível em: <https://ac.els-cdn.com/S0957417413009615/1-s2.0-S0957417413009615-main.pdf?_tid=c8ead551-00f0-42e2-8702-9ee84eb2e187&acdnat=1536401078_ccb7df23e54deed5564655806cd6fd66>.

PALMER, J.; WANG, X.; SHOURESHI, R.; MANDER, A.; TORGERSON, D.; RICH, C. Effect of aging on the spectral response of transformer oil. 2000. ISSN 1089-084X. Disponível em: <<http://ieeexplore.ieee.org/document/845548/>>.

PECEGO, J. G. **Implementação do Método Laboratorial de Determinação do Teor de Clorofila em Óleos Vegetais e Otimização do Processo de Refinação para a sua Remoção**. Dissertação (Mestrado) — Faculdade de Ciências e Tecnologias – Universidade Nova de Lisboa, Lisboa, Outubro 2014. RUM, Repositório Universidade Nova. Disponível em: <<http://hdl.handle.net/10362/25013>>.

- PEDRINI, H.; SCHWARTZ, W. R. **Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações**. [S.l.]: Prentice Hall, 1998.
- PEREIR, A. F. da C. **Determinação simultânea de acidez, índice de refração e viscosidade em óleos vegetais usando espectrometria NIR, calibração multivariada e seleção de variáveis**. Dissertação (Mestrado) — Universidade Federal da Paraíba, João Pessoa, 2007. Dissertação de Mestrado. Disponível em: <http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=92638>.
- PETERS, E. E. **Fractal Market Analysis: Applying Chaos Theory to Investment and Economics**. 1. ed. Canada: Wiley, 1994. ISBN 978-0471585244.
- PLATANIOTIS, K.; VENETSANOPOULOS, A. N. **Color Image Processing and Applications**. [S.l.]: Digital Signal Processing, 2000. ISBN 978-3-662-04186-4.
- PLOTNICK, R.; HARGROVE, R. G. and W.; K.PRESTEGAARD; PERLMUTTER, M. Lacunarity analysis: a general technique for the analysis of spatial patterns. *PHYSICAL REVIEW*, 1996. Disponível em: <<http://research.mssm.edu/cnic/pdfs/PlotnickLacunarity.pdf>>.
- PLOTNICK, R. E.; GARDNER, R. H.; O'NEILL, R. V. Lacunarity indices as measures of landscape texture. *Landscape Ecology*, 1993. Disponível em: <<https://link.springer.com/article/10.1007/BF00125351>>.
- PONTES, M. **Algoritmo das projeções sucessivas para seleção de variáveis espectrais em problema de classificação**. Dissertação (Mestrado) — Universidade Federal da Paraíba, João pessoa - PB, 2009. Programa de Pós-Graduação em Química.
- PUPIN, J. R. **Introdução às Séries e Transformadas de Fourier e Aplicações no Processamento de Sinais e Imagens**. Dissertação (Mestrado) — Universidade Federal de São Carlos, São Carlos, 2011. Licenciatura em Matemática.
- QUEIROZ, J. E. R. de; GOMES, H. M. Introdução ao processamento digital de imagens. *Revista Rita*, 2014. Departamento de Sistemas e Computação. Disponível em: <<http://www.dsc.ufcg.edu.br/~hmg/disciplinas/graduacao/vc-2014.1/Rita-Tutorial-PDI.pdf>>.
- RAO, A. R.; SCHUNCK, B. G. Computing oriented texture fields. *CVGIP: Graphical Models and Image Processing*, Issue 2, 1991. Disponível em: <[https://doi.org/10.1016/1049-9652\(91\)90059-S](https://doi.org/10.1016/1049-9652(91)90059-S)>.
- REZENDE, A. Análise da fibra do algodão. *O Algodoeiro - Informativo Técnico da AMIPA*, 2010.
- RIBEIRO, L. N. **Rede Neural com Retropropagação: uma Aplicação na Classificação de Imagens de Satélite**. Dissertação (Mestrado) — Universidade Federal de Mato Grosso do Sul, Campo Grande, 2003.
- RIBEIRO, P. P. E.; PAIVA, M. S. V. de; JORGE, L. A. de C. Application of texture analysis for differentiation of the greening from others pests. *X Workshop de Visão Computacional - WVC*, Uberlândia, 2014. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/wvc/2014/0035.pdf>>.
- RODRIGUES, B. U. **Reconhecimento do tipo de cachaça utilizando visão computacional e reconhecimento de padrões**. Dissertação (Mestrado) — UNIVERSIDADE FEDERAL DE GOIÁS INSTITUTO DE INFORMÁTICA, Goiânia, 2015.

RUDEK, M.; COELHO, L. dos S.; JUNIOR, O. C. Visão computacional aplicada a sistemas produtivos: Fundamentos e estudo de caso. Universidade Católica do Paraná, Curitiba - PR, 2011. Disponível em: <<https://slidex.tips/download/visao-computacional-aplicada-a-sistemas-produtivos-fundamentos-e-estudo-de-caso>>.

SALDANHA, M. F. S.; FREITAS, C. da C. Segmentação de imagens digitais: Uma revisão. Divisão de Processamento de Imagens – Instituto Nacional de Pesquisas Espaciais (INPE), 2010. Disponível em: <<http://mtc-m16c.sid.inpe.br/col/sid.inpe.br/mtc-m18@80/2010/06.22.18.13/doc/106003.pdf>>.

SCHINDEWOLF, T.; SCHIFFNER, R.; STOLZ, W.; ALBERT, R.; ABMAYR, W.; HARMS, H. Evaluation of different image acquisition techniques for a computer vision system in the diagnosis of malignant melanoma. *Journal of the American Academy of Dermatology*, 1994.

SCHUTTE, A. F. D.; MOTTA, G. L. de S.; CARVALHO, L. C. de; LEAL, M. A. J.; OLIVEIRA, M. I. L.; SOARES, P. B.; M.F.COSTA, P. R.; SANT'ANA, M. C. de; ROCHA, M. E. G.; GOMIDE, R. B. Boletim dos biocombustíveis. 10 2017. Disponível em: <<http://www.mme.gov.br/web/guest/secretarias/petroleo-gas-natural-e-combustiveis-renovaveis/publicacoes>>. Acesso em: 19 Março. 2018.

SILVA, O. R. R. F. da; SOFIATTI, V.; SANTANA, J. C. F. de; WANDERLEY, M. J. R.; SANTOS, J. W. dos. Impacto do beneficiamento sobre o número de neps e quantidade de impurezas da fibra do algodão. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 14, n. 1, p. 107 – 112, 2010. ISSN 1415-4366. Disponível em: <<http://dx.doi.org/10.1590/S1415-43662010000100015>>.

SKLANSKY, J. Image segmentation and feature extraction. *IEEE Transactions on Systems, Man, and Cybernetics*, 1978.

SKOOG, D. A.; HOLLER, F. J.; CROUCH, S. R. **Principles of Instrumental Analysis**. 6 ed. ed. [S.l.]: Thomson Brooks/Cole, 2007.

SKOOG, D. A.; M., W. D.; HOLLER, F. J.; STANLEY, R. **Fundamentos de Química Analítica**. 8. ed. São Paulo: Cengage Learning, 2007.

SMITH, T.; LANGE, G. D.; MARKS, W. B. Fractal methods and results in cellular morphology-dimensions, lacunarity and multifractals. *Journal of Neurosci Methods*, 1996.

SOARES, H. B. **Análise e Classificação de Imagens de lesões da pele por atributos de cor, forma e textura utilizando máquina de vetor de suporte**. Dissertação (Mestrado) — UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE, Natal- RN – Brasil, 2008. Disponível em: <http://www.repositorio.ufrn.br:8080/jspui/bitstream/123456789/15118/1/HelianaBS_TESE.pdf>.

SOBARWIK. File:schematic of uv- visible spectrophotometer. Setembro 2013. Disponível em: <https://commons.wikimedia.org/wiki/File:Schematic_of_UV-_visible_spectrophotometer.png>. Acesso em: 25 Março. 2018.

SOLOMON, C.; BRECKON, T. **Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab**. [S.l.]: John Wiley and Sons, Ltd, 2011. ISBN 9780470689776.

- SOUSA, W. S. **Operação Remota de Mesa Móvel Utilizando Visão Computacional**. Dissertação (Mestrado) — UNIVERSIDADE FEDERAL DO ABC, Santo André, 2015. Trabalho de Graduação.
- SOUZA, L.; CARNEIRO, L.; CANDEIAS, A. L. B. Modelos de cores cmy e yiq no spring e no matlab para análise visual de imagens ikonos e cbers. XIV Simpósio Brasileiro de Sensoriamento Remoto, 2009. 25-30 abril, INPE. Disponível em: <<http://marte.sid.inpe.br/col/dpi.inpe.br/sbsr@80/2008/11.15.05.47/doc/7173-7180.pdf>>.
- SRIVASTAVA, M.; CARTER, E. **Introduction to Applied Multivariate Statistics**. First. [S.l.]: Elsevier Science Ltd, 1983.
- TAPPI. **Size Estimation Chart (Transparency)**. [s.n.], 2000. Disponível em: <<https://imisrise.tappi.org//TAPPI/Products/01/DIR/0109DIRTT.aspx>>. Acesso em: 02 Abril. 2018.
- TEIXEIRA, W. S.; NETO, S. A. dos R. Perspectivas para a agropecuária / companhia nacional de abastecimento. Agosto 2017. ISSN 2318-3241. Disponível em: <<http://www.conab.gov.br>>. Acesso em: 19 Março. 2018.
- THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern Recognition**. 4. ed. [S.l.]: Academic Press, 2008.
- TRAVAINI, J. N. **Descritores de Textura Local para Reconhecimento Biométrico da Íris Humana**. Dissertação (Mestrado) — Universidade de São Paulo, São Carlos, 2015. Departamento de Engenharia Elétrica e de Computação. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/18/18152/tde-09112015-161059/pt-br.php>>.
- TRICOT, C. **Curves and Fractal Dimension**. New York: Springer-Verlag; Topics: Dynamical Systems and Ergodic Theory, 1995. ISBN 978-0-387-94095-3.
- TRUCCO, E.; VERRI, A. **Introductory Techniques for 3-D Computer Vision**. Edinburgh, UK: Prentice Hall, 1998. ISBN 978-0132611084.
- UNIVERSOAGRO. **PIB do setor do algodão atingiu 74,11 bilhões de dólares na safra 2016/17**. [s.n.], 2017. Disponível em: <<http://www.uagro.com.br/editorias/agricultura/algodao/2017/12/07/pib-do-setor-do-algodao-atingiu-us-74-11-bi-na-safra-2016-17.html>>. Acesso em: 19 Março. 2018.
- USDA. Guidelines for hvi testing. United States Department of Agriculture, Agricultural Marketing Service, Cotton Program, 3275 Appling Road, Memphis, Tennessee 38133, 2005. Disponível em: <<http://cotton.tamu.edu/Harvest/Guidelines\%20for\%20HVI\%20Testing.pdf>>.
- VAINSENER, S. A. Algodão. Agosto 2009. Dados publicados na Pesquisa Escolar Online, Fundação Joaquim Nabuco, Recife. Disponível em: <<http://basilio.fundaj.gov.br/pesquisaescolar/>>. Acesso em: 28 Mar. 2018.
- VAINSENER, S. A. **História do Algodão**. AMPA, 2012. Dados publicados pela Pesquisa Escolar Online, Fundação Joaquim Nabuco, Recife. Disponível em: <http://www.sincti.com/clientes/ampa/site/qs_historia.php>. Acesso em: 28 Mar. 2018.
- VIEIRA, R. T.; FERRAZ, C. T.; LANGONI, V. de M.; GONZAGA, A. Uma nova proposta para análise de textura local por conjuntos fuzzy. Anais do VIII Workshop de Visão Computacional, 2012. Disponível em: <https://www.researchgate.net/publication/234059469_Uma_nova_proposta_para_analise_de_textura_local_por_conjuntos_fuzzy>.

VÉRAS, G.; BRITO, A. L. B. de; SILVA, A. C. da; SILVA, P. da; COSTA, G. B. da; FÉLIX, L. C. N.; FERNANDES, D. D. de S.; FONTES, M. M. de. Biodiesel classification in the visible region. *Química Nova*, n. 2, 2012. ISSN 0100-4042. Disponível em: <<http://dx.doi.org/10.1590/S0100-40422012000200015>>.

WEBB, A. R. **Statistical Pattern Recognition**. 2. ed. Malvern, UK: Wiley, 2002. ISBN 978-0470845134.

WESZKA, J. S.; DYER, C. R.; ROSENFELD, A. A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 1976.

YANG, X.; YEO, S. Y.; WONG, S. T.; LEE, G.; SU, Y.; HONG, J. M.; CHOO, A.; CHEN, S. Color-texture based extreme learning machines for tissue tumor classification. *SPIE Medical Imaging 2016: Digital Pathology*, 2016.

YIMING, J.; KAIHSIUNG, C.; CHICHENG, H. Efficient edge detection and object segmentation using gabor filters. *Proceeding ACM-SE 42 Proceedings of the 42nd annual Southeast regional conference*, Huntsville, Alabama, 2004.