



*Universidade Federal da Paraíba*

*Centro de Tecnologia*

*Programa de Pós-Graduação em Engenharia Mecânica*

*Mestrado-Doutorado*



**DETECÇÃO DE ADULTERAÇÃO DE  
GASOLINA POR ETANOL POR MEIO DA  
ANÁLISE DE IMAGENS INFRAVERMELHAS  
POR ALGORITMOS DE APRENDIZADO DE  
MÁQUINA**

por

Pedro Jorge Martins Bichara Sobreira

*Dissertação de Mestrado apresentada à Universidade Federal da Paraíba  
para obtenção do grau de Mestre.*

**PEDRO JORGE MARTINS BICHARA SOBREIRA**

**DETECÇÃO DE ADULTERAÇÃO DE GASOLINA POR ETANOL  
POR MEIO DA ANÁLISE DE IMAGENS INFRAVERMELHAS POR  
ALGORITMOS DE APRENDIZADO DE MÁQUINA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Mecânica do Centro de Tecnologia da Universidade Federal da Paraíba, em cumprimento às exigências para obtenção do Grau de Mestre em Engenharia Mecânica – área de concentração: Dinâmica e Controle de Sistemas Mecânicos.

**Orientador: Prof. Dr. Abel Cavalcante Lima Filho**

**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

S677d Sobreira, Pedro Jorge Martins Bichara.

Detecção de adulteração de gasolina por etanol por meio da análise de imagens infravermelhas por algoritmos de aprendizado de máquina / Pedro Jorge Martins Bichara Sobreira. - João Pessoa, 2024.  
128 f. : il.

Orientação: Abel Cavalcante Lima Filho.  
Dissertação (Mestrado) - UFPB/CT.

1. Combustíveis líquidos. 2. Motores de combustão interna. 3. Aprendizado de máquina. 4. Classificação de imagens. 5. Imagens infravermelhas - Análise de componentes. I. Lima Filho, Abel Cavalcante. II. Título.

UFPB/BC

CDU 662.75(043)

# DETECÇÃO DE ADULTERAÇÃO DE GASOLINA POR ETANOL POR MEIO DA ANÁLISE DE IMAGENS INFRAVERMELHAS POR ALGORITMOS DE APRENDIZADO DE MÁQUINA

por

**PEDRO JORGE MARTINS BICHARA SOBREIRA**  
*Dissertação/aprovada em 16 de dezembro de 2024*

Documento assinado digitalmente



**Prof. Dr. Abel Cavalcante Lima Filho**  
Orientador/a – Universidade Federal da Paraíba



**Prof. Dr. Alisson Vasconcelos de Brito**  
Examinador Interno – Universidade Federal da Paraíba

Documento assinado digitalmente

gov.br

KOJE DANIEL VASCONCELOS MISHINA

Data: 17/12/2024 09:51:56-0300

Verifique em <https://validar.iti.gov.br>

**Prof. Dr. Koje Daniel Vasconcelos Mishina**  
Examinador Externo – Universidade Federal de Joazeiro

João Pessoa - PB  
2024

## **DEDICATÓRIA**

A Deus pelo dom da existência. A meus pais, Marcus e Denise, pelo incentivo, pela paciência e pelo apoio incondicional, À minha irmã, Fernanda, pelas conversas e pela amizade, que eu trago para sempre no coração. À minha namorada, Nathália, por todo o apoio e auxílio nos momentos em que mais precisei. À minha madrinha, Maria Lúcia, por todo o apoio e por ser sempre tão presente, participando de cada batalha e celebrando cada vitória. A todos os professores que um dia participaram da minha formação. Sem vocês esse trabalho não seria possível.

## **AGRADECIMENTOS**

A Deus, pela oportunidade de ser quem eu sou e pelas condições de concretizar esse trabalho.

A meus pais, Marcus e Denise, que sempre lutaram para que eu pudesse chegar onde estou hoje e que nunca mediram esforços para garantir a realização dos meus sonhos. É uma honra ser seu filho.

À minha namorada, Nathália de Vasconcelos Gomes, que esteve ao meu lado nos momentos de preocupação, e que nunca me deixou esquecer que eu era capaz de conquistar qualquer coisa.

Ao meu orientador, Abel Cavanlante Lima Filho, pela paciência e disponibilidade, por todas as conversas e orientações, e por todo o apoio, sem o qual esse trabalho não seria possível.

A meus colegas de laboratório, Anselmo e Júlio César pelo apoio e amizade, pelas valiosas dicas e sugestões, e por terem tornado o dia a dia da pesquisa mais agradável.

À CAPES pelo apoio e pelo financiamento, extremamente úteis e necessários à concretização desse trabalho.

Tudo o que temos que decidir é o que fazer com o tempo que nos é dado.

J. R. R. Tolkien

# **DETECÇÃO DE ADULTERAÇÃO DE GASOLINA POR ETANOL POR MEIO DA ANÁLISE DE IMAGENS INFRAVERMELHAS POR ALGORITMOS DE APRENDIZADO DE MÁQUINA**

## **RESUMO**

Veículos que utilizam combustíveis fósseis representam o maior percentual da frota ativa no mundo, por isso apesar de todos os malefícios atrelados ao uso da gasolina, essa ainda está diretamente envolvida no dia-a-dia das pessoas. Há um esforço por parte de empresas para reduzir o custo de produzir e distribuir gasolina no intuito de maximizar o lucro e, com esse fim, o uso de técnicas ilegais não é incomum. Um exemplo é a adulteração desse combustível com etanol, que pode causar danos aos componentes do motor quando essas máquinas não estão equipadas para lidar com o etanol. Há a necessidade de um método fácil, rápido e efetivo de detectar esse solvente em misturas de gasolina e etanol; o presente trabalho tem como objetivo apresentar um método baseado em processamento de imagens infravermelhas (IR), Análise de Componentes Principais (PCA), e algoritmos de aprendizado de máquina (ML) que possa classificar imagens de motores em operação no que diz respeito ao conteúdo de etanol que foi introduzido na gasolina utilizada como combustível. O estudo concluiu que Redes neurais artificiais (ANR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN) e Árvore de Decisão (DT) são capazes de classificar as imagens com acurácias de respectivamente 94,6%, 90,8%, 94,7% e 97,2%.

Palavras-chave: Imagens Infravermelhas, Motores de Combustão Interna, Aprendizado de Máquina, Classificação de Imagens, Análise de Componentes Principais.

# **DETECTION OF ADULTERATION OF GASOLINE WITH ETHANOL THROUGH INFRARED IMAGE ANALYSIS BY MACHINE LEARNING ALGORITHMS.**

## **ABSTRACT**

Fossil fuels-based vehicles represent the largest percentage of the world's active fleet, therefore despite all of the issues gasoline is linked to, it is directly involved in people's daily lives. There is effort made by the companies to reduce the cost of producing and distributing gasoline to maximize profit, and the use of illegal techniques is not uncommon. One example is the adulteration of that fuel with ethanol, which can lead to damage to the engine's components when the machines are not properly equipped to handle ethanol. There is a need for an easy, fast and effective way to detect this solvent in gasoline-ethanol blends; the present paper aims to present a method based on infrared imaging (IR) processing, principal component analysis (PCA) and machine learning (ML) algorithms to classify images of operating engines in regards to the content of ethanol that was introduced as fuel blended with gasoline. The study concluded that Artificial Neural Networks (ANR), Support Vector Machines (SVM), K-Nearest Neighbors (KNN) and Decision Tree (DT) algorithms are capable of classifying the images with accuracies of, respectively, 94,6%, 90,8%, 94,7% and 97,2%.

**Keywords:** Infrared Images, Internal Combustion Engine, Machine Learning, Image Classification, Principal Component Analysis.

## SUMÁRIO

1	INTRODUÇÃO .....	1
1.1	ADULTERAÇÃO POR ETANOL E MÉTODOS DE MEDIÇÃO .....	1
1.2	MOTIVAÇÃO .....	4
1.3	OBJETIVO .....	7
1.3.1	Objetivo Geral .....	7
1.3.2	Objetivos Específicos .....	8
1.4	APRESENTAÇÃO DA DISSERTAÇÃO .....	8
2	GASOLINA .....	10
2.1	COMPOSIÇÃO E PRINCIPAIS PROPRIEDADES .....	10
2.2	COMBUSTÃO DA GASOLINA .....	10
2.3	LEGISLAÇÃO BRASILEIRA .....	13
3	MOTORES A COMBUSTÃO INTERNA .....	14
3.1	CLASSIFICAÇÃO E ASPECTOS CONSTITUTIVOS E FUNCIONAIS .....	14
4	PROCESSAMENTO DE IMAGENS .....	19
4.1	LUZ E O ESPECTRO ELETROMAGNÉTICO .....	19
4.2	IMAGENS DIGITAIS .....	21
4.2.1	Amostragem e Quantização .....	22
4.3	APRIMORAMENTO DE IMAGENS .....	23
4.3.1	Filtros no Domínio Espacial .....	25
4.3.2	Filtros de Suavização .....	27
4.4	HISTOGRAMAS DE IMAGENS DIGITAIS .....	28
4.5	CONVERSÃO DE ESCALA DE CORES PARA ESCALA DE CINZA .....	31
4.6	REGISTRO DE IMAGEM .....	35
4.6.1	Interpolação de Intensidades .....	37
4.7	SEGMENTAÇÃO DE IMAGENS .....	38
4.7.1	Thresholding .....	39
4.7.2	Método de Otsu Para Thresholding .....	41
5	INTELIGÊNCIA ARTIFICIAL .....	46
5.1	APRENDIZADO DE MÁQUINA .....	46
5.2	K-NEAREST NEIGHBORS (KNN) .....	49
5.3	SUPPORT VECTOR MACHINE (SVM) .....	52

5.4	REDES NEURAIS ARTIFICIAIS.....	57
5.5	ÁRVORE DE DECISÃO.....	65
6	ANÁLISE DE COMPONENTES PRINCIPAIS.....	69
6.1	DECOMPOSIÇÃO DE VALOR SINGULAR (SVD).....	69
6.2	PRÉ-REQUISITOS PARA APLICAÇÃO DO MÉTODO .....	70
6.3	ENCONTRANDO OS COMPONENTES .....	71
7	TEORIA DA INFORMAÇÃO .....	73
8	MATERIAIS E MÉTODO .....	79
8.1	MATERIAIS UTILIZADOS .....	79
8.2	MEDIÇÃO DO PERCENTUAL DE ETANOL .....	81
8.2.1	Medição do percentual de etanol pelo método da ABNT NBR 13992:2015 81	
8.2.2	Medição do percentual de etanol pela utilização do sensor.....	82
8.3	AQUISIÇÃO DE FRAMES .....	83
8.4	PROCESSAMENTO DE IMAGEM .....	83
8.5	SELEÇÃO E ANÁLISE DE FEATURES .....	86
8.6	CONFIGURAÇÃO DOS MODELOS DE ML .....	86
9	RESULTADOS E DISCUSSÃO .....	88
10	CONCLUSÃO .....	102

## LISTA DE FIGURAS

Figura 2.1 – Representação em etapas do processo de “spark knock”, ou batida de pino (RICHARDS, BARKER, 2023). .....	11
Figura 2.2 - Diagrama pressão no cilindro x ângulo do virabrequim que compara os processos de combustão descritos. Pre-ignition Combustion refere-se à condição de pré-ignição, Knocking Combustion refere-se à condição em que há batida de pino e Normal Combustion refere-se à combustão que ocorre normalmente. TDC significa Top-dead Center ou ponto morto superior e Point of Ignition corresponde ao ponto em que ocorre a ignição (RICHARDS, BARKER, 2023). .....	12
Figura 3.1 - Os quatro tempos de um motor. Intake se refere à admissão, Compression se refere à compressão, Expansion se refere à expansão e Exhaust se refere à exaustão (HEYWOOD, 1988). .....	16
Figura 3.2 – a) Diagrama pressão x volume para ciclo padrão a ar Otto; b) Diagrama temperatura x entropia para o ciclo padrão a ar Otto. Adaptado de FERRARI e MARIN (2018). .....	17
Figura 3.3 - Diagrama pressão x volume de um ciclo Otto real (DA SILVEIRA, 2008). ....	18
Figura 4.1 – Espectro eletromagnético com destaque para o intervalo de comprimentos de onda visíveis por humanos a olho nu ( <a href="https://www.todamateria.com.br/espectro-eletromagnetico/">https://www.todamateria.com.br/espectro-eletromagnetico/</a> ). .....	20
Figura 4.2 – Operador de formato quadrado comumente utilizado em operações de aprimoramento de imagem centrado no ponto (x,y) (GONZALES e WOODS, 2018). ....	24
Figura 4.3 – Detalhamento de uma máscara/filtro utilizada para aprimoramento de imagens (GONZALES e WOODS, 2018). .....	25
Figura 4.4 – Exemplo de aplicação de filtro de suavização. a) Imagem original; b) aplicação do filtro; c) aplicação de um método de thresholding na imagem filtrada. Adaptado de GONZALES e WOODS (2018). .....	28
Figura 4.5 – Histograma relativo à imagem 4.7 a). .....	29
Figura 4.6 – Exemplos do comportamento de distribuições de probabilidade com assimetria à direita e à esquerda (SPIEGEL et al, 2013). .....	30

Figura 4.7 – Exemplo de distribuições de probabilidade com alta curtose (SPIEGEL et al, 2013).....	31
Figura 4.8 – espectro de absorção dos receptores para vermelho, verde e azul dos cones. Adaptado de NELSON e COX (2014). .....	32
Figura 4.9 – Diagrama ilustrativo da relação entre as escalas de cinza e RGB ( <a href="https://goo.gl/bebjMD">https://goo.gl/bebjMD</a> ).....	33
Figura 4.10 – Diagrama de cores na escala YQI (YAGLE e ULABY, 2018). .....	34
Figura 4.11 Exemplo de histograma de imagem em que a presença de dois picos em níveis de cinza diferentes facilita a determinação de um threshold para diferenciação de objetos e plano de fundo (GONZALES e WOODS, 2018). .....	40
Figura 4.12 – Efeitos do ruído na determinação do threshold para determinação de objetos e plano de fundo em uma imagem (GONZALES e WOODS, 2018). .....	41
Figura 5.1 - Ilustração das situações de a) overfitting e b) underfitting. Adaptado de MOHRI et al (2018).....	49
Figura 5.2 – Fronteiras de decisão obtidas pelo método K-Nearest Neighbors para valores de K iguais a 1,3 e 31 (BISHOP, 2006).....	52
Figura 5.3 – Representação visual das camadas de uma rede neural com destaque para as entradas (inputs), saídas (outputs) e os pesos utilizados (BISHOP, 2006).....	58
Figura 5.4 – representação da função erro Ew no espaço de pesos (BISHOP, 2006). .....	62
Figura 5.5 – representação da estrutura de uma árvore de decisão. Adaptado de ZHOU (2016). .....	66
Figura 6.1 – Representação geométrica da relação entre os dois primeiros componentes principais (eixos em laranja) e as variáveis originais (eixos em preto).....	72
Figura 7.1 – Principais componentes de um sistema de comunicação. Adaptado de SHANNON e WEAVER (1949).....	74
Figura 7.2 – Diagrama que relaciona entropia (H) e a probabilidade (p) de um sistema se encontrar em um dado estado de um universo de dois estados possíveis (SHANNON e WEAVER, 1949). .....	76
Figura 8.1 - Câmera térmica utilizada para realização do experimento. ....	80
Figura 8.2 - Setup utilizado para o experimento. ....	80
Figura 8.3 – Procedimento para leitura da fase aquosa (ABNT NBR 13992:2015) .....	82
Figura 8.4 – Calibração do sinal de saída do sensor flex (CONTINENTAL, 2014).....	83

Figura 8.5 Representação gráfica do processo de correção da imagem obtida após as transformações geométricas. ....	85
Figura 8.6 – Resumo do processo de processamento de imagem utilizado nesse estudo. ..	85
Figura 8.7 – Resumo do processo de coleta, processamento e análise de dados utilizado nesse estudo.....	86
Figura 9.1 – a) Seção de imagem antes do processo de filtragem; b) mesma seção pós-filtragem. ....	89
Figura 9.2 - Resultado do processo de registro da imagem, com linhas vermelhas indicando os limites das correções. ....	89
Figura 9.3 – Comparação entre duas imagens a) antes do processo de registro e b) após o processo de registro. As partes em cinza são comuns às duas imagens e as partes em verde e magenta correspondem às particularidades de cada uma. ....	90
Figura 9.4 - Diagrama das entropias dos frames dos vídeos utilizados no estudo. ....	91
Figura 9.5 – Diagrama da luminância média dos pixels dos frames dos vídeos utilizados no estudo.....	91
Figura 9.6 - Diagrama do desvio padrão de luminância dos pixels dos frames dos vídeos utilizados no estudo. ....	92
Figura 9.7 - Diagrama da curtose dos histogramas dos frames dos vídeos utilizados no estudo.....	92
Figura 9.8 - Diagrama da curtose dos histogramas dos frames dos vídeos utilizados no estudo.....	93
Figura 9.9 – Matriz de confusão para a rede neural artificial.....	94
Figura 9.10 – Matriz de confusão para K-nearest neighbors.....	94
Figura 9.11 – Matriz de confusão para support vector machine.....	95
Figura 9.12 – Matriz de confusão para a árvore de decisão. ....	95
Figura 9.13 – Fronteiras de decisão para a rede neural artificial plotadas sobre o gráfico que relaciona os dois componentes principais mais relevantes.....	96
Figura 9.14 - Fronteiras de decisão para K-nearest neighbors plotadas sobre o gráfico que relaciona os dois componentes principais mais relevantes.....	96
Figura 9.15 - Fronteiras de decisão para support vector machine plotadas sobre o gráfico que relaciona os dois componentes principais mais relevantes.....	97
Figura 9.16 - Fronteiras de decisão para a árvore de decisão plotadas sobre o gráfico que relaciona os dois componentes principais mais relevantes.....	97

Figura 9.17 Curva de aprendizado para o algoritmo árvore de decisão. ....	99
Figura 9.18 Curva de aprendizado para o algoritmo K-Nearest neighbors. ....	99
Figura 9.19 Curva de aprendizado para o algoritmo rede neural artificial.....	100
Figura 9.20 Curva de aprendizado para o algoritmo support vector machine.....	100

## LISTA DE QUADROS E TABELAS

Quadro 1.1 – Resumo dos trabalhos relacionados: principais métodos de medição do percentual de etanol em mistura com gasolina. ....	4
Quadro 1.2 – Resumo dos trabalhos selecionados: aplicações de imagens infravermelhas para análise de sistemas mecânicos. ....	7
Tabela 4.1 - Exemplos das principais transformações geométricas aplicadas em imagens. Adaptado de (GONZALES e WOODS, 2018). ....	37
Tabela 8.1 – Ficha técnica com as informações mais relevantes a respeito do veículo utilizado para testes. ....	80
Tabela 9.1 – Percentuais de etanol medidos em cada uma das amostras analisadas. ....	88
Tabela 9.2 – Percentual da variância explicado por cada componente principal. ....	93
Tabela 9.3 – Acurácias resultantes para cada um dos algoritmos estudados. ....	98

## LISTA DE ABREVIATURAS E SIGLAS

ABNT – Associação Brasileira de Normas Técnicas

ANP – Agência Nacional do Petróleo, Gás Natural e Biocombustíveis

ASTM – Abreviatura do inglês *American Society for Testing and Materials* [=Sociedade Americana de Testes e Materiais]

BTD – Benzotiazóis

CFD – Abreviatura do inglês *Computational Fluid Dynamics* [=Dinâmica de Fluidos Computacional]

CIE – Abreviatura do francês *Comission Internationale de l'Eclairage* [=Comissão Internacional de Iluminação]

E85 – Mistura com 85% de etanol

EAC – Etanol Anidro Combustível

EPA – Abreviatura do inglês *United States Environmental Protection Agency* [=Agência de Proteção Ambiental dos Estados Unidos]

IEA – Abreviatura do inglês *International Energy Agency* [=Agência Internacional de Energia]

IR – Abreviatura do inglês *Infrared* [=Infravermelho]

KKT – Karush-Kuhn-Tucker

MAPA – Ministério da Agricultura e Pecuária e Abastecimento

MCI – Motor de Combustão Interna

ML – Abreviatura do inglês *Machine Learning* [=Aprendizado de Máquina]

NBR – Norma Brasileira Regulamentadora

NOx – Número de Oxidação

NTSC – Abreviatura do inglês *National Television System Commitee* [=Comitê do Sistema Nacional de Televisão]

PCA – Abreviatura do inglês *Principal Component Analysis* [=Análise de Componentes Principais]

PMI – Ponto Morto Inferior

PMS – Ponto Morto Superior

RGB – Abreviatura do inglês *Red Green Blue* [=Vermelho verde azul]

TBM – Abreviatura do inglês *Total Base Number* [=Número de Base Total]

TC – Taxa de Compressão

## CAPÍTULO I

### 1 INTRODUÇÃO

#### 1.1 ADULTERAÇÃO POR ETANOL E MÉTODOS DE MEDIÇÃO

De acordo com a INTERNATIONAL ENERGY AGENCY (IEA) (2023), derivados de petróleo ainda exercem um papel crucial no cenário de transporte do século XXI, de modo que esse setor corresponde a cerca de 45% da demanda global de petróleo; além disso, mais de 90% da demanda energética do setor ainda é suprida por combustíveis fósseis e estima-se que até 2030 esse número ainda se mantenha acima de 80%. A UNITED STATES ENVIRONMENTAL PROTECTION AGENCY (EPA) (2023) afirma ainda que apesar da tendência de migração dos consumidores para veículos híbridos e elétricos, esses só correspondem, até 2022, a apenas respectivamente 10% e 7% da produção global. Pode-se concluir, portanto que a frota internacional ainda é majoritariamente baseada em combustíveis derivados do petróleo, especialmente a gasolina.

De acordo com a AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS (ANP) (2020), no Brasil, a gasolina é o combustível leve de maior comercialização utilizado em máquinas térmicas que operam por Ciclo Otto. Devido à posição privilegiada que ocupa no comércio de combustíveis brasileiros, a gasolina está sujeita a práticas ilegais para aumentar o lucro de empresas que realizam sua venda para o consumidor final. É válido destacar a adulteração de gasolina por solventes como etanol, querosene, diesel e thinner. TAKESHITA (2006) afirma que a adulteração é motivada principalmente por quatro fatores:

- Grande diferença de preços entre solventes e gasolina;
- Solventes são de fácil aquisição, e não há limites legais para sua obtenção;
- Solventes são solúveis na gasolina, o que torna difícil a sua detecção visual;

- Alta incidência de impostos sobre a gasolina, o que diminui a margem de lucro possível para o produto.

Apesar da difusão da tecnologia FLEX em veículos, que garantiria a segurança de se usar etanol como combustível sem grandes danos ao maquinário do veículo. A adulteração traz malefícios aos carros que não estão propriamente equipados para lidar com essa composição de combustível. A utilização de misturas de etanol e gasolina está associada à formação de depósitos nas válvulas injetoras, o que afeta o fluxo e padrão de spray de combustível. As consequências disso são piora na dirigibilidade, redução na potência do motor e na economia de combustível e aumento da emissão de gases na exaustão (STEPIEN, 2016). Além disso, de acordo com GONÇALVES, *et al* (2022), a adulteração de gasolina por etanol está atrelada a fenômenos relacionados à degradação dos óleos lubrificantes do motor:

- Oxidação: causa a formação de um lodo, ou borra, que pode entupir os canais de distribuição de lubrificante; apesar de não causar dano imediato ao motor, esse fenômeno pode causar a redução de vida útil desse;
- Sulfatação: refere-se à formação de depósitos de sulfato em componentes e é indesejável em qualquer nível; é responsável pelo consumo de aditivos dos lubrificantes e pela corrosão química dos dispositivos que entram em contato com esse lubrificante alterado;
- Total Base Number (TBN): indica a capacidade de um óleo lubrificante neutralizar compostos ácidos formados durante a combustão, prevenindo a corrosão; a presença de etanol na gasolina tem efeitos negativos sobre esse parâmetro.

Em função da relevância econômica e das consequências negativas associadas à adulteração de gasolina por etanol, diversos métodos para identificar e quantificar essa adulteração foram propostos.

É possível preparar uma solução composta por diclorometano de potássio e ácido sulfúrico para atuar como um indicador da presença de etanol na gasolina através da mudança de cor oriunda de reação química, que se dá através do contato da mistura de combustíveis a ser analisada e um papel indicador que possua a solução preparada (FADAIRO *et al*, 2020).

Além disso, já foi demonstrado que é possível utilizar 2,1,3-benzotiazóis (BTD) como base para criação de compostos que, em função de suas propriedades fotofísicas, podem ser usados como sensores para uma série de substâncias como fluoreto, cobalto, cianeto, mercúrio e DNA. Todas essas substâncias possuem grupos polares em suas estruturas, e essa característica limitava o uso de BTDs; entretanto, ISOPPO *et al* (2020) conseguiram sintetizar compostos capazes de detectar substâncias de caráter apolar, possibilitando a detecção de etanol em gasolina convencional em proporções que variam de 5-50% (v/v).

PEDROSO *et al* (2008) mostraram que é possível também identificar o percentual de diversos solventes em gasolina, incluindo etanol, com um erro que variou de 3,3% a 8,2%, a depender da mistura, através de um método que combina cromatografia gasosa bidimensional e detecção por ionização em chama. Além disso, BEZERRA *et al* (2019) mostraram ser possível quantificar a adulteração da gasolina por etanol anidro através da aplicação da espectroscopia de Raman; os espectros da gasolina pura, do etanol puro e da gasolina comercial foram comparados e analisados à luz do método de Análise de Componentes Principais (PCA).

ROMANEL *et al* (2018) propôs um novo método para identificar e quantificar adulteração de gasolina por etanol utilizando ressonância magnética nuclear no domínio do tempo a partir da comparação dos tempos de relaxamento transversal da gasolina, do etanol e das misturas analisadas.

De maneira sucinta, métodos disponíveis para identificar e/ou quantificar a adulteração por etanol anidro estão dispostos no Quadro 1.1.

Autor	Método
FADAIRO <i>et al</i> (2020)	Uso de papel indicador com solução de diclorometano de potássio e ácido sulfúrico.
ISOPPO <i>et al</i> (2020)	Observação das propriedades fotoquímicas de compostos baseados em 2,1,3-benzotiazóis quando em contato com misturas de gasolina e etanol.

PEDROSO <i>et al</i> (2008)	Cromatografia gasosa bidimensional e detecção por ionização em chama
BEZERRA <i>et al</i> (2019)	Espectrografia de Raman e Análise de Componentes Principais.
ROMANEL <i>et al</i> (2018)	Análise dos tempos de relaxamento transversal a partir da ressonância magnética nuclear no domínio do tempo.

Quadro 1.1 – Resumo dos trabalhos relacionados: principais métodos de medição do percentual de etanol em mistura com gasolina.

## 1.2 MOTIVAÇÃO

Apesar de esses métodos serem de fato capazes de realizar aquilo que se propõem, a identificação e quantificação da adulteração da gasolina por etanol, esses possuem aspectos que podem dificultar sua utilização:

- Necessidade de amostragem da mistura gasolina/etanol;
- Necessidade de aparelhagem e profissionais altamente especializados, sendo necessários em alguns casos laboratórios para realização das análises;
- Tempo relativamente longo para coleta e análises de amostras;
- Contato com substâncias tóxicas.

Por outro lado, a análise de imagens infravermelhas vem se mostrando um método útil para situações que envolvem classificação e análise de fenômenos, especialmente no que diz respeito à identificação e quantificação de anormalidades. UTH *et al* (2019) desenvolveram um método para identificar indivíduos em cenários de mineração; seu estudo mostrou que é possível, a partir de imagens infravermelhas de minas, localizar pessoas que estejam presas em decorrência de um acidente, por exemplo. Para isso, ele equipou um robô com um dispositivo termográfico, que pode auxiliar nos resgates sem colocar mais pessoas em risco na busca por feridos.

A observação dos padrões de temperatura também é útil no âmbito da medicina, com uma de suas aplicações sendo a detecção de COVID-19 a partir de imagens infravermelhas das costas de pacientes com até 92% de sensibilidade. Apesar de o fundamento fisiológico

para essa descoberta ainda não ter sido proposto, os autores sugerem a variação de temperatura na região de interesse se dá por disfunções microvasculares causadas pela COVID-19 (BRZEZINSKI *et al*, 2021).

Mais especificamente, a análise de imagens infravermelhas é amplamente usada para identificar o mau funcionamento de componentes mecânicos atrelado a variações nos padrões de temperatura do dispositivo em questão. As variações de temperatura podem fornecer informações importantes para predição de falhas em sistemas de engrenagens fornecendo a possibilidade de se visualizar os efeitos que as condições de operação têm no dispositivo (BŁAŻEJ *et al*, 2016).

No âmbito dos motores de combustão interna (MCIs), essas imagens podem ser usadas para monitorar a vida útil do motor em função do combustível utilizado levando em conta os valores de intensidade dos pixels comparados com a taxa de liberação de calor e o NO<sub>x</sub> (GONÇALVES *et al*, 2022). Ainda sobre motores, é possível utilizar as imagens infravermelhas para avaliar sua performance e funcionamento, uma vez que os seus padrões de temperatura estão associados ao NO<sub>x</sub> (PEER, 2019, SHAMEER, 2017). Por fim, há muitos benefícios em se estudar o processo de combustão dentro dos cilindros à luz da termografia infravermelha (MANCARUSO *et al*, 2020).

Uma maneira de expandir ainda mais as possibilidades da análise de imagens termográficas é uni-las a algoritmos de Machine Learning (ML), especialmente adequados para situações de classificação e amplamente utilizados em diversos cenários, incluindo o de análise de funcionamento de sistemas mecânicos. WILSON *et al* (2023) identificaram ainda aplicação de ML aliada a termografia na medicina, no controle de fronteiras, na análise predial, na inspeção de redes elétricas, na análise de obras de arte entre outros.

SHAMMI *et al* (2022) demonstraram a possibilidade de utilizar algoritmos de ML para estudar a termodinâmica do congelamento na superfície de plantas, fenômeno que pode trazer consequências negativas irreversíveis para as mesmas; foi possível identificar pontos de nucleação e locais de temperatura de congelamento a partir de imagens IR das plantas analisadas, o que possibilita a prevenção do congelamento e a adoção de medidas para proteção de plantações, por exemplo.

PARK *et al* (2022) utilizaram algoritmos de ML para determinar a profundidade de rachaduras no concreto, o que possibilita uma melhor proteção das estruturas analisadas,

uma vez que com o tempo essas rachaduras pequenas podem aumentar e causar danos estruturais. Além disso, NAGORI *et al* (2018) mostraram que é possível identificar, de modo proativo e preditivo, o risco de choques hemodinâmicos em crianças a partir das imagens IR de pacientes em uma unidade pediátrica.

Já no que diz respeito a sistemas mecânicos, foi demonstrado que é possível identificar vazamentos em compressores de ar através da análise das imagens por um algoritmo de Support Vector Machine aliado com um algoritmo de otimização do tipo Improved Firefly (TONG *et al*, 2020). Além disso, redes neurais convolucionais aliadas ao método bag-of-visual-words podem ser usadas para identificar defeitos em máquinas rotativas; mais especificamente, a pesquisa de JIA *et al* (2023) foram capazes de identificar: defeitos na pista externa, na pista interna, defeito no rolete ou rub impact. YOUNUS *et al* (2010), também obtiveram resultados similares em suas análises de máquinas rotativas ao utilizar uma série de algoritmos de ML aliados a métodos estatísticos como PCA para, a partir de imagens IR, identificar o tipo de defeito presente no momento da captura da imagem. Por fim, JANSSENS *et al* (2018) também foram capazes de identificar a condição de funcionamento de máquinas rotativas a partir da análise realizada por métodos de deep learning.

YU *et al* (2022) demonstraram que com a análise da textura de imagens IR à luz de algoritmos de ML aliados ao estudo dos padrões binários locais, é possível prever a profundidade de soldas feitas a arco com eletrodo consumível e proteção gasosa, métrica crucial para a avaliação dessas, e que é altamente dependente da geometria e posicionamento dos chanfros das peças a serem soldadas.

LI *et al* (2021) demonstraram que é possível aliar PCA a algoritmos de ML para estimar a profundidade da cavidade de combustão em fornalhas usadas para produção de ferro a partir de termogramas gerados em simulações CFD.

De modo sucinto, o Quadro 1.2 exhibe trabalhos que aliaram algoritmos de aprendizado de máquina à análise de imagens infravermelhas para avaliar e/ou diagnosticar sistemas de engenharia. Apesar de não estar especificado na tabela, todos os trabalhos citados possuem sua própria metodologia para realizar um processamento das imagens antes da análise propriamente dita.

Autor	Objetivo	Técnica
TONG <i>et al</i> (2020)	Identificar vazamentos em compressores de ar.	Análise de imagens térmicas via Support Vector Machine e otimização via Improves Firefly algorithm.
JIA <i>et al</i> (2023)	Identificar defeitos em máquinas rotativas.	Análise de imagens térmicas via redes neurais convolucionais e bag-of-visual-words.
YOUNUS <i>et al</i> (2010)	Identificar defeitos em máquinas rotativas.	Análise de imagens térmicas via algoritmos de aprendizado de máquina e análise de componentes principais.
JANSSENS <i>et al</i> (2018)	Identificar defeitos em máquinas rotativas.	Análise de imagens térmicas via redes neurais convolucionais.
YU <i>et al</i> (2022)	Previsão da profundidade de soldas.	Análise de texturas de imagens térmicas via algoritmos de aprendizado de máquina e estudo de padrões binários locais.
LI <i>et al</i> (2021)	estimar a profundidade da cavidade de combustão em fornalhas usadas para produção de ferro	Análise de imagens térmicas geradas em simulações por algoritmos de aprendizado de máquina e Análise de Componentes Principais.

Quadro 1.2 – Resumo dos trabalhos selecionados: aplicações de imagens infravermelhas para análise de sistemas mecânicos.

### 1.3 OBJETIVO

#### 1.3.1 Objetivo Geral

O objetivo geral desse trabalho é propor um método de análise de imagens IR de um MCI em funcionamento que possa estimar o percentual de etanol anidro misturado na

gasolina utilizada pelo veículo durante o período de coleta das imagens, dentro do intervalo limitado pelos percentuais de adulteração utilizados para realização do estudo.

### 1.3.2 Objetivos Específicos

- Determinar as etapas de processamento de imagem necessárias para posterior análise dessas;
- Selecionar as métricas adequadas para análise, bem como os métodos adequados para tratamento estatístico dos dados;
- Demonstrar a aplicabilidade de algoritmos de ML para estimativa do percentual de etanol na gasolina.

## 1.4 APRESENTAÇÃO DA DISSERTAÇÃO

Esta dissertação foi dividida em dez capítulos, assim organizados:

- **Capítulo I:** introduz o problema da adulteração de gasolina por etanol, ressaltando os motivos que levam isso a acontecer, bem como as consequências para os componentes mecânicos de veículos não preparados para a operação com esse tipo de substância. Apresenta os métodos já existentes para aferição do percentual de adulteração e justifica a escolha do método proposto. Por fim, formaliza os objetivos gerais e específicos deste trabalho;
- **Capítulo II:** fornece o arcabouço teórico para compreensão da utilização da gasolina pela humanidade, explicando as vantagens dessa substância. Debate aspectos físico-químicos da gasolina e descreve como a legislação brasileira trata a gasolina comercial, bem como a adição de etanol nessa;
- **Capítulo III:** fornece os fundamentos teóricos para compreensão da utilização de MCIs. Debate o funcionamento dos motores dos pontos de vista mecânico e termodinâmico;
- **Capítulo IV:** aborda toda a fundamentação teórica necessária para a compreensão dos conceitos relevantes a respeito de imagens digitais. Define imagens digitais e

discute suas principais características e propriedades. Por fim, debate as etapas de processamento utilizadas nesse estudo;

- **Capítulo V:** discute toda a fundamentação teórica para compreensão dos algoritmos de ML utilizados na realização desse estudo. De modo mais específico são debatidas as bases matemáticas desses métodos, mostrando como esses são capazes de realizar tarefas de classificação;
- **Capítulo VI:** aborda as bases matemáticas do método de PCA, utilizado no desenvolvimento desse trabalho;
- **Capítulo VII:** discute conceitos ligados à teoria da informação, cuja compreensão é necessária para aplicação do método proposto;
- **Capítulo VIII:** apresenta os materiais e métodos, explicando a maneira como as imagens foram obtidas e processadas, bem como as estratégias utilizadas para tratar e analisar os dados resultantes desse processo;
- **Capítulo IX:** apresenta e analisa os resultados obtidos através do método proposto no capítulo VIII. Mostra os dados obtidos a partir da extração de *features* das imagens, as matrizes de confusão e as fronteiras de decisão de cada algoritmo de ML aplicado, bem como suas acurácias;
- **Capítulo X:** conclui o estudo proposto explicando as vantagens e as limitações do método estudado, bem como as contribuições que esse traz para a comunidade científica. As considerações desse capítulo também indicam um norte para pesquisadores que desejem continuar nessa área de estudo.

## CAPÍTULO II

### 2 GASOLINA

#### 2.1 COMPOSIÇÃO E PRINCIPAIS PROPRIEDADES

De acordo com GARY *et al* (2007), gasolinas são misturas complexas de hidrocarbonetos que possuem pontos de ebulição que variam de 38 a 205°C, conforme determinado pela *AMERICAN SOCIETY FOR TESTING AND MATERIALS* (ASTM). TAKESHITA *et al* (2008) complementa essa definição afirmando que a composição exata da gasolina depende da natureza do petróleo que lhe deu origem, do processo de obtenção da gasolina e do uso que essa receberá.

Os seus componentes são misturados no intuito de se obter alta qualidade de auto detonação, aquecimento rápido e baixa tendência à formação de depósitos. Ainda de acordo com os autores, as três principais propriedades de uma gasolina que devem ser levadas em conta no que diz respeito à performance do motor são: pressão de vapor de Reid (pressão de vapor a 38°C), intervalo de ebulição e características de *antiknocking*.

#### 2.2 COMBUSTÃO DA GASOLINA

RICHARDS, BARKER (2023) discutem o processo de combustão da gasolina em MCI afirmando que quando o combustível é misturado com o ar, as chamadas reações de pré-chama ocorrem antes mesmo de a mistura alcançar a câmara de combustão, e continuam até que todo o combustível seja completamente consumido pelo frente de chama. A duração desse processo depende de fatores como a composição química da mistura ar-combustível,

bem como a temperatura e a pressão em que o processo ocorre. A combustão pode ser considerada normal ou anormal (pré-ignição, misfire ou knocking).

A combustão normal ocorre quando o frente de chama, após iniciado por uma fonte de ignição positiva (como uma centelha) se move fluida e irregularmente ao longo da câmara de combustão até que o processo esteja concluído. No processo normal, a pressão na câmara começa a aumentar com a compressão da mistura e aumenta rapidamente após a ignição; a pressão máxima ocorre após o pistão ter alcançado o ponto morto superior (PMS) (RICHARDS, BARKER, 2023).

Já nos fenômenos de caráter anormal, um dos fenômenos passíveis de ocorrer é o *spark knock*, ou batida de pino, que corresponde à autoignição descontrolada da mistura ar-combustível antes do frente de chama em função das altas temperatura e pressão. Essa autoignição irá gerar uma nova frente de chama que colidirá com a frente gerada pela centelha, o que gera ondas de choque que tem como consequência um barulho metálico vindo do motor. Batidas de pino prolongadas podem causar perda de potência e danos aos componentes internos do motor, podendo atingir níveis catastróficos. O processo de batida de pino pode ser visualizado na Figura 2.1 (RICHARDS, BARKER, 2023).

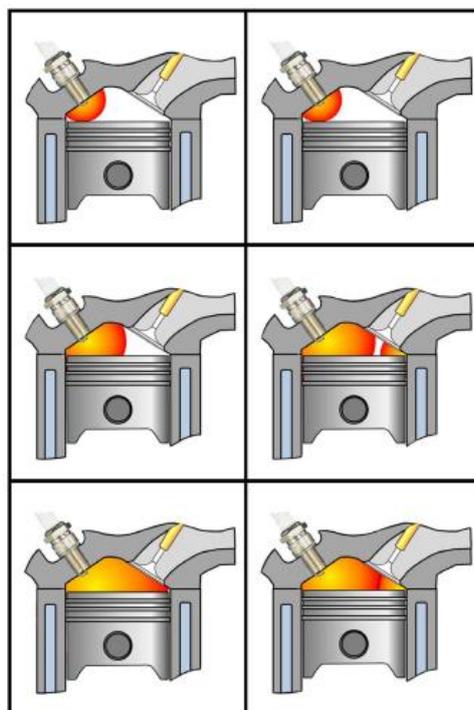


Figura 2.1 – Representação em etapas do processo de “spark knock”, ou batida de pino (RICHARDS, BARKER, 2023).

Além da batida de pino, outro fenômeno que caracteriza a combustão anormal é a pré-ignição, que pode ser definida como a iniciação de uma frente de chama em qualquer ponto na mistura antes da formação da frente oriunda da centelha. Alguns conceitos que se enquadram como pré-ignições são as ignições de superfície e por ponto quente. A comparação entre os processos de combustão descritos pode ser vista na Figura 2.2 (RICHARDS, BARKER, 2023).

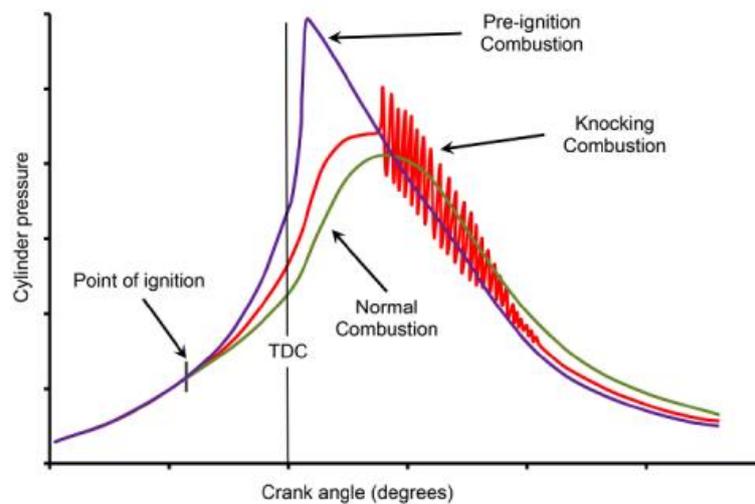


Figura 2.2 - Diagrama pressão no cilindro x ângulo do virabrequim que compara os processos de combustão descritos. *Pre-ignition Combustion* refere-se à condição de pré-ignição, *Knocking Combustion* refere-se à condição em que há batida de pino e *Normal Combustion* refere-se à combustão que ocorre normalmente. TDC significa *Top-dead Center* ou ponto morto superior e *Point of Ignition* corresponde ao ponto em que ocorre a ignição (RICHARDS, BARKER, 2023).

Por fim, a combustão anormal também engloba o fenômeno do *misfire*, que pode ser definido como a falha em promover a ignição da mistura ar-combustível em um ou mais cilindros sem afogar o motor. *Misfire* pode estar associado à qualidade da gasolina, e nesse aspecto tem como principais causas (RICHARDS, BARKER, 2023):

- Razão ar-combustível fora dos limites da combustão, em geral, atribuída a misturas pobres;
- Formação de depósitos nos eletrodos das velas de ignição ou danos causados a esses.

### 2.3 LEGISLAÇÃO BRASILEIRA

As especificações da gasolina de uso automotivo e as obrigações quanto ao controle da qualidade a serem atendidas pelos agentes econômicos que comercializam o produto em território brasileiro são estabelecidas pela RESOLUÇÃO ANP N° 807, DE 31.01.2020. A resolução afirma que as gasolinas automotivas classificam-se em:

- Gasolina A comum: produzida a partir dos processos utilizados em refinarias, nas centrais de matérias-primas e nos formuladores, destinada a veículos automotivos dotados de motores de ignição por centelha, isentos de componentes oxigenados;
- Gasolina A premium: combustível de elevada octanagem, produzido a partir de processos utilizados em refinarias, nas centrais de matérias-primas petroquímicas e nos formuladores, destinada a veículos automotivos dotados de motores de ignição por centelha cujo projeto exija maior octanagem, isento de componentes oxigenados;
- Gasolina C comum: combustível obtido a partir da mistura da gasolina A comum e de etanol anidro combustível, nas proporções definidas pela legislação em vigor;
- Gasolina C Premium: Combustível obtido a partir da mistura da gasolina A premium e de etanol anidro combustível, nas proporções definidas pela legislação em vigor.

Aqui, é válido salientar que apenas os distribuidores de combustível líquido podem realizar a adição de etanol à gasolina tipo A para formação do tipo C.

Apesar de a ANP estabelecer essas diretrizes, a definição do percentual de etanol permitido na gasolina é competência do Ministério da Agricultura e Pecuária (MAPA), que especificou na PORTARIA MAPA N° 75/2015, os valores máximos são de 25% de etanol (em volume) para a gasolina C premium e 27% (em volume) para a gasolina C comum. A venda de gasolina fora dessas especificações constitui adulteração, crime em que há a intenção de se obter benefícios financeiros com a comercialização de produto modificado por solventes ou contaminantes em geral. A determinação do teor de etanol na gasolina deve ocorrer, de acordo com a ANP, através do processo descrito pela ABNT NBR 13992.

## CAPÍTULO III

### 3 MOTORES A COMBUSTÃO INTERNA

#### 3.1 CLASSIFICAÇÃO E ASPECTOS CONSTITUTIVOS E FUNCIONAIS

HEYWOOD (1988) afirma que há diversos modos de se classificar motores de combustão interna:

- **APLICAÇÃO:** motores de automóveis, de caminhões, de locomotivas, de aeronaves, de geração de potência, dentre outros.
- **DESIGN BÁSICO:** motores alternativos ou rotativos.
- **CICLO DE FUNCIONAMENTO:** motores de quatro tempos, que podem ser naturalmente aspirados, superalimentados e turboalimentados, e motores de dois tempos, que podem ser varridos pelo carter, superalimentados ou turboalimentados.
- **LOCALIZAÇÃO DAS VÁLVULAS:** válvulas no cabeçote (ou válvulas *I-head*), válvulas sob o cabeçote (ou válvulas *L-head*), válvulas rotativas, portas de varredura cruzada (portas de admissão e escape em lados opostos do cilindro em uma extremidade), portas de varredura em laço (portas de admissão e escape no mesmo lado do cilindro em uma extremidade), varredura unidirecional ou em fluxo direto (portas ou válvulas de admissão e escape em extremidades opostas do cilindro).
- **COMBUSTÍVEL:** motores a gasolina, a óleo diesel, a gás natural, a álcool, a hidrogênio ou de duplo combustível.
- **MÉTODO DE PREPARAÇÃO DA MISTURA:** motores carburados ou com injeção eletrônica.

- MÉTODO DE IGNIÇÃO: motores de ignição por centelha ou de ignição por compressão.
- DESIGN DA CÂMARA DE COMBUSTÃO: motores de câmara aberta ou dividida.
- MÉTODO DO CONTROLE DE CARGA: motores que utilizam limitadores de fluxo de ar e de combustível (para que a mistura permaneça a mesma), motores que controlam apenas combustível ou uma combinação desses métodos.
- MÉTODO DE ARREFECIMENTO: motores arrefecidos a água, a ar ou por convecção e radiação naturais.

O autor elabora a respeito dos motores alternativos, ou seja, aqueles em que os pistões movem-se num movimento de vaivém num cilindro e nos quais a transmissão de potência ocorre através de um mecanismo pistão-manivela para o eixo virabrequim. O pistão fica em repouso em duas posições: os pontos mortos superior (PMS) e inferior (PMI), nos quais o volume do cilindro é mínimo e máximo respectivamente. O volume mínimo é chamado de volume de folga; já a diferença entre o volume máximo e o volume de folga é chamado de volume deslocado (HEYWOOD, 1988).

A maioria dos motores convencionais de veículos opera um ciclo de quatro tempos, o que significa que cada cilindro requer quatro etapas diferentes para que ocorra a geração de potência; essas quatro etapas ocorrem em duas rotações do eixo virabrequim. O ciclo mencionado pode ser visualizado na figura 3.1 (HEYWOOD, 1988).

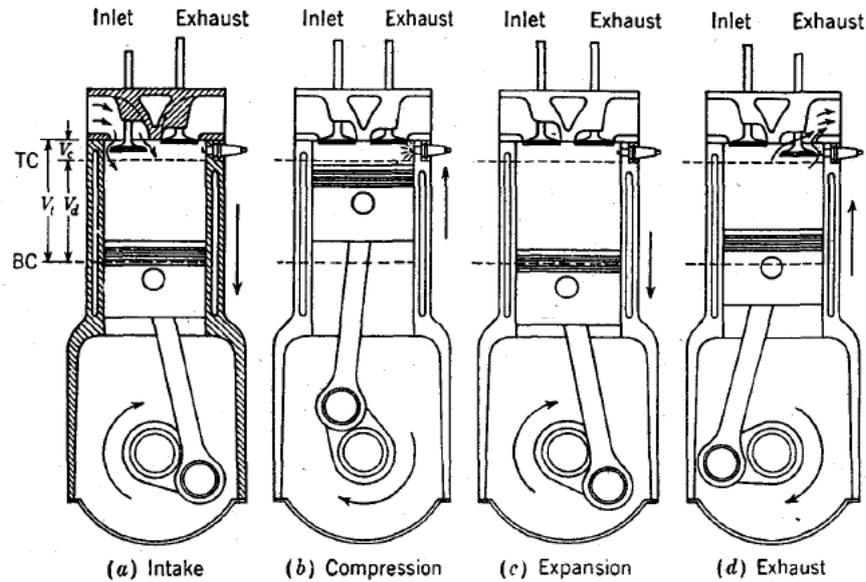


Figura 3.1 - Os quatro tempos de um motor. *Intake* se refere à admissão, *Compression* se refere à compressão, *Expansion* se refere à expansão e *Exhaust* se refere à exaustão (HEYWOOD, 1988).

- **ADMISSÃO:** Inicia-se com o pistão no PMS e termina com ele no PMI. Nessa etapa, ocorre a entrada do ar na câmara de combustão através da válvula de admissão, cuja abertura ocorre antes do início dessa etapa e o fechamento ocorre logo após seu término.
- **COMPRESSÃO:** ocorre quando ambas as válvulas estão fechadas e resulta na diminuição do volume inicial. No final dessa etapa, o mecanismo que originará a combustão é utilizado.
- **EXPLOSÃO:** ou expansão, corresponde ao processo em que o pistão inicia no PMS e termina no PMI, visto que os gases a altas temperatura e pressão oriundas da combustão empurram o pistão para baixo e forçam a rotação do virabrequim.
- **EXAUSTÃO:** os gases remanescentes saem do cilindro pela válvula de exaustão em função da diferença de pressão dentro e fora do cilindro e porque nessa etapa ocorre movimentação do pistão no sentido do PMS. Assim que o pistão alcança esse ponto, a válvula de exaustão se fecha e o ciclo reinicia.

De um ponto de vista termodinâmico, VAN WYLEM *et al* (1988) o ciclo teórico que se aproxima do funcionamento dos motores a combustão é o ciclo padrão a ar Otto. Os diagramas p-v e T-s para esse ciclo estão expostos na Figura 3.2.

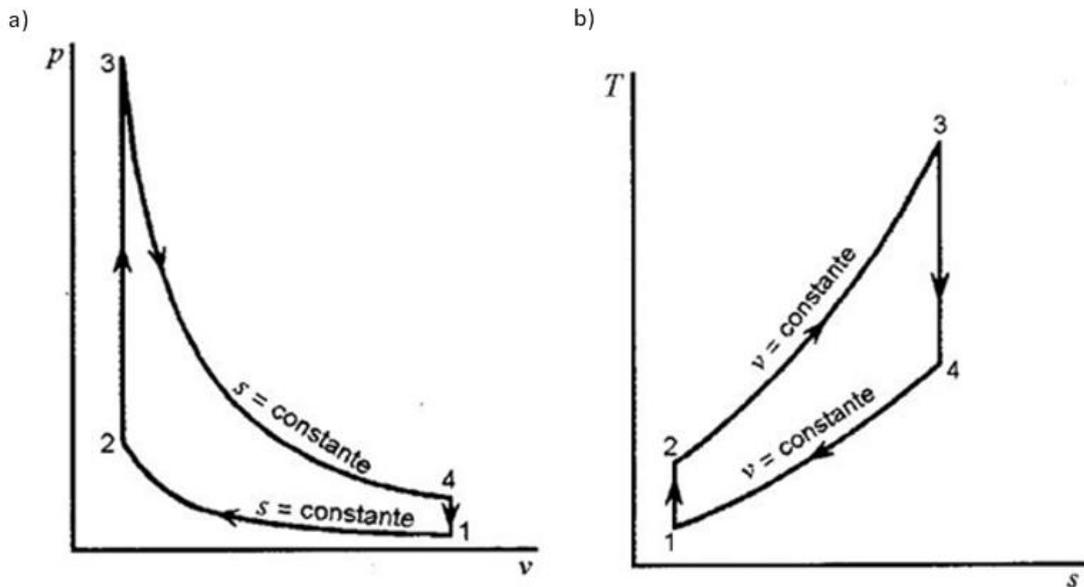


Figura 3.2 – a) Diagrama pressão x volume para ciclo padrão a ar Otto; b) Diagrama temperatura x entropia para o ciclo padrão a ar Otto. Adaptado de FERRARI e MARIN (2018).

- PROCESSO 1-2: Corresponde a uma compressão isentrópica, quando o pistão se move do PMI ao PMS;
- PROCESSO 2-3: corresponde a uma transferência de calor para o ar a volume constante enquanto o pistão está momentaneamente em repouso no PMS;
- PROCESSO 3-4: corresponde a uma expansão isentrópica, quando o pistão se move do PMS ao PMI;
- PROCESSO 4-1: rejeição de calor ao ar, enquanto o pistão está no ponto morto inferior.

É válido mencionar que o rendimento do ciclo-padrão Otto é função apenas da relação de compressão, sendo essas duas grandezas diretamente proporcionais; isso também é verdade para um motor real de ignição por centelha. Entretanto, num motor real, quando se aumenta a relação de compressão, ocorre o aumento da tendência de detonação do combustível, logo, é necessário que os combustíveis utilizados tenham bons valores de octanagem para que se possa aumentar a relação de compressão (VAN WYLEM *et al*, 1988).

Para traçar a relação entre o ciclo Otto e o processo real que acontece dentro dos cilindros do motor, é necessário incrementar o gráfico com um novo processo

correspondente à admissão do ar e considerar que o “tempo” associado à explosão acontece em duas etapas do ciclo termodinâmico. A relação descrita pode ser vista na Figura 3.3 (VAN WYLEM *et al*, 1988).

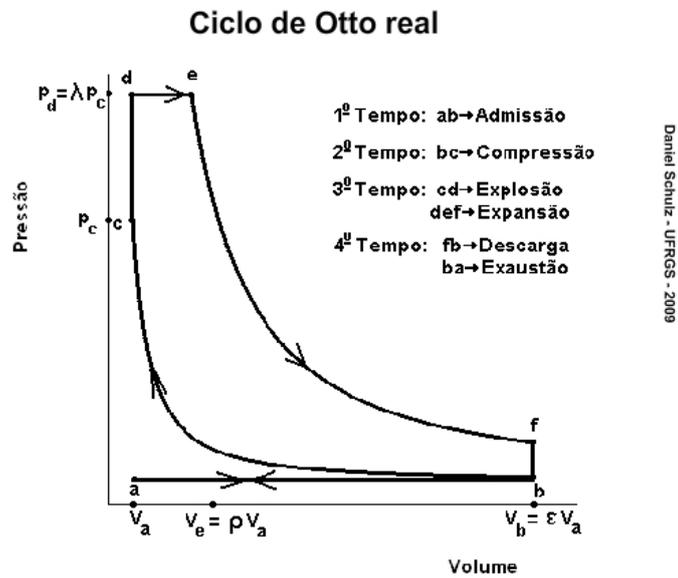


Figura 3.3 - Diagrama pressão x volume de um ciclo Otto real (DA SILVEIRA, 2008).

## **CAPÍTULO IV**

### **4 PROCESSAMENTO DE IMAGENS**

#### **4.1 LUZ E O ESPECTRO ELETROMAGNÉTICO**

Em 1666 Isaac Newton lançou um feixe de luz solar contra um prisma e o resultado desse experimento foi a decomposição dessa luz em um espectro de cores que iniciavam no violeta e terminavam no vermelho. Hoje já é sabido que essas cores representam apenas uma fração do espectro eletromagnético, chamada de espectro visível, pois é a única parte daquele que pode ser observada por seres humanos a olho nu. Entretanto, o espectro se estende, com comprimentos de onda muito maiores e muito menores que aqueles presentes na porção visível do espectro, que pode ser visto na Figura 4.1 (GONZALES e WOODS, 2018).

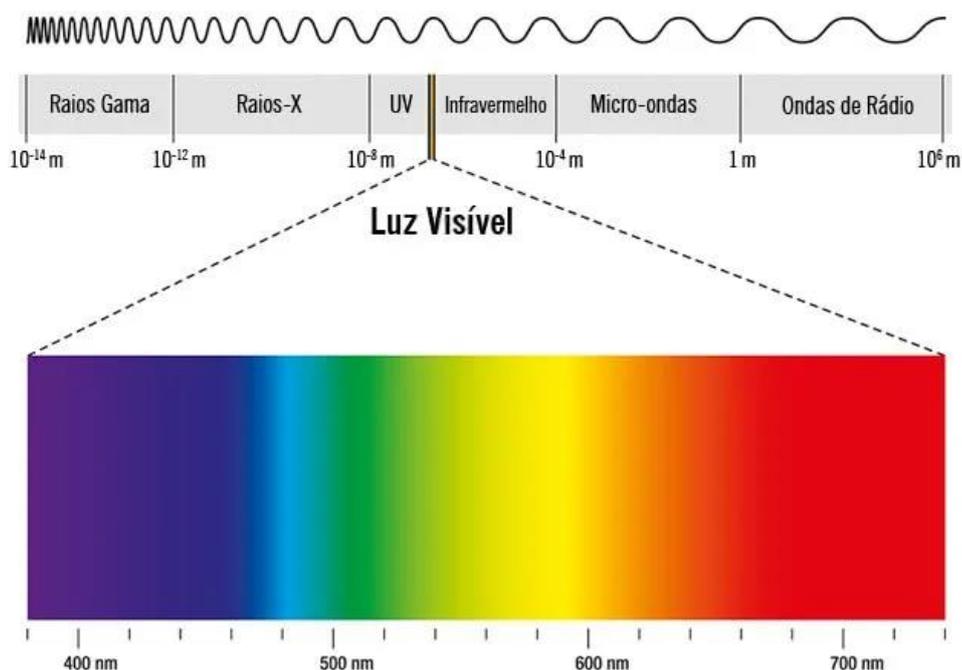


Figura 4.1 – Espectro eletromagnético com destaque para o intervalo de comprimentos de onda visíveis por humanos a olho nu (<https://www.todamateria.com.br/espectro-eletromagnetico/>).

Para entender melhor a natureza do espectro, HALLIDAY *et al* (2016) nos afirmam que James Clerk Maxwell demonstrou que um raio luminoso nada mais é do que uma propagação, no espaço, de campos elétricos e magnéticos, sendo, portanto uma onda eletromagnética. Sendo assim, o que Newton observou ao realizar seu experimento foi a decomposição do feixe de luz em seus diversos comprimentos de onda.

Uma onda eletromagnética pode ser visualizada como ondas senoidais que se propagam com um comprimento de onda  $\lambda$  que HALLIDAY *et al* (2016) definem como a distância entre repetições de forma da onda. Contudo, a luz também pode ser visualizada, em função da sua natureza dual, como um fluxo de partículas sem massa que se movimentam como uma onda na velocidade da luz. Cada uma dessas partículas contém uma certa quantidade de energia que é chamada de fóton.

As componentes do espectro eletromagnético podem ser expressas em função do seu comprimento de onda ( $\lambda$ ), da sua frequência ( $\nu$ ) ou da sua energia ( $E$ ). Essas grandezas se

relacionam através das Equações 4.1 e 4.2. Aqui,  $c$  corresponde à velocidade da luz,  $2,998 \times 10^8(m/s)$  e  $h$  corresponde à Constante de Planck (GONZALES e WOODS, 2018).

$$\lambda = \frac{c}{\nu} \quad (4.1)$$

$$E = h\nu \quad (4.2)$$

Vê-se que a energia é diretamente proporcional à frequência logo, quanto maior a frequência, maior a energia carregada pela componente eletromagnética. Por isso, ondas de rádio possuem menos energia que ondas infravermelhas; as ondas do espectro que possuem maior energia são os raios gama, o que as torna tão perigosas para os organismos vivos (GONZALES e WOODS, 2018).

A luz é um tipo particular de radiação eletromagnética que pode ser vista pelo olho humano, com comprimentos de onda que variam de  $0,43\mu m$  (violeta) até  $0,79\mu m$  (vermelho). As cores que os seres humanos veem são determinadas por quais comprimentos de onda são refletidos pelos objetos, logo, um objeto verde reflete comprimentos de onda verde e absorve os outros componentes, e um objeto que aparenta ser branco reflete todos os componentes eletromagnéticos de forma balanceada. Já a luz que não possui cores é chamada de acromática e seu único atributo é a intensidade, que geralmente é denominada nível de cinza, pois ela varia de preto a cinzas e finalmente branco (GONZALES e WOODS, 2018).

## 4.2 IMAGENS DIGITAIS

De acordo com GONZALES e WOODS (2018), uma imagem pode ser definida como uma função bidimensional  $f(x,y)$  em que  $x$  e  $y$  são coordenadas espaciais, ou seja, coordenadas no plano e a amplitude  $f$  de um par de coordenadas  $(x,y)$  é a intensidade ou nível de cinza da imagem no ponto considerado. Quando  $x,y$  e os valores de  $f$  são finitos, dizemos que a imagem é digital. Logo, uma imagem digital é composta por um número finito de elementos, cada um com sua localização e valor de intensidade. O processamento de imagens

digitais nada mais é o conjunto de operações realizadas nessas imagens por computadores digitais.

Quando uma imagem é gerada por um processo físico, seus valores são proporcionais à energia irradiada pela fonte física, logo  $f(x,y)$  deve ser um valor não nulo e finito para todos os pontos da imagem. Essa função será caracterizada por dois componentes: a quantidade de iluminação que incide na cena visualizada e a quantidade de iluminação refletida pelos objetos na cena, que são denotados por  $i(x,y)$  e  $r(x,y)$  respectivamente (GONZALES e WOODS, 2018).  $f(x,y)$ ,  $i(x,y)$  e  $r(x,y)$  são iguais às equações 4.3, 4.4 e 4.5 respectivamente.

$$f(x, y) = i(x, y)r(x, y) \quad (4.3)$$

Em que são válidas as Equações 4.4 e 4.5.

$$0 < i(x, y) < \infty \quad (4.4)$$

$$0 < r(x, y) < 1 \quad (4.5)$$

Essas equações mostram que a intensidade de uma imagem monocromática em qualquer coordenada deve estar dentro de uma faixa de valores finitos  $[L_{min}, L_{max}]$  chamada de escala de cinza. É prática comum representar essa escala como  $[0, L-1]$ , em que 0 corresponde ao preto,  $L-1$  ao branco e todos os valores intermediários a tonalidades de cinza (GONZALES e WOODS, 2018).

#### 4.2.1 Amostragem e Quantização

A saída de sensores que captam imagens é um sinal de voltagem contínuo, logo, para criar a imagem digital a partir desse sinal é necessário realizar dois processos: a amostragem e quantização, que se referem respectivamente aos processos de digitalizar as coordenadas e as amplitudes, ou seja, transformá-las em quantidades discretas (GONZALES e WOODS, 2018).

O resultado desse processo é uma matriz de números reais; uma imagem  $f(x,y)$  torna-se uma matriz de  $M$  linhas e  $N$  colunas com os valores relativos às coordenadas  $(x,y)$  sendo

valores discretos, como pode ser visto na Equação 4.6, que é a definição de imagem digital, em que cada elemento é chamado de pixel (GONZALES e WOODS, 2018).

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, N - 1) \\ f(1,0) & f(1,1) & \cdots & f(1, N - 1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M - 1,0) & f(M - 1,1) & \cdots & f(M - 1, N - 1) \end{bmatrix} \quad (4.6)$$

De modo formal, seja  $Z$  e  $R$  os conjuntos dos números inteiros e reais respectivamente. Diz-se que  $f(x,y)$  é uma imagem digital se  $(x,y)$  forem inteiros pertencentes a  $Z^2$  e se  $f$  for uma função que associa níveis de cinza, ou seja, valores reais, a cada par de elementos  $(x,y)$  (GONZALES e WOODS, 2018).

Vê-se que não há restrição aos valores de  $M$  e  $N$ , exceto que eles sejam inteiros positivos. Já o número  $L$  de níveis de cinza, em função de questões associadas a armazenamento e de hardware, é geralmente considerada como uma potência de 2, conforme a Equação 4.7 (GONZALES e WOODS, 2018).

$$L = 2^k \quad (4.7)$$

Logo, o número  $b$  de bits necessários para armazenar uma imagem digital é dado pela Equação 4.8.

$$b = M \times N \times k \quad (4.8)$$

Se uma imagem possui  $2^k$  níveis de cinza, diz-se que essa imagem é uma imagem de  $k$  bits. Por exemplo, uma imagem com 256 níveis de cinza será uma imagem de 8 bits (GONZALES e WOODS, 2018).

### 4.3 APRIMORAMENTO DE IMAGENS

Aprimoramento de imagem corresponde à série de procedimentos utilizados para processar a imagem de modo que ela se torne mais adequada para uma situação específica. Os métodos de aprimoramento de imagem se enquadram em duas categorias: de domínio espacial e de domínio da frequência. A primeira se refere ao plano da imagem em si, e aqui

as abordagens todas se baseiam na manipulação dos pixels diretamente. Já a segunda se refere a técnicas de processamento que se baseiam na modificação da Transformada de Fourier da imagem (GONZALES e WOODS, 2018). Processos do domínio espacial são denotados pela Equação 4.15:

$$g(x, y) = T[f(x, y)] \quad (4.15)$$

Em que  $f(x, y)$  é a imagem de entrada, ou seja, antes da transformação,  $g(x, y)$  é a imagem já processada e  $T$  é um operador em  $f$  definido nas vizinhanças de  $(x, y)$ . A principal abordagem para definir a vizinhança de um ponto  $(x, y)$  é utilizar um quadrado como uma sub imagem centrado em  $(x, y)$  como pode ser visto na Figura 4.2 essa subimagem se movimenta ao longo dos pixels sendo aplicado em cada localização  $(x, y)$ . A forma mais simples de  $T$  é uma região  $1 \times 1$ , caso no qual  $g$  depende apenas do valor de  $f$  em  $(x, y)$ . Nessa situação,  $T$  é considerado uma transformação (mapeamento) de escala de cinza (GONZALES e WOODS, 2018).

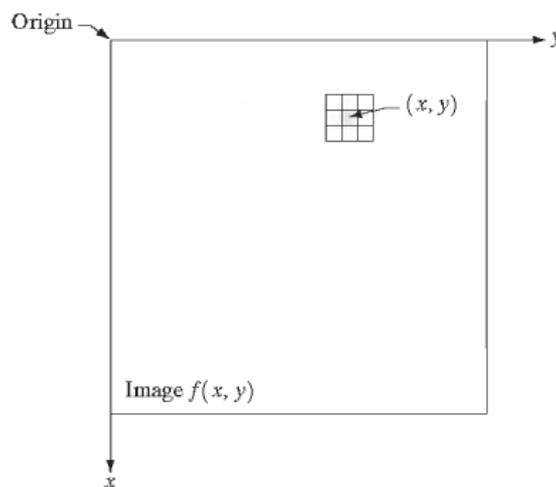


Figura 4.2 – Operador de formato quadrado comumente utilizado em operações de aprimoramento de imagem centrado no ponto  $(x, y)$  (GONZALES e WOODS, 2018).

Vizinhanças maiores permitem maior flexibilidade; a técnica padrão é utilizar uma função dos valores de  $f$  numa vizinhança predefinida de  $(x, y)$ . Uma das abordagens principais é utilizar uma formulação baseada em máscaras, também chamadas de filtros, *kernels* ou janelas, que correspondem a matrizes pequenas, cujos coeficientes determinam a

natureza do processo. Operações de aprimoramento que utilizam essa técnica são comumente chamadas de processo de filtragem (GONZALES e WOODS, 2018).

### 4.3.1 Filtros no Domínio Espacial

O processo de filtragem corresponde ao ato de percorrer a máscara (filtro) ao longo dos pixels da imagem de modo que em cada ponto  $(x, y)$ , uma resposta é calculada utilizando uma relação predefinida. No caso de filtragem espacial linear, essa resposta é dada pela soma dos produtos dos coeficientes do filtro pelos pontos correspondentes na imagem na área abrangida pelo filtro, conforme ilustra a Figura 4.3 (GONZALES e WOODS, 2018).

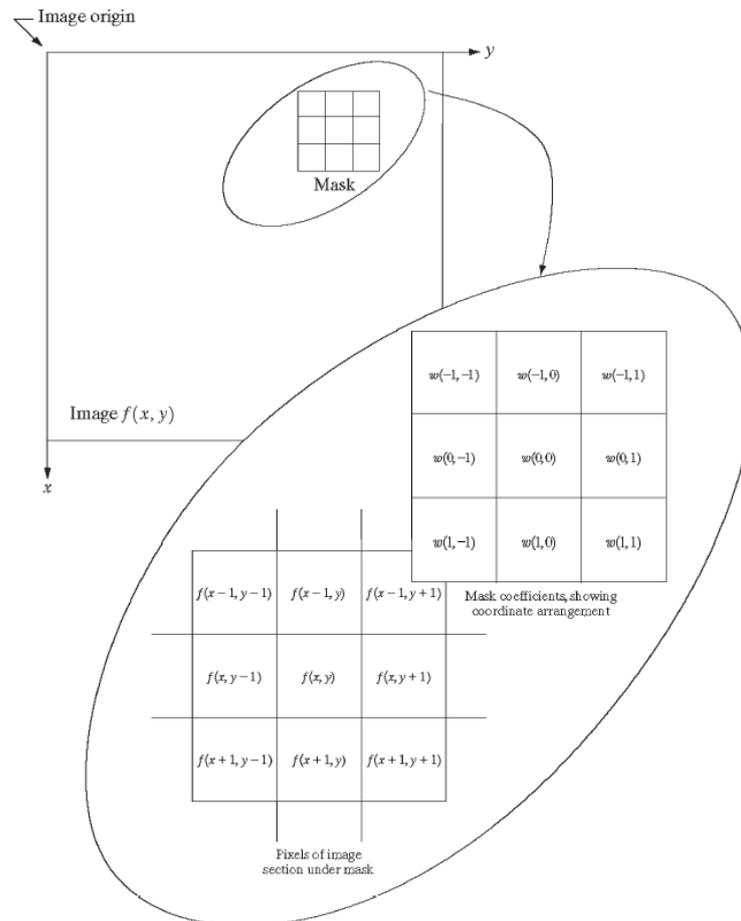


Figura 4.3 – Detalhamento de uma máscara/filtro utilizada para aprimoramento de imagens (GONZALES e WOODS, 2018).

Considerando o filtro 3x3 utilizado na Figura 4.3, a resposta  $R$  da máscara ao ponto  $(x, y)$  utilizado como exemplo é dada pela Equação 4.16:

$$R = w(-1, -1)f(x - 1, y - 1) + w(-1, 0)f(x - 1, y) + \dots + w(0, 0)f(x, y) \\ + \dots + w(1, 0)f(x + 1, y) + w(1, 1)f(x + 1, y + 1) \quad (4.16)$$

De modo geral, o processo de filtragem linear de uma imagem  $f$  de dimensões  $M \times N$  com um filtro de dimensões  $m \times n$  é dado pela Equação 4.17 em que  $a = (m - 1)/2$  e  $b = (n - 1)/2$ :

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t)f(x + s, y + t) \quad (4.17)$$

Quando apenas a resposta  $R$  de um filtro de dimensões  $m \times n$  em um ponto  $(x, y)$  é relevante, é comum simplificar a Equação 4.16 usando a notação exibida na Equação 4.18, em que os  $w$ 's correspondem aos coeficientes da máscara e os  $z$ 's correspondem aos valores, em escala de cinza, dos pixels correspondentes na imagem (GONZALES e WOODS, 2018):

$$R = w_1z_1 + w_2z_2 + \dots + w_{mn}z_{mn} = \sum_{i=1}^{mn} w_i z_i \quad (4.18)$$

Filtragem não-linear também opera utilizando o conceito de vizinhança de um pixel, e a mecânica de convolução do filtro pela imagem é a mesma; entretanto, nesses casos a operação de filtragem é baseada em condições relativas aos valores dos pixels na vizinhança considerada e as operações utilizadas não são necessariamente as de soma e produto como mostradas na Equação 4.17 (GONZALES e WOODS, 2018).

No caso de o centro do filtro se aproximar muito das bordas da imagem de modo que parte da máscara encontre-se fora de seus limites, há soluções que podem ser implementadas (GONZALES e WOODS, 2018):

- Limitar o deslocamento do filtro de tamanho  $n \times n$  para uma distância maior ou igual a  $(n - 1)/2$  pixels da borda. Parte da imagem será perdida com esse método, mas a imagem resultante terá apenas pixels que foram devidamente filtrados pela máscara aplicada;

- Filtrar apenas os pixels para os quais a máscara está totalmente contida na imagem. Com esse método, haverá pixels não filtrados próximos às bordas, mas a imagem manterá seu tamanho original;
- Adicionar colunas ou linhas, a depender da necessidade, compostas inteiramente por zeros, ou repetir colunas e linhas; esses acréscimos são excluídos da imagem após a filtragem. Esse método mantém o tamanho original da imagem e consegue filtrar todos os pixels, entretanto, os valores próximos às margens serão artificiais.

### 4.3.2 Filtros de Suavização

Filtros de suavização são usados para desfoque e redução de ruídos. Geralmente são usados em etapas de pré-processamento anteriores à extração de objetos e em operações de conexão de pontos e cuvas. Esses filtros podem ser lineares e não-lineares (GONZALES e WOODS, 2018).

A resposta de um filtro linear de suavização é a média da intensidade de cada pixel contido na vizinhança definida pela máscara utilizada, logo, são frequentemente chamados de filtros de média. A substituição dos valores pela média das intensidades das vizinhanças resulta em transições menos abruptas, que normalmente são os componentes de ruídos aleatórios, facilmente eliminados por esse tipo de filtragem. Entretanto, as bordas das imagens também são caracterizadas por essas transições abruptas e por consequência, o uso desse filtro pode resultar em contornos desfocados. É possível também aplicar esses filtros para a eliminação de detalhes irrelevantes nas imagens (GONZALES e WOODS, 2018). Uma máscara de dimensões  $3 \times 3$  em que todos os coeficientes têm valores iguais pode ser representada pela Equação 4.19:

$$R = \frac{1}{9} \sum_{i=1}^9 z_i \quad (4.19)$$

Contudo, a utilização de filtros com coeficientes diferentes, o que corresponderia a uma média ponderada, é perfeitamente plausível. A estratégia de diminuição desses coeficientes à medida que se distanciam do centro do filtro é inclusive uma maneira de reduzir o desfoque oriundo da aplicação de filtragem. Um procedimento de filtragem genérico com média

ponderada aplicado a uma imagem  $M \times N$  com um filtro de dimensões  $m \times n$  é dado pela expressão 4.20. Os efeitos dessa modalidade de filtragem em uma imagem podem ser vistos na Figura 4.4 (GONZALES e WOODS, 2018).

$$g(x, y) = \frac{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t)}{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t)} \quad (4.20)$$

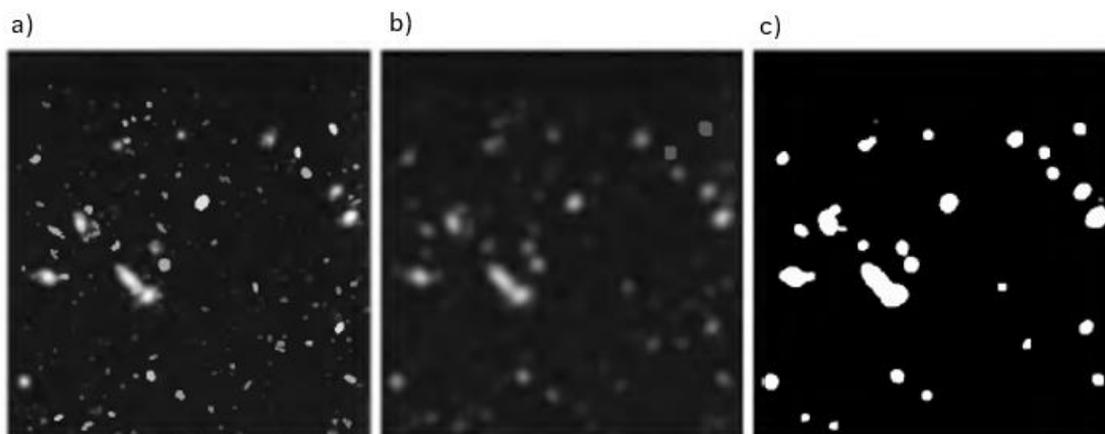


Figura 4.4 – Exemplo de aplicação de filtro de suavização. a) Imagem original; b) aplicação do filtro; c) aplicação de um método de thresholding na imagem filtrada. Adaptado de GONZALES e WOODS (2018).

#### 4.4 HISTOGRAMAS DE IMAGENS DIGITAIS

O histograma de uma imagem digital com escala de cinza de níveis no intervalo de  $[0, L - 1]$  é uma função discreta  $h(r_k) = n_k$  em que  $k$  corresponde ao  $k$ -ésimo nível de cinza e  $n_k$  é o número de pixels na imagem que tem o nível de cinza  $r_k$ . É uma prática comum normalizar o histograma através da divisão desse pelo número total de pixels na imagem, denotado por  $n$ . O histograma normalizado pode ser representado por  $p(r_k) = n_k/n$  para  $k = 0, 1, 2, \dots, L - 1$ . É possível interpretar o histograma normalizado como uma distribuição da probabilidade da ocorrência dos níveis de cinza  $r_k$  (GONZALES e WOODS, 2018). Um exemplo de histograma pode ser visto na Figura 4.5.

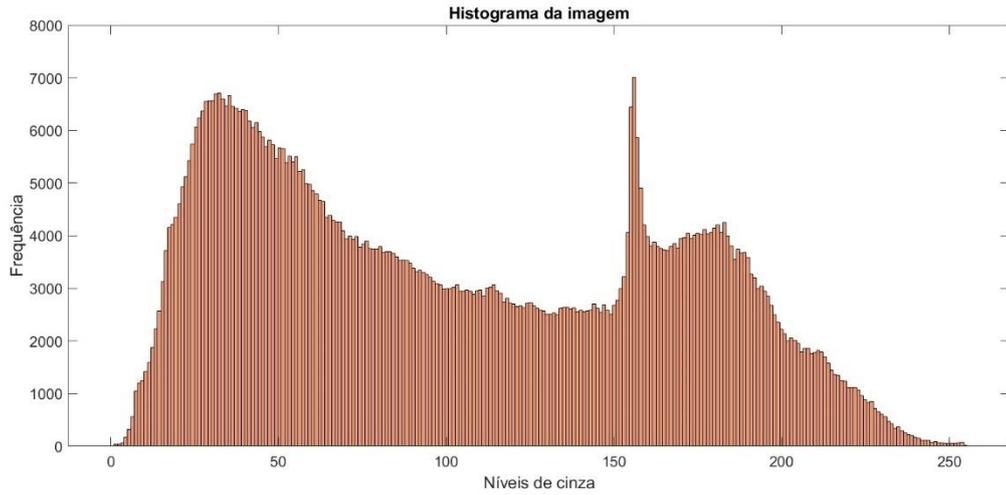


Figura 4.5 – Histograma relativo à imagem 4.7 a).

Histogramas podem fornecer parâmetros estatísticos muito úteis para o processamento e análise de imagens. Seja  $r$  uma variável aleatória discreta que representa os níveis de cinza discretos num intervalo de  $[0, L - 1]$  e seja  $p(r_i)$  a componente correspondente ao  $i$ -ésimo valor de  $r$  do histograma normalizado, que pode ser interpretada como a probabilidade de ocorrência do nível de cinza  $r_i$ . Logo, o  $n$ -ésimo momento de  $r$  com relação à sua média é definido pela Equação 4.23 (GONZALES e WOODS, 2018):

$$\mu_n(r) = \sum_{i=0}^{L-1} (r_i - m)^n p(r_i) \quad (4.23)$$

A média, aqui denotada por  $m$ , pode ser descrita pela Equação 4.24:

$$m = \sum_{i=0}^{L-1} r_i p(r_i) \quad (4.24)$$

A partir disso, é possível definir que  $\mu_0 = 1$  e  $\mu_1 = 0$ . Já o segundo momento, que corresponde à variância das intensidades dos pixels, é dado pela Equação 4.25:

$$\mu_2(r) = \sum_{i=0}^{L-1} (r_i - m)^2 p(r_i) \quad (4.25)$$

É válido também debater os conceitos de assimetria e curtose, que correspondem respectivamente ao terceiro e quarto momentos, dados pelas Equações 4.26 e 4.27 (BRUCE e BRUCE, 2019).

$$\mu_3(r) = \sum_{i=0}^{L-1} (r_i - m)^3 p(r_i) \quad (4.26)$$

$$\mu_4(r) = \sum_{i=0}^{L-1} (r_i - m)^4 p(r_i) \quad (4.27)$$

O conceito de assimetria está associado ao fato de uma variável estar ligada a valores maiores ou menores em uma distribuição. Já a curtose indica a propensão dos dados a terem valores extremos; visualmente, pode-se pensar nisso como a tendência de uma curva ser ou não achatada (BRUCE e BRUCE, 2019). Uma curva simétrica possui média e moda iguais, já uma curva que tem assimetria negativa possui moda maior que a média e uma curva com assimetria positiva possui média maior que a moda (DE OLIVEIRA, 1999). Assimetria negativa também pode ser chamada de assimetria à esquerda e assimetria positiva pode ser chamada de assimetria à direita, como mostra a Figura 4.6.

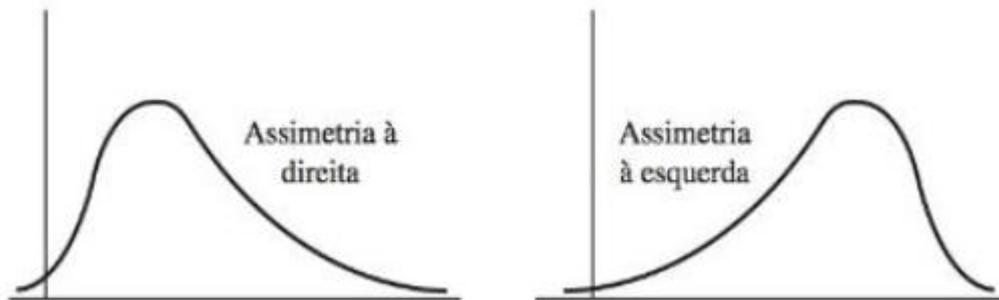


Figura 4.6 – Exemplos do comportamento de distribuições de probabilidade com assimetria à direita e à esquerda (SPIEGEL *et al*, 2013).

Distribuições que tenham alta curtose são aquelas que apresentam picos acentuados, com seus valores concentrados próximos da média, tal qual a distribuição exibida na Figura 4.7.

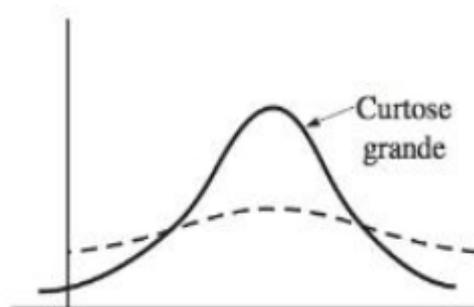


Figura 4.7 – Exemplo de distribuições de probabilidade com alta curtose (SPIEGEL *et al*, 2013).

#### 4.5 CONVERSÃO DE ESCALA DE CORES PARA ESCALA DE CINZA

No olho humano, estruturas chamadas de cones são os responsáveis pela visão das cores. Em geral, um olho humano possui cerca de 6 a 7 milhões de cones e foi descoberto que desses, 65% são sensíveis à cor vermelha, 33% à cor verde e 2% à cor azul. Em função disso, todas as cores visíveis ao ser humano são concebidas como combinações dessas cores primárias em diferentes proporções: vermelho, verde e azul, em inglês, red, green e blue, cujas iniciais são RGB, nome do sistema de classificação de cores definido pela CIE (Comission Internationale de l’Eclairage – Comissão Internacional de Iluminação). Os comprimentos de onda definidos como vermelho, verde e azul são respectivamente 445 nm, 535nm e 575nm; entretanto, para gerar todas as cores, é necessário considerar que os cones podem absorver variações nesses comprimentos de onda, visto que os três sozinhos não conseguiriam gerar todas as cores, como mostra a Figura 4.8 (GONZALES e WOODS, 2018).

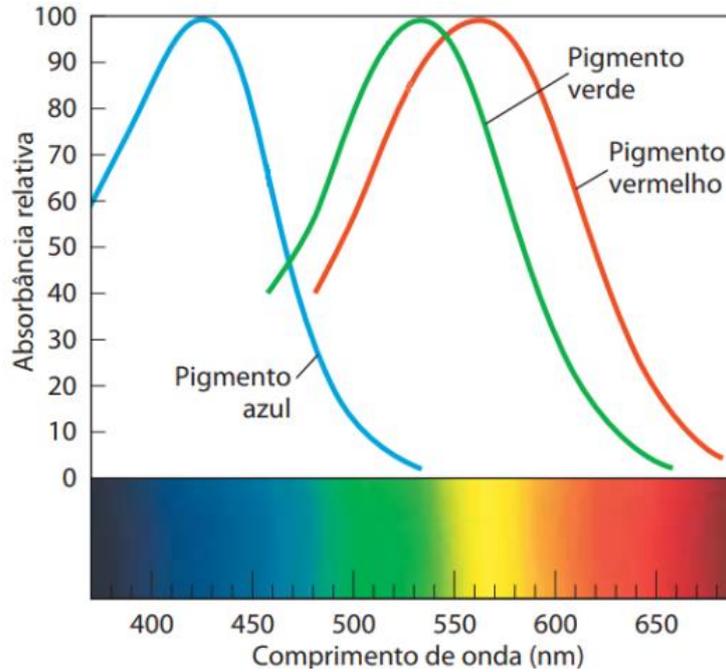


Figura 4.8 – espectro de absorção dos receptores para vermelho, verde e azul dos cones. Adaptado de NELSON e COX (2014).

É necessário diferenciar as cores primárias no âmbito da luz e no âmbito dos pigmentos. No segundo caso, cores primárias podem ser definidas como aquelas que subtraem ou absorvem uma cor primária da luz e reflete ou transmite as outras duas. Assim, as cores primárias nos pigmentos são magenta, ciano e amarelo, e as secundárias são vermelho, verde e azul (GONZALES e WOODS, 2018).

As características utilizadas para distinguir cores são:

- **Brilho:** corresponde à noção cromática de intensidade, de modo que uma cor menos brilhosa parecerá mais escura e uma mais brilhosa parecerá mais clara;
- **Matiz:** é um atributo associado ao comprimento de onda dominante numa mistura de ondas de luz, representa a cor dominante na percepção do observador;
- **Saturação:** representa a pureza relativa ou quantidade de luz branca misturada à matiz, sendo inversamente proporcional a essa. Cores saturadas aparentam ser mais vividas e cores menos saturadas têm uma aparência desbotada.

Como já citado, no modelo RGB, cada cor é representada pelos seus componentes espectrais de vermelho, verde e azul; esse modelo é baseado num sistema de coordenadas cartesianas. Nesse sistema, as cores estão localizadas no espaço definido por um cubo de

arestas de comprimento igual a 1 em que vermelho, verde e azul ocupam 3 dos vértices, magenta, ciano e amarelo ocupam mais três vértices e preto e branco ocupam os outros dois vértices (GONZALES e WOODS, 2018), tal qual a Figura 4.9.

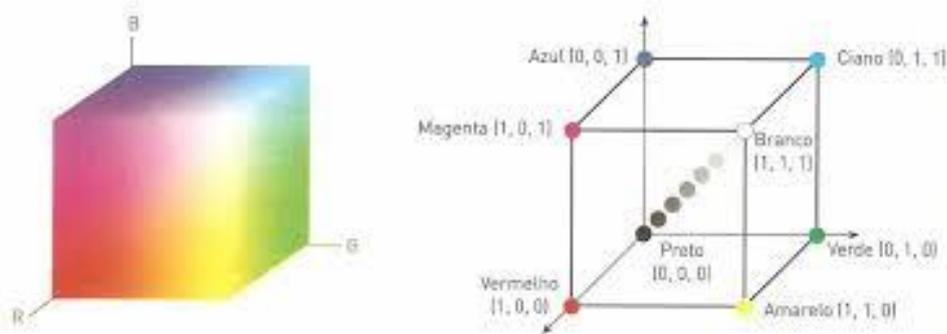


Figura 4.9 – Diagrama ilustrativo da relação entre as escalas de cinza e RGB (<https://goo.gl/bebjMD>).

Imagens representadas nesse sistema constituem-se de três componentes que, quando combinados num monitor RGB, são combinados e produzem a imagem colorida desejada. Considerando uma imagem em que cada um dos três planos de cores principais é uma imagem de 8-bits, diz-se que a profundidade de pixel dessa imagem é igual ao número de planos multiplicado pelo número de bits de cada plano, nesse caso 24. O número total de cores que podem estar presentes em uma imagem de 24-bits é  $(2^8)^3 = 16.777.216$ . Apesar disso, vários sistemas comumente desempenham sua função de modo satisfatório exibindo apenas 256 cores (GONZALES e WOODS, 2018).

De acordo com a imagem 4.12, para identificar a intensidade de uma cor, em outras palavras seu nível de cinza, bastaria traçar um plano perpendicular ao eixo preto-branco correspondente à diagonal do cubo e que passasse pelo ponto que define a cor em questão; essa abordagem está correta, mas não é a mais comumente utilizada. Para entender melhor como essa conversão funciona, é necessário conhecer o modelo YQI para representação de cores (GONZALES e WOODS, 2018).

O modelo Y (luminância/intensidade), I (em-fase), Q (quadratura) de representação de cores foi desenvolvido para televisão pelo NTSC (National Television System Committee) em 1953. Nesse sistema, a intensidade é dissociada da cor, o que possibilitou a transmissão de cores em televisores que exibiam apenas imagens em preto e branco (YAGLE e ULABY, 2018). Nesse modelo, a representação de uma cor é dada pela Equação 4.28:

$$f(x, y) = \{f_Y(x, y), f_I(x, y), f_Q(x, y)\} \quad (4.28)$$

Aqui, o componente de luminância  $f_Y(x, y)$  corresponde à intensidade de  $f(x, y)$  em  $(x, y)$ . Os outros dois componentes  $\{f_I(x, y), f_Q(x, y)\}$  correspondem à crominância, que especifica a cor  $f(x, y)$  em  $(x, y)$  utilizando a Figura 4.10 em que o eixo horizontal, que vai do azul ao laranja, corresponde às cores às quais o olho humano é mais sensível. Já o eixo vertical, que vai do roxo ao verde, corresponde às cores às quais o olho humano é menos sensível. Esse modelo permitiu que empresas definissem quanta largura de banda deveria ser destinada a cada componente  $f_Y(x, y), f_I(x, y), f_Q(x, y)$ , sendo a maior parte destinada à transmissão da intensidade (YAGLE e ULABY, 2018).

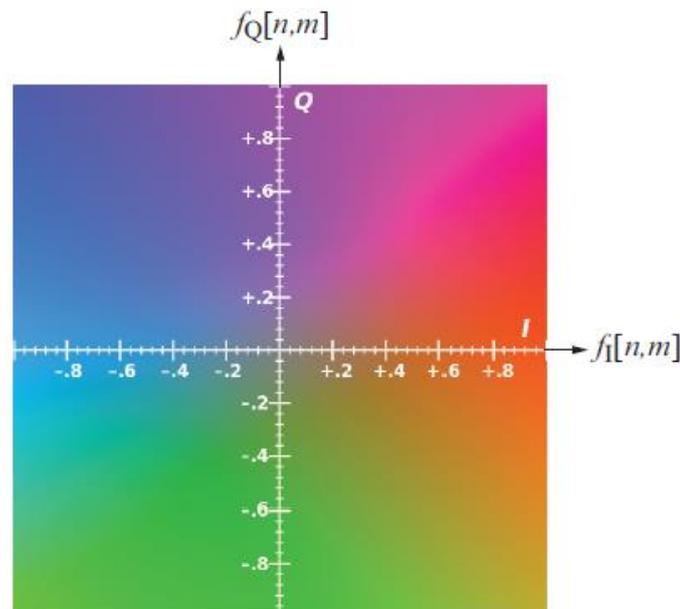


Figura 4.10 – Diagrama de cores na escala YQI (YAGLE e ULABY, 2018).

A conversão propriamente dita entre as escalas RGB e YIQ da representação de uma imagem  $f(x, y)$  se dá pela Equação 4.29:

$$\begin{bmatrix} f_Y(x, y) \\ f_I(x, y) \\ f_Q(x, y) \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,596 & -0,275 & -0,321 \\ 0,212 & -0,523 & 0,311 \end{bmatrix} \begin{bmatrix} f_R(x, y) \\ f_G(x, y) \\ f_B(x, y) \end{bmatrix} \quad (4.29)$$

A primeira linha dessa equação corresponde à fórmula para transformar a imagem colorida  $f(x, y)$  para escala de cinza, que é dada pela Equação 4.30:

$$f_{cinza}(x, y) = 0,299f_R(x, y) + 0,587f_G(x, y) + 0,114f_B(x, y) \quad (4.30)$$

Esse método de conversão é mais comumente utilizado por levar em consideração a sensibilidade do olho humano para cores (YAGLE e ULABY, 2018).

#### 4.6 REGISTRO DE IMAGEM

Registro de imagem corresponde ao processo de alinhar espacialmente duas imagens de modo que pontos correspondentes ocupem as mesmas coordenadas (GOSHTASBY, 2012). Nesse tipo de processo, tem-se disponível uma imagem de referência e uma imagem a ser alinhada com respeito à primeira. Para isso, são usadas transformações geométricas, que são um tipo de transformações espaciais e correspondem às operações utilizadas para modificar o arranjo espacial de pixels em uma imagem. Nessa etapa, é relevante apenas o reposicionamento de pixels e não a alteração das suas intensidades. Entretanto, essas transformações são compostas por duas etapas (GONZALES e WOODS, 2018):

- A transformação espacial das coordenadas propriamente dita;
- Interpolação para determinar os valores de intensidade desses pixels após a etapa anterior.

A transformação de coordenadas pode ser expressa pela Equação 4.31, também chamada de função de mapeamento:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (4.31)$$

Em que  $(x, y)$  são as coordenadas da imagem original e  $(x', y')$  correspondem às coordenadas da imagem pós transformação. A transformação  $(x', y') = (x/2, y/2)$  por

exemplo encolhe a imagem para metade de seu tamanho original (GONZALES e WOODS, 2018).

No âmbito do registro de imagens, é de particular interesse a compreensão das transformadas afins, que incluem as transformações elementares de translação, rotação, esticamento, escala e cisalhamento, podendo ser expressa por uma adição de vetores e multiplicação de matrizes (JÄHNE, 2002). Em coordenadas homogêneas, uma transformação pode ser representada pela Equação 4.32:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4.32)$$

Essa transformação pode ser usada de duas maneiras. Na primeira, chamada de *forward mapping*, os pixels da imagem de entrada são computados para suas novas localizações na imagem de saída; já no processo de *inverse mapping*, os pixels da imagem de saída são mapeados na imagem de entrada através de uma transformação reversa. O problema principal do primeiro método é a possibilidade da existência de lacunas na imagem de saída, além da possível superposição de pixels diferentes da imagem de origem. Já o segundo método, que é inclusive o mais utilizado em plataformas como *MATLAB*, pode se tornar um pouco mais complexo, mas garante que cada ponto da imagem de saída tenha pelo menos um correspondente na imagem de entrada. Entretanto, os dois métodos exigem, em geral, interpolação de intensidades, que será discutida na seção 4.7.1 (GONZALES e WOODS, 2018).

A transformação afim tem seis graus de liberdade: 2 para translação ( $t_x, t_y$ ), um para rotação, um para escala, um para esticamento e um para cisalhamento ( $a_{11}, a_{12}, a_{21}, a_{22}$ ) (JÄHNE, 2002). Esse tipo de transformação preserva pontos, linhas retas e planos. Para se obter os efeitos mencionados, basta escolher os valores corretos para a matriz A, conforme a Tabela 4.1

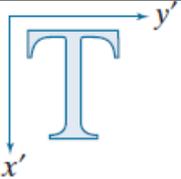
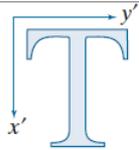
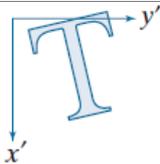
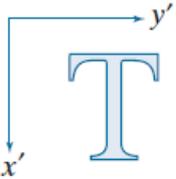
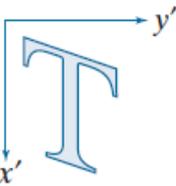
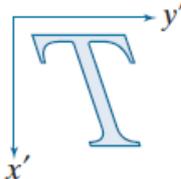
Nome da Transformação	Matriz Afim A	Equação Resultante	Exemplo
<b>Identidade</b>	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x \\ y' &= y \end{aligned}$	
<b>Escala</b>	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= c_x x \\ y' &= c_y y \end{aligned}$	
<b>Rotação</b>	$\begin{bmatrix} \cos\theta & -\text{sen}\theta & 0 \\ \text{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x \cos\theta - y \text{sen}\theta \\ y' &= x \text{sen}\theta + y \cos\theta \end{aligned}$	
<b>Translação</b>	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x + t_x \\ y' &= y + t_y \end{aligned}$	
<b>Cisalhamento (vertical)</b>	$\begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x + s_v y \\ y' &= y \end{aligned}$	
<b>Cisalhamento (horizontal)</b>	$\begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x \\ y' &= s_h x + y \end{aligned}$	

Tabela 4.1 - Exemplos das principais transformações geométricas aplicadas em imagens. Adaptado de (GONZALES e WOODS, 2018).

#### 4.6.1 Interpolação de Intensidades

Interpolação de imagens é um processo utilizado em aplicações como zoom, diminuição, rotação e correção geométrica de imagens. Supor que se quer aumentar uma imagem de tamanho 500x500 para um tamanho de 750x750. Uma maneira de visualizar esse problema é imaginar uma grade de pixels com 750x750 pixels e espaçamento igual à grade de 500x500; se a primeira grade for encolhida e sobreposta sobre a segunda, naturalmente haverá pixels que não têm correspondente na grade original. Uma maneira de sanar esse problema, chamada de interpolação pelo vizinho mais próximo, é atribuir a esses pixels a intensidade do seu pixel vizinho mais próximo, mas isso pode trazer distorção de linhas retas (GONZALES e WOODS, 2018).

Uma alternativa a essa abordagem é a utilização de interpolação bilinear, que leva em conta as intensidades dos quatro vizinhos mais próximos para definir o valor para o pixel de interesse. Considerar que  $(x, y)$  denote as coordenadas do local cuja intensidade  $v(x, y)$  se quer definir. Esse valor pode ser obtido utilizando-se a equação 4.33, em que os coeficientes são funções das intensidades dos pixels vizinhos (GONZALES e WOODS, 2018).

$$v(x, y) = ax + by + cxy + d \quad (4.33)$$

Por fim, é válido citar a interpolação bi cúbica, que leva em consideração os 16 pixels mais próximos do local cuja intensidade se quer determinar. Nesse caso, a intensidade pode ser determinada pela Equação 4.34, em que os coeficientes são determinados em função das intensidades dos pixels vizinhos (GONZALES e WOODS, 2018).

$$v(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j \quad (4.34)$$

## 4.7 SEGMENTAÇÃO DE IMAGENS

Seja  $R$  a totalidade da região espacial ocupada por uma imagem. O processo de segmentação pode ser compreendido como a partição de  $R$  em  $n$  sub-regiões  $R_1, R_2, \dots, R_n$  tal que (GONZALES e WOODS, 2018):

- a)  $\bigcup_{i=1}^n R_i = R$ ;
- b)  $R_i$  é um conjunto conectado para  $i = 0, 1, 2, \dots, n$ ;
- c)  $R_i \cap R_j = \emptyset$  para todo  $i$  e  $j$  tal que  $i \neq j$ ;
- d)  $Q(R_i) = VERDADEIRO$  para  $i = 0, 1, 2, \dots, n$ ;
- e)  $Q(R_i \cup R_j) = FALSO$  para quaisquer regiões adjacentes  $R_i$  e  $R_j$ .

Aqui,  $Q(R_k)$  é um predicado lógico definido sobre os pontos no conjunto  $R_k$  e  $\emptyset$  é o conjunto vazio. Além disso, duas regiões são ditas adjacentes se sua união formar um conjunto conectado; se o conjunto formado pela união das duas regiões não for conectado, essas são ditas disjuntas (GONZALES e WOODS, 2018).

A condição a) informa que todos os pixels devem estar em alguma região. A condição b) afirma que todos os pontos numa região devem estar conectados. A condição c) informa que as regiões oriundas da segmentação devem ser disjuntas. A condição d) lida com propriedades que devem ser satisfeitas pelos pixels de uma região segmentada. Por fim, a condição e) informa que regiões diferentes devem diferir em relação ao predicado  $Q$  (GONZALES e WOODS, 2018).

Algoritmos de segmentação geralmente são divididos em duas categorias: descontinuidade ou similaridade. Nos casos envolvendo descontinuidades, supõe-se que as fronteiras das regiões são suficientemente diferentes entre si e em relação ao plano de fundo, de modo que essas possam ser detectadas a partir de descontinuidades na intensidade da imagem. Já a segunda abordagem atua dividindo a imagem em regiões que são similares de acordo com algum critério estabelecido (GONZALES e WOODS, 2018).

#### 4.7.1 Thresholding

O método de *thresholding* envolve a divisão da imagem em regiões a partir do estabelecimento de um valor limite, chamado de *threshold*, que classificará os pontos que compõem uma imagem. Supor uma imagem  $f(x, y)$ , de histograma representado na Figura

4.11 composta por objetos claros num fundo escuro, de modo que os pixels possam ser segregados em dois grupos distintos (GONZALES e WOODS, 2018).

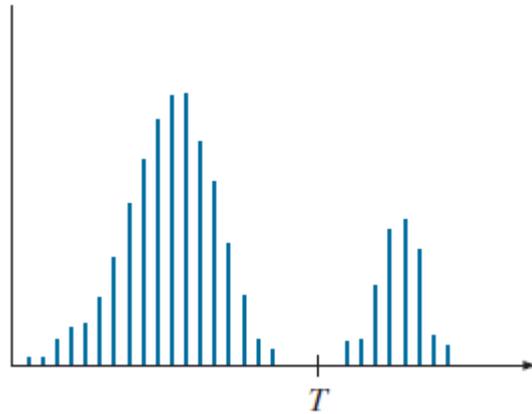


Figura 4.11 Exemplo de histograma de imagem em que a presença de dois picos em níveis de cinza diferentes facilita a determinação de um *threshold* para diferenciação de objetos e plano de fundo (GONZALES e WOODS, 2018).

Para extrair objetos do plano de fundo escuro, basta escolher um valor de *threshold*  $T$ , que possa dividir esses grupos de pixels que compõem a imagem. Desse modo, qualquer ponto  $(x, y)$  da imagem tal que  $f(x, y) > T$  será considerado como um ponto de objeto, do contrário, será chamado de ponto de plano de fundo. A imagem segmentada  $g(x, y)$  pode ser representada pela Equação 4.35 (GONZALES e WOODS, 2018):

$$g(x, y) = \begin{cases} 1 & \text{se } f(x, y) > T \\ 0 & \text{se } f(x, y) \leq T \end{cases} \quad (4.35)$$

O ruído presente numa imagem é um fator determinante para a definição de um *threshold*, como pode ser visto na Figura 4.12. A ausência de ruído torna simples a análise do histograma e a consequente definição do valor de *threshold*; no entanto, à medida que o ruído aumenta, a análise do histograma torna-se mais complicada, e a definição de um valor de *threshold*, mais difícil (GONZALES e WOODS, 2018).

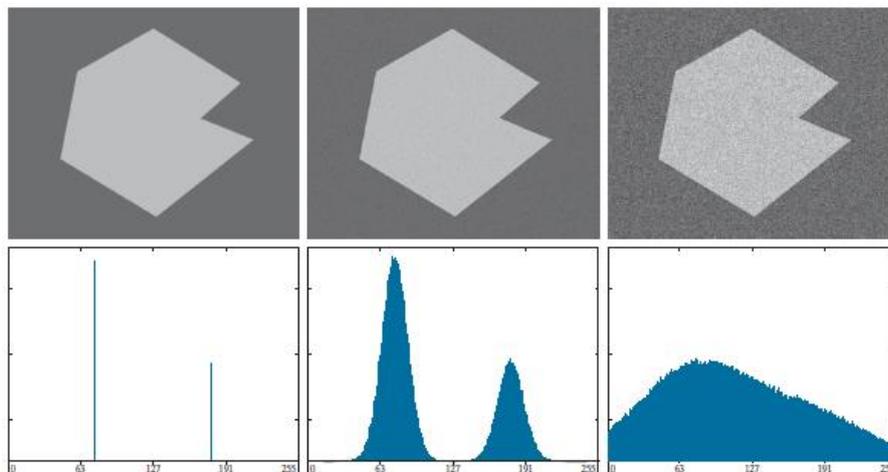


Figura 4.12 – Efeitos do ruído na determinação do *threshold* para determinação de objetos e plano de fundo em uma imagem (GONZALES e WOODS, 2018).

Quando as distribuições de intensidades entre o objeto (região de interesse) e o plano de fundo são suficientemente diferentes, pode-se usar um valor de *threshold* global aplicável a toda a imagem. Quando, na análise de imagens, houver, no entanto, alguma variabilidade entre as imagens, o *threshold* global pode não ser o método mais adequado, sendo válida a utilização de um algoritmo para determinar esse valor (GONZALES e WOODS, 2018):

1. Estimar um valor inicial para o *threshold* global,  $T$ ;
2. Segmentar a imagem usando o valor  $T$ , o que produzirá dois grupos de pixels:  $G_1$  e  $G_2$  com valores maiores que  $T$  e menores ou iguais a  $T$  respectivamente;
3. Calcular as médias  $m_1$  e  $m_2$  para os pixels em  $G_1$  e  $G_2$ ;
4. Calcular um novo *threshold* utilizando a média  $T = \frac{1}{2}(m_1 + m_2)$ ;
5. Repetir os passos 1 a 4 até que a diferença entre valores de  $T$  após sucessivas iterações seja menor que uma métrica definida previamente.

#### 4.7.2 Método de Otsu Para Thresholding

O método de Otsu é utilizado para determinação do *threshold* para segmentação de imagens baseado na maximização da variância entre-classes; aqui, classe corresponde à noção de grupos de pixels que se obtém após a segmentação (OTSU, 1979).

Sejam os pixels de uma imagem representados pelos  $L$  níveis de cinza  $[1, 2, \dots, L]$ ; o número de pixels em cada nível  $i$  é denotado por  $n_i$  e o número total de pixels na imagem é  $N = n_1 + n_2 + \dots + n_L$ . Ao normalizar-se o histograma, esse pode ser interpretado como uma distribuição de probabilidades, como mostra a Equação 4.36 (OTSU, 1979).

$$p_i = \frac{n_i}{N}, \quad p_i \geq 0, \quad \sum_{i=1}^L p_i = 1 \quad (4.36)$$

Supor que uma imagem foi segmentada a partir de um *threshold*  $k$  de modo a dividir os seus pixels em duas classes  $C_0$  e  $C_1$ , que denotam respectivamente os níveis de cinza  $[1, 2, \dots, k]$  e  $[k + 1, \dots, L]$ . Dessa forma, as probabilidades da ocorrência de cada classe são definidas pelas Equações 4.37 e 4.38 (OTSU, 1979):

$$\omega_0 = p(C_0) = \sum_{i=1}^k p_i = \omega(k) \quad (4.37)$$

$$\omega_1 = p(C_1) = \sum_{i=k+1}^L p_i = 1 - \omega(k) \quad (4.38)$$

Além disso, as médias das intensidades de cada uma dessas classes são definidas pelas Equações 4.39 e 4.40:

$$\mu_0 = \sum_{i=1}^k i p(i|C_0) = \sum_{i=1}^k \frac{i p_i}{p(C_0)} = \frac{\mu(k)}{\omega(k)} \quad (4.39)$$

$$\mu_1 = \sum_{i=k+1}^L i p(i|C_1) = \sum_{i=k+1}^L \frac{i p_i}{p(C_1)} = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \quad (4.40)$$

Em que  $\omega(k)$ ,  $\mu(k)$  e  $\mu_T$  são respectivamente o momento de número zero, o momento de número 1 e o nível de cinza médio total da imagem analisada, dados pelas Equações 4.41 a 4.43 (OTSU, 1979).

$$\omega(k) = \sum_{i=1}^k p_i \quad (4.41)$$

$$\mu(k) = \sum_{i=1}^k i p_i \quad (4.42)$$

$$\mu_T = \mu(L) = \sum_{i=1}^L i p_i \quad (4.43)$$

A partir disso, é possível concluir que é válida a relação exibida na Equação 4.44.

$$\omega_0 \mu_0 + \omega_1 \mu_1 = \mu_T, \quad \omega_0 + \omega_1 = 1 \quad (4.44)$$

As variâncias das classes são dadas pelas Equações 4.45 e 4.46.

$$\sigma_0^2 = \sum_{i=1}^k (i - \mu_0)^2 p(i|C_0) = \sum_{i=1}^k (i - \mu_0)^2 \frac{p_i}{\omega_0} \quad (4.45)$$

$$\sigma_1^2 = \sum_{i=k+1}^L (i - \mu_1)^2 p(i|C_1) = \sum_{i=k+1}^L (i - \mu_1)^2 \frac{p_i}{\omega_1} \quad (4.46)$$

No intuito de medir a qualidade de um *threshold*, é estabelecido uma métrica que servirá de base para a análise, definida na Equação 4.47 (OTSU, 1979).

$$\eta = \frac{\sigma_B^2}{\sigma_T^2} \quad (4.47)$$

Em que  $\sigma_B$  e  $\sigma_T$  correspondem à variância inter-classe e à variância total de níveis de cinza respectivamente (OTSU, 1979). Esses valores podem ser obtidos pelas Equações 4.48 e 4.49.

$$\sigma_B^2 = \omega_0 \omega_1 (\mu_1 - \mu_0)^2 \quad (4.48)$$

$$\sigma_T^2 = \sum_{i=1}^L (i - \mu_T)^2 p_i \quad (4.49)$$

Em função disso, o problema de definição de um valor de *threshold* torna-se apenas um problema de otimização, mais especificamente de maximização, do discriminante  $\eta$ , o que corresponde à maximização de  $\sigma_B^2$  (OTSU, 1979).

Assim, o valor ótimo para o threshold  $k^*$  será dado pela Equação 4.50.

$$\sigma_B^2(k^*) = \max_{1 < k < L} \sigma_B^2(k) \quad (4.50)$$

Para encontrar esse valor, utiliza-se as Equações 4.51 e 4.52 para calcular os valores de  $\sigma_B^2$  associados a cada  $k$  e para avaliar o discriminante respectivamente (OTSU, 1979).

$$\sigma_B^2(k) = \frac{[\mu_T \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (4.51)$$

$$\eta(k) = \frac{\sigma_B^2(k)}{\sigma_T^2} \quad (4.52)$$

É necessário, além disso, então definir o intervalo  $S^*$  de valores  $k$  válidos, ou seja, as restrições para o problema de otimização, descrito na Equação 4.53. Essa Equação nos informa que o intervalo de valores viáveis é aquele em que a probabilidade de um pixel

pertencer a uma classe é algum valor entre 0 e 1, porém diferente de 0 e 1; isso garante que não haverá nenhuma classe sem pixels, visto que não faria sentido segmentar uma imagem em apenas uma classe (OTSU, 1979).

$$S^* = \{k; \omega_0 \omega_1 = \omega(k) [1 - \omega(k)] > 0 \text{ ou } 0 < \omega(k) < 1\} \quad (4.53)$$

## CAPÍTULO V

### 5 INTELIGÊNCIA ARTIFICIAL

#### 5.1 APRENDIZADO DE MÁQUINA

Aprendizado de máquina pode ser definido como o conjunto de métodos computacionais que utilizam a experiência para aprimorar a performance e realizar previsões com acurácia. Nesse caso, experiência se refere à informação passada, que geralmente assume a forma de dados eletrônicos coletados e tornados aptos à análise; sua quantidade e qualidade são fatores decisivos no sucesso da aplicação do método de aprendizado de máquina. Como o sucesso de um algoritmo de aprendizado de máquina está atrelado à informação utilizada, essa área da ciência se relaciona intimamente à análise de dados e à estatística (MOHRI *et al*, 2018).

Aprendizado de máquina possui um amplo conjunto de aplicações, tais quais (MOHRI *et al*, 2018):

- Classificação de textos ou documentos. Como por exemplo designar um tópico a um documento ou fazer inferências a respeito do conteúdo de um texto;
- Processamento de discurso, que envolve o trabalho com falas e reconhecimento de seus conteúdos, além da geração do discurso em si; aqui também estão incluídos a identificação de quem está falando algo através da fala e modelagem acústica;
- Visão computacional, que envolve reconhecimento e identificação de objetos, detecção de faces, reconhecimento de caracteres e obtenção de informações baseada em imagem;

- Biologia computacional, que envolve a predição da função de proteínas ou análise de genes;
- Além de aplicações diversas como detecção de fraudes em cartões de crédito, direção autônoma de veículos, auxílio em diagnósticos médicos, design de mecanismos de busca dentre outras.

No âmbito de ML, há alguns conceitos de particular relevância para compreensão de suas aplicações e aspectos técnicos (MOHRI *et al*, 2018):

- Classificação: refere-se ao problema de determinar uma categoria a cada item de um grupo; por exemplo, determinar o assunto de que se trata um conjunto de textos, ou determinar a cor de veículos em um grupo de imagens;
- Regressão: refere-se ao problema de prever um valor real para um item; um exemplo de problema de regressão seria o de prever o valor de uma ação em dado momento. Diferentemente dos problemas de classificação, aqui a diferença entre o valor predito e o valor verdadeiro de um item é relevante, e é de interesse do pesquisador reduzir essa diferença;
- Clustering: corresponde ao problema de particionar um conjunto de itens em subconjuntos homogêneos sem que se saiba previamente como essa classificação deve ser feita; em geral é função do próprio algoritmo definir essa diferenciação.

O objetivo principal do ML é a geração de predições com acurácia para itens ainda não vistos através de algoritmos robustos e eficientes que funcionem mesmo em problemas de larga escala (MOHRI *et al*, 2018).

O desenvolvimento desses algoritmos envolve o processo de aprendizado em si, que pode ser de diversos tipos a depender dos dados disponíveis para treinamento, a ordem e o método pelo qual esses dados são utilizados pelo algoritmo (MOHRI *et al*, 2018):

- Aprendizado supervisionado: o algoritmo recebe exemplos rotulados de dados de treinamento e faz predições para pontos ainda não conhecidos;
- Aprendizado não supervisionado: o algoritmo recebe exemplos não rotulados e faz predições para pontos ainda não conhecidos; como exemplos, pode-se citar o clustering e redução de dimensionalidade;
- Aprendizado semi-supervisionado: o algoritmo recebe uma amostra composta por dados rotulados e não-rotulados, a partir das quais faz predições para dados ainda

não conhecidos. Essa metodologia é útil em situações em que dados não-rotulados são de fácil acesso mas os rótulos são difíceis de se obter;

- Inferência transdutiva: tal qual o método anterior, aqui o algoritmo recebe um conjunto de dados rotulados e um conjunto de dados não rotulados; contudo, aqui o objetivo é realizar inferência apenas em relação a esses pontos e não a respeito de dados ainda não conhecidos;
- Aprendizado on-line: envolve múltiplas etapas de treino e de teste que se misturam; em cada etapa, o algoritmo recebe um ponto não rotulado, faz a predição, recebe o resultado correto para fins de comparação e calcula sua perda. O objetivo dessa estratégia é reduzir a perda cumulativa;
- Aprendizado por reforço: aqui, as etapas de treino e teste também se misturam, mas para coletar informação, o algoritmo que está aprendendo precisa interagir com o ambiente, recebendo uma recompensa imediata por cada ação tomada.
- Aprendizado ativo: nesse caso, o algoritmo coleta ativamente amostras de treino; a ideia aqui é se conseguir uma performance similar à do aprendizado supervisionado mas com menos exemplos na etapa de treino.

Independentemente do modelo de aprendizado escolhido, o algoritmo deve ser capaz de realizar generalização, ou seja, a partir de exemplos conhecidos, a função selecionada precisa estar apta a fazer predições para exemplos ainda não conhecidos. É válido questionar quão complexa essa função deve ser para que se tenha uma capacidade de generalização adequada (MOHRI *et al*, 2018).

Em geral, o melhor preditor para um conjunto de dados de teste não é a melhor escolha para uma situação mais geral. Escolher uma função que seja muito complexa e que se adapte perfeitamente a um conjunto de dados pode gerar pouca generalização, pois considera que todos os dados seguirão o padrão apresentado pelos casos de teste, num fenômeno chamado de *overfitting*; já escolher uma função muito simples pode trazer problemas relativos à acurácia, visto que o algoritmo pode não estar apto a sequer compreender as características do conjunto analisado, num fenômeno chamado de *underfitting* (MOHRI *et al*, 2018). Essas duas situações podem ser vistas respectivamente na Figura 5.1.

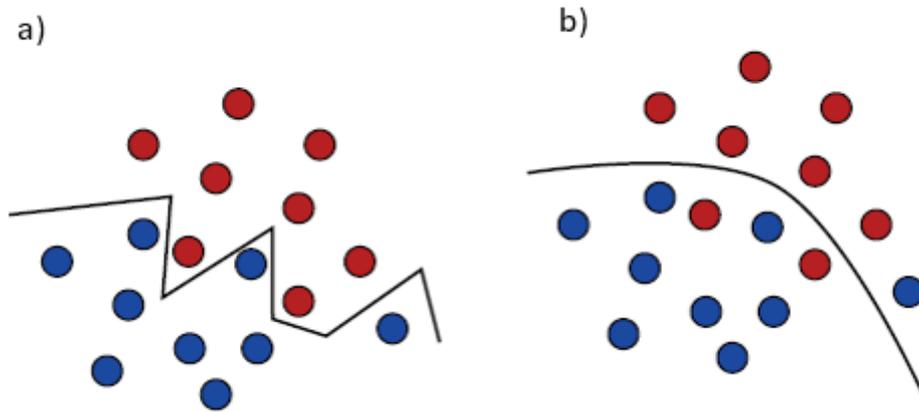


Figura 5.1 - Ilustração das situações de a) *overfitting* e b) *underfitting*. Adaptado de MOHRI et al (2018).

## 5.2 K-NEAREST NEIGHBORS (KNN)

Um histograma padrão tem a característica de particionar uma variável contínua  $x$  em classes de largura  $\Delta_i$ ; a quantidade de observações que recaem em uma classe  $i$  será denominada  $n_i$ . No intuito de interpretar esse problema como uma distribuição de probabilidades, é necessário dividir  $n_i$  conforme a equação 5.1, a partir da qual é possível perceber que  $\int p(x)dx = 1$ . Essa equação fornece um modelo para cálculo de densidade de probabilidade  $p(x)$  que é constante ao longo de cada classe (BISHOP, 2006).

$$p_i = \frac{n_i}{N\Delta_i} \quad (5.1)$$

A análise dos histogramas permite concluir que para estimar a densidade de probabilidade de um ponto, é necessário considerar as localidades próximas a esse ponto, o que por sua vez requer a consideração da forma de se medir distância; por padrão, será considerada a distância euclidiana (BISHOP, 2006).

Supor observações sendo feitas a partir de alguma densidade de probabilidade  $p(x)$ , que se quer conhecer, num espaço de  $D$  dimensões euclidiano. É necessário considerar uma

pequena região  $R$  que contenha  $x$ , cuja probabilidade associada será dada pela Equação 5.2 (BISHOP, 2006).

$$P = \int_R p(x) dx \quad (5.2)$$

Agora supor que um conjunto de  $N$  observações foi feito de  $p(x)$ . Como cada ponto observado tem uma probabilidade  $P$  de estar dentro de  $R$ , o número total de pontos  $K$  dentro dessa região será dado pela distribuição binomial exibida na Equação 5.3 (BISHOP, 2006).

$$Bin(K|N, P) = \frac{N!}{K!(N-K)!} P^K (1-P)^{1-K} \quad (5.3)$$

Para um número de observações suficientemente grande, essa distribuição terá a média dada pela Equação 5.4 (BISHOP, 2006).

$$K \simeq NP \quad (5.4)$$

Se  $R$  for considerada pequena o bastante para que a densidade  $p(x)$  seja constante ao longo da região, é válida a Equação 5.5.

$$P \simeq p(x)V \quad (5.5)$$

Em que  $V$  é o volume da região considerada. Combinando as duas equações, obtém-se a densidade de probabilidade exibida na Equação 5.6.

$$p(x) = \frac{K}{NV} \quad (5.6)$$

Uma abordagem para utilizar esse resultado é considerar o valor de  $K$  constante e determinar o valor de  $V$  correspondente. Para isso, considera-se uma pequena esfera centrada no ponto de interesse  $x$  em que se quer estimar a densidade de probabilidade; o raio dessa esfera será aquele que engloba precisamente  $K$  pontos vizinhos. Com isso, é possível utilizar a equação 5.6 considerando o volume da esfera obtida; essa técnica é conhecida como K-Nearest Neighbors (K-vizinhos mais próximos) (BISHOP, 2006).

Essa técnica pode ser expandida para problemas de classificação; para realizar isso, basta aplicar a equação para estimação da densidade de probabilidade para cada uma das classes separadamente e utilizar o teorema de Bayes. Por exemplo, supor um conjunto de dados composto por  $N_k$  pontos em classes  $C_k$  com  $N$  pontos totais de modo que  $\sum_k N_k = N$ . Para classificar um novo ponto  $x$ , basta considerar uma esfera centrada nesse ponto contendo  $K$  pontos. Supondo que essa esfera possui um volume  $V$  e contém  $K_k$  de uma classe  $C_k$ , pode-se estimar a densidade de probabilidade associada a cada classe através da Equação 5.7 (BISHOP, 2006).

$$p(x|C_k) = \frac{K_k}{N_k V} \quad (5.7)$$

Além disso, sabe-se que, pela Equação 5.8:

$$p(C_k) = \frac{N_k}{N} \quad (5.8)$$

Combinando esses resultados com a Equação 5.6 utilizando o Teorema de Bayes, pode-se obter a propriedade a posteriori de pertencimento a uma classe, dada pela Equação 5.9 (BISHOP, 2006)

$$p(C_k|x) = \frac{p(x|C_k)p(C_k)}{p(x)} = \frac{K_k}{K} \quad (5.9)$$

Para minimizar os erros de classificação, basta alocar cada ponto de teste para a classe que possui maior probabilidade a posteriori, ou seja, aquela cujo valor de  $K_k/K$  será maior. A classificação de um ponto envolve a identificação dos  $K$  pontos mais próximos e o cálculo das probabilidades a posteriori. Em geral, um valor pequeno de  $K$  produz várias pequenas regiões de cada classe, enquanto valores maiores produzem menos regiões maiores, conforme ilustra a Figura 5.2 (BISHOP, 2006).

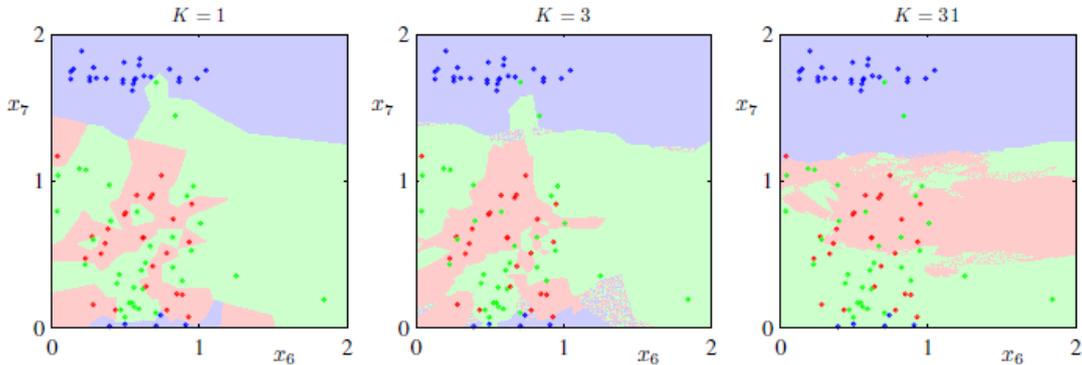


Figura 5.2 – Fronteiras de decisão obtidas pelo método K-Nearest Neighbors para valores de  $K$  iguais a 1,3 e 31 (BISHOP, 2006).

### 5.3 SUPPORT VECTOR MACHINE (SVM)

Considerar um problema de classificação de duas classes utilizando um modelo linear de forma descrita pela Equação 5.10.

$$y(x) = w^T \phi(x) + b \quad (5.10)$$

Em que  $\phi(x)$  representa uma transformação no espaço de *features* (características). O conjunto de treino é composto por  $N$  vetores  $x_1, \dots, x_N$ , cujos valores alvos de predição são  $t_1, \dots, t_N$  em que  $t_n \in \{-1, 1\}$  e novos valores são classificados de acordo com  $y(x)$ . Além disso, supõe-se que o conjunto de dados de treinamento é linearmente separável, de modo que existe ao menos uma escolha de parâmetros  $w$  e  $b$  tais que a Equação 5.10 satisfaça  $y(x_n) > 0$  para  $t_n = +1$  e  $y(x_n) < 0$  para  $t_n = -1$ . Pode haver diversas soluções plausíveis, mas a mais adequada será aquela que apresente o menor erro de generalização.

O algoritmo SVM encara esse problema utilizando o conceito de margem, definida como a menor distância entre a fronteira de decisão e qualquer uma das amostras. No SVM, as fronteiras de decisão são escolhidas de modo que o valor da margem seja maximizado (BISHOP, 2006).

A distância de um ponto  $x$  a um hiperplano definido por  $y(x) = 0$  é dada por  $|y(x)|/\|w\|$ . Além disso, são relevantes apenas as soluções para as quais os pontos estejam corretamente classificados, ou seja,  $t_n y(x_n) > 0$  para todo  $n$ . A distância de um ponto  $x_n$  à superfície de decisão é dada pela Equação 5.11 (BISHOP, 2006).

$$\frac{t_n y(x_n)}{\|w\|} = \frac{t_n (w^T \phi(x_n) + b)}{\|w\|} \quad (5.11)$$

A margem é dada pela distância do ponto mais próximo  $x_n$ , e o objetivo aqui é otimizar os parâmetros  $w$  e  $b$  de modo a maximizar essa distância; portanto, a solução será dada pela Equação 5.12 (BISHOP, 2006).

$$\arg \max_{w,b} \left\{ \frac{1}{\|w\|} \min_n [t_n (w^T \phi(x_n) + b)] \right\} \quad (5.12)$$

Se alterarmos a escala de valores de  $w$  e  $b$ , é possível definir o ponto mais próximo da superfície como sendo dado pela Equação 5.13 (BISHOP, 2006).

$$t_n (w^T \phi(x_n) + b) = 1 \quad (5.13)$$

Logo, todos os outros pontos em questão devem obedecer a restrição conhecida como representação canônica do hiperplano, exibidas na Equação 5.14. Quando valer a igualdade, a restrição é dita ativa, do contrário, inativa (BISHOP, 2006).

$$t_n(w^T \phi(x_n) + b) \geq 1 \quad n = 1, \dots, N \quad (5.14)$$

Sempre haverá um ponto mais próximo, logo, o problema de otimização apenas requer a maximização de  $\|w\|^{-1}$ , que é equivalente à minimização de  $\|w\|^2$ . Mais especificamente, o problema de otimização torna-se a Equação 5.15, em que o fator  $\frac{1}{2}$  é adicionado apenas por conveniência (BISHOP, 2006).

$$\arg \min_{w,b} \frac{1}{2} \|w\|^2 \quad (5.15)$$

Para resolver esse problema de otimização, pode-se utilizar multiplicadores Lagrangianos  $a_n \geq 0$ , sendo um multiplicador para cada restrição, o que fornece a função Lagrangiana exibida na Equação 5.16, em que  $a = (a_1, \dots, a_N)^T$  (BISHOP, 2006).

$$L(w, b, a) = \frac{1}{2} \|w\|^2 - \sum_{n=1}^N a_n \{t_n(w^T \phi(x_n) + b) - 1\} \quad (5.16)$$

O cálculo das derivadas de  $L(w, b, a)$  com relação a  $w$  e  $b$  fornece as Equações 5.17 e 5.18.

$$w = \sum_{n=1}^N a_n t_n \phi(x_n) \quad (5.17)$$

$$0 = \sum_{n=1}^N a_n t_n \quad (5.18)$$

Essas equações podem ser reorganizadas e o problema agora passa a ser o de maximizar a Equação 5.19 utilizando as relações descritas nas Equações 5.20 e 5.21.

$$\tilde{L}(a) = \sum_{n=1}^N a_n - \frac{1}{2} \sum_{n=1}^N a_n a_m t_n t_m k(x_n, x_m) \quad (5.19)$$

$$a_n \geq 0, \quad n = 1, \dots, N \quad (5.20)$$

$$\sum_{n=1}^N a_n t_n = 0 \quad (5.21)$$

A representação da equação 5.19 permite que o modelo seja reformulado usando *kernels* e a consequência disso é que o classificador baseado em margem máxima pode ser aplicado em espaços de *features* em que a dimensionalidade é maior que o número de pontos, incluindo espaços de *features* infinitas (BISHOP, 2006).

Para classificar novos dados utilizando o modelo treinado, utiliza-se  $y(x)$  escrito em termos dos parâmetros  $\{a_n\}$  através da Equação 5.22.

$$y(x) = \sum_{n=1}^N a_n t_n k(x, x_n) + b \quad (5.22)$$

Em que  $k(x, x_n)$  é a função kernel dada pela Equação 5.23.

$$k(x, x_n) = \phi(x)^T \phi(x_n) \quad (5.23)$$

Um problema de otimização dessa forma satisfaz as condições de Karush-Kuhn-Tucker (KKT), que exigem que as seguintes propriedades, descritas nas Equações 5.24, 5.25 e 5.26 sejam válidas (BISHOP, 2006).

$$a_n \geq 0 \quad (5.24)$$

$$t_n y(x_n) - 1 \geq 0 \quad (5.25)$$

$$a_n \{t_n y(x_n) - 1\} = 0 \quad (5.26)$$

Logo, para todos os pontos correspondentes a informações, ou  $a_n = 0$  ou  $t_n y(x_n) = 1$ . Os pontos para os quais  $a_n = 0$  não têm relevância na soma descrita na Equação 5.22; já os outros pontos são chamados de *support vectors*, que correspondem a pontos que existem nos hiperplanos de margem máxima no espaço de features. A vantagem de utilizar essa representação é que uma vez que o modelo for treinado, grande parte dos dados utilizados pode ser descartada, de modo que apenas os *support vectors* são relevantes na predição de novos pontos (BISHOP, 2006).

Uma vez que a Equação Lagrangiana foi resolvida e os *support vectors* encontrados, pode-se determinar o valor de  $b$  visto que qualquer *support vector* deve satisfazer a condição  $t_n y(x_n) = 1$ . É possível obter a equação 5.27, em que  $S$  denota o conjunto de índices dos *support vectors* (BISHOP, 2006).

$$t_n \left( \sum_{m \in S} a_m t_m k(x_n, x_m) + b \right) = 1 \quad (5.27)$$

O cálculo de  $b$  pode ser obtido ao utilizar  $t_n^2 = 1$  e ao calcular a média em relação a todos os *support vectors*, o que resulta na Equação 5.28, em que  $N_S$  é o número total de *support vectors* (BISHOP, 2006).

$$b = \frac{1}{N_S} \sum_{n \in S} \left( t_n - \sum_{m \in S} a_m t_m k(x_n, x_m) \right) \quad (5.38)$$

Para fins de classificação com outros modelos, o SVM pode ser expresso em termos da minimização de uma função erro descrita na Equação 5.39, em que  $E_\infty(z)$  é uma função que é zero se  $z \geq 0$  e  $\infty$  caso contrário, e em que  $\lambda > 0$  (BISHOP, 2006).

$$\sum_{n=1}^N E_{\infty}(t_n y(x_n) - 1) + \lambda \|w\|^2 \quad (5.39)$$

#### 5.4 REDES NEURAI ARTIFICIAIS

Um modelo básico de uma rede neural pode ser descrito como uma série de transformações funcionais. Pode-se construir  $M$  combinações lineares das variáveis de entrada  $x_1, \dots, x_D$  que possuem a forma da Equação 5.40 (BISHOP, 2006).

$$a_j = \sum_{i=1}^D w_{ji}^{(1)} x_i + w_{j0}^{(1)} \quad (5.40)$$

Aqui, “(1)” indica que os parâmetros em questão correspondem à primeira camada da rede neural. Os termos  $w_{ji}^{(1)}$  correspondem aos pesos e  $w_{j0}^{(1)}$  correspondem aos vieses (*biases*); já as quantidades  $a_j$  são chamadas de ativações, que são transformadas utilizando um função não-linear diferenciável chamada de função de ativação  $h(\cdot)$  (que geralmente é a função sigmoide ou a tangente hiperbólica) para originar a Equação 5.41 (BISHOP, 2006).

$$z_j = h(a_j) \quad (5.41)$$

Esses valores  $z_j$ , chamados de unidades escondidas, são novamente linearmente combinados para gerar os valores  $a_k$  na Equação 5.42.

$$a_k = \sum_{j=1}^M w_{ji}^{(2)} x_i + w_{j0}^{(2)} \quad (5.42)$$

Em que  $k = 1, \dots, K$  é o número total de saídas. Essa transformação corresponde à segunda camada da rede neural, e de modo similar à camada anterior,  $w_{ji}^{(2)}$  corresponde aos pesos e  $w_{j0}^{(2)}$  corresponde aos vieses dessa camada. Do mesmo modo, os valores de  $a_k$  são transformados utilizando uma função de ativação apropriada para gerar uma série de saídas  $y_k$ , descritas pela Equação 5.43 que considera a função de ativação sigmoide  $\sigma$ . A rede neural descrita até aqui pode ser visualizada na Figura 5.3 (BISHOP, 2006).

$$y_k = \sigma(a_k) \quad (5.43)$$

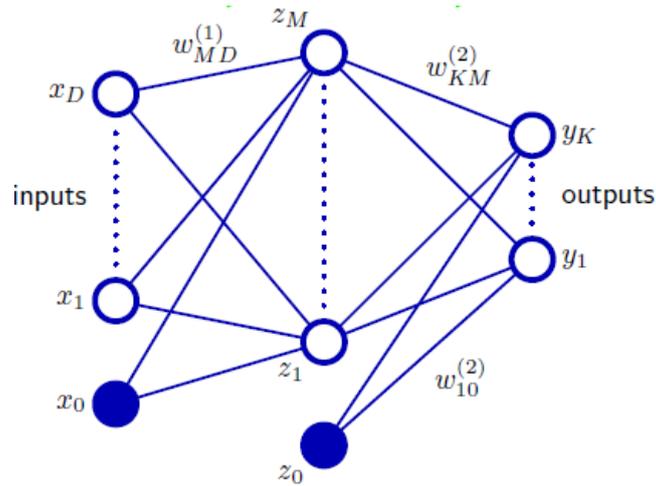


Figura 5.3 – Representação visual das camadas de uma rede neural com destaque para as entradas (inputs), saídas (outputs) e os pesos utilizados (BISHOP, 2006).

Os estágios descritos podem ser combinados para descrever a função geral da rede neural que, ao se considerar a função de ativação sigmoide, toma a forma da Equação 5.44 (BISHOP, 2006).

$$y_k(x, w) = \sigma \left( \sum_{j=1}^M w_{kj}^{(2)} h \left( \sum_{i=1}^D w_{ji}^{(1)} x_i + w_{j0}^{(1)} \right) + w_{k0}^{(2)} \right) \quad (5.44)$$

Vê-se que a rede neural nada mais é do que uma função não-linear de um conjunto de variáveis de entrada  $\{x_i\}$  para um conjunto de variáveis de saída  $\{y_k\}$  controlada por um vetor  $w$  de parâmetros ajustáveis. Essa equação descreve o fluxo de informação conhecido como *forward propagation*. Se for definido um parâmetro extra  $x_0 = 1$ , a equação pode tomar a forma simplificada da Equação 5.45 (BISHOP, 2006).

$$y_k(x, w) = \sigma \left( \sum_{j=1}^M w_{kj}^{(2)} h \left( \sum_{i=1}^D w_{ji}^{(1)} x_i \right) \right) \quad (5.45)$$

Se as funções de ativação de todas as unidades escondidas forem lineares, então sempre poderá ser definida uma rede neural equivalente sem camadas escondidas, visto que a composição de diversas combinações lineares em sequência é em si uma combinação linear. Entretanto, se o número de unidades escondidas for menor que o número de entradas ou saídas, então a transformação promovida pela rede não será a mais geral possível. A rede neural representada na Figura 5.3 pode ser considerada uma rede de duas camadas, visto que esse é o número de camadas que de fato possui pesos adaptativos, que é o critério relevante para determinar as propriedades da rede (BISHOP, 2006).

Nesse tipo de algoritmo, o treinamento em si pode ser melhor compreendido ao se considerar uma interpretação probabilística do mesmo. Ao se considerar o problema de regressão, toma-se uma variável  $t$  que pode assumir qualquer valor real, que tenha uma distribuição gaussiana com média dependente de  $x$  tal que tem-se a Equação 5.46, em que  $\beta$  é a precisão do ruído gaussiano (BISHOP, 2006).

$$p(t|x, w) = \mathcal{N}(t|y(x, w), \beta^{-1}) \quad (5.46)$$

Dado um conjunto de  $N$  observações independentes  $X = \{x_1, \dots, x_N\}$  correspondentes aos valores alvo  $t = \{t_1, \dots, t_N\}$ , pode-se construir a função de verossimilhança descrita pela equação 5.47 (BISHOP, 2006).

$$\frac{\beta}{2} \sum_{n=1}^N \{y(x_n, w) - t_n\}^2 - \frac{N}{2} \ln \beta + \frac{N}{2} \ln \ln (2\pi) \quad (5.47)$$

Maximizar essa função é equivalente a minimizar a função erro descrita pela Equação 5.48. Nela, o valor de  $w$  encontrado será denotado por  $w_{ML}$  (BISHOP, 2006).

$$E(w) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, w) - t_n\}^2 \quad (5.48)$$

Uma vez que o valor de  $w_{ML}$  for obtido, o valor de  $\beta$  pode ser obtido através da minimização da Equação 5.49 (BISHOP, 2006).

$$\frac{1}{\beta_{ML}} = \frac{1}{N} \sum_{n=1}^N \{y(x_n, w) - t_n\}^2 \quad (5.49)$$

Considerar agora o problema de classificação binária em que se tem uma variável alvo  $t$  que pode assumir o valor 1, que denota a classe  $C_1$  ou o valor 0, que denota a classe  $C_2$ . Considerar que a rede neural possui apenas uma saída governada pela função de ativação sigmoide descrita pela Equação 5.50, em que  $0 \leq y(x, w) \leq 1$  (BISHOP, 2006).

$$y = \sigma(a) = \frac{1}{1 + \exp(-a)} \quad (5.50)$$

$y(x, w)$  pode ser interpretado como a probabilidade condicional  $p(C_1|x)$ , em que  $p(C_2|x)$  é dado por  $1 - y(x, w)$ . A distribuição condicional de alvos é uma distribuição de Bernoulli com a forma da Equação 5.51 (BISHOP, 2006).

$$p(t|x, w) = y(x, w)^t \{1 - y(x, w)\}^{1-t} \quad (5.51)$$

Se as observações consideradas forem independentes, a função erro terá a forma da Equação 5.52, em que  $y_n$  corresponde a  $y(x_n, w)$  (BISHOP, 2006).

$$E(w) = - \sum_{n=1}^N \{t_n \ln y_n + (1 - t_n) \ln(1 - y_n)\} \quad (5.52)$$

Se a classificação for do tipo multiclasse, com K classes mutualmente exclusivas, a função erro e a função de ativação sigmoide, comumente utilizada nesse tipo de problema, estão exibidas nas Equações 5.53 e 5.54 (BISHOP, 2006).

$$E(w) = - \sum_{n=1}^N \sum_{k=1}^K t_{kn} \ln y_k(x_n, w) \quad (5.53)$$

$$y_k(x, w) = \frac{\exp(a_k(x, w))}{\sum \exp(a_j(x, w))} \quad (5.54)$$

Agora que se sabe o modelo para rede neural e também se sabe como medir o erro atrelado à mesma, é necessário conhecer o método para se escolher um vetor de parâmetros  $w$  que minimize a função erro  $E(w)$ . Geometricamente, a função erro pode ser concebida como uma superfície no espaço de pesos, como mostra a Figura 5.4. Nessa representação, se houver uma pequena modificação no valor de  $w$  para  $w + \delta w$ , a variação correspondente no erro será  $\delta E \simeq \delta w^T \nabla E(w)$ , em que  $\nabla E(w)$  corresponde ao vetor gradiente, que aponta na direção da maior taxa de aumento da função erro. O menor valor de erro ocorrerá no ponto em que  $\nabla E(w) = 0$ , também chamado de ponto estacionário (BISHOP, 2006).

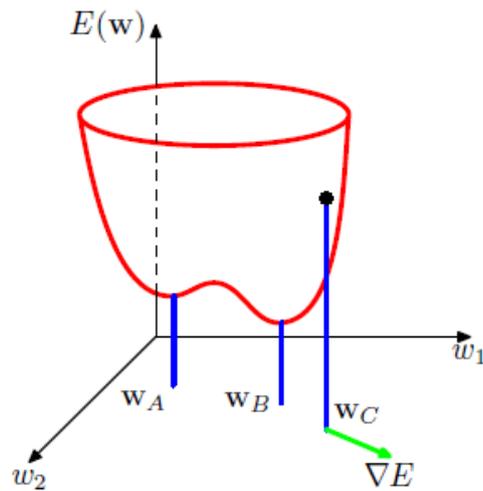


Figura 5.4 – representação da função erro  $E(w)$  no espaço de pesos (BISHOP, 2006).

Encontrar uma solução analítica para  $\nabla E(w) = 0$  pode não ser viável, o que traz à tona a necessidade da utilização de métodos numéricos. A maioria das técnicas envolve a seleção de um parâmetro inicial  $w^{(0)}$ , e a movimentação desse ao longo do espaço de pesos em passos sucessivos com a forma da Equação 5.55, em que  $\tau$  corresponde ao número da iteração (BISHOP, 2006).

$$w^{(\tau+1)} = w^{(\tau)} + \Delta w^{(\tau)} \quad (5.55)$$

Uma maneira simples de utilizar as informações do gradiente para guiar o processo de movimentação ao longo do espaço de pesos é utilizar a Equação 5.56, que leva em conta o sentido negativo do gradiente, ou seja, na direção em que o erro é minimizado. Esse método é conhecido como gradient descent (gradiente descendente) e  $\eta$  é a taxa de aprendizado (BISHOP, 2006).

$$w^{(\tau+1)} = w^{(\tau)} + \eta \nabla E(w^{(\tau)}) \quad (5.56)$$

Esse método, no entanto, é pouco prático, pois envolve o cálculo de todos os pontos de treino a cada iteração. Logo, uma técnica mais eficiente é necessária.

Uma maneira de utilizar o gradiente descendente para minimizar o erro associado a um algoritmo é aplicar a técnica de *backpropagation* (retroalimentação), em que informação é enviada no sentido da rede mas também no sentido oposto. Considerar uma função para cálculo de erro de uma rede neural definida pela equação 5.57, que leva em conta todos os pontos do conjunto de treino (BISHOP, 2006).

$$E(w) = \sum_{n=1}^N E_n(w) \quad (5.57)$$

Considerar um modelo linear simples em que as saídas são combinações lineares das entradas, tal qual a Equação 5.58.

$$y_k = \sum_i w_{ki} x_i \quad (5.58)$$

Unindo essas duas equações, é possível calcular o erro associado a um padrão de entrada  $n$ , em que  $y_{nk} = y_k(x_n, w)$  através da Equação 5.59.

$$E_n = \frac{1}{2} \sum_k (y_{nk} - t_{nk})^2 \quad (5.59)$$

O gradiente desse erro em função de um peso  $w_{ji}$  é dado pela Equação 5.60.

$$\frac{\partial E_n}{\partial w_{ji}} = (y_{nj} - t_{nj}) x_{ni} \quad (5.60)$$

Numa rede neural convencional, cada unidade computa a soma de todas as suas entradas com a forma da Equação 5.61.

$$a_j = \sum_i w_{ji} z_i \quad (5.61)$$

Além disso, nesse tipo de rede, a soma é transformada por uma função de ativação não-linear  $h(\cdot)$  conforme a Equação 5.41 (BISHOP, 2006).

Utilizando a regra de cadeia é possível demonstrar que é verdadeira a Equação 5.62.

$$\frac{\partial E_n}{\partial w_{ji}} = \frac{\partial E_n}{\partial a_j} \frac{\partial a_j}{\partial w_{ji}} \quad (5.62)$$

Além disso, é válido utilizar a notação expressa na Equação 5.63.

$$\delta_j = \frac{\partial E_n}{\partial a_j} \quad (5.63)$$

Lembrando que é válida a Equação 5.64

$$\frac{\partial a_j}{\partial w_{ji}} = z_i \quad (5.64)$$

é possível unir as Equações 5.63 e 5.64 para obter a Equação 5.65.

$$\frac{\partial E_n}{\partial w_{ji}} = \delta_j z_i \quad (5.65)$$

Essa Equação mostra que obter a derivada objetivo exige apenas a multiplicação de  $\delta_j$  por  $z_i$ . Para as unidades de saída, é válida a Equação 5.66 (BISHOP, 2006).

$$\delta_k = y_k - t_k \quad (5.66)$$

Já para as unidades escondidas, é necessário aplicar mais uma vez a regra da cadeia, conforme a Equação 5.67, em que a soma passa por todas as unidades  $k$  para quem a unidade  $j$  envia conexões (BISHOP, 2006).

$$\delta_j = \frac{\partial E_n}{\partial a_j} = \sum_k \frac{\partial E_n}{\partial a_k} \frac{\partial a_k}{\partial a_j} \quad (5.67)$$

A partir do exposto, é possível definir o valor de  $\delta_j$  para unidades escondidas a partir da Equação 5.68.

$$\delta_j = h'(a_j) \sum_k w_{kj} \delta_k \quad (5.68)$$

## 5.5 ÁRVORE DE DECISÃO

São algoritmos de tomada de decisão baseados numa estrutura de árvore, que é um método de escolha também aplicado pelos humanos, em que são realizadas diversas perguntas que classificam o ponto analisado; cada pergunta corresponde a uma *feature* específica e relevante para a classificação. Cada teste ou leva à conclusão ou a um novo questionamento, de modo que o final no processo, a conclusão corresponde à classe em que o alvo está incluído (ZHOU, 2016).

Em geral, uma DT consiste em um nodo raiz, nodos internos e nodos folha; esses últimos correspondem à classe final, enquanto os outros dois tipos correspondem a um teste de feature. O objetivo final desse método é desenvolver uma árvore de testes que consiga generalizar o processo de classificação, mesmo para pontos ainda não conhecidos (ZHOU, 2016). A estrutura básica de uma DT pode ser vista na Figura 5.5.

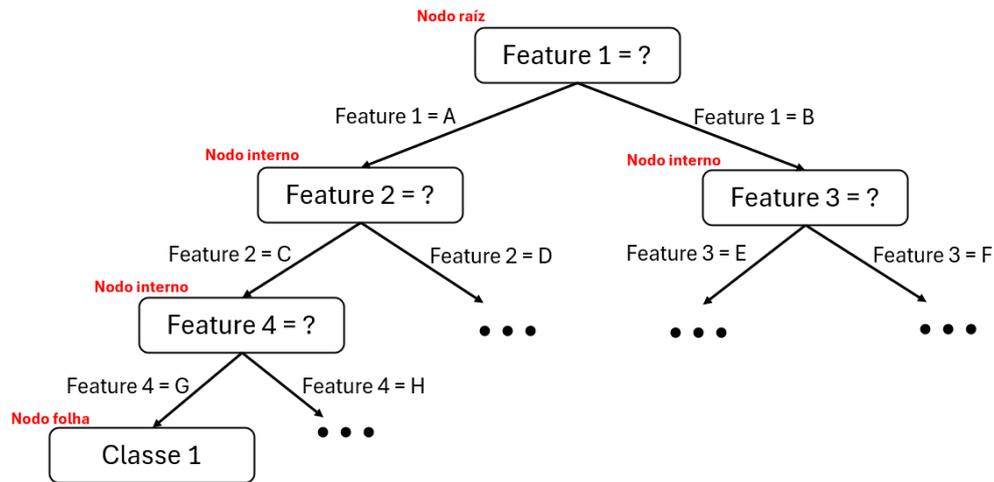


Figura 5.5 – representação da estrutura de uma árvore de decisão. Adaptado de ZHOU (2016).

A lógica básica de uma DT é (ZHOU, 2016):

1. Partir de um conjunto de treinamento  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$  e um conjunto de features  $A = \{a_1, a_2, \dots, a_d\}$ ;
2. Utilizar uma função  $f$  de criação de árvore que consiste em:
  - 2.1. Gerar um nodo  $i$ , que será classificado como um nodo folha se todos as amostras de  $D$  pertencem à mesma classe  $C$  e encerrar;
  - 2.2. Se  $A = \emptyset$  ou se todos as amostras tiverem o mesmo valor em  $A$ , definir o nodo  $i$  como um nodo folha cujo rótulo será o rótulo da maioria dos dados e encerrar;
  - 2.3. Selecionar a feature ótima  $a_*$  do conjunto  $A$ 
    - 2.3.1. Criar uma bifurcação para cada valor  $a_*^v$  possível da feature  $a_*$ , em que  $D_v$  é o subconjunto de pontos que possuem esse valor  $a_*^v$  para a feature  $a_*$ ;
    - 2.3.2. Se  $D_v$  for um conjunto vazio para algum valor  $a_*^v$ , marcar esse nodo como um nodo folha cujo rótulo será o rótulo majoritário dos pontos em  $D$  e encerrar;
    - 2.3.3. Se ainda houver dados a classificar, repetir recursivamente a função  $f$  com os dados restantes, do contrário, encerrar;

Um fator decisivo na qualidade de um algoritmo de árvore de decisão é decidir quais são as *features* ótimas para se utilizar como critério de classificação em cada caso; para isso, é necessário ter em mente que o objetivo de um algoritmo desse tipo é fazer com que os pontos dentro de cada nó pertençam à mesma classe, ou seja, é aumentar a pureza desse nodo (ZHOU, 2016).

Uma das medidas mais utilizadas para se medir a pureza é a entropia. Seja  $p_k$  a proporção da classe  $k = 1, 2, \dots, |y|$  no conjunto de dados  $D$ . Dessa forma, a entropia pode ser definida como na Equação 5.69, em que quanto menor a entropia, maior a pureza do conjunto  $D$  (ZHOU, 2016).

$$Ent(D) = - \sum_{k=1}^{|y|} p_k \log_2 p_k \quad (5.69)$$

Supor que a feature discreta  $a$  possui  $V$  valores possíveis  $\{a^1, a^2, \dots, a^V\}$ , logo, dividir o conjunto de dados de acordo com a feature  $a$  irá gerar  $V$  nodos, em que  $v$ -ésimo nodo  $D^v$  possuirá todas as amostras com o valor  $a^v$ . Como pode haver diferentes quantidades de pontos em cada nodo, é relevante estabelecer uma métrica que leve em conta essa diferença, ou seja, quanto maior o número de elementos em um nodo, maior a relevância desse nodo. Logo, o ganho de informação obtido por se dividir o conjunto de acordo com a feature  $a$  é calculado de acordo com a Equação 5.70, de modo que quanto maior o ganho, maior aumento de pureza pode-se esperar vindo dessa operação. Logo, deve-se selecionar a feature que promover o maior ganho (ZHOU, 2016).

$$Gain(D, a) = Ent(D) - \sum_{v=1}^V \frac{|D^v|}{|D|} Ent(D^v) \quad (5.70)$$

Contudo, esse método é tendencioso ao priorizar features com mais valores possíveis. Por isso, ao invés de utilizar o ganho de informação em si, é mais vantajoso utilizar a taxa de ganho, dada pela Equação 5.71 (BISHOP, 2006).

$$Gain_{ratio}(D, a) = \frac{Gain(D, a)}{IV(a)} \quad (5.71)$$

Em que  $IV(a)$  é chamado de valor intrínseco da feature  $a$  e definido na Equação 5.72; seu valor será maior quando a feature possuir vários valores possíveis (BISHOP, 2006).

$$IV(a) = - \sum_{v=1}^V \frac{|D^v|}{|D|} \log_2 \frac{|D^v|}{|D|} \quad (5.72)$$

Como esse método é tendencioso para features com menos valores possíveis, é utilizada uma heurística que seleciona aquela que possua a maior taxa de ganho dentro de um subconjunto de features que possuem ganho acima da média (BISHOP, 2006).

## CAPÍTULO VI

### 6 ANÁLISE DE COMPONENTES PRINCIPAIS

A análise de componentes principais (PCA) é provavelmente a técnica estatística multivariável mais famosa possivelmente a mais antiga, com suas origens ligadas a Pearson ou Cauchy. PCA analisa uma tabela de dados que representam observações descritas por variáveis que são, em geral, inter-relacionadas. Seu objetivo é extrair informações importantes a respeito dos dados fornecidos na tabela e expressar essa informação como um novo conjunto de variáveis ortogonais chamadas de componentes principais, sendo capaz também de representar esses pontos num espaço cartesiano (ABDI e WILLIAMS, 2010). Mais especificamente, os objetivos do método PCA são:

- Extrair as informações mais importantes do conjunto de dados;
- Comprimir o tamanho dos dados mantendo apenas as informações relevantes;
- Simplificar a descrição do conjunto de dados;
- Analisar a estrutura das observações e das variáveis.

#### 6.1 DECOMPOSIÇÃO DE VALOR SINGULAR (SVD)

É um método altamente relevante para a aplicação do PCA; ele tem a capacidade de decompor uma matriz retangular em 3 matrizes: duas ortogonais e uma diagonal. Se uma matriz  $A$  é retangular, seu SVD resulta na Equação 6.1 (ABDI e WILLIAMS, 2010).

$$A = P\Delta Q^T \quad (6.1)$$

- $P$ : autovetores normalizados da matriz  $AA^T$ ; as colunas de  $P$  são chamadas de vetores singulares esquerdos de  $A$ . É válido salientar que  $P^T P = I$ .
- $Q$ : autovetores normalizados da matriz  $A^T A$ ; as colunas de  $Q$  são chamadas de vetores singulares de  $A$ . É válido salientar que  $Q^T Q = I$ .
- $\Delta$ : matriz diagonal dos valores singulares;  $\Delta = \Lambda^{\frac{1}{2}}$ , em que  $\Lambda$  corresponde a matriz diagonal dos autovalores da matriz  $AA^T$  (e da matriz  $A^T A$ , visto que são os mesmos)

## 6.2 PRÉ-REQUISITOS PARA APLICAÇÃO DO MÉTODO

Para ser analisado pelo método PCA, a tabela com os dados deve se composta por  $I$  observações descritas por  $J$  variáveis organizadas em uma matriz  $X$  de dimensões  $I \times J$ , com elemento genérico  $x_{ij}$  e posto  $L$  em que  $L \leq \min \{I, J\}$ . Além disso, em geral as colunas de  $X$  deverão ser centralizadas de modo que a média de cada coluna seja igual a 0. Por fim, se cada elemento da matriz for dividido por  $\sqrt{I}$ , a análise é chamada de PCA de covariância, visto que a matriz  $X^T X$  corresponde à matriz de covariância. Já se as variáveis possuem unidades diferentes, se os valores forem normalizados, a análise é dita PCA de correlação, visto que a matriz  $X^T X$  será a matriz de correlação (ABDI e WILLIAMS, 2010).

A matriz  $X$  possui a seguinte decomposição de valor singular

$$X = P\Delta Q^T$$

Em que  $P$  é uma matriz de dimensões  $I \times J$  de vetores singulares esquerdos,  $Q$  é a matriz de dimensões  $J \times L$  de vetores singulares direitos e  $\Delta$  é a matriz diagonal de valores singulares (ABDI e WILLIAMS, 2010).

Além disso, a inércia de uma coluna é definida como a soma dos quadrados dos elementos que a compõem, conforme a Equação 6.2. Nessa, a soma de todos os valores  $\gamma_i^2$  é denotada por  $\Gamma$  e é chamada de inércia total, ou inércia da tabela (ABDI e WILLIAMS, 2010).

$$\gamma_i^2 = \sum_j x_{i,j}^2 \quad (6.2)$$

O centro de gravidade das linhas, também chamado de centroide ou baricentro, é denotado por  $g$  e é o vetor das médias de cada coluna de  $X$ . Quando a matriz já foi centralizada, seu centro de gravidade é igual ao vetor  $0^T$  de dimensões  $1 \times J$  (ABDI e WILLIAMS, 2010).

A distância euclidiana da  $i$ -ésima observação ao centroide é dada pela Equação 6.3, que se reduz à Equação 6.4 caso os dados estejam centralizados (ABDI e WILLIAMS, 2010).

$$d_{i,g}^2 = \sum_j^J (x_{i,j} - g_j)^2 \quad (6.3)$$

$$d_{i,g}^2 = \sum_j^J x_{i,j}^2 \quad (6.4)$$

### 6.3 ENCONTRANDO OS COMPONENTES

No PCA, os componentes são obtidos a partir da decomposição de valor singular; mais especificamente, a matriz de scores denotada por  $F$  de dimensões  $I \times J$  é dada pela Equação 6.5 (ABDI e WILLIAMS, 2010).

$$F = P\Delta \quad (6.5)$$

É válido, a partir dessa definição, a Equação 6.6.

$$F = XQ \quad (6.6)$$

Essa matriz  $Q$  é chamada de matriz de projeção, devido ao fato de que multiplicar  $X$  por  $Q$  fornece as projeções dos valores das observações na direção dos componentes principais. Esses componentes podem ser compreendidos geometricamente como o

resultado da rotação dos eixos das variáveis. Essa relação pode ser vista na Figura 6.1 (ABDI e WILLIAMS, 2010).

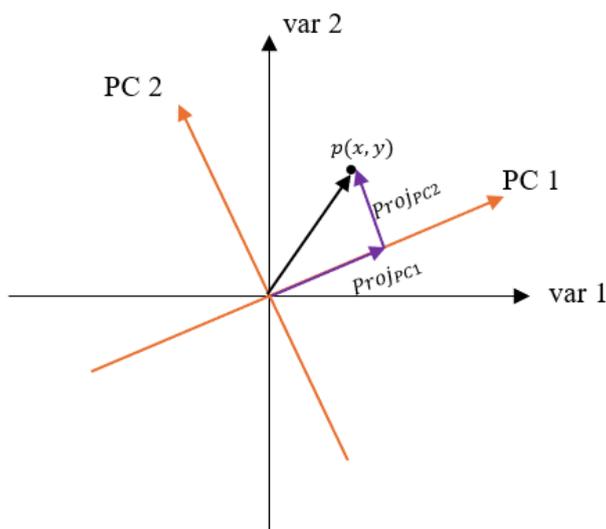


Figura 6.1 – Representação geométrica da relação entre os dois primeiros componentes principais (eixos em laranja) e as variáveis originais (eixos em preto).

Nesse contexto, a matriz  $Q$  é chamada de matriz de carregamento, e a matriz  $X$  pode ser interpretada como o produto da matriz de *scores* por essa matriz de carregamento. Ou seja, a matriz  $Q$ , composta pelos autovetores que definem os componentes principais pode ser compreendida como uma matriz de mudança de base (ABDI e WILLIAMS, 2010).

A importância de cada componentes principal é refletida por sua inércia que, quando comparada à inércia total, traduz o percentual da variação entre os dados que é explicado pelo componente principal em questão. Os autovalores ficam localizados na matriz diagonal  $\Delta$  (ABDI e WILLIAMS, 2010).

## CAPÍTULO VII

### 7 TEORIA DA INFORMAÇÃO

O problema fundamental da comunicação é o de reproduzir, em um ponto, uma mensagem que foi selecionada em outro ponto, seja de modo exato ou aproximado; frequentemente essas mensagens possuem significado, mas do ponto de vista da engenharia, o seu aspecto semântico é irrelevante. O ponto relevante é que a mensagem seja selecionada a partir de um conjunto de mensagens possíveis e o sistema de comunicação deve estar apto a trabalhar com todas as mensagens pensadas durante sua concepção (SHANNON e WEAVER, 1949).

Se o número de mensagens possíveis for finito, então esse valor ou qualquer função monotônica desse número pode ser considerada uma medida de informação produzida quando uma mensagem é escolhida desse conjunto, considerando que todas as escolhas são igualmente possíveis. No âmbito da Teoria da Informação, a escolha mais natural é a da função logarítmica em função da sua adequação às medidas utilizadas na engenharia. A escolha da base do logaritmo corresponde à escolha da unidade de medida de informação; se a base for 2, as unidades de informação são chamadas de bits ou seja, um dispositivo que pode ocupar duas posições (como um interruptor) é capaz de armazenar uma unidade de informação (SHANNON e WEAVER, 1949).

Por sistema de comunicação, entenda-se um sistema similar ao exibido na Figura 7.1 composto primordialmente por cinco componentes:

1. Fonte de informação: unidade responsável pela produção de uma mensagem ou sequência de mensagens que serão comunicadas ao terminal de recebimento; as mensagens podem ser de vários tipos, como: sequências de letras e funções;
2. Transmissor: dispositivo que realiza operações na mensagem de modo a produzir um sinal adequado para transmissão através do canal;

3. Canal: corresponde ao meio utilizado para transmitir a mensagem do transmissor ao receptor; a etapa na qual o sinal atravessa o canal é susceptível à presença de ruídos;
4. Receptor: unidade responsável por realizar a operação inversa do transmissor, reconstruindo a mensagem a partir do sinal recebido;
5. Destinatário: pessoa ou coisa a quem se destina a mensagem enviada.

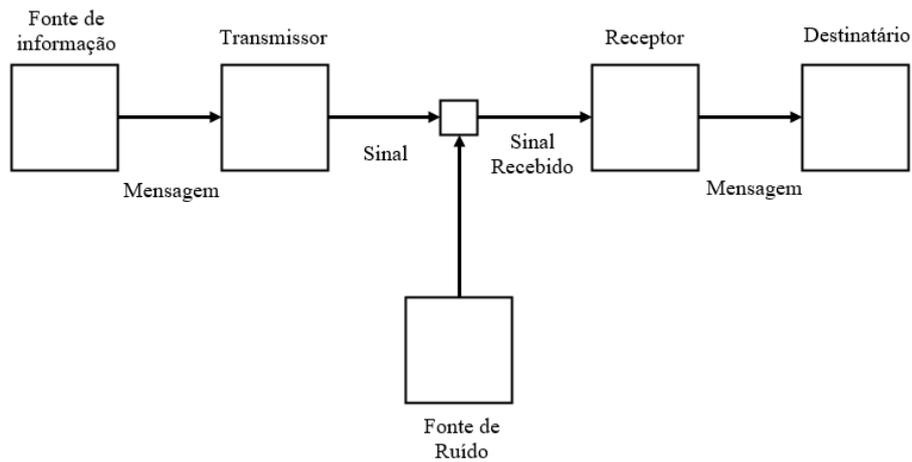


Figura 7.1 – Principais componentes de um sistema de comunicação. Adaptado de SHANNON e WEAVER (1949).

Em geral, sistemas de comunicação podem ser classificados como discretos, contínuos ou mistos. Os sistemas discretos são aqueles em que as mensagens e os sinais são seqüências de símbolos discretos. Já os sistemas contínuos são aqueles em que a mensagem e o sinal são funções contínuas. Por fim, um sistema misto é aquele em que a mensagem e o sinal podem ser variáveis discretas ou contínuas (SHANNON e WEAVER, 1949).

Pode-se pensar em uma fonte discreta como aquela que gera uma mensagem símbolo a símbolo; a escolha desses símbolos é feita de acordo com certas probabilidades, logo, esse sistema é governado por um processo estocástico. De fato, qualquer processo estocástico que produz uma série de símbolos escolhidos a partir de um conjunto finito pode ser considerado uma fonte discreta de informação (SHANNON e WEAVER, 1949).

A definição de um sistema de informação traz consigo questionamentos a respeito da possibilidade de se medir a quantidade de informação produzida, bem como a taxa com a qual isso ocorre. Para avaliar isso, considerar um conjunto de eventos possíveis cujas

probabilidades de ocorrência são  $p_1, p_2, \dots, p_n$ ; é possível medir a quantidade de “escolha” ou liberdade de escolha que se tem a respeito da saída fornecida por essa fonte de informação? Se essa medida, denotada por  $H(p_1, p_2, \dots, p_n)$  de fato, existir, ela deve ser dotada das seguintes características (SHANNON e WEAVER, 1949)

1.  $H$  deve ser contínua em  $p_i$ ;
2. Se todos os  $p_i$  forem iguais,  $H$  deve ser uma função monotonicamente crescente de  $n$ , ou seja, se todos os eventos são igualmente prováveis, quanto mais eventos forem possíveis, maior a liberdade de escolha;
3. Se uma escolha for desmembrada em duas escolhas consecutivas, o valor original de  $H$  deve ser a soma ponderada dos valores individuais de  $H$ .

A equação que satisfaz essas 3 condições é a Equação 7.1, em que  $K$  é uma constante positiva (SHANNON e WEAVER, 1949).

$$H = -K \sum_{i=1}^n p_i \log p_i \quad (7.1)$$

Expressões nesse formato indicam a entropia  $H$  do conjunto de probabilidades  $p_1, p_2, \dots, p_n$  em que  $p_i$  indica a probabilidade de o sistema se encontrar no estado  $i$ ; em outras palavras, pode-se entender  $p_i$  como sendo a probabilidade de a mensagem gerada por uma fonte de informação ser  $i$ . De acordo com essa equação, a entropia no caso de haver dois estados possíveis, com probabilidades iguais a  $p$  e  $q = p - 1$  é igual a  $H = -(p \log p + q \log q)$  (SHANNON e WEAVER, 1949).

O valor de  $H$  em função da probabilidade  $p$  pode ser visto na Figura 7.2, o que indica que  $H$  será máximo quando as probabilidades dos eventos forem iguais (SHANNON e WEAVER, 1949):

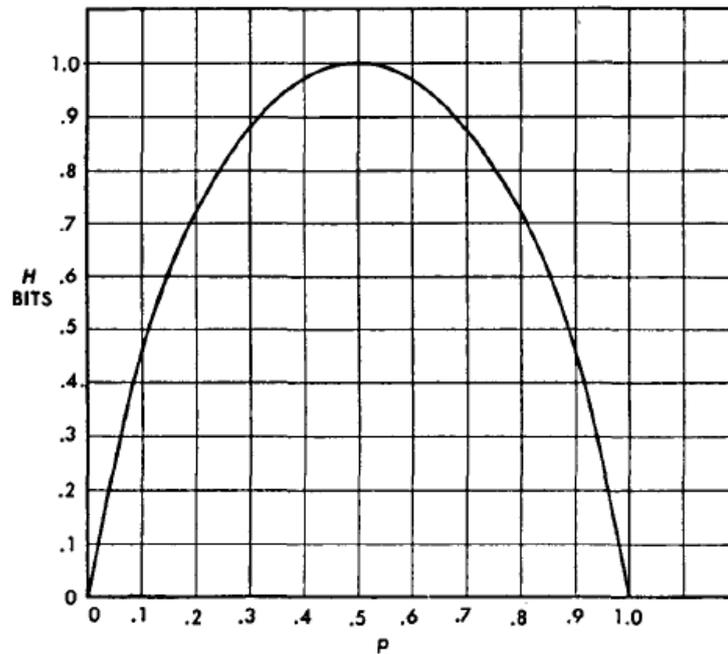


Figura 7.2 – Diagrama que relaciona entropia (H) e a probabilidade (p) de um sistema se encontrar em um dado estado de um universo de dois estados possíveis (SHANNON e WEAVER, 1949).

A entropia possui propriedades que comprovam o fato de que sua escolha como medida da quantidade de informação é adequada (SHANNON e WEAVER, 1949):

1.  $H = 0$  se e somente se todos os valores  $p_i$  forem iguais a 0 com exceção de um; ou seja, a entropia só será 0 quando se tiver absoluta certeza de saída informada por um sistema;
2. Para uma dada quantidade de estados possíveis  $n$ ,  $H$  será máximo e igual a  $\log n$  quando todos os valores de  $p_i$  forem iguais a  $\frac{1}{n}$ , o que corresponde à situação em que a incerteza é a maior possível, ou seja, em que a previsão da próxima saída do sistema é a mais difícil;
3. Supor dois eventos  $x, y$  com  $m$  possibilidades para o primeiro e  $n$  possibilidades para o segundo. Seja  $p(i, j)$  a probabilidade da ocorrência conjunta de  $i$  e  $j$ . Logo, a entropia para esse evento será dada pela Equação 7.2.

$$H(x, y) = - \sum_{i,j} p(i, j) \log p(i, j) \quad (7.2)$$

Em que é válida a Equação 7.3.

$$H(x, y) \leq H(x) + H(y) \quad (7.3)$$

4. Qualquer mudança no sentido de aproximar ou igualar os valores de  $p_i$  resultam no aumento da entropia  $H$ ;
5. Para qualquer valor  $i$  que  $x$  possa assumir, há uma probabilidade condicional  $p_i(j)$  de que  $y$  terá o valor igual a  $j$ , que é dado pela Equação 7.4;

$$p_i(j) = \frac{p(i, j)}{\sum_i p(i, j)} \quad (7.4)$$

A entropia condicional de  $y$ , representada por  $H_x(y)$ , pode ser definida como a média das entropias de  $y$  para cada valor de  $x$  ponderada pela probabilidade da ocorrência de cada  $x$ , e é dada pela Equação 7.5;

$$H_x(y) = - \sum_{i,j} p(i, j) \log p_i(j) \quad (7.5)$$

Que indica que é válida a Equação 7.6

$$H(x, y) = H(x) + H_x(y) \quad (7.6)$$

Dos tópicos 3 e 5, é possível concluir que é válida a Equação 7.7

-

$$H(y) \geq H_x(y) \quad (7.7)$$

Que quer dizer que a incerteza a respeito de  $y$  nunca aumenta quando se tem conhecimento de  $x$ , ela diminuirá a não ser que  $x$  e  $y$  sejam variáveis independentes, caso em que ela não muda (SHANNON e WEAVER, 1949).

## CAPÍTULO VIII

### 8 MATERIAIS E MÉTODO

#### 8.1 MATERIAIS UTILIZADOS

Inicialmente gasolina foi adquirida de um posto de combustível e fornecida ao veículo de teste para posterior mistura com etanol. A concentração inicial de etanol na gasolina foi de 29% em volume e o veículo utilizado foi um Renault Logan com características exibidas na Tabela 8.1. É importante ressaltar que esse carro estava equipado com tecnologia Flex, o que permite que o mesmo opere tanto com gasolina quanto com álcool, de modo que a adulteração não representaria riscos à integridade dos componentes do sistema de powertrain.

<b>Versão</b>	<b>1.6 16V Flex Manual 5 gears</b>
<b>Type of engine</b>	H4M
<b>Admissão de Ar</b>	Natural
<b>Cilindros</b>	4 unidades em linha
<b>Tipo de combustível</b>	FLEXFUEL System
<b>Cilindrada (cm<sup>3</sup>)</b>	1,598
<b>Diâmetro do Cilindro (mm)</b>	78,0
<b>Curso do Pistão (mm)</b>	83,6
<b>Torque Máximo (kgf ↔ rpm) gasolina</b>	16 ↔ 4000
<b>Torque Máximo (kgf ↔ rpm) etanol</b>	16 ↔ 4000
<b>Potência Máxima (kgf ↔ rpm) gasolina</b>	115 ↔ 5500
<b>Potência Máxima (kgf ↔ rpm) etanol</b>	118 ↔ 5500
<b>Velocidade angular do motor em marcha lenta (rpm)</b>	750 ± 50

Tabela 8.1 – Ficha técnica com as informações mais relevantes a respeito do veículo utilizado para testes.

Uma câmera térmica de modelo Fluke Ti300 com emissividade definida em 0,95, mostrada na Figura 8.1, foi posicionada a uma distância de 2m do motor, apontada para esse, tal qual exibido na Figura 8.2.



Figura 8.1 - Câmera térmica utilizada para realização do experimento.



Figura 8.2 - Setup utilizado para o experimento.

## **8.2 MEDIÇÃO DO PERCENTUAL DE ETANOL**

Além disso, o método utilizado para definir o percentual de etanol na gasolina é aquele descrito na ABNT NBR 13992. Os resultados obtidos por esse método também foram comparados com os valores obtidos a partir da medição realizada com o auxílio de um sensor flex fuel da Continental de modelo MPL-3303M. A utilização desses dois métodos garante a consistência dos resultados e permite a identificação de erros de medição. Entretanto, para fins práticos, os valores obtidos pelo método descrito pela ABNT NBR 13992:2015 foram utilizados como base para aplicação dos algoritmos de ML.

### **8.2.1 Medição do percentual de etanol pelo método da ABNT NBR 13992:2015**

A ABNT NBR 13992:2015 prescreve o método para determinação do teor de etanol anidro combustível (EAC) a partir de 1% em volume em gasolinas automotivas. Para isso, utiliza-se como reagentes cloreto de sódio, água destilada. Já a aparelhagem utilizada consiste de uma proveta de vidro de 100mL, graduada, com tampa e boca esmerilhadas, calibrada nos pontos 50 mL, 60 mL, 62 mL e 100 mL.

O método em si consiste em colocar 50 mL da amostra a ser analisada na proveta limpa, desengordurada e seca e em seguida adicionar solução de cloreto de sódio 10% m/v até o volume completar 100 mL. Em seguida, deve-se tampar a proveta e invertê-la 10 vezes para promover a extração de etanol. Aqui é importante não agitar a proveta no intuito de evitar a formação de emulsão.

Por fim, deixa-se a proveta em repouso por 10 minutos a fim de permitir a separação completa das duas fases e realiza-se a leitura da fase aquosa tomando como referência a base inferior do menisco formado na interface entre a fase aquosa e a gasolina, com os olhos posicionados na altura da linha da interface, tal qual ilustra a Figura 8.3. Deve-se anotar o volume final da fase aquosa em mililitros, com aproximação de 0,5 mL.

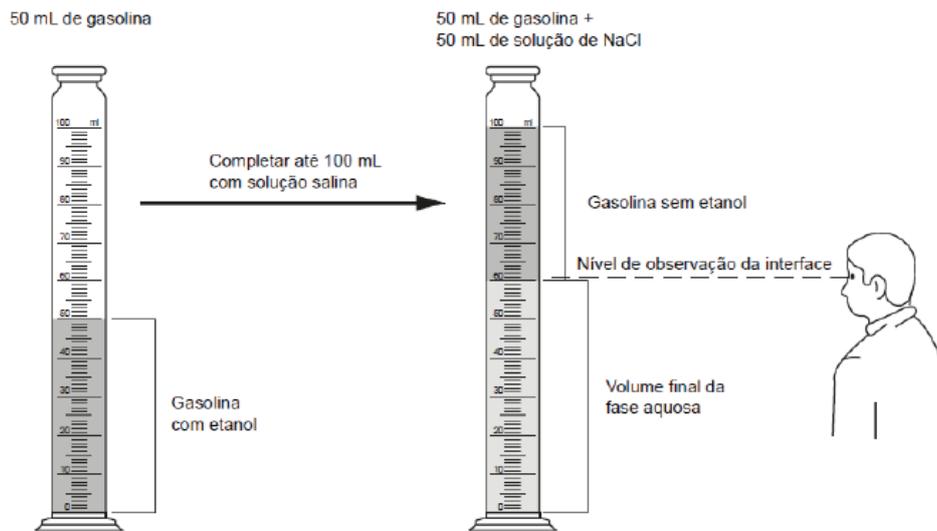


Figura 8.3 – Procedimento para leitura da fase aquosa (ABNT NBR 13992:2015)

Para calcular o teor de EAC presente na amostra, utiliza-se a Equação 8.1:

$$\% \text{ em volume de EAC} = [(volum\ e\ final\ da\ fase\ aquosa - 50) \times 2] + 1 \quad (8.1)$$

### 8.2.2 Medição do percentual de etanol pela utilização do sensor

Esse sensor é um dispositivo eletrônico com atualização contínua de leitura a cada 225ms capaz de medir a porcentagem de etanol total bem como a temperatura do combustível analisado. O sinal de saída do sensor é caracterizado por uma onda quadrada com frequências que variam de 50 Hz a 150 Hz, que correspondem respectivamente aos valores de 0% e de 100% de etanol na mistura. No caso de valores fora desse intervalo, é apresentada uma frequência de erro entre 170 Hz e 190 Hz. A Figura 8.4 ilustra a relação entre a frequência fornecida pelo sensor e o teor correspondente de álcool anidro na gasolina comum.

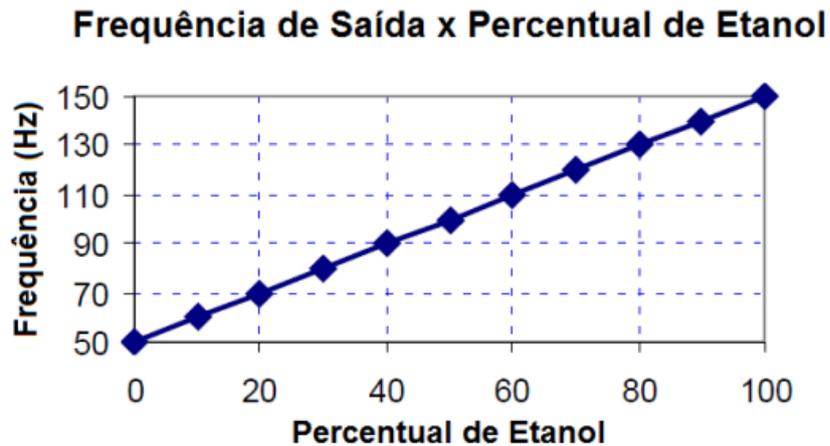


Figura 8.4 – Calibração do sinal de saída do sensor flex (CONTINENTAL, 2014)

### 8.3 AQUISIÇÃO DE FRAMES

A câmera infravermelha foi utilizada para gravar vídeos de cerca de 4 minutos de duração utilizando um colormap arco-íris do motor operando em neutro; o primeiro vídeo corresponde à gasolina tal qual essa foi obtida do posto de combustível. Em seguida, etanol foi adicionado à gasolina em quantidades de 500ml, com vídeos sendo gravados entre esses acréscimos; no total, 6 vídeos foram gravados. A câmera conseguia captar esse vídeos numa taxa de 9Hz, o que totaliza cerca de 2100 *frames* por vídeo obtido. No intuito de utilizar imagens que refletissem uma condição mais estável de operação do motor, apenas *frames* pertencentes à segunda metade de cada vídeo foram selecionados para esse experimento; mais especificamente, 500 *frames* de cada vídeo foram escolhidos, totalizando 3000 frames utilizados no processo de treinamento e teste dos algoritmos de ML.

### 8.4 PROCESSAMENTO DE IMAGEM

O método selecionado para processamento dos *frames* é baseado em técnicas consagradas na literatura referente a imagens térmicas. O primeiro passo utilizado foi a aplicação de um filtro de média em função de suas propriedades de redução de ruído, tal qual mencionado na seção 4.4.2. Em seguida, os frames foram convertidos da escala RGB para a escala de cinza no intuito de tornar possível a análise das intensidades dos *frames* e não as

cores dos seus pixels; essa conversão foi realizada seguindo a Equação 4.30 cuja elaboração foi descrita na Seção 4.5 e teve o intuito de facilitar a definição da região de interesse para análise bem como a de reduzir o esforço computacional para análise de cada frame.

O próximo desafio é lidar com o fato de que as imagens apresentavam pequenas variações de posição em função do vento presente no local de coleta dos vídeos e do manuseio da câmera em si. A consequência disso é que algumas porções das imagens, especialmente aquelas próximas às bordas, apareciam em alguns vídeos e não em outros; isso poderia gerar uma classificação falsa, pois o critério dos algoritmos poderia ser as mudanças nos descritores das imagens oriundos dessas mudanças de posição da câmera no momento de captura dos vídeos e não diferenças associadas à adulteração da gasolina no motor. Em função disso, os *frames* analisados foram submetidos a um processo de registro, baseado em transformações geométricas tal qual descrito na Seção 4.7. Aqui é válido salientar que apenas rotação, translação e interpolação para cálculo das intensidades foram utilizados. Além disso, todos os *frames* foram alinhados utilizando como base um *frame* manualmente selecionado em função de seu posicionamento adequado, que possibilitou o enquadramento dos componentes analisados.

Após o processo de registro das imagens estar finalizado, é necessário aplicar correções como a remoção de áreas pretas deixadas após o alinhamento. Isso foi feito a partir da localização dos pixels a partir dos quais não mais seria possível ver essas áreas indesejáveis que foram removidas da imagem obtida pós-registro; esse processo pode ser exemplificado na Figura 8.5.

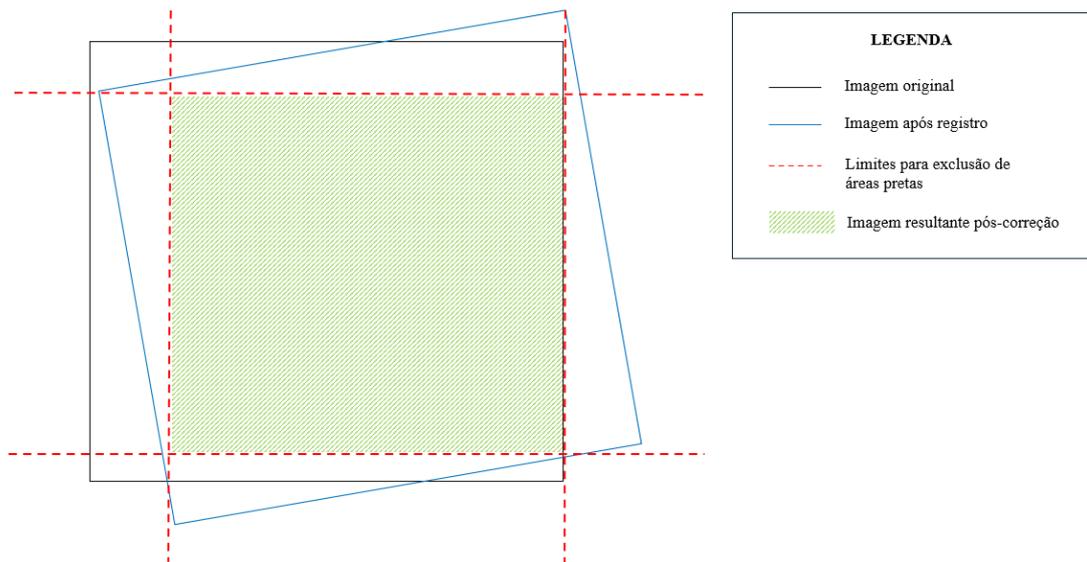


Figura 8.5 Representação gráfica do processo de correção da imagem obtida após as transformações geométricas.

A última etapa do processamento dos *frames* é a segmentação desses; mais especificamente, deseja-se definir uma região de interesse para análise que se diferencie do plano de fundo. Para isso, foi utilizada a estratégia de definir um valor limite de intensidade, ou *threshold*, a partir do qual os pixels são considerados como objeto e abaixo do qual são considerados plano de fundo. A definição desse limite se deu através do método de Otsu, conforme definido na seção 4.8.2. O esquema que ilustra o método adotado para processamento de imagens pode ser visto na Figura 8.6.

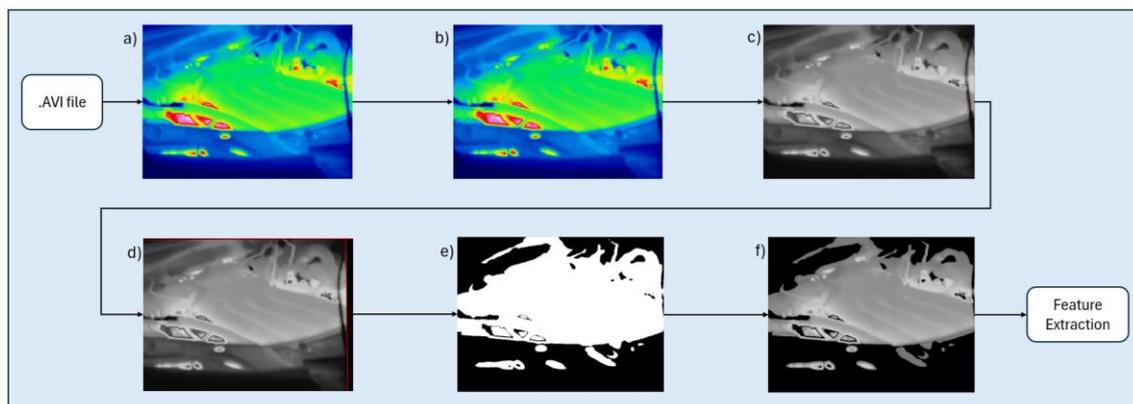


Figura 8.6 – Resumo do processo de processamento de imagem utilizado nesse estudo.

## 8.5 SELEÇÃO E ANÁLISE DE FEATURES

Uma vez que as imagens foram processadas, *features* que as representassem, também chamadas de descritores, foram selecionadas para análise. A literatura mostra que algumas das *features* mais relevantes na análise de imagens IR são: entropia, luminância média, desvio padrão da luminância e curtose e assimetria do histograma das imagens. Essas *features* correspondem a um espaço de 5 dimensões, logo, no intuito de simplificar as operações computacionais envolvidas no processamento desses dados, o método PCA tal qual descrito no Capítulo 6, foi aplicado no intuito de reduzir a dimensionalidade desses dados, traduzindo-os em novos componentes que melhor traduzem a variabilidade entre as amostras.

O conjunto de dados foi normalizado e em seguida o método PCA foi aplicado, de modo que o espaço pentadimensional foi reduzido a um novo espaço bidimensional. Esses novos componentes foram utilizados como dados de entrada para algoritmos de aprendizado de máquina: SVM, DT, KNN e RNA; a tarefa designada a esses algoritmos foi a de estimar a concentração de etanol na gasolina a partir dos componentes fornecidos. O processo utilizado no desenvolvimento desse estudo pode ser visualizado, de maneira sucinta, na Figura 8.7.

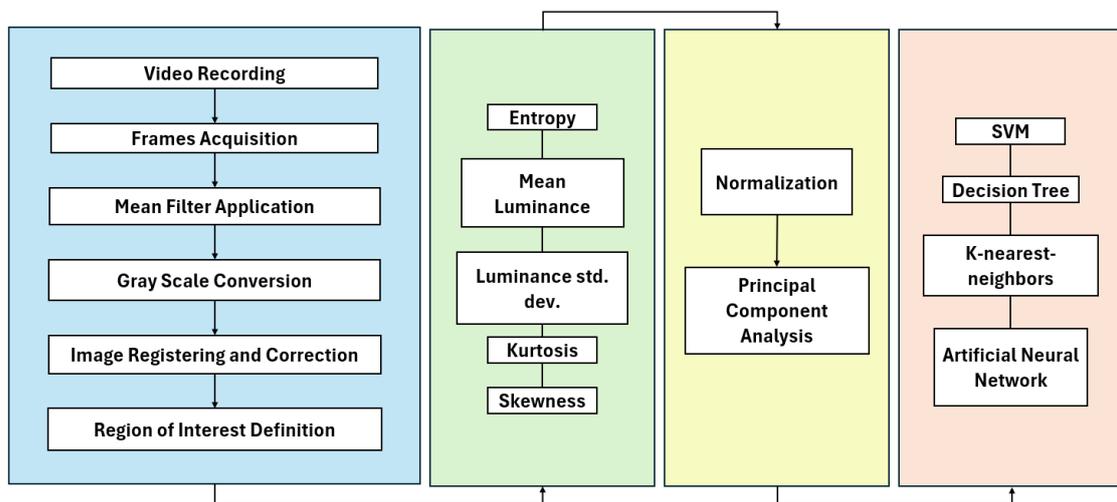


Figura 8.7 – Resumo do processo de coleta, processamento e análise de dados utilizado nesse estudo.

## 8.6 CONFIGURAÇÃO DOS MODELOS DE ML

O modelo de rede neural artificial utilizado contou com 3 camadas escondidas, cada uma com 64 neurônios e função de ativação ReLu; além disso, foi utilizada uma camada de saída com o número de neurônios igual ao número de classes e função de ativação softmax. Foram utilizadas 60 épocas para análise, bem como batches de tamanho 32; além disso, 80% dos dados foram utilizados para treinamento e 20% para teste. Por fim, o otimizador utilizado foi o Adam, a função de perda selecionada foi a `sparse_categorical_crossentropy` e a métrica selecionada foi a acurácia.

Já o modelo de support vector machine utiliza kernel linear para encontrar o hiperplano que melhor separa as classes nos dados. Além disso, 80% dos dados foram utilizados para treinamento e 20% para teste. Por fim, a métrica utilizada foi a acurácia.

No caso do algoritmo de K-nearest neighbors, foi utilizado um número  $k = 5$  de vizinhos e o classificador foi ajustado para identificar as  $k$  observações mais próximas em relação a cada amostra para fins de classificação. Além disso, 80% dos dados foram utilizados para treinamento e 20% para teste. Por fim, a métrica selecionada foi a acurácia.

O algoritmo decision tree também foi construído de modo que 80% dos dados foram utilizados para treinamento e 20% para teste. Além disso, a métrica utilizada para avaliação do modelo também foi a acurácia.

## CAPÍTULO IX

### 9 RESULTADOS E DISCUSSÃO

Conforme mencionado no Capítulo 8, os algoritmos selecionados para estudo foram treinados e testados com os dois componentes principais mais relevantes obtidos de um conjunto de dados composto por 3000 imagens com o objetivo de classificar esses *frames* em função do percentual de etanol em mistura com a gasolina. Os percentuais obtidos através do método descrito pela ABNT NBR 13992, que foram utilizados como referência para avaliar o desempenho dos algoritmos, estão expostos na Tabela 9.1.

<b>Nome do Vídeo</b>	<b>Adulteração (v/v)</b>
<b>MEA-1</b>	33%
<b>MEA-2</b>	35%
<b>MEA-3</b>	37%
<b>MEA-4</b>	41%
<b>MEA-5</b>	42%
<b>MEA-6</b>	45%

Tabela 9.1 – Percentuais de etanol medidos em cada uma das amostras analisadas.

O *kernel* escolhido para a aplicação do filtro de média foi de  $2 \times 2$ , o que resultou em bordas mais suaves e contornos mais orgânicos, que podem ser vistos na Figura 9.1. O processo seguinte, de conversão das imagens para escala de cinza, é necessário para aplicação do algoritmo de segmentação de Otsu e viabiliza a operação de registro, bem como as correções posteriores.

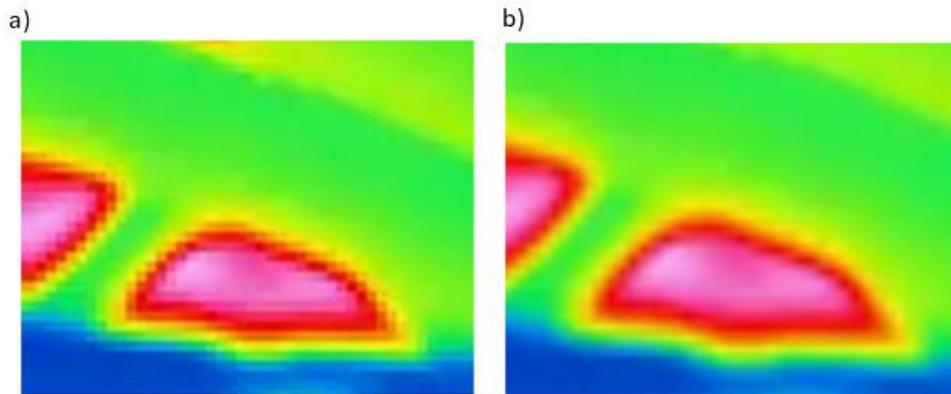


Figura 9.1 – a) Seção de imagem antes do processo de filtragem; b) mesma seção pós-filtragem.

Conforme esperado, o processo de registro de imagem e a correção desse traz consigo a perda de informação em função das partes das imagens que foram removidas junto às seções escuras geradas pelo alinhamento. Isso pode ser melhor visualizado na Figura 9.2, em que as linhas vermelhas correspondem aos limites da nova imagem após a implementação do algoritmo de correção. Já a Figura 9.3 exemplifica as diferenças entre uma imagem base e outra imagem antes e depois do registro respectivamente; as cores verde e roxo ressaltam as diferenças entre os *frames* e as partes em escala de cinza correspondem às semelhanças entre elas.

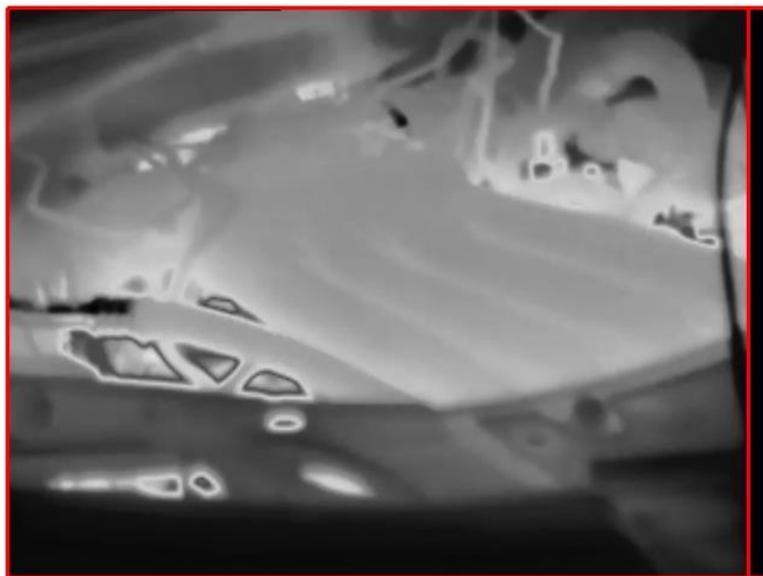


Figura 9.2 - Resultado do processo de registro da imagem, com linhas vermelhas indicando os limites das correções.

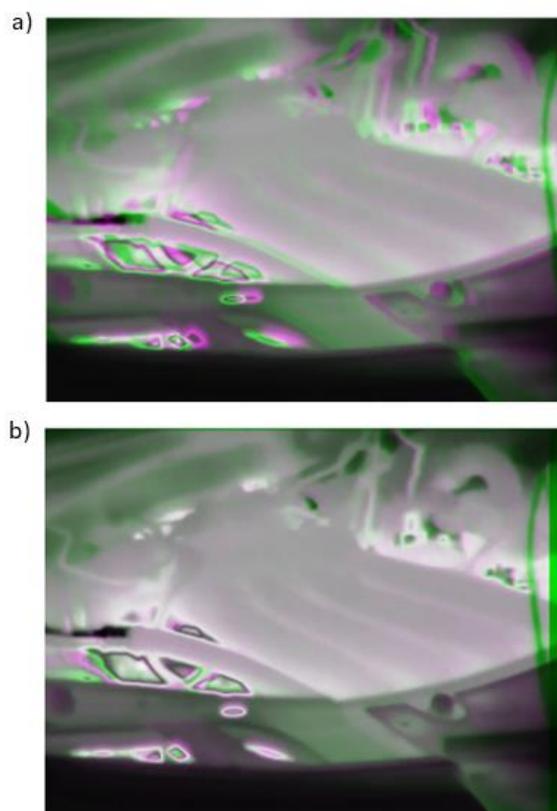


Figura 9.3 – Comparação entre duas imagens a) antes do processo de registro e b) após o processo de registro. As partes em cinza são comuns às duas imagens e as partes em verde e magenta correspondem às particularidades de cada uma.

A extração das features selecionadas gerou os resultados expostos nas Figuras 9.4 a 9.8. Esses gráficos corroboram com a escolha de algoritmos de ML para esse estudo, uma vez que não há um padrão facilmente identificável para um humano. Além disso, as curvas se interceptam, o que elimina a possibilidade de curvas diferentes ocuparem regiões diferentes do plano cartesiano. Nessas imagens, a nomenclatura “MEA-X” refere-se ao vídeo de número X; os percentuais de etanol correspondente a cada vídeo foram especificados anteriormente neste capítulo.

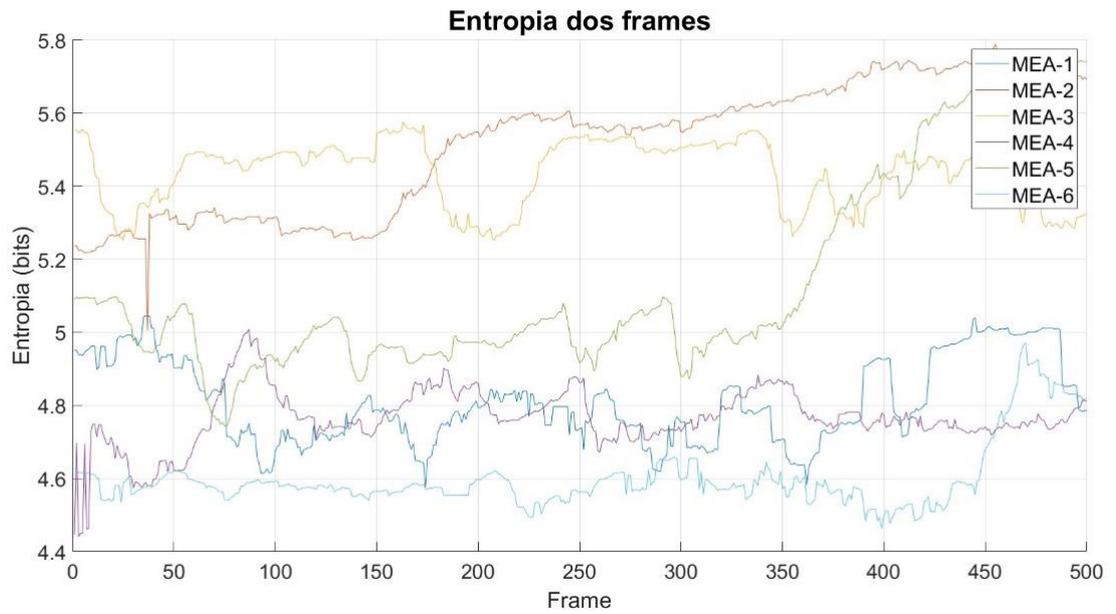


Figura 9.4 - Diagrama das entropias dos *frames* dos vídeos utilizados no estudo.

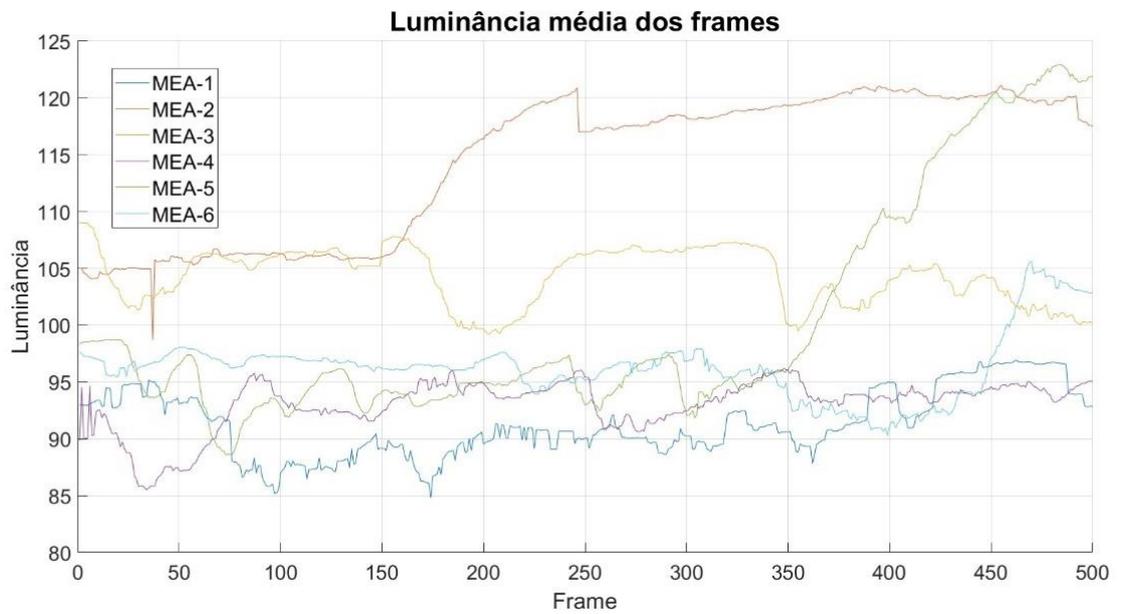


Figura 9.5 – Diagrama da luminância média dos pixels dos *frames* dos vídeos utilizados no estudo.

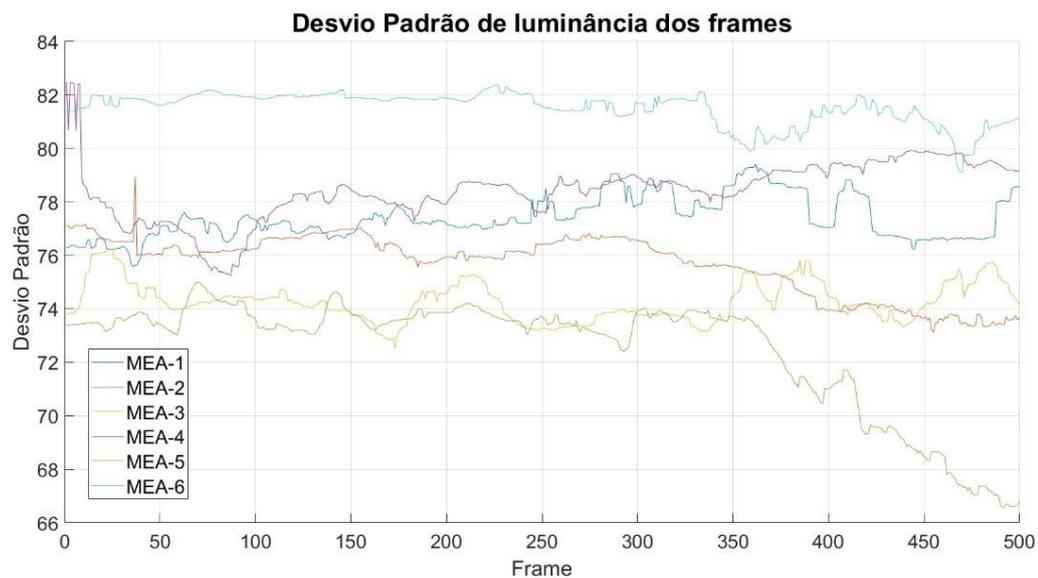


Figura 9.6 - Diagrama do desvio padrão de luminância dos pixels dos *frames* dos vídeos utilizados no estudo.

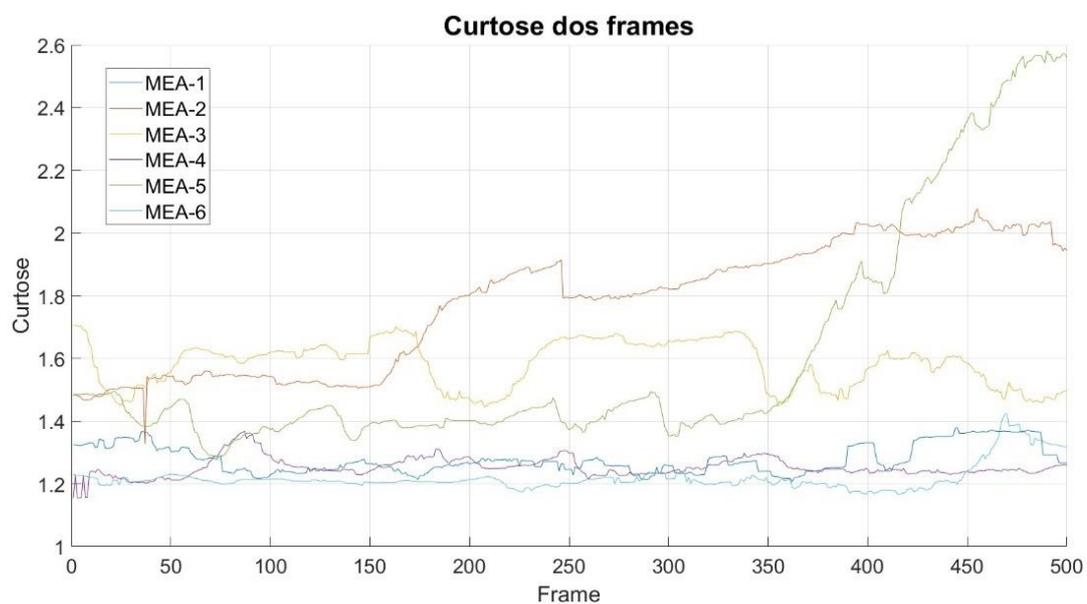


Figura 9.7 - Diagrama da curtose dos histogramas dos *frames* dos vídeos utilizados no estudo.

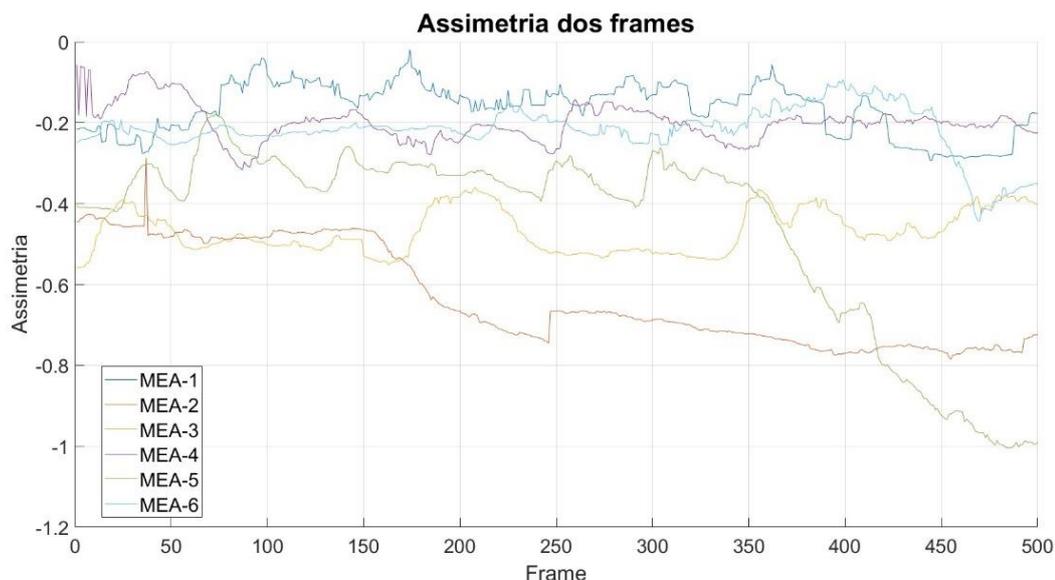


Figura 9.8 - Diagrama da curtose dos histogramas dos *frames* dos vídeos utilizados no estudo.

Após a extração das *features*, os dados foram normalizados e o método PCA foi aplicado; os resultados para os cinco componentes podem ser vistos na Tabela 9.2. Os dois componentes escolhidos explicam aproximadamente 97,66% da variação das amostras.

Componente	Percentual de Variância (%)
<b>Componente Principal 1</b>	88,341
<b>Componente Principal 2</b>	9,2401
<b>Componente Principal 3</b>	2,1124
<b>Componente Principal 4</b>	0,2883
<b>Componente Principal 5</b>	0,0111

Tabela 9.2 – Percentual da variância explicado por cada componente principal.

Uma vez que isso foi realizado, os componentes para todas as 3000 imagens foram tabelados e utilizados para treinar os algoritmos de ML, cujas matrizes de confusão e diagramas de fronteiras de decisão podem ser vistos nas Figuras 9.9 a 9.16. A métrica escolhida para avaliar a qualidade dos modelos foi a acurácia, e os resultados para cada um dos métodos podem ser vistos na Tabela 9.3. Apesar de que todos os modelos tiveram acurácia superior a 90%, é evidente que o SVM teve a pior performance, com acurácia de 90,8%, enquanto DT obteve o melhor resultado com acurácia de 97,2%.

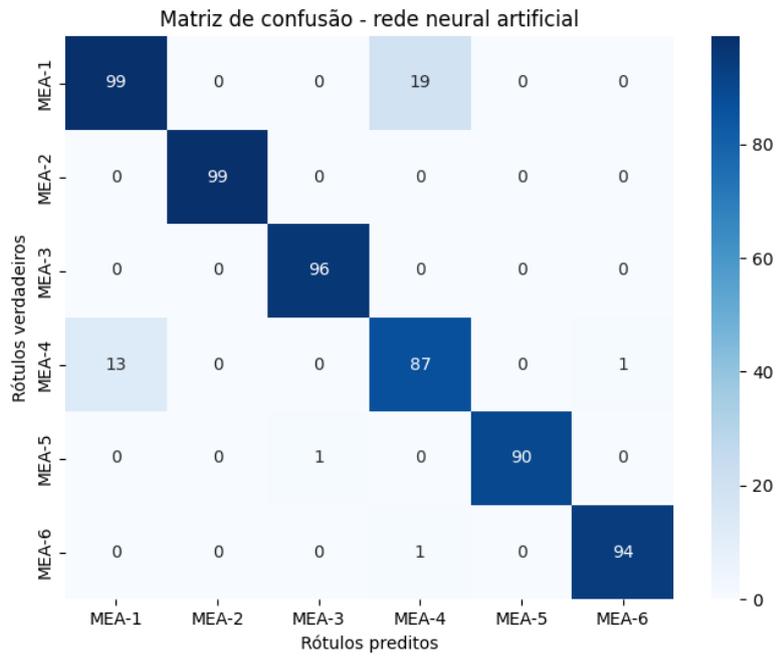


Figura 9.9 – Matriz de confusão para a rede neural artificial.

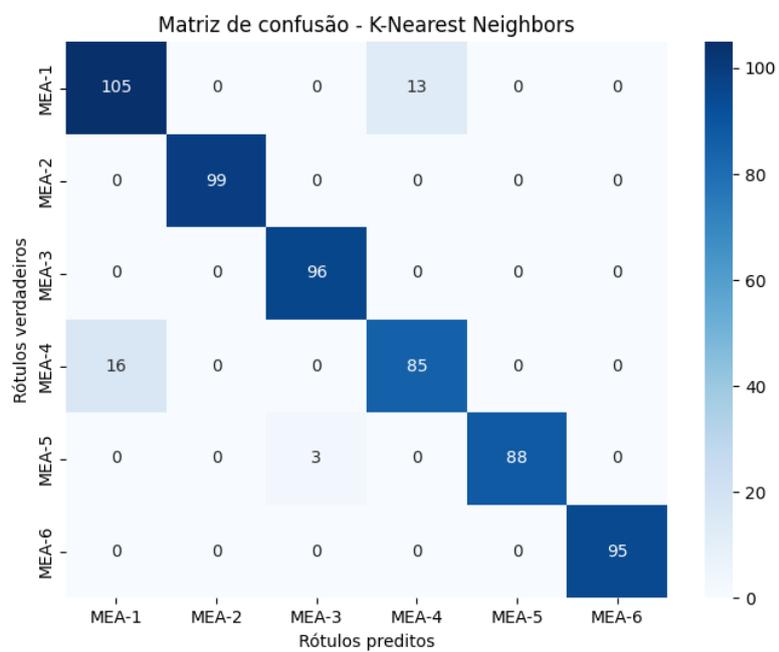


Figura 9.10 – Matriz de confusão para K-nearest neighbors.

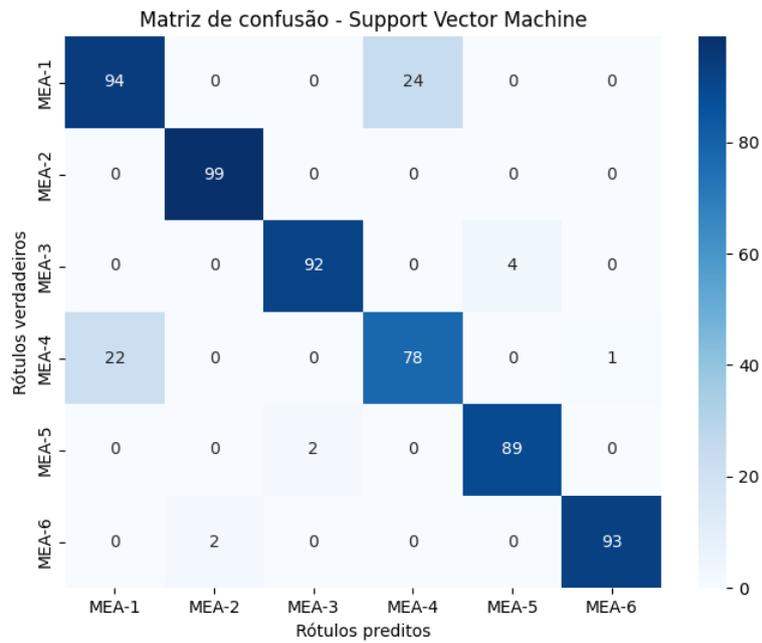


Figura 9.11 – Matriz de confusão para support vector machine.

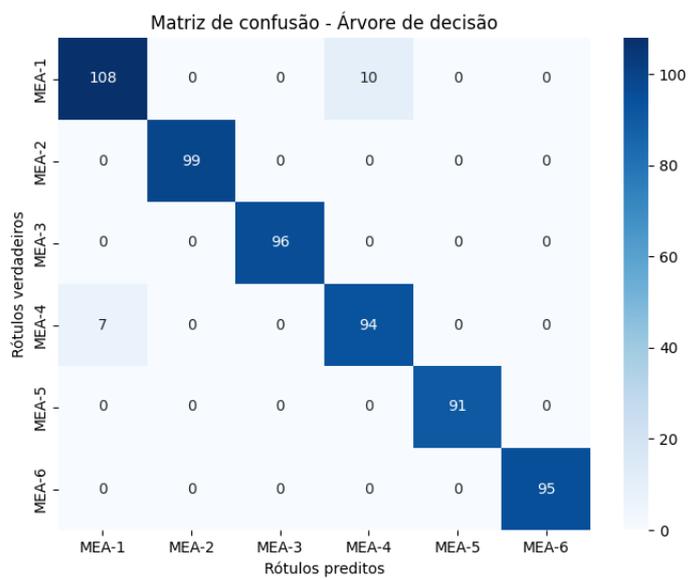


Figura 9.12 – Matriz de confusão para a árvore de decisão.

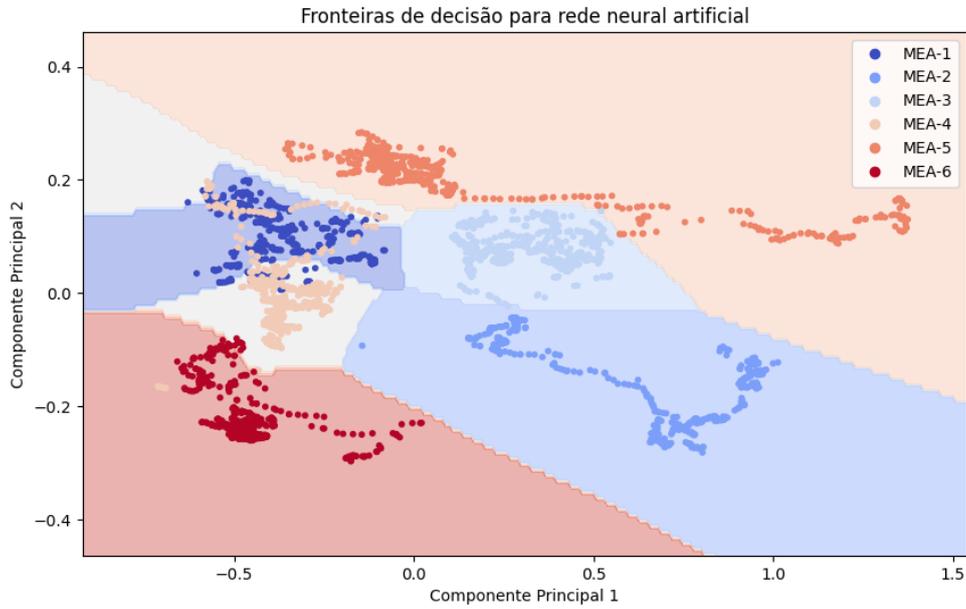


Figura 9.13 – Fronteiras de decisão para a rede neural artificial plotadas sobre o gráfico que relaciona os dois componentes principais mais relevantes.

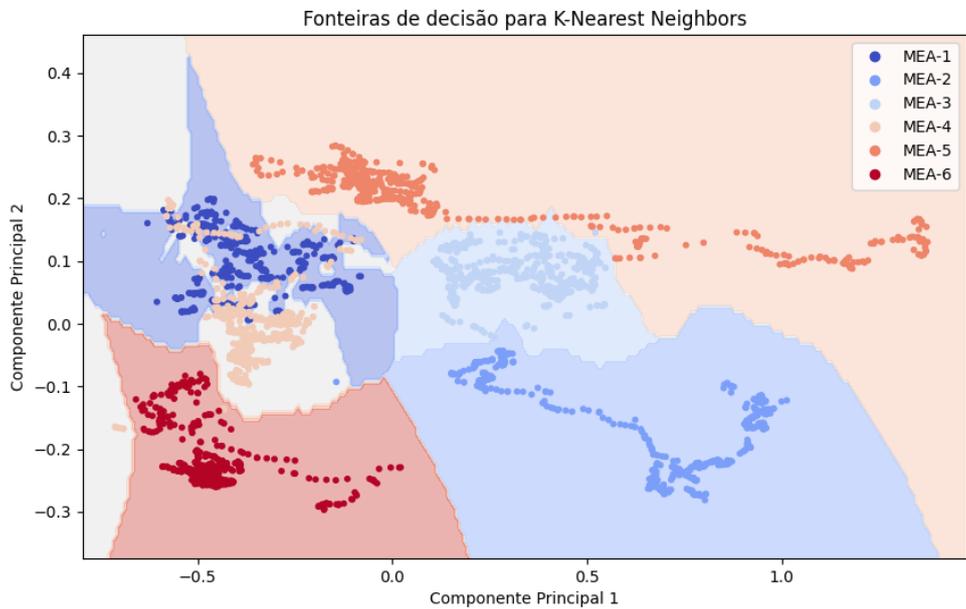


Figura 9.14 - Fronteiras de decisão para K-nearest neighbors plotadas sobre o gráfico que relaciona os dois componentes principais mais relevantes.

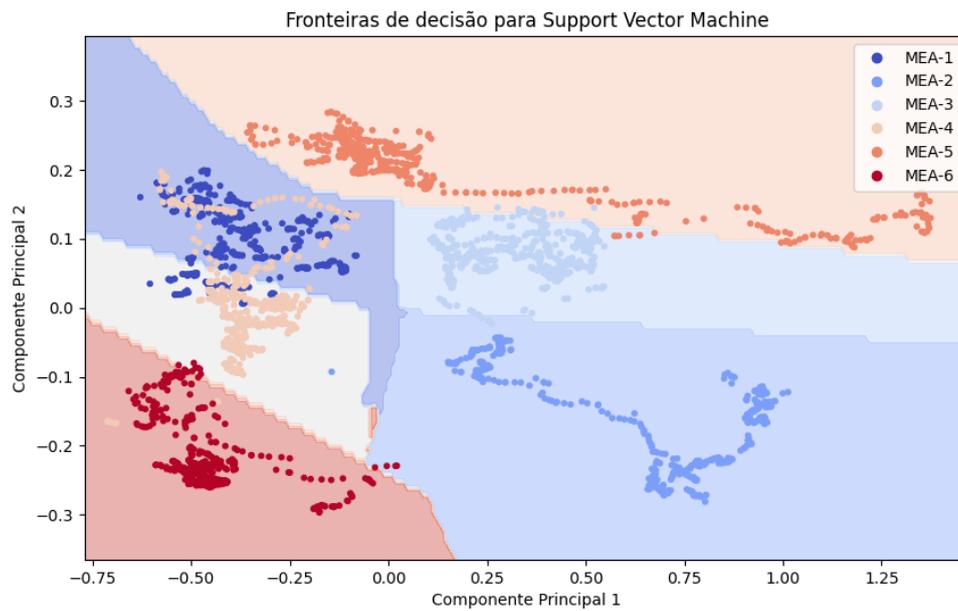


Figura 9.15 - Fronteiras de decisão para support vector machine plotadas sobre o gráfico que relaciona os dois componentes principais mais relevantes.

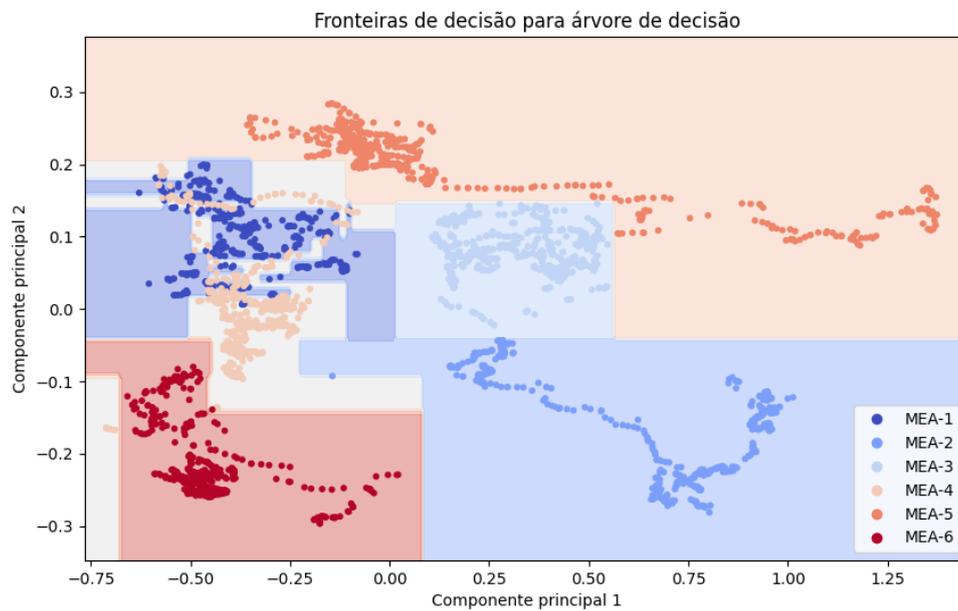


Figura 9.16 - Fronteiras de decisão para a árvore de decisão plotadas sobre o gráfico que relaciona os dois componentes principais mais relevantes.

<b>Modelo</b>	<b>Acurácia (%)</b>
<b>Rede Neural Artificial</b>	94,6
<b>Support Vector Machine</b>	90,8
<b>K-Nearest Neighbors</b>	94,7
<b>Árvore de Decisão</b>	97,2

Tabela 9.3 – Acurácias resultantes para cada um dos algoritmos estudados.

Os diagramas mostram que os pontos representado os *frames* relativos a cada percentual de adulteração ocupam áreas relativamente diferentes do espaço a depender da classe à qual pertencem, e esse fato torna possível a classificação pelos algoritmos de aprendizado de máquina com altas acurácias. Entretanto, é possível ver que apesar de classes diferentes estarem separadas, elas estão distribuídas de maneira irregular pelo diagrama, com algumas ocupando mais espaços do que outras, o que torna difícil estimar com precisão quais seriam de fato as fronteiras de cada classe, caso mais *frames* tivessem sido usados para treinar os algoritmos ou se mais percentuais de adulteração tivessem sido considerados.

Isso é relevante pois essa distribuição irregular pode fazer com que classes diferentes se interceptem, como nas classes MEA-1 e MEA-4, em que a maior taxa de erros de classificação ocorreu. Isso mostra que apesar de ser possível separar os *frames* desse estudo com base no percentual de etanol considerado, ainda há dificuldades no que diz respeito à confiabilidade do método, visto que algumas classes não estão completamente separadas no diagrama que considera os componentes principais.

Por fim, acurácias altas trazem consigo a preocupação quanto à possibilidade de overfitting. Para garantir que esse fenômeno não ocorreu, foram traçadas curvas de aprendizado para cada um dos algoritmos de ML utilizados, dispostos nas Figuras 17 a 20. Para traçar essas curvas, foi utilizada a estratégia de validação cruzada; aqui, os dados foram divididos em 5 grupos, e desses, quatro foram utilizados como conjuntos de treino e o quinto como conjunto de teste. As acurácias são calculadas 5 vezes, de modo que cada um dos 5 conjuntos atue como teste. Por fim, a média e o desvio padrão das acurácias são calculados para os conjuntos de treino e de validação; a primeira métrica é representada pelas linhas e a segunda é representada pelas regiões sombreadas em torno das linhas.

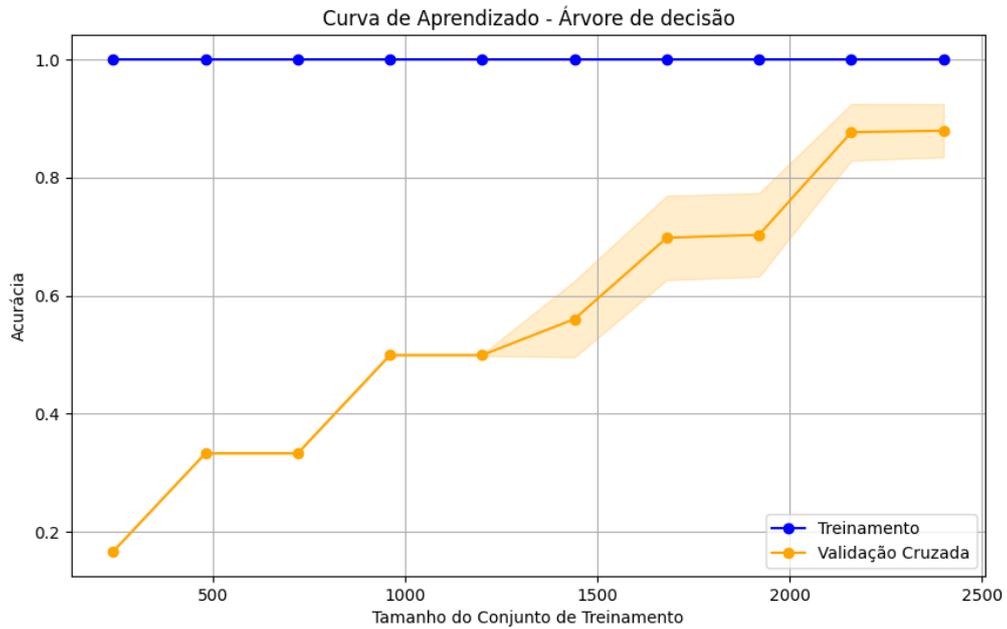


Figura 9.17 Curva de aprendizado para o algoritmo árvore de decisão.

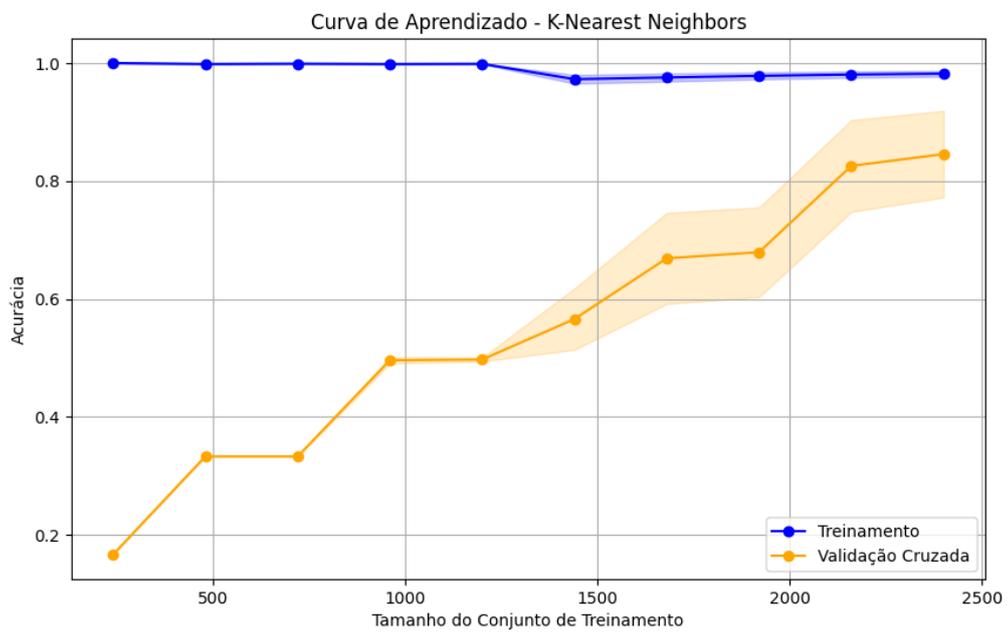


Figura 9.18 Curva de aprendizado para o algoritmo K-Nearest neighbors.

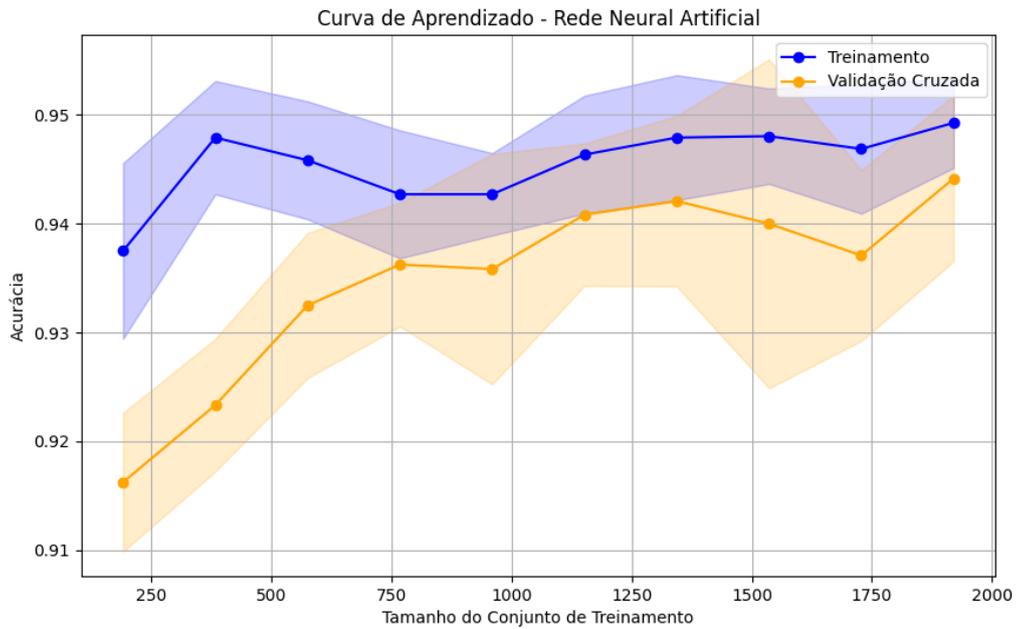


Figura 9.19 Curva de aprendizado para o algoritmo rede neural artificial.

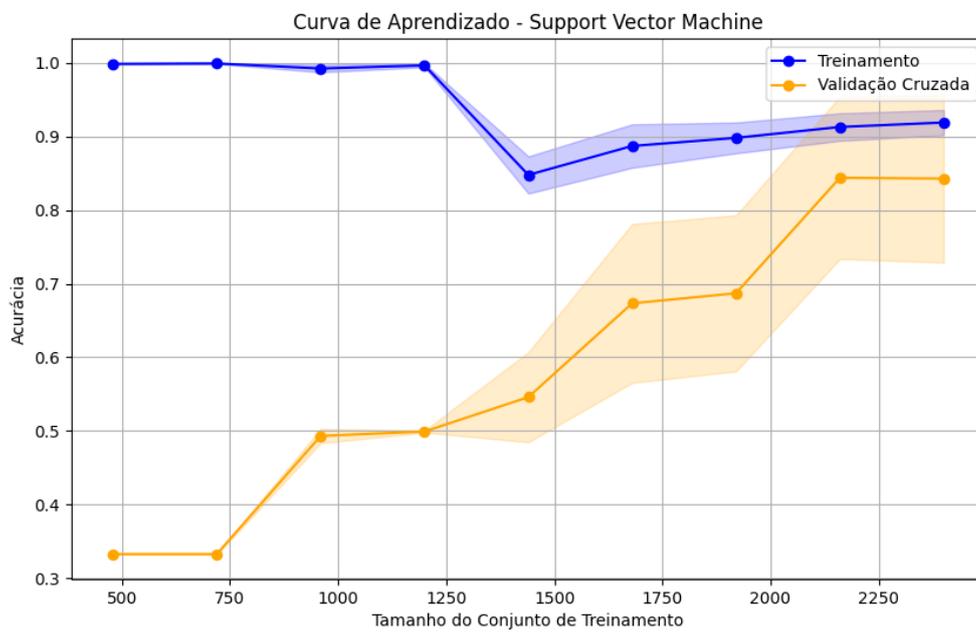


Figura 9.20 Curva de aprendizado para o algoritmo support vector machine.

As quatro curvas mostram que o overfitting não é um fenômeno significativo nos modelos em questão, uma vez que as curvas laranjas, de validação cruzada, indicam que a acurácia desses modelos aumenta com o tamanho dos conjuntos de treinamento. Isso mostra que à medida que os conjuntos aumentam, também aumenta a capacidade de generalização dos modelos; além disso, apesar de apresentarem alguma oscilação, as curvas azuis, referentes ao treinamento, se mantêm altas, o que nos permite também descartar o fenômeno do underfitting, pois vê-se que os modelos apresentam alta acurácia em relação aos conjuntos de treino.

## CAPÍTULO X

### 10 CONCLUSÃO

Um novo método para estimar o percentual de adulteração de gasolina por etanol baseado na análise de imagens IR do motor operando com a mistura de combustíveis foi proposto. Uma câmera térmica foi montada em um tripé virado para o motor operando e vídeos correspondentes às misturas com as concentrações analisadas foram gravados. Esses vídeos foram separados em *frames*, dos quais 3000 foram submetidos a uma série de métodos de processamento de imagem: aplicação de filtro de média para remoção de ruído, conversão de escala RGB para escala de cinza, registro de imagem e correção que levou em conta variações no alinhamento das imagens, e segmentação através do método de Otsu para definição de regiões de interesse. Após isso, a entropia, luminância média, desvio padrão da luminância, a curtose e a assimetria de cada um desses frames foram calculados e essas métricas foram aplicadas num algoritmo de PCA, o que resultou em dois principais componentes que representavam 97,66% da variação encontrada nos dados analisados. Esses dados foram usados para treinar quatro algoritmos de aprendizado de máquina: RNA, SVM, KNN e DT, que tiveram acurácias de 94,6%, 90,8%, 94,7% e 97,2% respectivamente. Por fim, a análise das curvas de aprendizado demonstrou a ausência de overfitting e de underfitting; pelo contrário, essas curvas permitem concluir que os modelos possuem boa capacidade de generalização.

O estudo mostrou que é possível classificar imagens IR de um motor com base na adulteração do combustível (gasolina) por etanol, o que indica que os perfis de temperatura criados durante a operação das máquinas mudam no que diz respeito aos parâmetros analisados. O estudo expande as possibilidades da detecção de comportamento anormal em sistemas mecânicos através da termografia IR ao mostrar a possibilidade de se identificar não apenas falhas em componentes, mas mudanças na eficiência e funcionamento das

máquinas; além disso, o trabalho fornece mais informações a respeito de como a variação da composição dos combustíveis pode afetar a performance do motor.

Entretanto, utilizar apenas esse método para quantificar a adulteração de gasolina por etanol em postos de combustível com objetivo de impor a lei, possivelmente através de mecanismos judiciais, ainda não é uma opção segura em função de algumas variáveis que são levadas em conta nessa análise, como o ângulo em que o vídeo é coletado, o modelo do veículo analisado, o efeito da temperatura inicial do motor e comportamento ainda não completamente compreendido das classes no diagrama de fronteiras de decisão, em que algumas classes se interceptam, o que poderia levar a classificações erradas. Apesar de ser uma ferramenta poderosa e importante para a compreensão dos efeitos da composição do combustível no funcionamento do veículo, sua utilização na identificação de adulteração ainda exige maior robustez e aprimoramento.

## REFERÊNCIAS

ABDI, H., WILLIAMS, L. J., 2010, “Principal component analysis”, *Wiley interdisciplinary reviews: computational statistics*, 2(4), pp. 433-459.

ABNT NBR 13992, *Gasolina automotiva — Determinação do teor de etanol anidro combustível (EAC)*. Rio de Janeiro: Associação Brasileira de Normas Técnicas, 2015.

ABUSALEEM, A., et al, 2023, “Combined mean median filter for image enhancement” In AIP Conference Proceedings, v. 2979, n. 1, AIP Publishing, October.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS (ANP). Gasolina. Disponível em: <https://www.gov.br/anp/pt-br/assuntos/producao-de-derivados-de-petroleo-e-processamento-de-gas-natural/producao-de-derivados-de-petroleo-e-processamento-de-gas-natural/gasolina>. Acesso em: 21 out. 2024.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS (ANP). Resolução nº 807/2020. Disponível em: <https://atosoficiais.com.br/anp/resolucao-n-807-2020>. Acesso em: 21 out. 2024.

BEZERRA, A. C. D. M., SILVA, D. D. O., MATOS, G. H. M. D., et al, 2019, “Quantification of anhydrous ethanol and detection of adulterants in commercial Brazilian gasoline by Raman spectroscopy”, *Instrumentation Science & Technology*, 47(1), pp. 90-106.

BISHOP, C. M., 2006. *Pattern Recognition and Machine Learning*. New York, Springer.

BŁAŻEJ, R., SAWICKI, M., KONIECZNA, M., et al, 2016, “Automatic analysis of thermograms as a means for estimating technical of a gear system”, *Diagnostyka*, v. 17.

BRASIL. Projeto de Lei nº 528/2020. Brasília: Senado, 2020. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/162696>. Acesso em: 21 out. 2024.

BRUCE, P., BRUCE, ANDREW, A., 2019, *Estatística Prática para Cientistas de Dados - 50 Conceitos Essenciais*. 1 ed. Alta Books.

BRZEZINSKI, R. Y., RABIN, N., LEWIS, N., et al, 2021, "Automated processing of thermal imaging to detect COVID-19", *Scientific Reports*, v. 11, n. 1, pp. 17489.

CHOUDHARY, A., GOYAL, D., & LETHA, S. S., 2020, “Infrared thermography-based fault diagnosis of induction motor bearings using machine learning” *IEEE Sensors Journal*, v. 21, n. 2, p. 1727-1734.

CONTINENTAL. Brazilian Flex Fuel Sensor (Br-FFS). Continental, U.S.A. 28 fev. 2014.

DA SILVEIRA, F. L., 2008, Máquinas térmicas à combustão interna de Otto e de Diesel. *Instituto de Física da Universidade Federal do Rio Grande Do Sul*.

DE OLIVEIRA, F. E. M, 1999, *Estatística e Probabilidade*. 2 ed. São Paulo, Editora Atlas S.A.

EPA (2023), The 2023 EPA Automotive Trends Report , <https://www.epa.gov/automotive-trends/download-automotive-trends-report#Full%20Report>

FADAIRO, A. S., EKOCHUKWUKALU, J., ADEYEMI, G. A., et al, 2020, A fast and cost-efficient method to detect ethanol as adulterant in gasoline. *MethodsX*, 7, 100974.

FERRARI, N. L., MARIN N. P. A., 2018, Modelo de um Ciclo Otto com um Simulador de Processos, *XXII Congresso Brasileiro de Engenharia Química*, São Paulo.

GARY, J. H., HANDWERK, G. E., KAISER, M. J., 2007, *Petroleum Refining Technology and Economics*. 5 ed. CRC Press, Florida.

GONÇALVES, A. C., FERREIRA R. L., PRETO E., et al, 2022, Analysis of the influence of gasoline adulteration on lubricant degradation and internal combustion engine performance. *ACS omega*, v. 7 n. 38, pp. 34012-34021.

GONZALES, R. C., WOODS, R. E., 2018, *Digital image processing*. 4 ed. Pearson.

GOSHTASBY, A. A., 2012, *Image Registration: Principles, Tools and Methods*. 1 ed. New York, Springer.

HALLIDAY, D., RESNICK, R., WALKER, J., 2016, *Fundamentos da Física: Gravitação, Ondas e Termodinâmica*. 10 ed. Rio de Janeiro, LTC.

HALLIDAY, D., RESNICK, R., WALKER, J., 2016, *Fundamentos da Física: Óptica e Física Moderna*. 10 ed. Rio de Janeiro, LTC.

HEYWOOD, J. B., 1988, *Internal Combustion Engine Fundamentals*. 1 ed. New York, McGraw-Hill, Inc.

IEA (2023), World Energy Outlook 2023, IEA, Paris <https://www.iea.org/reports/world-energy-outlook-2023>, Licence: CC BY 4.0 (report); CC BY NC SA 4.0 (Annex A).

ISOPPO, V. G., GIL E. S., GONÇALVES P. S. B., et al, 2020, Highly fluorescent lipophilic 2, 1, 3-benzothiadiazole fluorophores as optical sensors for tagging material and gasoline adulteration with ethanol, *Sensors and Actuators B: Chemical*, v. 309, pp. 127701.

JÄHNE, B., 2002. *Digital Image Processing*. 5 ed. New York, Springer.

JANSSENS, O., VAN DE WALLE, R., LOCCUFIER, M., et al, 2017, Deep learning for infrared thermal image based machine health monitoring. *IEEE/ASME Transactions on Mechatronics*, v. 23, n. 1, pp. 151-159.

JIA, Z., LIU, Z., VONG, C. M., et al, 2019, A rotating machinery fault diagnosis method based on feature learning of thermal images. *Ieee Access*, v. 7, pp. 12348-12359.

KHUDHAIR, Z. N., 2023, Color to grayscale image conversion based on singular value decomposition. *Ieee Access*, V. 11, pp. 54629-54638.

Li, W., Zhuo, Y., Bao, J., et al, 2021, A data-based soft-sensor approach to estimating raceway depth in ironmaking blast furnaces. *Powder Technology*, v. 390, pp. 529-538.

LOZANOV, Y., TZVETKOVA, S., & PELTESHKOV, A, 2020, Use of machine learning techniques for classification of thermographic images, *12th Electrical Engineering Faculty Conference (BULEF) IEEE* , pp. 1-4, September.

MANCARUSO, E., SEQUINO, L., & VAGLIECO, B. M, 2020, . Infrared Imaging in Internal Combustion Engines: Advanced Techniques for Vapor Phase Visualization and CO2 Detection. In *Journal of Physics: Conference Series*, v. 1589, n. 1, pp. 012018, IOP Publishing, July.

MARQUES, O., 2011, *Practical Image and Video Processing Using MATLAB*. 1 ed. New Jersey, Wiley-IEEE Press.

Ministério da Agricultura, Pecuária e Abastecimento. (2015). Portaria nº 75/2015. Diário Oficial da União, Brasil. Disponível em: <https://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?data=06/03/2015&jornal=1&pagina=17&totalArquivos=200>.

MOHAMAD, B., SZEPESEI, G. L., & BOLLO, B, 2018, Effect of ethanol-gasoline fuel blends on the exhaust emissions and characteristics of SI engines, *Vehicle and Automotive Engineering 2: Proceedings of the 2nd VAE2018, Miskolc, Hungary*, pp. 29-41.

MOHAMMED, M. K., 2021, Effect of ethanol-gasoline blends on SI engine performance and emissions. *Case studies in thermal engineering*, v. 25, pp. 100891.

MOHRI, M., ROSTAMIZADEH, A., TALWALKAR, A., 2018, *Foundations of Machine Learning*. 2 ed. Massachussets, The MIT Press.

NAGORI, A., DHINGRA, L. S., BHATNAGAR, A., et al, 2019, Predicting hemodynamic shock from thermal images using machine learning. *Scientific reports*, v. 9 , n. (1), pp. 91.

NELSON, D. L., COX, M. M., 2014, *Princípios de bioquímica*, 6 ed. Artmed Editora.

OTSU, N. (1979). A threshold selection method from gray-level histograms, *Automatica*, v. 11, n. 285-296, pp. 23-27.

PARK, M. J., KIM, J., JEONG, S., et al, 2022, Machine learning-based concrete crack depth prediction using thermal images taken under daylight conditions. *Remote Sensing*, v. 14, n. 9, pp. 2151.

PEDROSO, M. P., DE GODOY, L. A. F., FERREIRA, E. C., 2008, Identification of gasoline adulteration using comprehensive two-dimensional gas chromatography combined to multivariate data processing, *Journal of Chromatography A*, v. 1201, n. 2, pp. 176-182.

PEER, M. S., PEER, M. N, 2019, Experimental investigation on engine characteristics fueled with waste HDPE oil and study on NO x emission variation using thermal imager, *Environmental Science and Pollution Research*, n. 26, pp. 3436-3446.

RICHARDS, P., BARKER, J., 2023, *Automotive Fuels Reference Book*. 4 ed. Pennsylvania, SAE International.

ROMANEL, S. A., CUNHA, D. A., CASTRO, E. V., et al, 2018, Time domain nuclear magnetic resonance (TD-NMR): A new methodology to quantify adulteration of gasoline, *Microchemical Journal*, n. 140, pp. 31-37.

SARKAR, T., et al, 2022, DETECTION OF EDIBILITY OF AMLA (*Emblica officinalis*) THROUGH PCA BASED IMAGE ANALYSIS, *Economic Computation & Economic Cybernetics Studies & Research*, v. 56, n. 2.

SHAMEER, P. M., & RAMESH, K, 2017, Experimental evaluation on performance, combustion behavior and influence of in-cylinder temperature on NO<sub>x</sub> emission in a DI diesel engine using thermal imager for various alternate fuel blends, *Energy*, n. 118, pp. 1334-1344.

SHAMMI, S., SOHEL, F., DIEPEVEEN, D., et al, 2022, Machine learning-based detection of freezing events using infrared thermography. *Computers and Electronics in Agriculture*, v. 198, pp. 107013.

SHANNON, C. E., WEAVER, W., 1949, *The Mathematical Theory of Communication*. Urbana, University of Illinois Press.

SPARAVIGNA, A. C., 2019, Entropy in image analysis, *Entropy*, v. 21, n. 5, pp. 502.

SPIEGEL, M. R., SCHILLER, J., SRINIVASAN, A., 2013, *Probabilidade e Estatística*, 3 ed, Porto Alegre, Bookman.

STEPIEN, Z., 2016, Study of the various factors influencing deposit formation and operation of gasoline engine injection systems. In *IOP Conference Series: Materials Science and Engineering*, v. 148, n. 1, pp. 012052, IOP Publishing, September.

TAKESHITA, E. V., REZENDE, R. D. P., DE SOUZA, S. G. U., et al, 2008, Influence of solvent addition on the physicochemical properties of Brazilian gasoline. *Fuel*, v. 87, n. 10-11, pp. 2168-2177.

TARUN S., SHASHI K., ANKIT G., 2022, Analysis on emissions of gasoline engine with different ethanol blends at different compression ratio. *YMER Digital*. V. 21, pp. 18–28.

TONG, K., WANG, Z., SI, L., et al, 2020, A novel pipeline leak recognition method of mine air compressor based on infrared thermal image using IFA and SVM. *Applied Sciences*, v. 10, n. 17, pp. 5991.

UDDIN, M. P., MAMUN, M. A., & HOSSAIN, M. A., 2021, PCA-based feature reduction for hyperspectral remote sensing image classification. *IETE Technical Review*, v. 38 n. 4, pp. 377-396.

UTH, F., POLINK, B., KURPIEL W., et al, 2019, An innovative person detection system based on thermal imaging cameras dedicate for underground belt conveyors. *Mining Science*, v. 26, pp. 275-288.

VAN WYLEN, G. J., BORGNARKKE, C., SONNTAGE, R. E., 2003, Fundamentos da Termodinâmica. 6 ed. Edgard Blücher LTDA.

WILSON, A. N., GUPTA, K. A., KODURU, B. H., et al, 2023, Recent advances in thermal imaging and its applications using machine learning: A review. *IEEE Sensors Journal*, v. 23, n. 4, pp. 3395-3407.

YAGLE, A. E., ULABY F. T., 2018, *Image Processing for Engineers*. Michigan Publishing.

YOUNUS, A. M., WIDODO, A., & YANG, B. S., 2010, Evaluation of thermography image data for machine fault diagnosis, *Nondestructive Testing and Evaluation*, v. 25, n. 3, pp. 231-247.

YU, R., GUO, S., HUANG, Y., et al, 2023, Prediction of variable-groove weld penetration using texture features of infrared thermal images and machine learning methods. *Journal of Materials Research and Technology*, v. 23, pp. 1039-1051.

ZHOU, Z., 2016, *Machine Learning*. Springer.