

**UNIVERSIDADE DO ESTADO DE SANTA CATARINA - UDESC  
CENTRO DE CIÊNCIAS TECNOLÓGICAS - CCT  
MESTRADO EM COMPUTAÇÃO APLICADA**

**JULIANA PATRÍCIA DETROZ**

**DETECÇÃO DE FRAUDE EM HIDRÔMETROS UTILIZANDO  
TÉCNICAS DE RECONHECIMENTO DE PADRÕES**

**JOINVILLE  
2016**



**UNIVERSIDADE DO ESTADO DE SANTA CATARINA - UDESC  
CENTRO DE CIÊNCIAS TECNOLÓGICAS - CCT  
MESTRADO EM COMPUTAÇÃO APLICADA**

**JULIANA PATRÍCIA DETROZ**

**DETECÇÃO DE FRAUDE EM HIDRÔMETROS UTILIZANDO  
TÉCNICAS DE RECONHECIMENTO DE PADRÕES**

Dissertação submetida ao Programa  
de Pós-Graduação em Computação  
Aplicada do Centro de Ciências Tec-  
nológicas da Universidade do Es-  
tado de Santa Catarina, para a ob-  
tenção do grau de Mestre em Com-  
putação Aplicada.

**Orientador:**

Prof. Dr. André Tavares da Silva

**JOINVILLE**

**2016**

D483d

Detroz, Juliana Patricia

Detecção de fraude em hidrômetros utilizando técnicas de reconhecimento de padrões / Juliana Patricia Detroz. – 2016.

117 p. : il. ; 21 cm

Orientador: André Tavares da Silva

Bibliografia: p. 109-117

Dissertação (mestrado) – Universidade do Estado Santa Catarina, Centro de Ciências Tecnológicas, Programa de Pós-Graduação em Computação Aplicada, Joinville, 2016.

1. Computação aplicada . 2. Reconhecimento de padrões. 3. Detecção de fraudes. 4. Visão computacional. I. Silva, André Tavares. II. Universidade do Estado de Catarina. Programa de Pós- Graduação Computação Aplicada. III. Detecção de fraude em hidrômetros utilizando técnicas de reconhecimento de padrões.

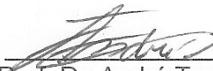
CDD 006.4 – 23.ed.

**JULIANA PATRÍCIA DETROZ**  
**DETECÇÃO DE FRAUDE EM HIDRÔMETROS UTILIZANDO**  
**TÉCNICAS DE RECONHECIMENTO DE PADRÕES**

Dissertação apresentada ao Curso de Mestrado Acadêmico Computação Aplicada como requisito parcial para obtenção do título de Mestra em Computação Aplicada na área de concentração “Ciência da Computação”.

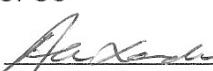
**Banca Examinadora**

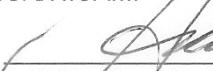
Orientador:

  
Prof. Dr. André Tavares da Silva  
CCT/UDESC

**Membros**

  
Prof. Dr. Alexandre Gonçalves Silva  
UFSC

  
Prof. Dr. Alexandre Xavier Falcão  
IC/UNICAMP

  
Prof. Dr. Marcelo da Silva Hounsell  
CCT/UDESC

Joinville, SC, 26 de fevereiro de 2016.



## **AGRADECIMENTOS**

À minha família, em especial a meus pais, Dorival e Sueli, e minha irmã Daniela por todo apoio e carinho.

Ao meu orientador, professor Dr. André Tavares da Silva, pela dedicação e apoio durante a realização deste trabalho e a todos professores do Programa de Pós-Graduação em Computação Aplicada por todo o conhecimento compartilhado. Ao professor Dr. Jefersson Alex dos Santos e Waner Miranda, da Universidade Federal de Minas Gerais, pela colaboração e auxílio na realização dos experimentos de combinação de descritores relacionados à programação genética.

À Companhia Águas de Joinville por apoiar e autorizar o uso dos dados do estudo de caso desta pesquisa e em especial aos meus gestores diretos, Aline S. G. Machado e Giovani Osmarini, pela compreensão e incentivo ao longo da realização deste trabalho.



*A mente que se abre a uma nova ideia jamais volta  
ao seu tamanho original.  
(Oliver Wendell Holmes, Sr)*



## RESUMO

DETROZ, Juliana P. **Detecção de fraude em hidrômetros utilizando técnicas de reconhecimento de padrões.** 117 p. Dissertação de Mestrado (Mestrado Acadêmico em Computação Aplicada - Área: Processamento Gráfico) - Universidade do Estado de Santa Catarina. Programa de Pós-Graduação em Computação Aplicada, Joinville, 2016.

Em tempos de racionamento dos recursos hídricos, o desperdício de água tem sido um tema de relevância mundial. Os vazamentos ocultos e aparentes são uma das principais causas dos elevados índices de perdas de água tratada. Esforços são despendidos pelas companhias de saneamento a fim de reduzir as perdas, sendo o combate às fraudes uma destas ações. Neste contexto, o uso da tecnologia para automatizar a identificação de fraude mostra-se uma importante ferramenta de apoio no combate ao desperdício. Esta pesquisa tem como objetivo aplicar técnicas de reconhecimento de padrões na detecção automatizada de casos suspeitos de irregularidades em hidrômetros. No escopo deste trabalho foram consideradas suspeitas de fraude as violações e ausências de lacres. A abordagem proposta visa, através de um sistema de visão computacional, auxiliar no combate a fraudes em hidrômetros e, consequentemente, evitar o desperdício de água associado a estas. Para isto, a execução do sistema proposto é dividida em três etapas: detecção do hidrômetro, fazendo uso do classificador OPF e descritor HOG; a detecção da área estimada dos lacres, obtida pela aplicação de métodos de processamento morfológico e segmentação; e a classificação das fraudes a partir da condição dos lacres do hidrômetro. A validação foi executada utilizando-se um conjunto de imagens de fiscalizações. Na primeira etapa, a solução utilizando o classificador OPF alcançou taxa de acerto média de 89,03%, sendo superior a resultados dos métodos SVM linear e RBF. Para a classificação da condição dos lacres, realizou-se uma análise comparativa de 12 descritores de imagem, de cor e textura, sendo avaliados os resultados individuais e combinados, atingindo taxas de acerto média de até 81,29%. Com isto, pode-se concluir que o uso de um sistema especialista de visão computacional para o problema de detecção de fraudes é uma estratégia promissora e com potencial para beneficiar a análise e o suporte à tomada de decisões.

**Palavras-chaves:** Reconhecimento de Padrões. Detecção de Fraudes. Visão Computacional. Classificação Multiclasses.



## ABSTRACT

With the emerging hydric crisis, water shortage has been a great global concern. Water supply companies have been increasingly looking for solutions to reduce water wastage and many efforts have been made aiming to promote a better management of this resource. Fraud detection is one of these actions, as the irregular violations are usually held precariously, thus, causing leaks. Hidden and apparent leakage is a major cause of the high water loss rates. In this context, the use of technology in order to automate the identification of potential frauds can be an important support tool to avoid water waste. In this sense, this research aims to apply pattern recognition techniques in the implementation of an automated detection of suspected irregularities cases in water meters, through image analysis. We considered as a potential fraud when there is evidences of violations and seals absences. The proposed computer vision system is composed by three steps: the detection of the water meter location, obtained by OPF classifier and HOG descriptor, detecting the seals through morphological image processing and segmentation methods; and the classification of frauds, in which the condition of the water meter seals is assessed. We validated the proposed framework using a dataset containing images of water meter inspections. The water meter detection solution (HOG+OPF) achieved an average accuracy of 89.03%, showing superior results than SVM (linear and RBF). A comparative analysis of 12 feature descriptors (color and texture) was performed on the classification of the seals condition step. The results of these methods were evaluated individually and also combined, reaching average accuracy up to 81.29%. We concluded that the use of a computer vision system is a promising strategy and has potential to benefit and support the analysis of fraud detection.

**Key-words:** Pattern Recognition. Fraud Detection. Computer Vision. Multiclass Classifier. OPF. HOG



## LISTA DE FIGURAS

Figura 2.1 –Ciclo de reconhecimento de padrões. . . . .	27
Figura 2.2 –Exemplo de imagem 64x128 do conjunto de treinamento e a representação do descritor HOG resultante do treinamento. . . . .	35
Figura 2.3 –Estrutura do descritor HOG . . . . .	37
Figura 2.4 –Exemplo de SVM em problema de espaço bidimensional	40
Figura 2.5 –Representação do funcionamento do classificador OPF	43
Figura 2.6 –Operações morfológicas. (a) imagem original A. (b) elemento estruturante B. (c) dilatação de A por B. (d) erosão de A por B. (e) elemento estrutura E. (f) abertura de A por E. (g) fechamento de A por E. . . . .	46
Figura 2.7 –Exemplificação do processo de segmentação por <i>Graph Cut</i> . . . . .	49
Figura 2.8 –Exemplo de aplicação de métodos de segmentação. . .	50
Figura 3.1 –Exemplo de anomalias em conjunto de dados bidimensional . . . . .	57
Figura 3.2 – <i>Framework</i> para detecção de fraude em consumo de eletricidade. . . . .	59
Figura 4.1 –Estrutura de hidrômetros . . . . .	64
Figura 4.2 –Exemplos de fraudes em hidrômetros . . . . .	65
Figura 4.3 –Visão macro do esquema da abordagem proposta . . . .	67
Figura 4.4 –Modelo proposto para detecção de hidrômetros . . . .	70
Figura 4.5 –Exemplo de aplicação de janela de detecção HOG . . .	71
Figura 4.6 –Exemplos de aplicação do HOG em amostra do conjunto de treinamento . . . . .	72
Figura 4.7 –Classes do conjunto de treinamento . . . . .	73
Figura 4.8 –Comparação de imagens do conjunto de treinamento .	74

Figura 4.9 –Rotação de imagem para otimizar detecção de hidrômetro . . . . .	76
Figura 4.10 –Incremento de margem do retângulo de detecção . . . . .	77
Figura 4.11 –Processamento de imagem para localização de lacres . . . . .	82
Figura 4.12 –Exemplos de amostras de conjunto de dados de classificação de lacres . . . . .	83
Figura 4.13 –Modelo proposto de classificação de lacres . . . . .	84
Figura 5.1 –Resultados de detecção de hidrômetros. . . . .	88
Figura 5.2 –Estrutura dos classificadores multiclasses . . . . .	91
Figura 5.3 –Resultados de detecção de lacres. . . . .	94
Figura 5.4 –Condição dos lacres - Exemplos de amostras classificadas . . . . .	99
Figura 5.5 –Visão macro da solução - técnicas aplicadas . . . . .	102

## **LISTA DE TABELAS**

Tabela 5.1 –Resultados - Detecção de hidrômetros . . . . .	92
Tabela 5.2 –Experimento 1 - Classificação da condição dos lacre s . . . . .	96
Tabela 5.3 –Experimento 2 - Comparação taxa média de acerto de métodos de validação cruzada . . . . .	97
Tabela 5.4 –Experimento 3 - Taxa média de acerto . . . . .	98
Tabela 5.5 –Matriz de confusão - Classificação da condição dos lacre s (OPF + ACC) . . . . .	99
Tabela 5.6 –Experimento 4 - Combinação de descritores . . . . .	101
Tabela 5.7 –Experimento 4 - Taxa de acerto de combinação de descritores . . . . .	101



## LISTA DE ABREVIATURAS E SIGLAS

ACC	<i>Auto Color Correlogram</i>
ANN-MLP	<i>Artificial Neural Network - Multilayer Perceptron</i>
BIC	<i>Border/Interior Pixel Classification</i>
CBIR	<i>Content-Based Image Retrieval</i>
CCH	<i>Code Chain Histogram</i>
CCV	<i>Color Coherence Vector</i>
CSD	<i>Color Structure Descriptor</i>
CSS	<i>Curvature Scale-Space</i>
GCH	<i>Global Color Histogram</i>
GLOH	<i>Gradient Local and Orientation Histograms</i>
GP	<i>Genetic Programming</i>
FN	<i>False Negatives</i>
FP	<i>False Positives</i>
HOG	<i>Histograms of Oriented Gradients</i>
HTD	<i>Homogeneous Texture Descriptor</i>
k-NN	<i>k-Nearest Neighbors</i>
IM	<i>Invariant Moments</i>
JAC	<i>Joint Auto-Correlogram</i>
LAS	<i>Local Activity Spectrum</i>
LBP	<i>Local Binary Patterns</i>
LCH	<i>Local Color Histogram</i>
OPF	<i>Optimum-Path Forest</i>
PHOG	<i>Pyramid of Histogram of Gradient Orientations</i>
RGB	Espaço de cores formado pelos componentes vermelho (Red), verde (Green) e azul (Blue)
RP	Reconhecimento de Padrões
SASI	<i>Statistical Analysis of Structural Information</i>
SID	<i>Invariant Steerable Pyramid Decomposition</i>

SOM	<i>Self-Organizing Map</i>
SIFT	<i>Scale-Invariant Feature Transform</i>
SVM	<i>Support Vector Machine</i>
SVM RBF	<i>Support Vector Machine - Radial Basis Function kernel</i>
SURF	<i>Speeded Up Robust Features</i>
TP	<i>True Positives</i>
ZM	<i>Zernike Moments</i>

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>21</b>
1.1	Objetivos	22
1.2	Metodologia	23
1.3	Escopo	24
1.4	Estrutura do documento	24
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>25</b>
2.1	Aprendizado de Máquina e Reconhecimento de Padrões	25
2.1.1	Extração de Características	27
2.1.1.1	Medidas de distância	28
2.1.1.2	Descritores de imagem	29
2.1.1.3	Descritores de cor	30
2.1.1.4	Descritores de textura	32
2.1.1.5	Descriptor de Forma Histogramas de Gradi- entes Orientados	35
2.1.2	Classificação	38
2.1.2.1	Máquina de Vetores de Suporte	39
2.1.2.2	Florestas de Caminhos Ótimos	42
2.2	Processamento Morfológico de Imagens	45
2.3	Segmentação	46
2.4	Considerações	50
<b>3</b>	<b>TRABALHOS RELACIONADOS</b>	<b>53</b>
3.1	Detecção de Objetos	53
3.2	Detecção de fraudes e anomalias	56
3.3	Considerações	61
<b>4</b>	<b>DESENVOLVIMENTO DA SOLUÇÃO</b>	<b>63</b>
4.1	Hidrômetro	63
4.2	Definição do projeto	64

4.2.1	Aquisição de imagens . . . . .	68
4.2.2	Detecção de hidrômetro . . . . .	69
4.2.3	Detecção de lacres . . . . .	77
4.2.4	Classificação de fraudes . . . . .	81
4.3	Considerações . . . . .	85
<b>5</b>	<b>EXPERIMENTOS E RESULTADOS . . . . .</b>	<b>87</b>
5.1	Detecção de hidrômetros . . . . .	87
5.2	Detecção de lacres . . . . .	93
5.3	Classificação de lacres . . . . .	93
5.3.1	Validação cruzada . . . . .	94
5.3.2	Medidas de distância . . . . .	97
5.3.3	Combinação de descritores . . . . .	100
5.4	Considerações . . . . .	101
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>105</b>
6.1	Trabalhos futuros . . . . .	107
<b>REFERÊNCIAS . . . . .</b>		<b>109</b>

## 1 INTRODUÇÃO

A preocupação com a escassez dos recursos hídricos tem levado a um crescente interesse na busca por soluções que minimizem o desperdício de água. O relatório do Sistema Nacional de Informações de Saneamento (SNIS) aponta que em 2013 o índice brasileiro de perdas de água na distribuição correspondia a 37% do total de água tratada. Este valor é 0,1% superior ao ano anterior (BRASIL, 2015). Isto indica que uma parcela considerável do volume de água potável distribuída não chega até o consumidor final. Estas perdas, denominadas reais ou físicas, podem ser ocasionadas por vazamentos ocultos ou aparentes, desvios fraudulentos de tubulação, a má condição devido à ocorrência de ligações clandestinas, entre outros fatores.

O combate a fraudes é uma das ações na busca por reduzir as perdas de água ocasionadas por desvios irregulares de ramal, violações em hidrômetros e ligações clandestinas. Desta forma, a detecção de fraudes em hidrômetros torna-se uma medida efetiva na redução do desperdício de água potável. Periodicamente as concessionárias de saneamento e distribuição de água realizam fiscalizações para identificar estas fraudes, sendo usualmente originadas pela análise de variação de consumo, observação dos leitistas ou por denúncias.

Mensalmente todas as unidades consumidoras hidrometradas são visitadas por leitistas, a fim de realizar a coleta do volume consumido no período. Devido ao grande número de unidades e o curto período do ciclo de leitura, realizar uma análise individual no local para detectar possíveis fraudes se torna inviável. Nesta ocasião, poderia ser realizado o registro por foto da situação atual do hidrômetro. Desta forma, uma solução que utilize-se de técnicas de reconhecimento de padrões para a detecção automática de fraudes em hidrômetros poderia ser adotada como ferramenta de suporte na identificação de suspeitas de fraudes, minimizando os custos e riscos associados ao deslocamento para realização de

fiscalização, otimizando a identificação dos casos inconsistentes e, consequentemente, reduzindo o desperdício de água associado a estes.

Um dos desafios desta aplicação trata-se da não uniformidade das condições das imagens utilizadas como entrada para a aplicação, visto que as fotos podem ser tiradas de diversos ângulos, escalas e sem capturar o hidrômetro de forma completa, ocasionando uma variação na forma do objeto de interesse representada na imagem. Além disto, também podem conter fundos heterogêneos e com grande quantidade de objetos em cena, assim como apresentar oclusão parcial do hidrômetro por objetos como, grades, cercas ou plantas, por exemplo. A detecção de objetos em imagens de fundo complexo é um problema bastante estudado, porém ainda sem uma solução única que aplique-se a problemas dos mais diversos contextos (JAIN; RATHA; LAKSHMANAN, 1997).

## 1.1 Objetivos

Este trabalho tem como objetivo geral aplicar técnicas de reconhecimento de padrões para realizar a detecção automatizada, por meio de análise de imagens, de casos suspeitos de irregularidades em hidrômetros. Desta forma, a abordagem proposta visa, através de um sistema de visão computacional, auxiliar no combate a fraudes em hidrômetros e, consequentemente, evitar o desperdício de água associado a estas. Destacam-se ainda os seguintes objetivos específicos:

- Desenvolver um método para detecção de hidrômetros capaz de encontrá-los em diferentes orientações e posições;
- Aplicar métodos de detecção de objetos a um problema multiclasse, possibilitando a localização de objetos a partir de diferentes visões, a fim de verificar sua eficiência e compará-los a outras abordagens propostas na literatura científica;
- Investigar e comparar um conjunto de descritores de imagem, a fim de verificar qual melhor se aplica ao problema de classificação da

condição dos lacres de hidrômetros;

## 1.2 Metodologia

Foi realizada uma pesquisa exploratória, a fim de conhecer o estado da arte da aplicação de técnicas de reconhecimento de padrões em detecção de fraudes, além de aprofundar o conhecimento sobre técnicas relacionadas à análise de imagens por visão computacional. Em seguida, projetou-se o modelo a ser utilizado no desenvolvimento da solução proposta.

O desenvolvimento foi dividido em três etapas: detecção do hidrômetro, através de implementação utilizando o descriptor HOG associado ao classificador OPF; aplicação de métodos de processamento morfológico e segmentação, com o objetivo de localizar e reduzir a região de interesse à área dos lacres; e por fim, a classificação da condição dos lacres, possibilitando estimar a ausência ou suspeita de fraude. Nesta última fase, foi empregado o classificador OPF em conjunto com diferentes descritores de cor e textura, para realização de uma análise comparativa.

Optou-se por adotar uma abordagem de detecção por contexto Torralba, Murphy e Freeman (2010), na qual realiza-se, a princípio, a identificação da posição do hidrômetro para que posteriormente sejam localizados e analisado o estado dos lacres do cavalete. Esta estratégia visa otimizar o resultado da detecção dos lacres, visto que os mesmos possuem forma simples e genérica, o que potencialmente levaria a um maior índice de falsos positivos, caso fossem procurados por toda a extensão da imagem. Realizando a associação ao objeto a que estes devem estar acoplados, pretende-se alcançar uma maior taxa de acerto ao reduzir a área de busca.

Após a construção do modelo proposto, foi feita uma pesquisa experimental com o intuito de verificar e validar a eficiência do mesmo. Para isto foi utilizada uma base de imagens privada, obtida a partir de fotos de fiscalizações realizadas e disponibilizadas para esta pesquisa.

pela Companhia Águas de Joinville.

### 1.3 Escopo

A fim de delimitar o escopo desta pesquisa, serão analisadas as fraudes a partir da condição dos lacres do hidrômetro. Serão considerados suspeitos de fraudes os casos em que não sejam localizados os lacres do cavalete do medidor ou quando estes encontram-se violados, isto é, rompidos.

As imagens utilizadas neste estudo formam uma base de dados privada, cedida pela Companhia Águas de Joinville, obtidas em ações de fiscalização. Estas imagens contêm estrutura complexa, sendo o objeto de interesse (hidrômetro), localizado em diferentes tipos de ambientes, podendo ainda estar representado em diferentes posições e com ocorrência de oclusão.

### 1.4 Estrutura do documento

Esta dissertação está organizada da seguinte forma: O Capítulo 2 apresenta a fundamentação teórica, detalhando os principais conceitos e definições relacionados à abordagem proposta. No Capítulo 3 é realizada uma revisão da literatura, na qual são expostos trabalhos relacionados à presente pesquisa. A metodologia aplicada e o desenvolvimento da solução proposta são descritos no Capítulo 4. Por fim, são detalhados os experimentos, resultados e conclusões no Capítulo 5 e no Capítulo 6, respectivamente.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são expostos os principais conceitos relacionados ao escopo desta pesquisa, tendo como intuito auxiliar na compreensão da problemática e da proposta apresentada neste trabalho.

### 2.1 Aprendizado de Máquina e Reconhecimento de Padrões

O termo aprendizado de máquina foi introduzido na década de 50 por Arthur Lee Samuel para designar "um abrangente conjunto de atividades inteligentes realizadas por humanos que poderiam ser transferidas para máquinas, de forma que estas tivessem capacidade de executá-las", isto é, reproduzir a capacidade humana de aprendizado para programas de computador (GUYON; ELISSEEFF, 2006). Mitchell (1997) define que:

(...) pode-se dizer que um programa de computador aprende a partir de uma experiência  $E$  a respeito de uma classe de tarefas  $T$  e medida de desempenho  $P$ , se seu desempenho nas tarefas  $T$ , segundo medido por  $P$ , melhora com a experiência  $E$ .

Isto é, a partir de um conhecimento *a priori* ( $E$ ), o programa de computador pode inferir e tomar decisões sobre uma atividade ( $T$ ), sendo o resultado avaliado a partir de  $P$ .

Guyon e Elisseeff (2006) afirmam que as pesquisas relacionadas a aprendizado de máquina apresentam maior ênfase em encontrar relações existente em dados e analisar os processos de extração destes relacionamentos. Assim, a partir de um conjunto de exemplos de treinamento associados a saídas relacionadas a cada amostra (rótulos), o sistema pode identificar a relação entre estes padrões, aprendendo e inferindo uma provável saída a dados não rotulados.

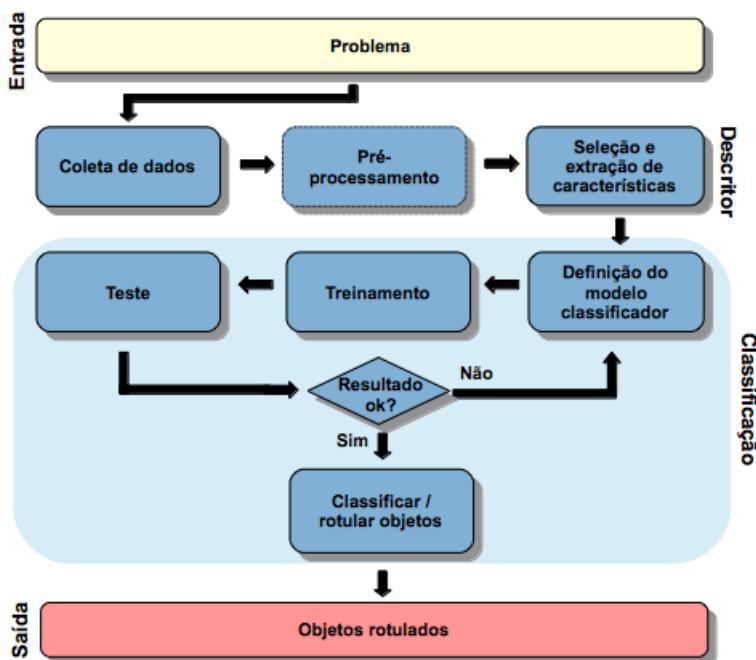
Kuncheva (2004) define reconhecimento de padrões (RP) como a tarefa de atribuir rótulos a objetos, sendo estes descritos através de medidas denominadas atributos ou características. Classificam-se os métodos de RP em duas categorias principais: aprendizado supervisionado e não supervisionado. No primeiro, tem-se *a priori* um conjunto de dados rotulados utilizado para treinar um classificador, de modo que este possa, posteriormente, identificar a que classe pertencem os objetos não rotulados. Usualmente, para esta decisão considera-se uma função que avalia a semelhança entre as características do objeto em questão com as amostras do conjunto de treinamento.

O aprendizado não supervisionado tem como objetivo descobrir a estrutura de um conjunto de dados não rotulados, agrupando-os de acordo com a similaridade de suas características, sem a necessidade de uma base de treinamento. Jafar e Sivakumar (2010), afirmam que estes métodos buscam organizar os dados em diferentes grupos, particionando o conjunto de dados não rotulados em subgrupos de objetos similares.

Para o escopo da presente pesquisa foram adotados métodos de aprendizado supervisionado, uma vez que tem-se o objetivo de identificar a qual classe determinados objetos pertencem.

Um sistema de RP comprehende a execução de diversas etapas. A Figura 2.1 mostra de forma sucinta os principais elementos a serem definidos. O pré-processamento refere-se a uma etapa opcional, podendo ser aplicados tratamentos de acordo com o tipo de dados e das informações que pretende-se extrair. Em aplicações de RP em imagens, por exemplo, é comum a utilização de filtros para ressaltar características ou eliminar ruídos. Em conjuntos de dados com valores de domínio contínuo, o pré-processamento poderia ser representado por uma função de normalização. As etapas posteriores, apresentadas na Figura 2.1, tratam-se da extração de características e classificação dos dados.

Figura 2.1: Ciclo de reconhecimento de padrões.



Fonte: Adaptado de Kuncheva (2004)

### 2.1.1 Extração de Características

Os autores Guyon e Elisseeff (2006) explicam que, para o escopo do aprendizado de máquina, os dados podem ser representados através de um conjunto fixo de atributos, podendo estes ser de diversos domínios como, por exemplo, binários, categóricos, discretos e contínuos. A seleção de características consiste em identificar quais atributos possuem informação relevante para possibilitar a correta classificação dos mesmos. Trucco e Verri (1998) afirmam que modelos baseados em características requerem que os atributos sejam identificados e descritos, a fim de que os dados possam ser comparados. Para isto, utiliza-se um

método descritor. Desta forma, os descritores podem ser definidos como uma função ou algoritmo que extrai um vetor de características a partir da codificação das propriedades de um dado de entrada.

Uma função de similaridade é utilizada para calcular a semelhança entre dados distintos, a partir da distância entre seus respectivos vetores de características (TORRES et al., 2009). Para isto, diferentes medidas de distância podem ser aplicadas. Segundo Torres et al. (2009), “diferentes tipos de vetores de características podem necessitar de diferentes funções de similaridade” e complementam que, substituindo a função de similaridade pode-se obter um conjunto de dados mais ou menos compacto, ou seja, pode-se aumentar ou reduzir a separação entre estes no espaço de distância.

Kuncheva (2004) sugere que nem todas as características possuem mesma relevância, sendo algumas destas importantes somente quando relacionadas a outras. Determinados atributos podem representar apenas “ruídos”, não tendo influência no resultado final da classificação. De forma geral, a definição e extração das características são tarefas dependentes do problema a ser tratado, ou seja, dois problemas distintos utilizando o mesmo conjunto de dados podem beneficiar-se pela escolha de características ou técnicas de extração diferentes. Sendo assim, a escolha dos atributos e métodos mais adequados para cada problema é de fundamental importância para o bom resultado da aplicação proposta.

### 2.1.1.1 Medidas de distância

A similaridade entre duas amostras de um conjunto de dados é obtida através do cálculo da distância dos seus respectivos vetores de características. Quanto menor a distância, maior a semelhança entre as amostras em questão. Diversas funções podem ser adotadas para medir a similaridade. As mais usuais se baseiam na métrica de Minkowsky  $L_q$  como, por exemplo, a  $L_1$  (também conhecida por distância *Manhattan*, geometria do *Taxi* ou *city-block*) e  $L_2$  (distância Euclidiana). O cálculo de  $L_q$  obedece à fórmula (MISSAOUI; SARIFUDDIN; VAILLANCOURT,

2005):

$$L_q = \left( \sum_c |V^X(c) - V^Y(c)|^q \right)^{(1/q)}$$

onde  $X$  e  $Y$  correspondem às amostras comparadas,  $V^X(c)$  e  $V^Y(c)$  representam os atributos dos respectivos vetores de características, considerando que estes possuem  $c$  atributos.

Outras funções de distância utilizadas incluem dLog (STEHLING; NASCIMENTO; FALCÃO, 2002), *Hamming*, *Mahalanobis*, *Jaccard*, Coeficiente de Correlação, *Canberra*, entre outros (GONÇALVES; NUNES; DELAMARO, 2011).

### 2.1.1.2 Descritores de imagem

Segundo Bernal, Vilariño e Sánchez (2010), no contexto devisão computacional, isto é, da aplicação de técnicas de aprendizado de máquina para a análise de imagens, existem três categorias principais de descritores de imagem: os descritores de cor, de textura e de forma. Manjunath et al. (2001) afirmam que a cor é provavelmente a característica visual mais expressiva em relação as demais, sendo o histograma de cor um dos descritores mais utilizados. Este tipo de método de representação de atributos pode ser aplicado a diferentes espaços de cor como, por exemplo, RGB, HSV, HSL e Lab. Dentre os descritores de cor podem ser citados *Border/Interior pixel Classification* (BIC), MPEG-7, *Global Color Histogram* (GCH), *Local Color Histogram* (LCH), *Color Coherence Vector* (CCV) (PENATTI; VALLE; TORRES, 2012).

Paget e Longstaff (1998) definem textura como uma característica visual usualmente utilizada para denominar a estrutura da superfície, podendo ser descrita através de cinco propriedades: regularidade, contraste, direcionalidade, delineamento e rugosidade/aspereza. Os autores (*op. cit.*) afirmam que este atributo pode ser definido em termos da distribuição ou interação espacial dos valores dos *pixels* dentro da imagem. São exemplos de descritores de textura: *Gradient Local and Orientation Histograms* (GLOH), Filtros de Gabor, *Homogeneous Texture Descriptor*

(HTD), *Haar-like*, *Local Binary Patterns* (LBP), entre outros (BERNAL; VILARIÑO; SÁNCHEZ, 2010).

Bernal, Vilariño e Sánchez (2010) citam que a forma é uma das características mais importantes de um objeto, sendo de grande valia para diferenciá-lo de outros artefatos contidos em uma imagem. Porém, apresenta maior complexidade e volume de informação. Zhang e Lu (2004) afirmam que os descritores de forma podem ser classificados ainda em dois grupos: descritores baseados em contorno (*contour-based descriptors*) e descritores baseados em região (*region-based descriptors*). São exemplos desta categoria de descritores: *Histogram of Oriented Gradients* (HOG), *Code Chain Histogram* (CCH), *Convex Hull*, *Invariant Moments* (IM), *Zernike Moments* (ZM), *Curvature Scale-Space* (CSS), *Fourier* e transformada *Wavelets* (YANG et al., 2008).

No desenvolvimento desta pesquisa, foram utilizados treze descritores, sendo um de forma, HOG, seis de cor e seis de textura. O primeiro foi aplicado na fase de localização do hidrônmetro e os demais considerados em um estudo comparativo realizado sobre a classificação da condição dos lacres. A definição e funcionamento destes são detalhados a seguir.

### 2.1.1.3 Descritores de cor

- *Border/Interior pixel Classification* (BIC)

BIC é um descritor de cor que foi proposto inicialmente para aplicações de Recuperação de Imagens por Conteúdo (CBIR, do inglês, *Content-Based Image Retrieval*). O método é formado por três componentes principais: um algoritmo de análise que classifica cada *pixel* como pertencente à borda ou ao interior, para isto a sua cor é comparada a de seus vizinhos; uma medida de distância logarítmica (dLog) utilizada na comparação dos histogramas; e uma representação compacta das características visuais extraídas da imagem.

Dois histogramas são calculados, um para *pixels* de interior e o outro para os de borda, e concatenados de forma a compor o vetor de características. A verificação da similaridade entre os vetores de duas imagens distintas é realizada através da aplicação da medida de distância dLog (STEHLING; NASCIMENTO; FALCÃO, 2002).

- *Auto Color Correlogram* (ACC)

O descritor de cor ACC realiza o mapeamento da correlação espacial de cores de *pixels* em diferentes distâncias. Para isto, um correlograma de cor da imagem é calculado, isto é, uma tabela indexada por pares de cores, na qual uma entrada  $k$  para as cores  $(i, j)$  corresponde à probabilidade de encontrar um *pixel* de cor  $j$  a uma distância  $k$  de um *pixel* de cor  $i$ . Como resultado tem-se um vetor de características de tamanho pequeno (HUANG et al., 1997). A medida de distância utilizada usualmente é a *Manhattan* (L1).

- *Joint Auto-Correlogram* (JAC)

O método de extração de características JAC calcula o autocorrelograma de diferentes propriedades da imagem como, por exemplo, cor, magnitude do gradiente, *rank* e a existência de textura, sendo calculada a probabilidade simultânea da ocorrência destas. Desta forma, busca capturar uma representação mais abrangente da imagem (WILLIAMS; YOON, 2007).

- *Color Bitmap*

*Color Bitmap* é um descritor que representa a característica de cor de uma imagem a partir da distribuição de cores. O método calcula a média e desvio padrão global de cada canal do modelo RBG. Após isto, a imagem é dividida em blocos, sendo calculadas a média para cada um deles. Para os blocos que possuem média maior do que a média global da imagem, atribui-se à posição correspondente do vetor de características o valor 1, caso contrário, a mesma recebe valor 0.

Desta forma, a imagem é representada por duas características globais de cor, média e desvio padrão, além das características locais calculadas bloco a bloco (LU; CHANG, 2007).

- *Global Color Histogram* (GCH)

O Histograma Global de Cor (GCH, do inglês, *Global Color Histogram*) é o descritor de cor mais comumente utilizado. O cálculo do histograma é obtido a partir da discretização das cores da imagem, sendo em seguida realizada a contagem de vezes que cada uma das  $n$  cores aparece na mesma. O vetor resultante possui, portanto,  $n$  características, cada uma contendo o total de ocorrências de cada cor. Desta maneira, tem-se a distribuição global de cores da imagem (SWAIN; BALLARD, 1991).

- *Color Structure Descriptor* (CSD)

Manjunath et al. (2001) afirmam que o Descritor de Estrutura de Cor - CSD - representa a estrutura local de cor de uma imagem a partir do uso de um elemento estruturante  $8 \times 8$ , realizando a contagem da ocorrência que uma determinada cor está contida dentro da posição onde o elemento estruturante encontra-se localizado. O mesmo percorre toda a extensão da matriz de *pixels* da imagem, compondo  $m$  histogramas da estrutura de cores ( $h(m)$ ), onde  $m$  representa o número de cores consideradas. Cada *bin* do histograma, ou seja, faixa de agrupamento de dados, representa o número de elementos estruturantes na imagem que contêm um ou mais *pixels* da cor em questão.

#### 2.1.1.4 Descritores de textura

- *Local Binary Pattern* (LBP)

Segundo Ojala, Pietikäinen e Mäenpää (2002), o descritor LBP “baseia-se na premissa de que determinados padrões binários locais são propriedades fundamentais da textura local de uma imagem”. Este método de extração de características define uma janela de

raio R que translada pela imagem quantificando a ocorrência de variações positivas e negativas de valores de níveis de cinza entre os *pixels* da vizinhança  $g_p$  ( $p = 0, \dots, P - 1$ , onde  $P$  = níveis de cinza da imagem) e o *pixel* central ( $g_c$ ) da posição onde encontra-se a janela.

Desta forma, quando a diferença resulta em um valor positivo, considera-se 1 e, atribui-se 0, quando esta variação é negativa. Em seguida, o número de transições 0-1 são quantificadas, a partir do qual obtém-se o valor LBP. O histograma é formado pelos valores LBP calculados a cada posição em que a janela foi deslocada pela imagem.

- *Homogeneous Texture Descriptor* (HTD)

HTD é um método de descrição extraído da imagem a partir da aplicação, no domínio de frequência, de um conjunto de filtros de Gabor sensíveis à escala e rotação. Após isto, sobre o resultado já filtrado, calcula-se a média e desvio padrão. A distância entre dois HTDs é calculada por (MANJUNATH et al., 2001):

$$D(HTD_1, HTD_2) = \sum_k \left| \frac{HTD_1(k) - HTD_2(k)}{a(k)} \right|,$$

onde  $a(k)$  representa o desvio padrão do conjunto de dados e aplica-se para fins de normalização.

- *Statistical Analysis of Structural Information* (SASI)

O descriptor SASI baseia-se em um conjunto de janelas de diferentes tamanhos e orientações para extrair as propriedades estruturais de textura da imagem. Cada janela percorre a imagem e a cada posição calcula-se um valor de autocorrelação. Este deslocamento pode ocorrer de diferentes formas para cada janela, definidas por vetores denominados *lag vector*, isto é, um vetor entre duas localizações de uma janela. O vetor de características extraído da imagem contém a média e o desvio padrão dos valores de autocorrelação calculados para cada configuração das janelas. Este descriptor utiliza uma medida própria para o cálculo de similaridade entre

duas texturas  $T_1$  e  $T_2$  denominada “distância SASI” e definida pela fórmula (ÇARKACIOGLU; YARMAN-VURAL, 2003):

$$S(T_1, T_2) = \frac{D'^{T_1} \cdot D'^{T_2}}{D'^{T_1} \cdot D'^{T_1} + D'^{T_2} \cdot D'^{T_2} - D'^{T_1} \cdot D'^{T_2}},$$

sendo  $\cdot$  o produto escalar.

- *Invariant Steerable Pyramid Decomposition* (SID)

A proposta do descriptor SID é representar a característica de textura com base na decomposição piramidal direcionável, a fim de facilitar a invariância à rotação e escala. Neste método a imagem é decomposta em múltiplas resoluções, sendo subdividida em um conjunto de sub-bandas de diferentes escalas e rotações. Inicialmente, a decomposição resulta em duas sub-bandas, obtidas através do uso de dois filtros passa-baixas e passa-altas ( $L_0$  e  $H_0$ , respectivamente). A banda originada pela aplicação de  $L_0$  é decomposta recursivamente em  $k$  sub-bandas ( $B_0, \dots, B_{k-1}$ ), resultantes da aplicação de filtros passa-banda, e uma banda obtida por passa-baixas  $L_1$ ). A cada chamada recursiva considera-se uma informação direcional diferente para a escala em questão (MONTOYA-ZEGARRA et al., 2007).

- *Steerable Mean Standard Deviation* (SMS)

Sendo este descriptor, uma variação do SID, proposto por Montoya-Zegarra et al. (2007), o vetor de características extraído por SMS contém a média, o desvio padrão da distribuição de energia das imagens filtradas a partir do mesmo processo do método anterior.

- *Local Activity Spectrum* (LAS)

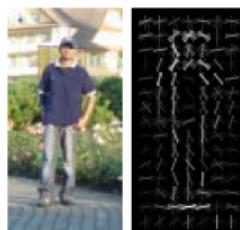
Segundo Tao e Dickinson (2000), LAS é um descriptor de textura que captura a atividade espacial em quatro direções diferentes: horizontal, vertical, diagonal e antidiagonal. A medida das quatro direções é calculada para cada *pixel*, a partir dos valores de sua vizinhança, a partir dos quais é computado um histograma.

### 2.1.1.5 Descritor de Forma Histogramas de Gradientes Orientados

Dalal e Triggs (2005) definem Histogramas de Gradientes Orientados (do inglês, *Histograms of Oriented Gradients* - HOG) como um método descritor que extrai um vetor de características de uma imagem a partir da avaliação dos histogramas locais normalizados de seu gradiente. Este método é baseado no conceito de que a forma de um objeto pode ser caracterizada através da distribuição da intensidade local do gradiente da imagem, ou seja, a informação concentra-se nos contornos dos objetos.

O método original, proposto por Dalal e Triggs (2005), caracteriza-se pela divisão da imagem em um *grid* de células de  $n \times n$  pixels, usualmente  $8 \times 8$ , sendo estas agrupadas em blocos de tamanho  $2 \times 2$ , havendo a possibilidade de sobreposição entre estes. Para cada pixel é calculado o vetor do gradiente, identificando-se a orientação do mesmo. A Figura 2.2 apresenta uma imagem de exemplo do conjunto de treinamento utilizado na pesquisa de Dalal e Triggs (2005) e a representação dos vetores dos gradientes orientados resultante da aplicação do descritor para todas as imagens positivas da mesma base, ou seja, aquelas que representam pessoas.

Figura 2.2: Exemplo de imagem 64x128 do conjunto de treinamento e a representação do descritor HOG resultante do treinamento.



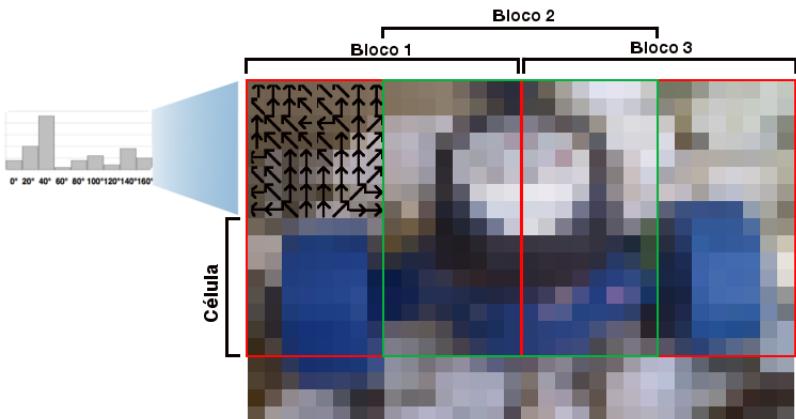
Fonte: Dalal e Triggs (2005)

A partir desta estrutura, o método HOG utiliza uma janela de detecção que percorre a imagem, sendo calculado um histograma do gradiente local de todas as células contidas em cada posição em que a janela encontra-se, à medida que esta se locomove pela imagem. Para a aplicação deste descritor ao problema de detecção de pessoas, Dalal e Triggs (2005) utilizaram uma janela de detecção de tamanho  $64 \times 128 \text{ pixels}$ . Consequentemente, todas as imagens da base de treinamento devem possuir esta dimensão. Cada histograma é formado por 9 *bins*, que referem-se às diferentes direções que o gradiente pode assumir em cada *pixel*, variando de  $20^\circ$  em  $20^\circ$  graus, em um intervalo entre  $0^\circ$  e  $180^\circ$  graus. Assim, as orientações que ocorrem com maior frequência dentro da célula possuem maior magnitude e maior impacto sobre o histograma calculado.

Desta forma, a dimensionalidade do vetor de características de cada imagem do conjunto de treinamento corresponde ao seu tamanho,  $64 \times 128 \text{ pixels}$ , dividido por células de  $8 \times 8 \text{ pixels}$ , agrupadas em blocos, duas a duas, e com ocorrência de sobreposição, o que resulta em 105 blocos ( $7 \times 15$  blocos). Como cada bloco contém os histogramas de 4 células e cada histograma é formado por 9 valores, um por *bin*, obtém-se um vetor de características com tamanho  $105 \times 4 \times 9$ , totalizando 3.780 atributos. Após extraídas as características, é possível normalizar os histogramas de cada bloco de células, com o objetivo de reduzir a variação de luminosidade e contraste. A Figura 2.3 demonstra de forma visual os componentes do descritor HOG. Na imagem, observa-se o destaque de três blocos, dos quais o segundo corresponde à sobreposição.

Originalmente, Dalal e Triggs (2005) aplicaram o descritor HOG combinado a um classificador SVM linear. A partir deste conjunto, a técnica proposta fornece, além de um modelo de descrição de características, um mecanismo de detecção de objetos por análise de imagem. Para isto, uma janela de detecção desloca-se pela imagem, aplicando ao classificador SVM o vetor de características resultante da região da foto contida nesta janela. O classificador é treinado com imagens de duas classes:

Figura 2.3: Estrutura do descriptor HOG



Fonte: do próprio autor

uma positiva, contendo exemplos do objeto a ser detectado (neste caso, pessoas) e outra contendo amostras negativas ("não pessoas").

Considera-se o objeto encontrado quando a região avaliada é rotulada como pertencente à primeira classe. Os autores (*op. cit.*) afirmam que esta solução apresenta melhor performance em relação a outros descriptores de imagem, como por exemplo, métodos baseados em *wavelets*, funções utilizadas para representar os componentes em uma imagem através da decomposição hierárquica de um sinal (MOHAN; PAPAGEORGIOU; POGGIO, 2001). Posteriormente, Enzweiler e Gavrila (2009) realizaram uma pesquisa bibliográfica do tipo *survey* e experimentos para avaliar o desempenho de diferentes métodos de detecção de pedestres. Os resultados indicaram que o HOG mostrou-se mais vantajoso do que os demais métodos analisados para detecção em imagens com alta resolução.

## 2.1.2 Classificação

Após a etapa de extração de características, realiza-se a classificação dos dados. Duda, Hart e Stork (2000) definem a etapa de classificação como uma tarefa de recuperação do modelo que originou os padrões, sendo que diferentes técnicas de classificação podem ser aplicadas dependendo do tipo de modelo a ser encontrado. Segundo Lugosi (2002), a classificação é a atividade de prever a classe a qual uma amostra pertence, sendo esta última um conjunto de medidas numéricas representadas por um vetor de dimensão variável. Considerando um problema binário, ou seja, com apenas duas classes, esta ação pode ser descrita matematicamente através da função:

$$g(x) : \mathbb{R}^d \rightarrow \{0, 1\}$$

este mapeamento representa o método classificador e expressa que  $g(x)$  corresponde à classe (0 ou 1) que o classificador atribui a um objeto  $x$ , com base no seu conhecimento prévio e no algoritmo de aprendizado.

Considerando um classificador supervisionado, este modelo compreende as etapas de definição do modelo classificador, treinamento, teste e rotulação dos objetos, tendo como dados de entrada os vetores de características das amostras de treinamento obtidos na etapa anterior. O treinamento pode ser realizado repetidas vezes com diferentes parâmetros, a fim de encontrar o conjunto de parâmetros que retorne um melhor resultado para o problema proposto. Outros métodos de controle e otimização de parâmetros podem ser aplicados para o mesmo fim. Com o classificador definido e treinado, o mesmo é utilizado para rotular os objetos e retorná-los como resposta ao usuário.

A seguir são apresentados alguns métodos de classificação pertinentes ao escopo desta pesquisa.

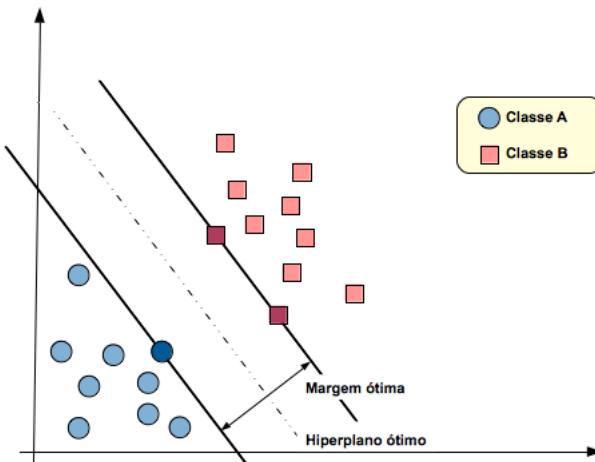
### 2.1.2.1 Máquina de Vetores de Suporte

Conforme Cortes e Vapnik (1995), máquina de vetor de suporte (SVM - *Support Vector Machine*) é um conjunto de métodos de aprendizado de máquina supervisionado, criado por Boser, Guyon e Vapnik (1992), para tratar problemas de classificação binários. O modelo SVM realiza o mapeamento dos vetores de entrada não lineares para um espaço de características de maior dimensão. Desta forma, busca encontrar o hiperplano que melhor separe as amostras de classes distintas, maximizando a margem entre os padrões do conjunto de treinamento e o limite de decisão, ou seja, uma superfície que maximize a distância entre os elementos mais próximos das duas classes.

Os vetores de suporte correspondem aos exemplos do conjunto de treinamento que encontram-se mais próximos do limite de decisão, determinando a margem com a qual as duas classes são separadas (BEN-HUR; WESTON, 2010). Na Figura 2.4, é apresentado um exemplo de problema linearmente separável em um espaço bidimensional. Os vetores de suporte são representados pelas amostras destacadas, sobre as linhas da margem ótima.

Papageorgiou e Poggio (2000) afirmam que uma das vantagens do SVM é o fato deste ser capaz de aprender em espaços de alta dimensionalidade mesmo com um número pequeno de exemplos de treinamento. Ainda segundo os autores, isto é possível, pois o método busca ao mesmo tempo minimizar o limite do erro empírico e a complexidade do classificador. Segundo Ben-Hur e Weston (2010), o modelo SVM pertence à categoria de métodos de *kernel*, sendo este um algoritmo que depende dos dados apenas através do produto interno de seus vetores. Nestes casos, o produto pode ser substituído por outra função de *kernel*, podendo ser, linear, gaussiano (*Radial Basis Function kernel* - RBF), polinomial, entre outras. Para um classificador SVM de *kernel* linear, sua função discriminante pode ser escrita na forma (PAPAGEORGIOU; POGGIO, 2000):

Figura 2.4: Exemplo de SVM em problema de espaço bidimensional



Fonte: Adaptado de Cortes e Vapnik (1995)

$$f(x) = \theta \left( \sum_{i=1}^l \alpha_i y_i K(x, x_i) + b \right),$$

onde  $l$  corresponde ao número de vetores de suporte,  $b$  ao valor de *bias*,  $\alpha$  ao vetor de pesos, ambos definidos na etapa de treinamento.  $K(x, x_i)$  representa a função de *kernel* adotada e  $x$  o vetor de características de entrada. O *kernel* Gaussiano é definido pela função:

$$K(x, x_i) = \exp(-\gamma \|x - x_i\|^2),$$

onde  $\gamma$  corresponde ao parâmetro que controla a largura da curva Gaussiana, definindo o alcance da influência de um exemplo de treinamento e influenciando assim na flexibilidade dos resultados do classificador. Um valor grande a  $\gamma$  tende a levar ao supertreino (BEN-HUR; WESTON, 2010).

A escolha pelo tipo de *kernel* adotado depende do problema que pretende-se solucionar, do conjunto de dados de treinamento e das carac-

terísticas utilizadas na classificação. Segundo Ben-Hur e Weston (2010), usualmente utiliza-se primeiro um *kernel* linear, já que o mesmo é de mais fácil parametrização. De forma geral, apenas definindo a constante de suavização de margem  $C$  já possibilita um resultado satisfatório para grande parte das aplicações. Esta constante representa a penalidade aplicada aos erros de classificação. Um valor pequeno para a variável de folga  $C$  possibilita uma maior generalização, uma vez que aplica menor penalização aos erros de classificação, permitindo que alguns dados permaneçam mais próximos da margem de decisão. Porém, esta suavização pode levar a uma maior incidência de erros.

Tendo como base o resultado obtido através da solução linear, aplica-se um *kernel* não-linear na tentativa de obter um melhor resultado. Contudo, a adoção deste tipo de *kernel* pode levar ao supertreino em conjuntos de treinamento com alta dimensionalidade e pequeno número de amostras.

Embora seja um método de classificação binário, podem ser adotadas diferentes estratégias para conduzir uma classificação multiclasse utilizando classificadores SVMs. Milgram, Cheriet e Sabourin (2006) citam a abordagem “dividir e conquistar” como uma das formas mais comuns para simular este comportamento. Esta consiste na decomposição do problema multiclasse em diversos subproblemas binários, sendo construído um classificador SVM para cada um deles. Dentre os métodos de decomposição mais utilizados destaca-se o “um contra todos”, para o qual define-se um classificador SVM por classe, sendo este treinado para distinguir as amostras da classe em questão das amostras de todas as demais classes. Outra estratégia de decomposição, denominada “um contra um” consiste em uma combinação de classificadores SVM, cada um sendo treinado para um par de classes do problema principal.

Segundo os autores (*op. cit.*), a decisão sobre qual estratégia adotar dependerá do problema a ser tratado. Os mesmos sugerem que, de forma geral, para problemas com poucas classes a estratégia “um contra todos” mostra-se mais precisa. No entanto, a abordagem “um contra

um" seria a solução mais adequada para problemas com um grande número de amostras de treinamento, por possibilitar um menor tempo de treinamento.

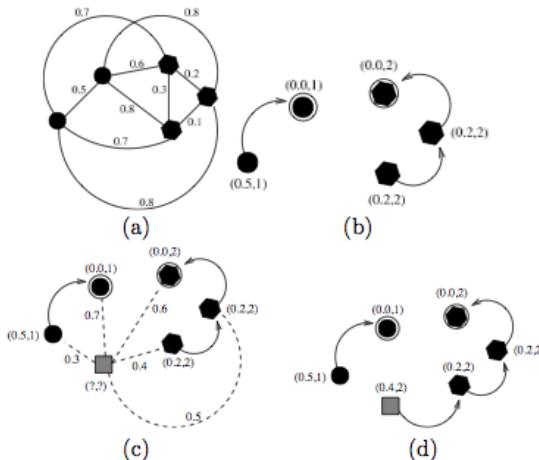
### 2.1.2.2 Florestas de Caminhos Ótimos

O método de classificação OPF (do inglês, *Optimum-Path Forest*), proposto por Papa, Falcão e Suzuki (2009), baseia-se na Teoria de Grafos para reduzir o problema de reconhecimento de padrões ao cálculo de uma floresta de caminhos mínimos. A classificação é tratada, portanto, como um problema combinatorial, onde cada classe é representada por uma ou mais árvores que têm como raiz uma amostra principal, denominada protótipo. As demais amostras do conjunto de treinamento pertencem à árvore do protótipo ao qual este objeto encontra-se mais fortemente conectado (PAPA et al., 2012). A Figura 2.5 exemplifica o processo de treinamento e classificação do OPF, desde a etapa inicial, representada pelo grafo completo (a), as árvores de caminhos mínimos (b) definidas na etapa de treinamento e o processo de classificação de uma nova amostra (c,d).

Segundo Papa, Falcão e Suzuki (2009), cada nó do grafo representa uma amostra do conjunto de treinamento  $z_1$ . Sobre cada uma destas tem-se um vetor de características  $v(s)$  através do qual é possível determinar a dissimilaridade  $d(s, t) \geq 0$  entre duas amostras  $s$  e  $t$ , obtida através de uma função de distância. Usualmente, utiliza-se a distância Euclidiana, porém o método possibilita a aplicação de outras medidas. Considerando que, para o escopo desta pesquisa, cada amostra corresponde a uma imagem, o vetor de características de cada uma destas pode ser codificado através de descritores de imagem, como visto em 2.1.1.

Cada nó conecta-se aos demais através de arcos definidos por uma relação de adjacência e valorados através da função de distância. A etapa de treinamento do classificador inicia-se com um grafo completo. A partir dos nós protótipos são calculadas as árvores de caminhos mí-

Figura 2.5: Representação do funcionamento do classificador OPF



Fonte: Papa, Falcão e Suzuki (2009)

nimos compostas pelas amostras mais fortemente conectadas, onde  $S$  é o conjunto de protótipos de todas as classes ( $S \subset Z_1$ ). Usualmente, os protótipos correspondem aos nós localizados nos limites entre duas classes distintas, sendo que, quando encontrado um arco que o conecte a um vértice de outra classe, esta aresta é eliminada. Neste contexto, denomina-se um caminho, a sequência de amostras distintas de  $\pi_t = \langle s_1, s_2, \dots, t \rangle$ , sendo este finalizado na amostra  $t$ . Cada caminho do grafo resultante possui um valor atribuído através da função de conectividade, que assume o valor do maior arco contido no trajeto entre uma amostra e o nó raiz (PAPA; FALCÃO; SUZUKI, 2009). Desta forma, a função de conectividade pode ser definida por:

$$f_{max} = (\langle s \rangle) = \begin{cases} 0 & \text{se } s \in S \\ +\infty & \text{caso contrário} \end{cases}$$

$$f_{max} = (\pi_s \cdot \langle s, t \rangle) = \max\{f_{max}(\pi_s), d(s, t)\}$$

ou seja, inicialmente todos os protótipos apresentam valor 0 e os demais nós,  $+\infty$ . Cada caminho recebe o maior valor entre o custo do maior arco de  $\pi_s$  e a distância entre as amostras  $s$  e  $t$ . A complexidade de tempo para o treinamento é de  $O(|Z_1|^2)$ , devido ao custo associado para definir um caminho ótimo partindo de  $S$  para cada amostra  $s \in Z_1$ .

A classificação de uma nova amostra  $t$  é realizada através da avaliação, de forma incremental, de todos os caminhos ótimos que conectem os protótipos do conjunto  $S$  a  $t$ , como se  $t$  pertencesse ao grafo de treinamento. Assim, de forma simplificada, pode-se dizer que os protótipos competem para verificar a qual deles a amostra encontra-se mais fortemente conectada. Para isto, utiliza-se uma lista dos vértices, ordenada pelo custo de seus caminhos, avaliando para cada um destes o custo ótimo de  $C(t)$ :

$$C(t) = \min\{\max\{C(s), d(s, t)\}\}, \forall s \in Z.$$

A nova amostra  $t$  será conectada ao nó  $s$  do caminho que obtiver menor  $C(t)$ , sendo  $s$  definido como predecessor de  $t$  e atribuído a esta o rótulo do protótipo da árvore a que  $s$  pertence. A complexidade de tempo para a classificação é de  $O(|Z_1||Z_2|)$ , ou seja, proporcional ao tamanho dos conjuntos de treinamento,  $Z_1$ , e teste,  $Z_2$ (PAPA et al., 2012).

Este classificador foi concebido originalmente como um método de classificação supervisionada, sendo implementada paralelamente por Rocha, Cappabianco e Falcão (2009) uma versão não supervisionada para tratar problemas de agrupamento. Dentre as vantagens do método OPF, destacam-se sua característica nativamente multi-classes, a independência de parâmetros, a possibilidade de tratar uma relativa sobreposição entre classes e o fato de não definir uma forma ou separabilidade do espaço de características (PAPA; FALCÃO; SUZUKI, 2009). Além disto, diversas pesquisas demonstraram que o OPF é um método robusto, podendo ser aplicado a problemas de diversos contextos e obter acurácia equivalente a de classificadores como, SVM, podendo ser superior à rede neural Perceptron Multi-Camadas (do inglês, Artificial Neural

Network Multilayer perceptron - ANN-MLP) e ao método baseado em vizinhança *k-Nearest Neighbor* (k-NN), porém apresentando menor tempo de treinamento (PAPA; FALCÃO; SUZUKI, 2009; RAMOS et al., 2011; SILVA, 2011; PAPA et al., 2012).

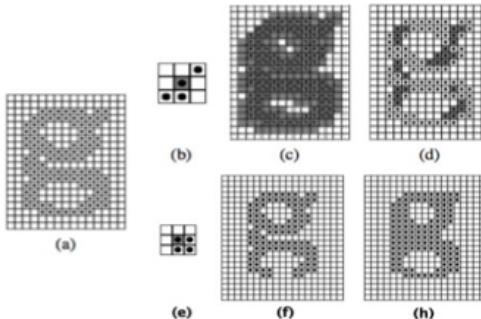
## 2.2 Processamento Morfológico de Imagens

Conforme Dougherty e Lotufo (2003), a “morfologia matemática refere-se ao ramo da análise e processamento de imagens que concentra-se na estrutura geométrica dos componentes contidos em uma imagem”. Haralick, Sternberg e Zhuang (1987) afirmam que estas técnicas usualmente implicam em uma simplificação da imagem, porém preservando as características essenciais de sua forma e eliminando elementos irrelevantes a esta. Os autores (*op. cit*) citam que esta torna-se a abordagem natural para a processamento de operações de detecção de objetos, características, anomalias de forma em aplicações de visão computacional, sendo amplamente utilizada na área de manufatura.

De forma geral, as técnicas morfológicas baseiam-se na realização de uma operação em uma imagem por um elemento estruturante (DOUGHERTY; LOTUFO, 2003). As operações fundamentais da morfologia matemática denominam-se dilatação e erosão, podendo estas serem aplicadas tanto em imagens binárias, quanto em níveis de cinza. Quando aplicada a imagens binárias, a erosão consiste em manter na imagem todos os pontos centrais do elemento estruturante pelos quais a translação do mesmo encaixa-se dentro da imagem. Esta operação resulta em uma imagem de tamanho menor ou igual ao da imagem original. A dilatação trata-se da operação dual à erosão e pode ser definida como o conjunto de *pixels* resultantes da translação do elemento estruturante por todos os pontos da imagem binária (DOUGHERTY; LOTUFO, 2003). A Figura 2.6 demonstra o efeito das operações descritas acima em uma imagem binária.

A partir da combinação dos operadores citados pode-se obter as demais operações morfológicas como, por exemplo, abertura, fechamento, etc.

Figura 2.6: Operações morfológicas. (a) imagem original A. (b) elemento estruturante B. (c) dilatação de A por B. (d) erosão de A por B. (e) elemento estrutura E. (f) abertura de A por E. (g) fechamento de A por E.



Fonte: Dougherty e Lotufo (2003)

mento, *top-hat* e gradiente. A aplicação da abertura tem como resultado a eliminação de objetos pequenos e de ruídos da imagem, mantendo seu tamanho original aproximado. Isto é obtido através da combinação de uma operação de erosão seguida por uma dilatação, sendo estas realizadas com elemento estruturante de mesmo tamanho e formato. Com a aplicação do operador dual da abertura, o fechamento, elimina-se os buracos de tamanho menor ao do elemento estruturante utilizado, sendo esta operação realizada através de uma erosão executada sobre a imagem dilatada. Mais detalhes sobre a teoria e técnicas de processamento morfológicos podem ser encontrados em (DOUGHERTY; LOTUFO, 2003; HARALICK; STERNBERG; ZHUANG, 1987; GONZALEZ; WOODS, 2002).

## 2.3 Segmentação

Pal e Pal (1993) afirmam que a segmentação é “o processo de partitionar a imagem em regiões sem intersecção, de forma que cada região é homogênea e que nenhuma união entre duas regiões adjacentes seja homogênea”. Sua definição formal pode ser expressada por um con-

junto  $F$  que contém todos os *pixels* da imagem e  $P()$ , um predicado de uniformidade entre grupos de *pixels* conectados. Através da segmentação  $F$  é dividido em  $n$  subconjuntos ( $S_1, \dots, S_n$ ):

$$\bigcup_{i=1}^n S_i = F \text{ sendo que } S_i \cap S_j = \emptyset, i \neq j$$

O predicado de uniformidade  $P(S_i)$  é verdadeiro para todas as regiões  $S_i$  e falso para  $P(S_i \cup S_j)$  quando  $S_i$  e  $S_j$  são adjacentes.

Usualmente, esta técnica é aplicada para simplificar a imagem, a fim de facilitar sua análise em processamentos posteriores. Outro objetivo consiste em localizar objetos a partir da divisão da imagem em dois conjuntos distintos: o de *pixels* que compõem o objeto e os pertencentes ao plano de fundo.

Diversos algoritmos de segmentação foram propostos pela comunidade científica, desde técnicas de limiarização em níveis de cinza resultando em uma imagem de saída binária (OTSU, 1975; RIDLER; CALVARD, 1978), assim como métodos por Contorno Ativo (KASS; WITKIN; TERZOPOULOS, 1988; CASELLES; KIMMEL; SAPIRO, 1997), *Watershed* (BEUCHER; LANTUÉJOUL, 1979), *Graph Cuts* (BOYKOV; JOLLY, 2001), entre outros.

Segundo Gonzalez e Woods (2002), a limiarização em níveis de cinza é um dos mais populares métodos de segmentação, devido à simplicidade de implementação. Esta pode ser realizada com base em uma informação global como, por exemplo, o histograma da distribuição de intensidade de cor dos *pixels* da imagem. Define-se um valor de limiar  $T$  que, idealmente, esteja entre os valores de intensidade dos *pixels* do objeto e do plano de fundo. Desta forma, todos os *pixels* com nível de cinza  $\leq T$ , enquadraram-se em uma região e os demais pertencem à outra. Isto é, considerando que a saída será uma imagem binária, atribui-se 1 aos *pixels*  $\leq T$  (objeto) e 0, aos *pixels*  $\geq T$  (plano de fundo) (PAL; PAL, 1993).

Porém, definir o limiar que melhor separa estas duas partições nem sempre é uma tarefa trivial. Otsu (1975) propõe um método não paramétrico e não supervisionado de seleção de limiar para segmentação

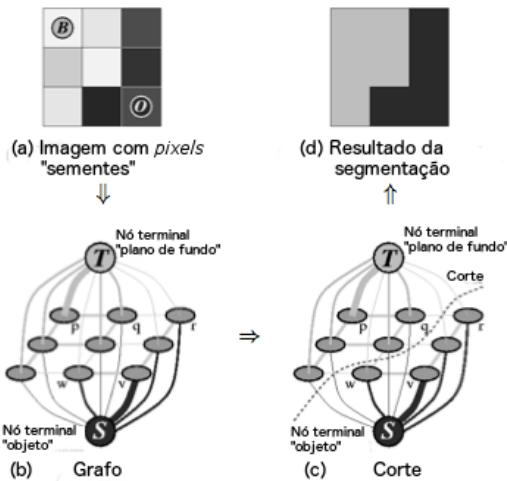
de imagens em níveis de cinza. A determinação do limiar ótimo visa maximizar a separabilidade das classes ( $C_0$  e  $C_1$ ).

Considerando um histograma contendo  $L$  níveis, um limiar  $k$  indica que a classe  $C_0$  é formada pelos *pixels* de níveis  $1, 2, \dots, k$  e,  $C_1 = (k + 1, \dots, L)$ . Sendo  $\sigma_B^2$  a variância entre as classes e  $\sigma_T^2$  a variância total da imagem, o limiar ótimo é obtido pela maximização de  $\frac{\sigma_B^2}{\sigma_T^2}$ . No caso ideal, o método propõe segmentar os componentes conexos principais do plano de fundo da imagem.

Utilizando outra abordagem, o método de cortes em grafos, do inglês, *graph cuts* é um algoritmo de segmentação interativo que divide a imagem em “objeto” e “plano de fundo”, satisfazendo restrições impostas pelo usuário.

Para isto, indica-se um conjunto de *pixels* que pertencem ao objeto a ser segmentado e um conjunto contido no plano de fundo (sementes). A partir destes *pixels* rotulados, o restante da imagem é segmentado computando-se um ótimo global entre todas as segmentações que satisfazam as restrições, com base em uma função de custo definida a partir das propriedades de região e contorno dos segmentos (BOYKOV; JOLLY, 2001). Este método é baseado no algoritmo de otimização combinatorial *max-flow*, proposto por Boykov e Kolmogorov (2001). A Figura 2.7 exemplifica o processo de segmentação aplicado a uma imagem 2D de tamanho  $3 \times 3$  *pixels*. À esquerda (a,b), demonstra-se a representação dos *pixels* da imagem no grafo. O resultado da segmentação (c,d), mostra o conjunto de *pixels* que permanecem vinculados ao nó terminal de plano de fundo (T) e do objeto (S), após definido o corte.

Em 2004, o algoritmo *GrabCut* surgiu como uma proposta de extensão do *Graph Cut*. Segundo os autores Rother, Kolmogorov e Blake (2004), a inovação reside em dois aprimoramentos aplicados à segmentação: a estimativa iterativa e o rotulamento incompleto que permitem alcançar resultados com qualidade satisfatória com um menor interação do usuário em relação ao método anterior.

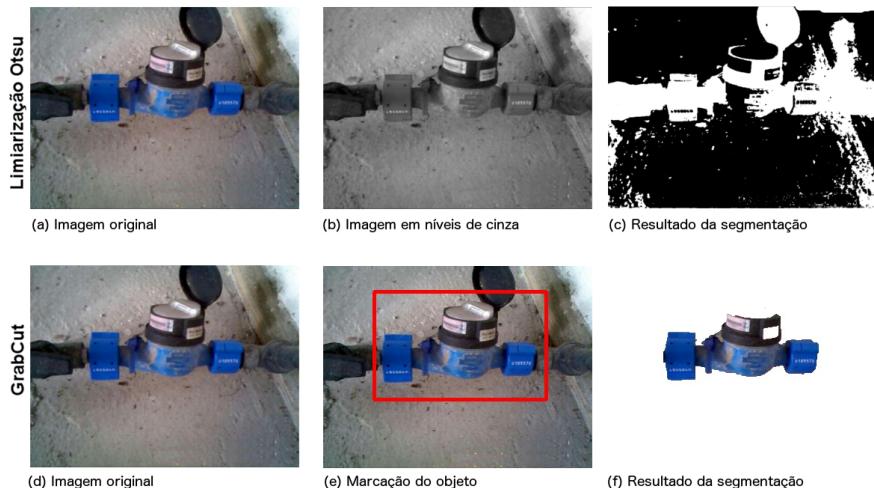
Figura 2.7: Exemplificação do processo de segmentação por *Graph Cut*.

Fonte: Boykov e Jolly (2001) (tradução nossa)

Para segmentar a imagem utilizando *GrabCut*, é necessário apenas que o usuário arraste um retângulo ao redor do objeto de interesse. Com isto, o algoritmo comprehende que toda a região externa ao retângulo é composta seguramente por *pixels* do plano de fundo. Desta forma, o rótulo destes não irão se alterar nas iterações seguintes. Dentro do retângulo, existem *pixels* que pertencem ao objeto de interesse e outros que possivelmente correspondem ao plano de fundo. Para estes é aplicado o conceito do caráter incompleto do rotulamento, atribuindo rótulos provisórios que poderão ser alterados durante a execução. Durante várias iterações estima-se o valor dos rótulos provisórios e a segmentação através da aplicação de cortes em grafos (ROTHER; KOLMOGOROV; BLAKE, 2004). A Figura 2.8 mostra exemplos do resultado da segmentação obtido através dos algoritmos de limiarização Otsu e *GrabCut*.

Observa-se que o primeiro método espera como entrada uma imagem em tons de cinza resultando, sem nenhuma intervenção do usuá-

Figura 2.8: Exemplo de aplicação de métodos de segmentação.



Fonte: do próprio autor

rio, em uma imagem binária contendo o hidrômetro e ruídos do plano de fundo. No segundo exemplo, a segmentação é aplicada e retorna uma imagem no modelo de cores RGB, sendo necessário delimitar a região onde encontra-se o objeto a ser segmentado.

## 2.4 Considerações

Neste capítulo foram apresentados conceitos importantes para auxiliar no entendimento do desenvolvimento da abordagem proposta. Foi possível observar que a definição dos métodos de extração de características e classificação adotados, assim como a identificação das características relevantes, são componentes fundamentais para o funcionamento de um método de detecção de fraudes e de detecção de objetos de forma geral. O pré-processamento de imagens através da aplicação de operações morfológicas e técnicas de segmentação auxilia na remo-

ção de ruídos, objetos indesejados e no destaque das características e regiões de interesse.



### 3 TRABALHOS RELACIONADOS

Neste capítulo serão apresentados trabalhos correlatos ao tema desta pesquisa. Não foram encontradas pesquisas que tratem especificamente da mesma área de aplicação, isto é, a aplicação de métodos de detecção de fraudes em hidrômetros através de análise de imagens. Portanto, a seguir serão mostrados trabalhos relacionados à detecção de objetos e detecção de fraudes ou anomalias.

#### 3.1 Detecção de Objetos

A detecção de objetos por análise de imagens é um tema de grande interesse na comunidade científica, sendo amplamente discutido. Diversos autores publicaram pesquisas relacionadas a propostas de técnicas para detecção de objetos (JAIN; RATHA; LAKSHMANAN, 1997; PAGEORGIOU; OREN; POGGIO, 1998; VIOLA; JONES, 2004; OPELT et al., 2006; BERGBOER; POSTMA; HERIK, 2006; LEIBE; LEONARDIS; SCHIELE, 2008; FELZENSZWALB; GIRSHICK; MCALLESTER, 2010), além de artigos do tipo *survey* sobre o assunto em questão (ROTH; WINTER, 2008; ANDREOPoulos; TSOTSOS, 2013). Áreas de aplicação bastante populares incluem a detecção de faces, sistemas de vigilância (detecção de veículos, detecção de pessoas/pedestres) (OREN et al., 1997; DALAL; TRIGGS, 2005) e aplicações médicas (detecção de tumores, análise de imagens de retina). Com a popularização de mídias de vídeo, surgiram também técnicas de detecção e rastreamento de objetos em movimento, o *object tracking* (YILMAZ; JAVED; SHAH, 2006).

De forma simplificada, Bergboer, Postma e Herik (2006) definem detecção de objetos por computador como "a determinação automática das localizações e tamanhos de objetos em uma imagem, onde os objetos pertencem a uma classe pré-definida."

Andreopoulos e Tsotsos (2013) afirmam que o reconhecimento

de objetos em bases de dados de dimensões relativamente pequenas, com reduzida semelhança entre objetos distintos e em condições de iluminação e em ambientes controlados é considerada pela comunidade científica, de forma geral, um problema da área de visão computacional já solucionado. Porém, questões como oclusão, variação de luminosidade, entre outros, ainda tornam a localização e detecção de objetos um desafio. Papageorgiou e Poggio (2000) complementam que um dos desafios da detecção de objetos em imagens reais, como rostos, pessoas e carros, trata-se da dificuldade em definir modelos destes com variedade significativa de cor e textura, além dos objetos de interesse estarem usualmente situados em fundos complexos, em condições de iluminação e tamanhos diversos e não uniformes. Felzenszwalb (2001) cita também que nem sempre é fácil definir qual característica ou conjunto de características melhor representa um bom modelo para determinado objeto.

Oliva e Torralba (2007) explicam que uma das características da capacidade humana de reconhecer uma grande quantidade de categorias de objetos, apesar da ocorrência de variações no seu posicionamento, ângulos, iluminação e da existência de oclusão, se baseia em informações extraídas da relação entre os objetos em cena, ou seja, no contexto da imagem apresentada. Muitos objetos do mundo real encontram-se usualmente associados a outros, e uma vez que o observador já tenha percebido esta associação anteriormente, a partir deste conhecimento *a priori*, o mesmo pode supor a existência de um deles quando identificar a presença do outro.

A fim de obter maior precisão e menor número de falsos positivos, alguns estudos propõem a detecção de objetos através da exploração do contexto e associação dos mesmos com outros artefatos na imagem. Esta estratégia utiliza a estrutura semântica da cena para reduzir o espaço de busca, procurando o objeto apenas nas regiões onde estejam localizados outros elementos aos quais espera-se o objeto de interesse esteja próximo. Isto beneficia principalmente a busca por objetos pequenos ou que possuam forma bastante genérica e que usualmente estejam

associados a objetos de forma mais distinta. Fazendo uso da metáfora apresentada por Torralba, Murphy e Freeman (2010) esta estratégia pode ser expressada como “utilizar a floresta para encontrar as árvores”.

Utilizando-se das mesmas premissas, Song et al. (2011) propõem um modelo denominado *Context-SVM* que consiste na classificação e detecção de objetos em cena através de um classificador SVM apoiado pelas informações de contexto. A técnica fundamenta-se na relação de dependência entre as características do contexto e as características do objeto de interesse. O modelo é composto por duas etapas, a classificação de objetos e a detecção, sendo os dados de saída sobre os objetos em cena (contexto) obtidos em uma etapa utilizados para ajustar dinamicamente o hiperplano de classificação. Assim, diz-se que o classificador é adaptativo a contexto e busca-se alcançar uma melhora na performance do classificador. Os autores sugerem que, considerando a base de imagens VOC 2007, a técnica proposta supera outros métodos do estado-da-arte em 12 das 20 classes existentes neste conjunto de dados, indicando que informações externas ao objeto procurado podem ser uma ferramenta de grande valia no aprimoramento de métodos de detecção e podendo levar a uma melhora na acertividade.

Diversas pesquisas tratam da aplicação do HOG, em sua versão canônica ou propostas de melhoria, ao problema de detecção de pessoas ou pedestres (BERTOZZI et al., 2007; YANG; WU; ZHANG, 2012; LEE; MIN; SUH, 2013; LIU; GUO; CHANG, 2014; WU; LAGANIÈRE; PAYEUR, 2015; HOANG; LE; JO, 2014), assim como no trabalho de Dalal e Triggs (2005). O método também tem sido amplamente utilizado em pesquisas relacionadas à detecção de veículos (MADHOGARIA et al., 2015; MORDUZZO; MELGANI, 2014) e a questões relativas a trânsito de forma geral (ZAKLOUTA; STANCIULESCU, 2014).

Llorca et al. (2007) empregaram o descritor HOG à detecção de diferentes posições das mãos. Para realizar a classificação multiclass, o mesmo foi associado ao classificador SVM linear, executando o método “um contra um”, ou seja, construiu-se  $k(k - 1)/2$  classificadores binários

diferentes, sendo cada um treinado com duas classes distintas.

Em uma proposta utilizando abordagem semelhante, Chen e Ellis (2011) aplicaram uma variação do descritor HOG, baseado em uma pirâmide de histogramas de gradientes orientados (PHOG), e o classificador SVM para a detecção e classificação de veículos em tráfego urbano. O objeto de interesse é classificado em quatro categorias: carro, van, ônibus e motocicleta. Para aplicar o classificador SVM ao problema multiclasse, foi adotada a estratégia “um contra todos”.

O classificador OPF foi aplicado a diversas áreas de estudo como, por exemplo, detecção de faces (CHIACHIA et al., 2009), em pesquisas de Recuperação de Imagens Baseada em Conteúdo (*Content-Based Image Retrieval - CBIR*) (SILVA; FALCÃO; MAGALHÃES, 2011), classificação de sinais de arritmia cardíaca por análise de eletrocardiograma (LUZ et al., 2013), detecção de nódulos mamários (RIBEIRO et al., 2014), diagnóstico de parasitos intestinais por análise de imagem (SAITO et al., 2015), entre outros.

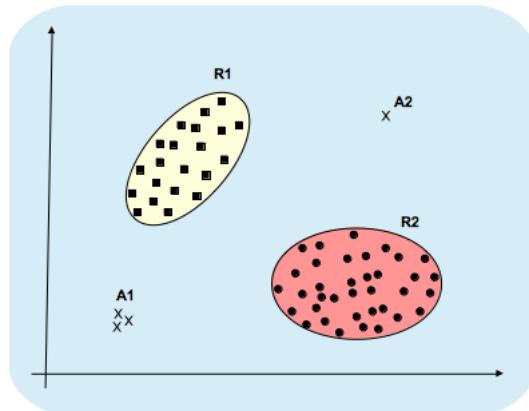
É de conhecimento da autora apenas um trabalho que menciona o uso associado de OPF e HOG (PEREIRA et al., 2013). Uma das abordagens propostas na pesquisa citada codifica as propriedades visuais das amostras, a serem classificadas por OPF, a partir de um modelo *Bag-of-Features* que combina as características extraídas pelos descritores *Spation-temporal Interest Points* (STIP) e HOG. Diferente do escopo da presente pesquisa, no estudo de Pereira et al. (2013) o objetivo principal consiste na proposta de um método de classificação multirrótulo, ou seja, que possibilite a associação de uma amostra a mais de uma classe.

### 3.2 Detecção de fraudes e anomalias

Segundo Chandola, Banerjee e Kumar (2009) denomina-se detecção de anomalias a ação de encontrar padrões em dados que não estão em conformidade com um comportamento esperado e bem definido. A detecção de anomalias é relevante para que possam ser tomadas

ações para minimizá-las, visto que estas usualmente causam danos, sejam financeiros ou, como no caso de fraudes em hidrômetros, a perda de água. A Figura 3.1 apresenta um exemplo de existência de anomalias em um conjunto de dados bidimensional. Nesta podem ser observadas duas regiões bem delimitadas, R1 e R2, que representam o comportamento normal. As amostras isoladas (A1 e A2), distantes destas duas regiões, são identificadas como anomalias.

Figura 3.1: Exemplo de anomalias em conjunto de dados bidimensional



Fonte: Adaptado de Chandola, Banerjee e Kumar (2009)

Uma abordagem simplificada para detecção de anomalias consiste em delimitar uma região que corresponde ao comportamento normal e considerar qualquer dado fora deste espaço como uma anomalia. Porém, o limite entre o comportamento anormal e o esperado nem sempre é preciso, desta forma amostras localizadas próximas às regiões de fronteira poderiam representar tanto uma fraude quanto uma observação normal.

Além disto, em diversos domínios de aplicação o comportamento dos dados observados pode não ser estacionário, ou seja, podem variar de acordo com o tempo, dificultando a definição correta de um comportamento regular. A dificuldade em obter dados rotulados para treinamento

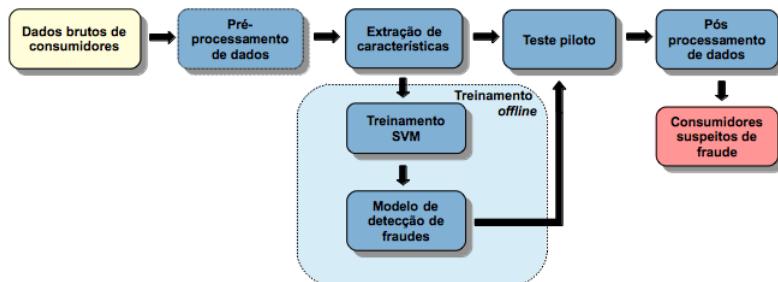
dos modelos, a ocorrência de ruídos em dados e a possibilidade de que, quando uma observação anormal é decorrente de ações maliciosas, tentar adaptar-se para simular uma conduta em conformidade, são outros desafios que compreende esta área de estudo (CHANDOLA; BANERJEE; KUMAR, 2009). Os autores (*op. cit.*) classificam a detecção de fraudes como um tipo de aplicação de detecção de anomalias, definindo-a como a "identificação de atividades criminosas executadas contra organizações comerciais, cujo executor consome de forma não autorizada recursos fornecidos por esta organização". Kou et al. (2004) conceitualizam a detecção de fraudes como o monitoramento do comportamento de grupos de usuários ou consumidores com o objetivo de mapear, estimar, detectar e evitar comportamentos indesejados.

Alguns trabalhos foram encontrados na literatura relacionados à utilização de métodos de reconhecimento de padrões para tratar problemas de fraudes em diversas áreas de aplicação: em consumo de energia elétrica (MONEDERO et al., 2012), telecomunicações (HILAS; MASTOROCOSTAS, 2008), cartões de crédito (GHOSH; REILLY, 1994), entre outros. Nestas pesquisas foram adotadas técnicas de classificação variadas como, por exemplo, redes bayesianas, SVM, algoritmo aglomerativo de clusterização hierárquica, SVM associado a algoritmos genéticos e redes neurais. Durante o desenvolvimento deste estudo não foi identificado nenhum trabalho que abordasse a aplicação de RP para tratar da problemática de detecção de fraudes em hidrômetros. A seguir são detalhados estudos relacionados à detecção de fraudes em um contexto próximo ao da presente pesquisa.

A fim de reduzir o índice de perdas não técnicas, a pesquisa de Nagi et al. (2008) visa localizar potenciais fraudes em unidades consumidoras de energia elétrica. Os autores propõem um *framework* para detecção que utiliza-se de um classificador SVM para identificação de irregularidades e comportamentos de consumo anormais. A Figura 3.2 mostra o esquema de funcionamento da abordagem proposta. Os experimentos foram realizados em um conjunto de dados contendo amostras

de 186.968 consumidores, constituídos de 25 características: 24 valores de consumo médio da unidade nos últimos 25 meses e um atributo adicional que indica se o cliente apresenta atrasos de pagamento. De forma geral, unidades com ocorrência de atraso de pagamento e com redução acentuada de consumo médio apresentam indícios de possível fraude.

Figura 3.2: Framework para detecção de fraude em consumo de eletricidade.



Fonte: Adaptado de Nagi et al. (2008)

Para o conjunto de treinamento foram consideradas 453 amostras, sendo que destas, 131 representavam fraudes confirmadas ou não. O classificador SVM, de *kernel* RBF, categoriza as amostras em quatro classes de acordo com o grau de disparidade do comportamento normal e atípico: inexistência de fraude confirmada, inexistência de fraude não confirmada, suspeita de fraude não confirmada e fraude confirmada. Foi adotada a estratégia “um contra um” para conduzir a classificação multi-classe utilizando SVM.

Para validar os resultados, foram realizadas inspeções presenciais em unidades consumidoras apontadas como suspeitas de fraude. Em relação ao classificador SVM adotado, para definir os valores dos parâmetros  $\gamma$  e  $C$  que retornem resultado mais acertivo, os autores utilizaram o método de otimização *Grid-Search*. Nesta técnica, busca-se a melhor parametrização através da análise dos resultados obtidos em sucessivas execuções do SVM, variando-se os parâmetros em um intervalo e passo

pré-definidos. A performance de cada combinação de parâmetros é avaliada através de uma medida de performance estabelecida (FRIEDRICHSS; IGEL, 2005). Para evitar o supertreinamento, utilizou-se validação cruzada *10-fold*, para a qual divide-se o conjunto de dados em 10 partições. Destas 9 são utilizadas para treinamento e uma para validação. Repete-se o experimento 10 vezes, sempre substituindo a partição de validação por uma ainda não utilizada previamente.

Durante a realização de teste piloto foi observada uma incidência de 78% de falsos positivos nas unidades apontadas com suspeita de fraude. Posteriormente, identificou-se a ocorrência de quatro eventos distintos que influenciavam a classificação incorreta, como por exemplo, a troca do medidor de energia. Foram adotadas ações para detectar estas condições, minimizando os erros de classificação. Os autores afirmam que, com a aplicação do sistema especialista proposto, a taxa de detecção de fraudes apresentou um aumento de 31% em relação ao resultado obtido anteriormente através da realização de inspeções presenciais.

De forma similar, o trabalho de Ramos et al. (2011) propõe a utilização do classificador OPF aplicado à detecção de fraudes, com finalidade de evitar perdas não técnicas de energia elétrica. Nesta pesquisa, as perdas de energia foram caracterizadas pela diferença entre a quantidade de energia gerada e o total faturado, sendo que parte desta diferença é originada do consumo ilegal.

Para a realização dos experimentos foi utilizado um conjunto de dados contendo o perfil de 5.190 unidades consumidoras industriais e 8.067 comerciais, dentre estes havendo consumidores caracterizados como legais (sem fraude) e ilegais (com fraude). Os resultados obtidos pela abordagem proposta, utilizando método OPF, foi comparado a classificadores SVM-Linear, SVM-RBF e às redes neurais ANN-MLP e Mapas Auto-Organizáveis (SOM, do inglês, *Self-Organizing Map*). Os autores afirmam que o OPF mostrou-se superior nos quesitos relacionados à acurácia e tempo de treinamento, apresentando taxa de acerto entre 82% e 85%. Porém, este classificador foi superado pelos demais quando o crité-

rio avaliado é o tempo de classificação.

Apesar de aplicados a diferentes tipos de fraudes, o estudo de Ramos et al. (2011) possui enfoque semelhante ao da presente pesquisa, porém sua classificação é realizada sobre dados numéricos de domínio contínuo, relativos ao perfil de consumo de unidades comerciais e industriais, sendo compostos por quatro características. Na proposta apresentada no presente trabalho, a detecção de fraudes é aplicada em imagens, utilizando técnicas de visão computacional, sendo os dados aplicados ao classificador OPF extraídos de fotos através do descritor de características HOG. Além disto, o estudo de Ramos et al. (2011) foi aplicado a um problema de classificação binária, ou seja, composto por apenas duas classes. O presente estudo pretende beneficiar-se do caráter multiclasse do OPF, aplicando-o a um problema composto por 5 classes distintas, conforme descrito na Subseção 4.2.2.

### 3.3 Considerações

Neste capítulo foram apresentados estudos relacionados a detecção de objetos através de análise de imagens e trabalhos que abordam à detecção de fraudes e anomalias, ambos relevantes para o entendimento e percepção da evolução das pesquisas realizadas pela comunidade científica da área, possibilitando assim o desenvolvimento solução proposta.

Apesar da detecção e reconhecimento de objetos ser um tema amplamente pesquisado pela comunidade científica, a formalização de um método que solucione-o de forma geral a diferentes escopos permanece uma questão em aberto. A detecção de fraudes por análise de imagens ainda é uma área pouco estudada, sendo que, durante a execução desta pesquisa, nenhum trabalho foi encontrado para tratar especificamente do problema de detecção de fraudes em hidrômetros.



## 4 DESENVOLVIMENTO DA SOLUÇÃO

Neste capítulo será detalhada a construção da solução proposta para detecção de fraudes em hidrômetros, assim como as técnicas adotadas e os quesitos considerados para a tomada de decisões.

Diversas ferramentas estão disponíveis para o processamento de imagens e desenvolvimento de aplicações de visão computacional. Para a escolha das bibliotecas utilizadas neste projeto, optou-se por ferramentas implementadas ou com interface para uma mesma linguagem de programação, a fim de otimizar a integração entre estas. Por este motivo, a solução proposta foi desenvolvida em C++. Outro aspecto considerado foi a forma de distribuição, sendo adotadas bibliotecas gratuitas e de código-fonte aberto, possibilitando adaptações necessárias.

Nesta pesquisa, utilizou-se a biblioteca OpenCV na extração de características, através do descritor de imagens HOG, para a aplicação de operações de morfologia matemática, algoritmos de segmentação e manipulação de imagens, como redimensionamento e rotação. Para as operações relacionadas à classificação pelo método OPF, utilizou-se a implementação disponibilizada através do *framework* LibOPF<sup>1</sup>. Para possibilitar a realização de experimentos comparativos do arcabouço proposto, além do classificador OPF, o *framework* de detecção de hidrômetros foi construído de forma a possibilitar a classificação por SVM linear e RBF, para os quais utilizou-se a biblioteca *SVMLight* (JOACHIMS, 1998)

### 4.1 Hidrômetro

Hidrômetros são medidores utilizados para registrar o volume de água consumido por uma unidade em um determinado período. Na Figura 4.1 são apresentados os principais componentes do hidrômetro, entre eles destacam-se os lacres anti-fraude.

---

<sup>1</sup> <http://www.ic.unicamp.br/~afalcao/libopf/node1.html>

Figura 4.1: Estrutura de hidrômetros



Fonte: do próprio autor

Os mecanismos mais simples e populares contêm uma turbina, a qual é rotacionada pela velocidade do fluxo de água que passa por este instrumento. A contagem de rotações da turbina é registrada através de um relógio analógico, possibilitando a medição.

## 4.2 Definição do projeto

Frequentemente são identificados, seja durante a realização da leitura mensal do consumo das unidades ou em fiscalizações periódicas, casos de violação em hidrômetros, efetuados com o objetivo de consumir água de forma irregular. Estas fraudes geram diversas consequências negativas, sendo a principal destas a perda de água, uma vez que usualmente a violação é realizada de forma precária, danificando a tubulação e ocasionando vazamentos. Visto que as tubulações são subterrâneas, muitas vezes estes vazamentos permanecem ocultos sob a terra, dificultando a sua identificação e o conserto, desperdiçando o recurso hídrico cada vez mais escasso. A Figura 4.2 apresenta alguns exemplos de fraude encontrados em fiscalizações realizadas por um empresa de

distribuição e saneamento básico brasileira. Percebem-se exemplos em que o medidor é totalmente retirado e outros no qual o mesmo é danificado ou rotacionado, com o intuito de reduzir a marcação de consumo ou impedindo a realização da mesma. Em grande parte dos casos, é possível notar o rompimento ou remoção dos lacres, o que já identifica uma suspeita de fraude.

Figura 4.2: Exemplos de fraudes em hidrômetros



Fonte: do próprio autor

Devido à grande variedade de possíveis tipos de irregularidades, para delimitar o escopo deste projeto, optou-se por identificar as suspeitas de fraudes através da observação do estado dos lacres do cavalete ou ausência destes, visto que em grande parte das ocorrências de violações há a incidência de rompimento ou retirada dos mesmos, isto é, a retirada dos lacres de cavalete é um forte indício de fraude, uma vez que somente retirando-os é possível efetuar alguns tipos de adulteração no hidrômetro.

Para o desenvolvimento desta proposta, decidiu-se por realizar

a detecção em três etapas principais. Na primeira é localizada a posição do hidrômetro na imagem, mais precisamente a área compreendida entre o medidor e os dois laces de cavalete. Para isto, utilizou-se o descritor HOG combinado ao classificador OPF para detectar o objeto de interesse. A definição pelo uso de um descritor de forma, neste caso HOG, devido ao fato desta ser a característica mais relevante do objeto de interesse. Como este descritor usualmente é utilizado para detecção de pessoas, com este estudo propõem-se também avaliar o seu funcionamento para objetos de formas diferentes. O classificador OPF foi adotado por apresentar características como por exemplo, a independência de parâmetros e ser nativamente multiclasses, o que beneficia a aplicação do mesmo, permitindo que a solução proposta se adapte a mudanças no conjunto de treinamento, permitindo um aprendizado constante do modelo.

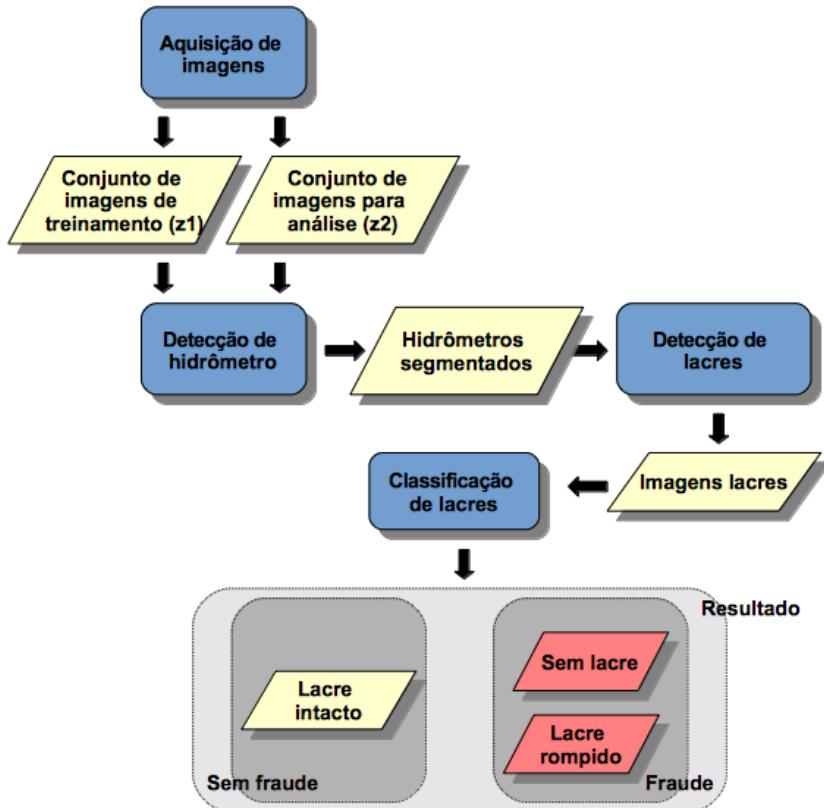
Na segunda etapa, a partir da área selecionada na tarefa anterior, estima-se a região onde deverão estar contidos os laces do cavalete, através de processamento morfológico de imagens e segmentação.

Por último, realiza-se a classificação do estado dos laces, identificando a existência de condição anormal dos mesmos. Nesta etapa, foi adotado novamente o classificador OPF, sendo executados experimentos com doze descritores de cor e textura, a fim de identificar a partir da análise comparativa qual o mais adequado ao problema proposto.

Optou-se por adotar a estratégia por contexto, dado que os laces possuem forma simples e genérica. Consequentemente, uma solução que buscasse detectá-los por toda a imagem poderia, possivelmente, apresentar um alto índice de falsos positivos, uma vez que diversos objetos encontrados no plano de fundo podem apresentar forma semelhante ao lacre. Visto que o lacre procurado sempre está vinculado a um hidrômetro e considerando que este possui uma silhueta mais singular, através da detecção por contexto, a área de interesse é reduzida, passando a ser considerada após a primeira etapa apenas a região que compreende o hidrômetro e, posteriormente, apenas a região do lacre para que possa então ser verificada a suspeita de fraude. A Figura 4.3 apresenta o

esquema da abordagem proposta. Os retângulos representam as ações executadas em cada fase, sendo as entradas e saídas destas retratadas por paralelogramos.

Figura 4.3: Visão macro do esquema da abordagem proposta



Fonte: do próprio autor

Desta forma, a solução proposta é composta pelas seguintes ações:

- **Aquisição de imagens:** obtenção das imagens que compõem o conjunto de treinamento ( $z_1$ ) do classificador da detecção de hidrô-

metros e do conjunto de teste ( $z2$ ), ou seja, as imagens que serão submetidas à análise para detecção das suspeitas de fraudes (Subseção 4.2.1);

- **Detectção de hidrômetro:** localização da posição onde encontra-se o hidrômetro na imagem, da qual extrai-se os vetores de características por um método de descrição, aplicando-os à classificação (Subseção 4.2.2);
- **Detectção de lacres:** identificação da região estimada nos lacres do hidrômetro, executada através de métodos de processamento morfológico de imagens e segmentação (Subseção 4.2.3);
- **Classificação de lacres:** categorização dos lacres para verificar a suspeita de fraude. Para esta etapa, utiliza-se o classificador OPF, sendo realizada a análise comparativa dos resultados obtidos a partir da extração de características por meio de 12 métodos descritores, a fim de avaliar qual o mais adequado para o problema em questão (Subseção 4.2.4).

Observa-se que, como resultado final, é alcançada uma das três condições: lacre encontrado e não violado, o que caracteriza o hidrômetro sem fraude; hidrômetro sem lacre e lacre rompido. Estes dois últimos representam a suspeita de fraude. A seguir são detalhadas as fases do modelo proposto.

#### 4.2.1 Aquisição de imagens

As imagens utilizadas para a realização deste trabalho pertencem a uma base privada cedida pela Companhia Águas de Joinville, uma empresa de saneamento básico e distribuição de água localizada no norte catarinense. Este conjunto é constituído por fotos obtidas através de fiscalizações presenciais realizadas para a averiguação de suspeitas de fraude, portanto não seguem um padrão. Em 2015, durante as ações de fiscalização foram registradas pouco mais de 52.000 fotos.

Deste conjunto foram selecionadas 704 imagens contendo hidrômetros em diferentes posições e rotações, as quais foram recortadas mantendo apenas a área contendo o hidrômetro, sendo redimensionadas para o tamanho da janela de detecção ( $128 \times 80$  pixels). Cada amostra uma das 704 foi rotulada manualmente em uma das possíveis classes de hidrômetro, conforme descrito na Subseção 4.2.2. Do fundo das mesmas imagens foram obtidos 700 recortes no tamanho  $128 \times 80$  pixels, formando a classe de não hidrômetros. As imagens resultantes constituem o conjunto de treinamento formado por 1.404 amostras.

Outras 744 imagens, distintas das anteriores, foram consideradas para formar o conjunto de dados para validação do método proposto.

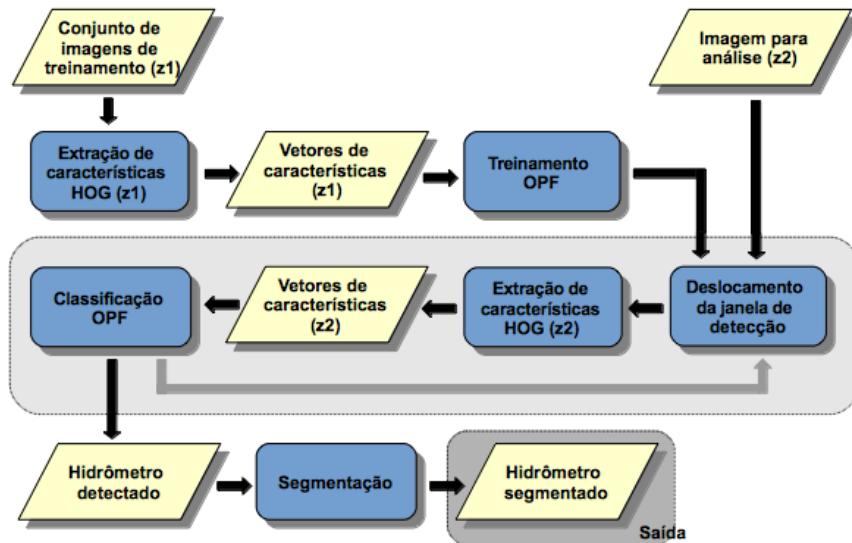
## 4.2.2 Detecção de hidrômetro

Para localizar o hidrômetro na imagem, a mesma é dividida em partes menores das quais extrai-se um vetor de características, sendo este submetido a uma classificação. Como resultado, são destacadas as partes rotuladas pelo classificador como pertencentes a classes que representam o hidrômetro. Esta operação foi realizada utilizando-se a combinação dos métodos de descrição e classificação HOG e OPF, respectivamente. Após encontrado o objeto em questão, a região selecionada é segmentada, resultando na separação do objeto de interesse do plano de fundo. A Figura 4.4 mostra o modelo proposto para a etapa de detecção de hidrômetro, detalhando a visão macro apresentada na Figura 4.3. As principais ações desta etapa são descritas a seguir.

**Extração de características:** observou-se que a característica que mais distingue o hidrômetro dos demais artefatos em cena é a forma. Por este motivo, optou-se por realizar a extração de características através do descriptor HOG, proposto por Dalal e Triggs (2005).

Seja  $z_1$  uma base de imagens, utilizada para treinamento. De cada imagem  $t \in z_1$  extrai-se um vetor de características  $\vec{v}_t$  através de um método descriptor  $(v, d)$ , neste caso, HOG. A similaridade entre duas

Figura 4.4: Modelo proposto para detecção de hidrômetros



Fonte: do próprio autor

imagens  $s$  e  $t$  é calculada através de uma função de distância  $d(s, t)$ . Para a detecção do hidrômetro, utilizou-se a distância Euclidiana.

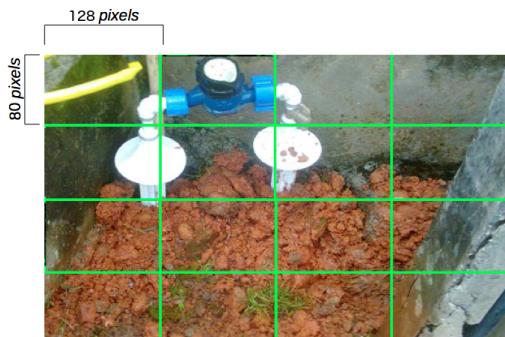
No método HOG, o tamanho das amostras do conjunto de treinamento  $z1$  define a dimensão da janela de detecção ( $n \times m$ ). Usualmente esta possui tamanho reduzido, uma vez que o número de características extraídas é proporcional ao seu tamanho e, consequentemente, influencia no tempo de classificação. Além disto, é necessário atentar que o método localiza objetos de tamanho igual ou maior do que o representado nas amostras do conjunto de treinamento. Contudo, deve-se ponderar que quanto menor o tamanho das amostras de  $z1$ , menos detalhes estas compreendem, o que refletirá na qualidade da classificação. Levando estas questões em consideração, para a solução proposta decidiu-se por um conjunto  $z1$  com amostras de dimensão  $128 \times 80$  pixels.

Considerando uma base de imagens de teste  $z2$ , onde cada ima-

gem  $s \in z2$  pode possuir dimensão variável, desde que esta seja maior ou igual a  $n \times m$ . No contexto desta pesquisa,  $z2$  corresponde ao conjunto de imagens que serão submetidas à análise para verificação da suspeita de fraude.

Sobre cada imagem  $s$  realiza-se uma varredura na qual uma janela de detecção  $n \times m$  percorre  $s$  e a cada posição que se encontra, é extraído um vetor de características  $\vec{v}_s$ , o qual será submetido à classificação. A Figura 4.5 exemplifica este procedimento. Para uma janela de detecção de tamanho  $128 \times 80$  pixels, uma imagem de  $z2$  de dimensão  $512 \times 320$ , seria dividida em 16 partes, desconsiderando a sobreposição de blocos, sendo extraídas as características de cada uma delas e atribuídas à classificação.

Figura 4.5: Exemplo de aplicação de janela de detecção HOG



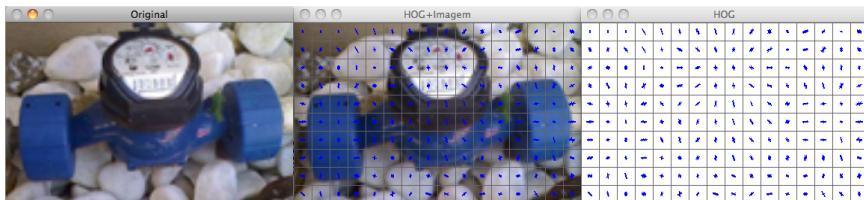
Fonte: do próprio autor

Na primeira posição, a janela de detecção encontra-se em uma área sem hidrômetro, logo, o classificador retornaria um resultado negativo para a detecção. Após o primeiro deslocamento, a janela posiciona-se na região onde o hidrômetro está localizado. O vetor de características da parte da imagem de  $z2$  contida nesta janela é extraído e aplicado à classificação. Caso o treinamento esteja adequado, este vetor será mais similar ao das amostras de uma das classes que representa o objeto de interesse, sendo este detectado. O deslocamento ocorre até que toda a

área da imagem tenha sido percorrida, sendo realizada em múltiplas escadas.

O vetor de características de cada amostra de  $z1$  e de cada parte das imagens de  $z2$  é formado por 4.680 atributos, todos pertencentes ao domínio contínuo. Esta dimensionalidade corresponde ao número de blocos resultante do tamanho da imagem ( $128 \times 80 \text{ pixels} = 15 \times 9$  blocos, com sobreposição), do número de células por bloco (4) e do número de *bins* por histograma (9). A Figura 4.6 apresenta uma amostra do conjunto de treinamento e a visualização da orientação dos gradientes da imagem extraídos pela aplicação do HOG.

Figura 4.6: Exemplos de aplicação do HOG em amostra do conjunto de treinamento



Fonte: do próprio autor

**Classificação:** Para treinar o classificador OPF, o conjunto de treinamento  $z1$  é formado por amostras de cinco classes distintas. As imagens utilizadas para compor esta base não foram incluídas no conjunto de teste  $z2$ , de forma a permitir verificar a generalização do modelo e evitar *overfitting*. As quatro primeiras classes representam as imagens positivas, ou seja, que contêm hidrômetro em diferentes posições, sendo estas: inclinado à direita ( $w_1$ ), horizontal ( $w_2$ ), superior ( $w_4$ ) e inclinado à esquerda ( $w_3$ ). A última classe ( $w_5$ ) corresponde às imagens negativas, que não possuem hidrômetro, sendo obtidas por recortes de fundos das imagens de fiscalização. A Figura 4.7 mostra exemplos de amostras de cada classe.

Utiliza-se o conjunto  $z1$  para formar a florestas de caminhos óti-

Figura 4.7: Classes do conjunto de treinamento



Fonte: do próprio autor

mos. A partir deste são estimadas as amostras mais representativas de cada classe, ou seja, os protótipos ( $S \subset z1$ ) e calcula-se a qual protótipo cada uma das amostras restantes encontra-se mais fortemente conectada. Para toda parte  $s_i$  contida na janela de detecção de cada imagem  $s$  do conjunto  $z2$ , submete-se a mesma ao classificador OPF resultante desta etapa de treinamento. Caso o classificador tenha atribuído  $s_i$  à classe  $w_5$  indica que não foi identificado um hidrômetro na mesma. Caso contrário, ou seja,  $s_i$  foi rotulada como pertencente a  $w_1, w_2, w_3$  ou  $w_4$ , significa que o método encontrou o objeto de interesse da detecção em  $s_i$ .

Inicialmente, o conjunto  $z1$  foi projetado para conter em suas amostras o hidrômetro completo, ou seja, incluindo as pernas do instrumento de medição, como mostra a Figura 4.8 (a). Porém, observou-se que desta maneira, as amostras continham quase metade de sua área cobertas pelo fundo da imagem, devido à concavidade do hidrômetro.

Além disto, em diversos casos, as pernas do medidor encontram-se enterradas sob o solo, o que dificultava a detecção, já que isto modifica a representação do modelo procurado. Por este motivo, optou-se por considerar apenas a região do hidrômetro compreendida entre os dois lacres do cavalete (b). Com esta decisão, notou-se uma melhora nos resultados obtidos em experimentos iniciais.

Figura 4.8: Comparação de imagens do conjunto de treinamento



(a) Hidrômetro completo

(b) Região entre lacres

Fonte: do próprio autor

Devido ao fato do classificador OPF ser um método não paramétrico, seu resultado não depende do ajuste de parâmetros. Entretanto, a formação do conjunto de treinamento é de fundamental importância para o sucesso da aplicação. Quanto mais distintas e separáveis as amostras das diferentes classes, maior será a probabilidade de alcançar uma taxa de acerto satisfatória.

Após a primeira fase de treinamento, foram observados e analisados os casos de falsos negativos (FN), ou seja, áreas de imagens que não continham o objeto de interesse e que foram classificadas como pertencentes a uma das classes de “hidrômetro”, sendo estes incluídos na base e rotulados como imagens negativas. Em seguida, repetiu-se novamente a etapa de treinamento do classificador OPF. Este passo, denominado “hard training”, tem como objetivo melhorar a taxa de acerto, otimizando a detecção e minimizando os erros de classificação.

Originalmente, o método acessa as informações do vetor de ca-

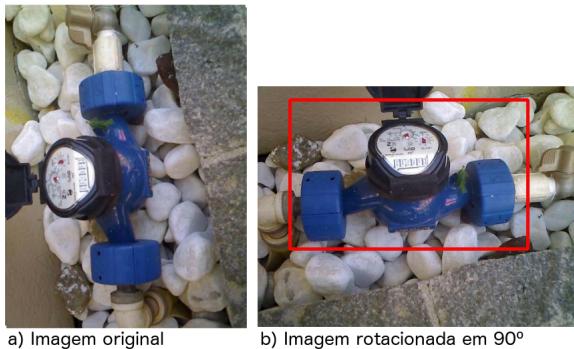
racterísticas de cada imagem a ser classificada a partir de um arquivo binário. A fim de obter um ganho de desempenho, o mesmo foi adaptado para receber como entrada os dados encapsulados em uma estrutura de dados do tipo vetor, a mesma que é retornada como saída do método descriptor. Desta forma, minimiza-se o acesso a disco, uma vez que o vetor já encontra-se em memória e, consequentemente, reduz o tempo de classificação. Para esta aplicação, observou-se que esta alteração impactou positivamente, visto que ao realizar a busca pelo objeto a ser detectado em uma imagem, a classificação é realizada diversas vezes, a cada deslocamento da janela de detecção, até percorrer toda a imagem.

Caso o hidrômetro não tenha sido localizado, a imagem  $s \in z2$  é rotacionada em 90 graus, sendo atribuída novamente ao detector. Esta decisão justifica-se pelo fato do *framework* do descriptor HOG requerer que todas as imagens do conjunto  $z1$  tenham as mesmas dimensões de largura e comprimento (janela de detecção). Neste caso, como o objeto de interesse possui maior medida de largura em relação à altura, a inclusão de amostras na base de treinamento onde o medidor é apresentado na posição vertical teria como consequência a alteração de proporção do mesmo e a inclusão de uma maior área de fundo nas laterais, o que prejudicaria a classificação.

Com a rotação da imagem  $s$  que está sendo analisada, evita-se este fator e possibilita a identificação de hidrômetros que encontram-se posicionados verticalmente, aumentando a eficiência do sistema de detecção. A Figura 4.9 exemplifica a situação exposta, demonstrando que na imagem original, o hidrômetro não seria identificado. Após a rotação, o objeto enquadra-se no modelo a ser detectado.

Outra alternativa seria a construção de dois modelos classificadores com conjuntos de treinamento distintos, uma para a identificação do objeto de interesse na horizontal e outro, na vertical. Porém, esta estratégia demanda um maior tempo de treinamento, uma vez que a operação deverá ser repetida para os dois conjuntos. Desta forma, optou-se pela primeira abordagem, simplificando o modelo.

Figura 4.9: Rotação de imagem para otimizar detecção de hidrômetro



Fonte: do próprio autor

**Segmentação:** Após a detecção da região de interesse, realiza-se a segmentação da imagem através do método *GrabCut*, a fim de separar o objeto de interesse, o hidrômetro, do plano de fundo. A escolha por este método justifica-se por possibilitar a segmentação com menor interação do usuário, sendo necessário apenas delimitar a região onde encontra-se o objeto a ser segmentado (retângulo envolvente). Para obter um resultado satisfatório é necessário manter todo o objeto a ser segmentado dentro do retângulo de detecção informado, que no caso desta pesquisa é definido automaticamente pela detecção HOG+OPF.

Para garantir que esta condição seja atendida, foi adicionada uma margem de 10% em cada lado do retângulo encontrado pela etapa anterior. Desta forma, pretende-se garantir que o objeto estará contido de forma completa no interior do retângulo de detecção, possibilitando assim a localização dos lacre, caso existam. Na Figura 4.10 é mostrado um exemplo da aplicação da margem no retângulo de detecção. O retângulo verde representa a região de detecção resultante da classificação através do método OPF. O retângulo vermelho corresponde à mesma região já contendo o incremento de 10% em todos os lados.

É possível observar que sem esta margem de segurança parte

Figura 4.10: Incremento de margem do retângulo de detecção



Fonte: do próprio autor

dos lacres não seriam considerados nas etapas posteriores à segmentação, prejudicando a análise e identificação das fraudes. Observou-se que este procedimento é vantajoso principalmente quando o objeto de interesse encontra-se em posição inclinada, isto é, com ângulo próximo a 45 graus em relação à horizontal.

O Algoritmo 1 sintetiza o procedimento de detecção de hidrômetros descrito nesta seção. Recebe-se como entradas o conjunto de imagens de treinamento  $z1$ , os rótulos destas amostras ( $\lambda(z1)$ ), ou seja, a classe  $(w_1, \dots, w_5)$  de cada imagem  $t \in z1$  conhecida *a priori*, e a imagem  $s \in z2$  na qual pretende-se detectar o hidrômetro.

Após a extração de características de  $z1$  e  $s$ , executado o treinamento do classificador e a classificação de cada região de  $s$  por onde a janela de detecção deslocou-se, o método obtém como retorno o conjunto  $HD'$  que contém a imagem, já segmentada, dos hidrômetros encontrados em  $s$ . Com esta informação, pode-se proceder para a segunda etapa, descrita na Subseção 4.2.3.

### 4.2.3 Detecção de lacres

Considerando que os lacres do cavalete encontram-se nas duas extremidades do corpo do hidrômetro e tendo como resultado das etapas

---

**Algoritmo 1** Detecção de hidrômetros
 

---

```

1: procedure DETECTAHIDROMETRO( $z1, \lambda(z1), s \in z2$ )
2:   for toda imagem  $t \in z1$  do
3:      $\vec{v}_t \leftarrow extraiCaracteristicasHOG(t)$ 
4:   end for
5:    $OPF \leftarrow$  treina classificador OPF a partir de  $\vec{v}(z1)$  e  $\lambda(z1)$ 
6:    $(x, y) \leftarrow$  posiciona janela de detecção na posição  $(0, 0)$  de  $s$ 
7:    $j \leftarrow 0, HD[] \leftarrow \phi, HD'[] \leftarrow \phi$ 
8:   while  $x < s.largura$  do
9:     while  $y < s.altura$  do
10:     $s_i \leftarrow$  região de  $s$  contida na janela de detecção de tamanho  $m \times$ 
11:     $n$ , posicionada em  $(x, y)$ 
12:     $\vec{v}_{s_i} \leftarrow extraiCaracteristicasHOG(s_i)$ 
13:     $\lambda(s_i) \leftarrow OPF(\vec{v}_{s_i})$  ▷ Classifica  $s_i$ 
14:    if  $\lambda(s_i) = w_1, w_2, w_3$  ou  $w_4$  then ▷ localizou hidrômetro
15:       $HD(j) \leftarrow s_i$ 
16:       $HD(j) \leftarrow$  acrescenta margem de 10% aos limites de  $HD(j)$ 
17:       $HD'(j) \leftarrow$  Segmenta  $s$  por de grabCut( $s, HD(j)$ )
18:       $j++$ 
19:    else
20:      Hidrômetro não localizado na posição  $(x, y)$ 
21:    end if
22:     $y \leftarrow y + m$ 
23:  end while
24: end while
25: Retorna  $HD'[]$ 
26: end procedure
  
```

---

anteriores o objeto já segmentado, para localizar a posição dos lacres foi implementada uma combinação de métodos de processamento de imagens. O Algoritmo 2 apresenta o procedimento executado. O mesmo representa a ação “Detecção de lacres” ilustrada na visão macro do modelo proposto (Figura 4.3).

A partir da imagem do hidrômetro, separam-se os canais de cor RBG, subtraindo os demais do canal azul. Com isto, pretende-se remover grande parte do fundo restante, mantendo o corpo do hidrômetro e lacres, visto que estes possuem predomínio desta cor. Do resultado desta obtém-se o gradiente, por morfologia matemática, o qual é adicionado à imagem de forma a evidenciar a delimitação do objeto. Aplica-se limiarização Otsu (OTSU, 1975), binarizando a imagem. Optou-se por este método de limiarização por ser não supervisionado, encontrando desta forma

---

**Algoritmo 2** Processamento de Imagem para Detecção de Lacres
 

---

```

1: procedure DETECTALACRES(imgOrig, imgHDSegm)
2:   imgRGB[]  $\leftarrow$  split(imgHDSegm);
3:   img  $\leftarrow$  imgRGB[2]  $-$  imgRGB[0]  $-$  imgRGB[1] ▷ B - R - G
4:   elemEstrut  $\leftarrow$  elipse()
5:   imgGrad  $\leftarrow$  gradiente(img, elemEstr)
6:   img  $\leftarrow$  img  $+$  imgGrad
7:   img  $\leftarrow$  limiarizacao(img, 0, 255, OTSU)
8:   img  $\leftarrow$  abertura(img, elemEstr)
9:   img  $\leftarrow$  dilatacao(img, elemEstr)
10:  contornos[]  $\leftarrow$  findContours(contornos)
11:  p1, p2  $\leftarrow$  extremidades(img)
12: ▷ Delimita retângulo Hidrômetro
13:  boundingRect.height  $\leftarrow$  defineRetanguloEnvolvente(p1, p2)
14:  rectangle(img, boundingRect.height) ▷ atribui retângulo à imagem
15:  imgLacreDir  $\leftarrow$  recortaLacres(imgOrig, img, boundingRect, 'D') ▷ retorna imagens contendo área estimada dos lacres
16:  imgLacreEsq  $\leftarrow$  recortaLacres(imgOrig, img, boundingRect, 'E')
17: end procedure
18: procedure RECORTALACRES(imgOrig, img, rectHD, flLado)
19:   if flLado = 'D' then ▷ Lacre direito
20:     rectLacre.x  $\leftarrow$  rectHD.x
21:   else ▷ lacre esquerdo
22:     rectLacre.x  $\leftarrow$  (rectHD.x  $+$  rectHD.width)  $-$  (rectHD.width/3.66)
23:   end if
24:   rectLacre.y  $\leftarrow$  rectHD.y
25:   rectLacre.width  $\leftarrow$  rectHD.width/3.66
26:   rectLacre.height  $\leftarrow$  rectHD.height
27:   imgLacre  $\leftarrow$  grabCut(img, rectLacre)
28:   imgLacre  $\leftarrow$  converteNiveisCinza(imgLacre)
29:   imgLacre  $\leftarrow$  limiarizacao(imgLacre, 0, 255, OTSU)
30:   contornos[]  $\leftarrow$  findContours(imgLacre)
31:   p1, p2  $\leftarrow$  retornaExtremos(contornos)
32:   rectLacre  $\leftarrow$  defineRetanguloEnvolvente(p1, p2)
33:   imgLacre  $\leftarrow$  imgOrig(rectLacre)
34:   return imgLacre;
35: end procedure
36: procedure DEFINERETANGULOENVOLVENTE(p1, p2)
37:   boundingRect.x  $\leftarrow$  p1.x
38:   boundingRect.y  $\leftarrow$  p1.y
39:   boundingRect.width  $\leftarrow$  p2.x  $-$  p1.x
40:   boundingRect.height  $\leftarrow$  p2.y  $-$  p1.y
41:   return boundingRect
42: end procedure
  
```

---

o limiar ótimo para separação entre o objeto de primeiro plano do fundo da imagem, que nesta etapa já encontra-se simplificado. Em seguida são empregadas novamente operações morfológicas: uma abertura, para remover ruídos e objetos pequenos de fundo, e uma dilatação. Esta última com o objetivo de conectar os componentes próximos e garantir que o hidrômetro esteja contido dentro desta máscara binária resultante.

Os contornos da máscara binária são encontrados com o uso do algoritmo proposto por Suzuki e Abe (1985) para delimitar o retângulo envolvente desta máscara. Este método consiste na análise topológica da estrutura de imagens binárias, permitindo obter informações relacionadas às bordas de contorno e de buracos dos componentes conexos contidos na imagem e sobre a relação de envolvimento entre estes. Os autores Suzuki e Abe (1985) definem borda externa de uma imagem binária como um conjunto de pontos que encontram-se no limite entre um componente conexo de *pixels* com valor 1 e um componente de *pixels* 0 que cercam o objeto anterior diretamente. De forma inversa, uma borda de buraco consiste do conjunto de pontos localizados entre um buraco e o componente de *pixels* 1 que o envolve completamente.

De forma geral, os pontos de borda são aqueles que possuem *pixels* de valor inverso, 1 ou 0, em sua vizinhança. Para fins de contextualização, são considerados *pixels* de fundo aqueles que possuem valor 0 e pertencentes aos componentes conexos, os pontos de intensidade 1, neste caso a máscara binária do hidrômetro.

A partir dos contornos, as coordenadas dos pontos extremos desta máscara são identificadas para então ser delimitado o retângulo envolvente da mesma. A partir desta, a imagem é fracionada em três partes. Considera-se como área provável do lacre à esquerda a primeira fração e do lacre direto, a última. Para extrair a imagem dos dois lacres, executa-se o método "recortarLacres", passando como parâmetros a imagem original submetida à detecção, a máscara binária contendo a região do hidrômetro segmentado, o retângulo envolvente do mesmo e por último informa-se qual dos lacres deverá ser retornado.

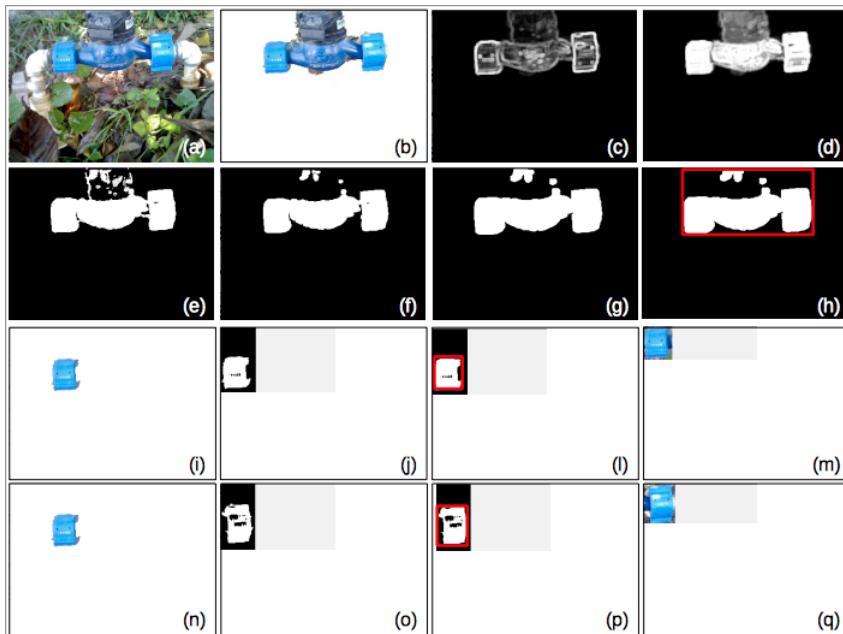
Com estas informações, define-se o retângulo envolvente do lacre procurado, sendo a região da máscara binária contida no mesmo atribuída ao método *GrabCut* (ROTHER; KOLMOGOROV; BLAKE, 2004), a fim de remover partes de objetos do hidrômetro conexos aos lacres. O resultado é binarizado por limiarização, sendo identificado o contorno e o retângulo envolvente do lacre através de método baseado no algoritmo proposto por (SUZUKI; ABE, 1985), desta vez removendo as áreas superior e inferior da região estimada que não correspondem ao lacre. A partir deste contorno, a imagem original é recortada, tendo como resultado final a imagem dos lacres que serão classificados na etapa seguinte.

A Figura 4.11 demonstra, passo a passo, a aplicação da estratégia adotada para a localização estimada dos lacres do cavalete, conforme exposto no Algoritmo 2: (a) imagem original, (b) imagem hidrômetro segmentado, resultante da etapa de detecção de hidrômetro, (c) gradiente do canal azul da imagem, (d) imagem canal azul + gradiente, (e) limiarização Otsu, (f) abertura, (g) dilatação, (h) retângulo envolvente da máscara do hidrômetro, (i) lacre segmentado por *GrabCut* a partir da região à esquerda, (j) máscara binária do lacres à esquerda, (l) retângulo envolvente do lacre, (m) imagem de saída - região do lacre esquerdo segmentado, (n, o, p, q) mesmo processo de (i, j, l, m) repetido para a região do lacre à direita.

#### 4.2.4 Classificação de fraudes

Como saída da etapa anterior (Subseção 4.2.3) tem-se um conjunto de imagens, de tamanhos diversos, contendo apenas a região estimada dos lacres do hidrômetro. Este conjunto é submetido a uma classificação a partir da qual as amostras serão rotuladas, de acordo com a condição do lacre, em três possíveis classes: lacre intacto (sem fraude), lacre rompido ou sem lacre, estas duas últimas representando os casos suspeitos de fraude. A Figura 4.12 apresenta exemplos de imagens das três classes. O método de classificador adotado para esta etapa foi novamente o OPF devido a este ser nativamente multiclassses.

Figura 4.11: Processamento de imagem para localização de lacre



Fonte: do próprio autor

Para extrair as características do conjunto de dados, foram utilizados 12 descritores, sendo 6 de cor e 6 de textura, a fim de possibilitar uma análise comparativa dos mesmos e avaliar qual o mais adequado para o problema em questão ou se há a necessidade de adotar, por exemplo, outros tipos de descritores, uma estratégia de combinação de descritores, entre outros. Os métodos de extração de características considerados foram: BIC (STEHLING; NASCIMENTO; FALCÃO, 2002), ACC (HU-ANG et al., 1997), CSD (MANJUNATH et al., 2001), GHC (SWAIN; BAL-LARD, 1991), *Color Bitmap* (LU; CHANG, 2007), JAC (WILLIAMS; YOON, 2007), LBP (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002), SID (MONTOYA-ZEGARRA et al., 2007), SMS (MONTOYA-ZEGARRA et al., 2007), HTD (MANJUNATH et al., 2001), SASI (ÇARKACIOGLU; YARMAN-VURAL, 2003) e LAS (TAO; DICKINSON, 2000).

Figura 4.12: Exemplos de amostras de conjunto de dados de classificação de lacres

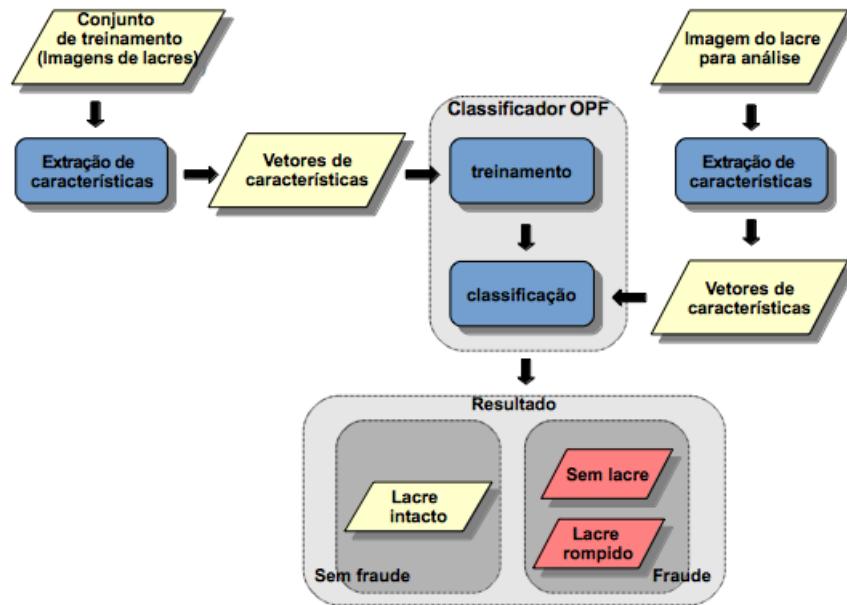


Fonte: do próprio autor

De forma semelhante à etapa de detecção de hidrômetro (Subseção 4.2.2), as principais ações da execução desta fase são a extração de características e a classificação, porém na classificação de fraudes o procedimento é mais simples, uma vez que a classificação é realizada uma única vez por imagem do conjunto de testes, já que aqui o objetivo não é localizar a posição de um objeto, mas sim, verificar a qual classe o mesmo pertence. A Figura 4.13 ilustra o modelo proposto para etapa a classificação de lacres, complementando a visão macro apresentada na Figura 4.3.

Seja  $Z$  uma base de imagens, da qual cada amostra  $r \in Z$  representa uma imagem contendo a região estimada do lacre do hidrômetro. Dado que, diferente do HOG, os descritores utilizados nesta etapa não apresentam como restrição a necessidade de todas as amostras possuírem mesma dimensão de largura e altura, para obter os conjuntos de treinamento  $z_1$  e teste  $z_2$ , divide-se aleatoriamente a base  $Z$ , considerando para cada partição a mesma proporção de imagens das classes do conjunto original. Optou-se pela divisão da base para possibilitar a avaliação

Figura 4.13: Modelo proposto de classificação de lacres



Fonte: do próprio autor

do desempenho do método proposto. Na aplicação prática, os conjuntos  $z_1$  e  $z_2$  possivelmente serão adquiridos em momentos distintos, sendo  $z_2$  representado por cada nova imagem que será submetida à detecção de suspeita de fraude.

O procedimento descrito no Algoritmo 3 é repetido para todos os descritores considerados nesta etapa. O método de classificação de fraudes recebe como entradas o conjunto de imagens  $z$ , os respectivos rótulos ( $\lambda(Z)$ ), o descritor que será utilizado. Para a realização dos experimentos, todo o conjunto  $Z$  foi previamente rotulado, uma vez que a base será dividida aleatoriamente. Na prática, apenas para as imagens que formarão o conjunto  $z_1$  há a necessidade de se conhecer *a priori* a classe a qual pertencem.

---

**Algoritmo 3** Classificação de fraudes
 

---

```

1: procedure CLASSIFICAFraude( $Z$ ,  $\lambda(Z)$ , descriptor)
2:    $\lambda(z1) \leftarrow \phi$ ,  $\lambda(z2) \leftarrow \phi$ 
3:    $(z1, z2) \leftarrow$  divisãoAleatoria( $Z$ )
4:    $\lambda(z1) \leftarrow$  rótulos  $\lambda$  das imagens  $\in z1$ 
5:   for toda imagem  $r \in z1$  do
6:      $\vec{v}_r \leftarrow$  extraiCaracteristicas( $r$ , descriptor)
7:   end for
8:    $OPF \leftarrow$  treina classificador OPF a partir de  $\vec{v}(z1)$  e  $\lambda(z1)$ 
9:   for toda imagem  $u \in z2$  do
10:     $\vec{v}_u \leftarrow$  extraiCaracteristicas( $u$ , descriptor)
11:     $\lambda(u) \leftarrow OPF(\vec{v}_u)$ 
12:     $\lambda(z2) \leftarrow \lambda(z2) \cup \lambda(u)$  ▷ Classifica  $u$ 
13:   end for
14:   retorna conjunto de rótulos de toda imagem  $u \in z2$  ( $\lambda(z2)$ ) atribuídos
      na classificação OPF
15: end procedure
  
```

---

A partir da informação retornada na classificação das fraudes pode-se identificar as imagens que foram atribuídas às classes que correspondem à suspeita de fraude, ou seja, lacre rompido ou sem lacre. Desta forma, o número de imagens que necessitam passar por uma análise individualizada é reduzido, facilitando a tomada de decisão por parte do usuário do sistema especialista de detecção de fraudes.

### 4.3 Considerações

Neste capítulo foi detalhado o desenvolvimento da solução proposta para a implementação do método de detecção de fraudes em hidrômetros. Como visto, o processo é composto por três etapas (detecção do hidrômetro, detecção do lacre e classificação da condição do lacre), que consistem da aplicação de métodos de reconhecimento de padrões e processamento de imagens.



## 5 EXPERIMENTOS E RESULTADOS

Nesta seção são apresentados e discutidos os experimentos realizados para verificar o funcionamento e desempenho da solução desenvolvida. Na etapa de detecção de hidrômetros, os resultados alcançados através da abordagem HOG+OPF são comparados aos obtidos por HOG+SVM linear. Uma análise comparativa dos resultados obtidos pelo método OPF associado a doze descritores aplicados ao problema de classificação de lacres é descrita.

### 5.1 Detecção de hidrômetros

A fim de validar o funcionamento da primeira etapa da abordagem proposta, o classificador OPF foi treinado com um conjunto  $z_1$  formado pelos vetores de características, extraídos pelo descritor HOG, de 1.404 amostras de tamanho  $128 \times 80$  pixels, com a seguinte distribuição entre classes: 116 pertencentes à classe  $w_1$  (hidrômetro inclinado à direita), 328 da classe  $w_2$  (horizontal), 136 correspondem a  $w_3$  (inclinado à esquerda),  $w_4 = 124$  amostras do hidrômetro na perspectiva superior e a classe negativa,  $w_5$ , contendo 700 amostras sem hidrômetro. Esta proporção visa manter um equilíbrio entre as classes positivas ( $w_1$  a  $w_4$ ) e a negativa. Foram consideradas as imagens de lacres resultantes das etapas anteriores, sendo excluídos somente os casos em que houve oclusão parcial do lacre.

Após o treinamento, um conjunto de teste  $z_2$  constituído por 744 imagens, de tamanhos entre  $480 \times 640$  e  $800 \times 600$  pixels, foi submetido à detecção de hidrômetros. Todas as fotos de  $z_2$  apresentam o hidrômetro no ambiente em que o mesmo está localizado. O mesmo conjunto de dados de teste foi aplicado ao *framework* de detecção associado aos classificadores SVM linear e RBF.

A Figura 5.1 apresenta alguns exemplos do resultado obtido ao

aplicar o método de detecção de hidrômetros utilizando o classificador OPF. Nota-se que a localização dos hidrômetros foi detectada, apesar da relativa variação de posição dos mesmos.

Figura 5.1: Resultados de detecção de hidrômetros.



Fonte: do próprio autor

Uma das principais vantagens do modelo de detecção HOG+OPF é o fato do mesmo ser nativamente multiclass, uma característica do classificador OPF. Como o objeto de interesse, neste caso o hidrômetro, pode estar representado em diferentes posições, rotações e perspectivas, isto torna a classificação mais complexa. Isto justifica a decisão de cada posição do objeto ter sido tratada como uma classe distinta, resultando, para a detecção de hidrômetros, em 4 classes positivas.

Esta abordagem pode auxiliar na tomada de decisões em eta-

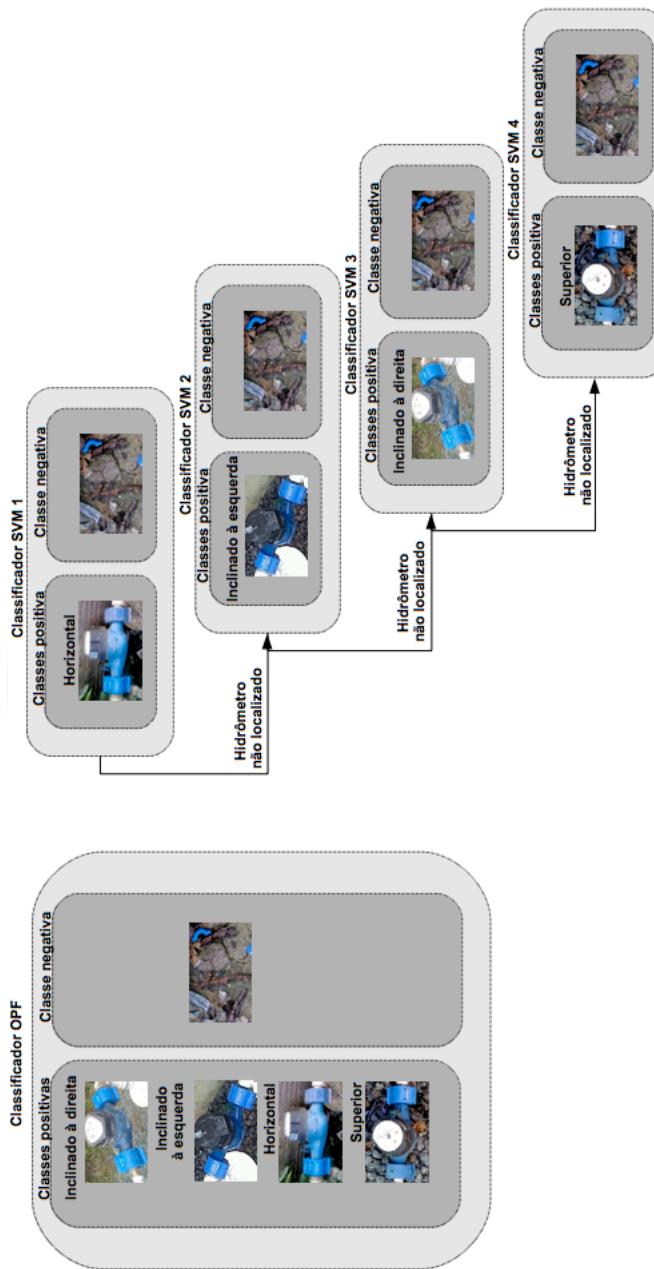
pas e processamentos posteriores, uma vez que tem-se uma informação adicional, um maior conhecimento sobre o posicionamento do objeto na cena. Além disto, possibilita a comparação dos resultados com o classificador SVM de forma mais justa, visto que este último é originalmente binário e assume existir uma forma de separabilidade entre as classes (linear ou gaussiana, por exemplo).

Para auferir um caráter multiclasses ao classificador binário SVM, a alternativa utilizada foi realizar o treinamento de múltiplos classificadores em cascata, um para cada classe do objeto a ser detectado. Em cada um destes, considera-se como classe positiva uma das posições (classe) do objeto e atribui-se a todos classificadores a mesma classe negativa, neste caso, não hidrômetros (recortes de fundos). Para realizar a detecção, cada imagem do conjunto de teste  $z2$  é submetida ao primeiro classificador SVM binário. Caso o mesmo não encontre o objeto de interesse, a imagem é aplicada ao próximo classificador e assim por diante. A Figura 5.2 demonstra e compara a estrutura do classificador OPF e da estratégia adotada para o SVM ao tratar problemas multiclasses.

A Tabela 5.1 apresenta os resultados obtidos pela aplicação dos três modelos classificadores: OPF, SVM linear e SVM RBF. Foi observada a ocorrência de verdadeiros positivos (TP, do inglês, *true positive*), ou seja, a correta detecção do objeto, falsos positivos (*false positive* - FP), isto é, o método apontar a existência do objeto em uma região onde este não se encontra, e falsos negativos (*false negative* - FN), quando o objeto de interesse não é detectado. Nas imagens de  $z2$  onde houve incidência de mais de uma destas condições simultaneamente (FP e TP ou FP e TN), a imagem foi considerada em ambas. Não foram computados os percentuais de verdadeiros negativos (*true negative* TN), uma vez que a busca em cada imagem ocorre diversas vezes, devido ao deslocamento da janela de detecção, e desta forma não é relevante contabilizar o número de regiões em que o hidrômetro não se encontra. Por exemplo, uma imagem de resolução  $800 \times 600$  e uma janela de detecção de tamanho  $128 \times 80$ , é possível encontrar até 2685 verdadeiros negativos, considerando a de-

tecção em múltiplas escalas.

Figura 5.2: Estrutura dos classificadores multiclasse



É possível observar que o modelo HOG+OPF obteve o menor percentual de FP, apontando em apenas 0,65% das imagens a existência de um hidrômetro em uma região onde o mesmo não se encontrava. Esta combinação alcançou uma maior taxa de TP detectando o objeto de interesse corretamente em 89,03% dos casos.

Tabela 5.1: Resultados - Detecção de hidrômetros

Classificador	FP(%)	TP(%)	FN(%)
SVM linear	15,19	82,40	<b>02,41</b>
SVM RBF	21,58	71,53	06,89
OPF	<b>00,65</b>	<b>89,03</b>	10,32

Apesar do OPF indicar um maior percentual de FN do que os demais classificadores, isto ocasiona um menor impacto do que a ocorrência de um maior percentual de FP, considerando-se que o mesmo será adotado em um sistema de detecção de fraudes. Isto por que as imagens em que o sistema especialista não conseguiu localizar o hidrômetro (FN) poderiam ser retornadas ao usuário para uma análise manual (interativa). Já para os casos de FP, ao contrário, o sistema especialista avaliaria a existência de fraude em uma região onde o medidor não está localizado, levando a uma análise incorreta.

Percebeu-se que grande parte dos casos de FN apresentaram condições de oclusão parcial, sendo partes do hidrômetro sobrepostas, por exemplo, por grades e plantas. Variações de iluminação também foram observadas. A luminosidade também foi vista como um fator de influência nas ocorrências de FP. Imagens que continham região significativa de *pixels* com intensidade de cor alta, ou seja, áreas muito claras, apresentaram maior incidência de FP. Esta condição pode ter relação com o fato do relógio dos hidrômetros conter significativa área desta cor, o que supõe-se influenciar neste resultado.

Por ser não paramétrico, o classificador OPF dispensa a necessidade de uma etapa de controle de parâmetros, como ocorre com os classificadores SVM linear e RBF, para os quais é preciso encontrar a me-

lhor configuração das variáveis  $C$  e  $C/\gamma$ , respectivamente, que retornem um resultado satisfatório. Além disto, mudanças no conjunto de treinamento podem implicar na necessidade de reajustar os parâmetros. Neste estudo, os parâmetros dos classificadores SVM foram obtidos de forma empírica, através de diversos experimentos consecutivos. Foram realizados diversos testes onde o critério de parada estabelecido foi a ausência de melhora nas taxas de acerto durante as sucessivas execuções. Métodos exploratórios como, por exemplo, *grid-search* podem ser aplicados de modo a estimar os valores de parâmetros mais adequados ao problema e que alcancem melhor taxa de acertos.

## 5.2 Detecção de lacres

Para validar a segunda etapa, foram utilizadas 662 imagens de hidrômetros segmentados, obtidos como resultado da correta detecção de hidrômetros por OPF da primeira etapa.

Considerando que cada hidrômetro idealmente possui 2 lacres, a partir deste conjunto foram encontradas, de forma acertiva, a região de 1206 lacres, alcançando taxa de acerto de 91,0876%. A Figura 5.3 mostra exemplos de lacres identificados, resultado da execução da segunda etapa.

Observa-se a existência de oclusões parciais por objetos do plano de fundo. Para evitar que esta variável indicasse a ocorrência de fraudes indevidamente e possibilitar avaliar a classificação da condição dos lacres de forma justa, foram desconsideradas as imagens com oclusão na formação do conjunto de amostras submetidos à terceira etapa, descrita na Seção 5.3.

## 5.3 Classificação de lacres

Para a avaliação da classificação dos lacres, considerou-se um conjunto de teste formado por 740 imagens da região esperada dos lacres. Desta base, foram extraídos os vetores de características utilizando-

Figura 5.3: Resultados de detecção de lacres.



Fonte: do próprio autor

se 12 descritores de imagem, 6 de cor e 6 de textura. Para possibilitar a classificação da suspeita de fraude de acordo com a condição do lacre, cada amostra foi rotulada e atribuída a uma das três classes: lacre intacto, sem lacre e lacre rompido.

Desta forma, o conjunto de dados resultante da extração de características dos 12 descritores de imagens foi submetido ao classificador OPF, a fim de compará-los e verificar qual o descritor que melhor se aplica ao problema em questão.

### 5.3.1 Validação cruzada

Com o objetivo de avaliar a capacidade de generalização do método proposto, utilizou-se validação cruzada, isto é, a cada execução o conjunto de dados é dividido aleatoriamente em duas partições,  $z1$  para treinamento e  $z2$  para teste, mantendo a mesma proporção das classes da base completa. Diferentes técnicas para a formação destes subconjuntos podem ser adotadas. Três formas de validação cruzada comumente

utilizadas são: *hold-out*, *k-fold* e *leave-one-out*.

- No primeiro, o conjunto de dados é dividido aleatoriamente em duas partições mutualmente exclusivas, a primeira é destinada para treinamento e a última, para teste. O particionamento pode ser realizado em proporções iguais, porém usualmente utiliza-se maior quantidade de amostras no conjunto de treinamento.
- No método *k-fold*, a base é dividida em  $k$  partições iguais ou aproximadas. A classificação ocorre  $k$  vezes, na qual  $k - 1$  partições são utilizadas como conjunto de treinamento e a restante compõe o conjunto de teste. A cada repetição a partição de teste é substituída, até que todas tenham sido utilizadas.
- A estratégia do método *leave-one-out* consiste em uma variação do *k-fold*, onde  $k$  corresponde ao número de amostras do conjunto de dados. Logo, o conjunto de teste contém apenas uma amostra, sendo todas as demais utilizadas para o treinamento (KOHAVI et al., 1995).

Em todos os métodos, recomenda-se manter a proporção entre classes, evitando assim que uma partição contenha, por exemplo, grande parte das amostras de uma única classe e pouca representatividade nas demais, o que influenciaria negativamente a classificação, uma vez que dificilmente uma amostra destas últimas classes seria rotulada corretamente.

Na Tabela 5.2 são apresentados os resultados obtidos no primeiro experimento, com validação cruzada do tipo *hold-out*. Os seis primeiros descritores correspondem aos de cor e os demais, os de textura. Foi considerada a proporção de 75% das amostras em  $z1$  e 25% em  $z2$ . O procedimento foi repetido 20 vezes, sendo calculada a média e o desvio padrão para cada descritor. É possível observar que, em sua maioria, os descritores de cor alcançaram melhores resultados do que os de textura, apresentando maior taxa de acerto em 19/20 casos. Individualmente, o

descritor BIC obteve o melhor resultado geral (77,8832), além de ser o descritor que obteve a maior taxa de acerto em mais repetições (7/20).

Tabela 5.2: Experimento 1 - Classificação da condição dos lacres

Descriptor	Tipo	Média (%)	Maior (%)	Menor (%)	Desvio padrão (%)
<b>BIC</b>	Cor	71,4998	<b>77,8832</b>	62,0386	$\pm 4,1616$
<b>GCH</b>	Cor	<b>71,9499</b>	76,4276	<b>66,7559</b>	$\pm 2,6971$
<b>ACC</b>	Cor	69,6505	77,1735	64,0527	$\pm 3,2521$
<b>CSD</b>	Cor	69,1645	74,5864	63,3544	$\pm 3,5255$
<b>Color Bitmap</b>	Cor	67,2615	72,5832	62,7083	$\pm 2,4247$
<b>JAC</b>	Cor	66,4643	70,2356	61,9502	$\pm 2,5937$
<b>LBP</b>	Textura	58,0573	61,6918	52,5695	$\pm 2,3999$
<b>SMS</b>	Textura	63,3269	70,0118	56,5498	$\pm 3,7553$
<b>SASI</b>	Textura	67,4626	73,3025	61,7053	$\pm 2,7510$
<b>HTD</b>	Textura	57,9176	64,6995	52,2684	$\pm 3,0603$
<b>SID</b>	Textura	54,8730	59,6100	50,5296	$\pm 2,6739$
<b>LAS</b>	Textura	58,7027	65,6600	53,4400	$\pm 2,8519$

O experimento foi repetido, desta vez utilizando o método de validação cruzada *leave-one-out*. Desta forma, foram realizadas 740 execuções, onde  $z1$  é composto por 739 amostras e  $z2$  contendo apenas uma. Como apenas uma imagem é submetida à classificação, para cada execução o classificador retorna taxa de acerto 100% ou 0%, isto é, atribui o rótulo correto à imagem ou não, respectivamente. Para este experimento, calculou-se a média das 740 execuções. A Tabela 5.3 compara a média dos dois experimentos realizados.

Nota-se que o segundo experimento (*leave-one-out*) apresenta melhor taxa de acerto para 8 descritores e também obteve o melhor resultado geral, 73,3783% (BIC). Isto pode ser explicado pelo fato deste conter maior quantidade de informação no conjunto de treinamento. Apesar disto, a técnica *hold-out* pode ser considerada mais confiável, já que o seu resultado é avaliado a partir de um conjunto maior de amostras de teste, enquanto no *leave-one-out*, o mesmo é estimado a partir da classificação de uma única amostra. Neste último,  $k - 1$  amostras compõem o conjunto de treinamento, o que pode levar ao supertreino, e a avaliação

Tabela 5.3: Experimento 2 - Comparação taxa média de acerto de métodos de validação cruzada

Descriptor	Tipo	Hold-out (%)	Leave-one-out (%)
<b>BIC</b>	Cor	71,4998	<b>73,3783</b>
<b>GCH</b>	Cor	<b>71,9499</b>	72,5675
<b>ACC</b>	Cor	69,6505	70,6756
<b>CSD</b>	Cor	69,1645	72,8378
<b>Color Bitmap</b>	Cor	67,2615	66,3513
<b>JAC</b>	Cor	66,4643	67,0270
<b>LBP</b>	Textura	58,0573	52,2972
<b>SMS</b>	Textura	63,3269	63,7837
<b>SASI</b>	Textura	67,4626	69,0540
<b>HTD</b>	Textura	57,9176	55,6756
<b>SID</b>	Textura	54,873	50,8108
<b>LAS</b>	Textura	58,7027	58,9189

aplicada somente a uma amostra. Caso esta possua características muito distintas das demais, possivelmente poderá ser classificada de forma incorreta. Apesar disto, o método *leave-one-out* é indicado para conjuntos de dados com poucas amostras.

### 5.3.2 Medidas de distância

Para os dois experimentos supracitados, a classificação foi realizada utilizando a distância Euclidiana ( $L_2$ ) como medida para calcular a dissimilaridade. O terceiro experimento foi realizado utilizando medidas distintas por descriptor, de acordo com a orientação descrita na proposta original destes. Quando não mencionada, adotou-se a distância Manhattan ( $L_1$ ). Os dados apresentados na Tabela 5.4 representam a taxa média de acertos de 20 execuções. A cada execução dividiu-se o conjunto de dados aleatoriamente de acordo com o método *hold-out*.

Percebe-se uma melhora em relação aos experimentos anteriores, o que corrobora com a afirmação de Torres et al. (2009) de que a função de similaridade tem papel fundamental no resultado obtido pelo descriptor.

Tabela 5.4: Experimento 3 - Taxa média de acerto

Descriptor	Tipo	Medida	Taxa média de acerto (%)
<b>BIC</b>	Cor	dLog	76,5960
<b>GCH</b>	Cor	L1	79,5025
<b>ACC</b>	Cor	L1	<b>81,2957</b>
<b>CSD</b>	Cor	L1	79,9080
<b>Color Bitmap</b>	Cor	Hamming	71,7540
<b>JAC</b>	Cor	L1	70,8300
<b>LBP</b>	Textura	L1	64,6825
<b>SMS</b>	Textura	L1	72,3843
<b>SASI</b>	Textura	SASI	65,5100
<b>HTD</b>	Textura	HTD	68,7369
<b>SID</b>	Textura	L1	73,3489
<b>LAS</b>	Textura	L1	66,0700

O descriptor ACC apresentou o melhor resultado, alcançando taxa média de acerto de 81,2957%. Assim como nos demais experimentos, é possível observar que os descritores de cor obtêm, de forma geral, resultados melhores que os de textura. Isto sugere que a característica de cor é mais significativa do que a textura no contexto de classificação de lacres, considerando o conjunto de dados utilizado nesta pesquisa. A Tabela 5.5 mostra a matriz de confusão dos resultados obtidos pelo classificador OPF combinado ao descriptor ACC.

Na linha diagonal podem ser observadas as quantidades de amostras classificadas corretamente. Nota-se que, proporcionalmente à quantidade de amostras de cada categoria, a classe “lacre intacto” alcançou maior percentual de acertos individual (90,7310%). O fato desta classe possuir uma quantidade maior de amostras tende a favorecer uma classificação mais acertiva da mesma, uma vez que o classificador possui maior informação para treinamento acerca desta classe. Um conjunto de dados com número equivalente ou aproximado de amostras em todas as classes beneficiaria-se desta condição.

A Figura 5.4 ilustra algumas amostras, sua classe real e a classe predita pelo classificador. Foram selecionadas imagens classificadas com

Tabela 5.5: Matriz de confusão - Classificação da condição dos lacre (OPF + ACC)

		Classe Real		
Classe predita		Lacre intacto	Lacre rompido	Sem lacre
Classe predita	Lacre intacto	1899	264	67
	Lacre rompido	158	955	151
	Sem lacre	36	23	184

mesmo rótulo a partir de todos ou da maioria dos 12 descritores.

Figura 5.4: Condição dos lacre - Exemplos de amostras classificadas

		Classe real		
Classe predita		Lacre intacto	Lacre rompido	Sem lacre
	Lacre intacto			
	Lacre rompido			
	Sem lacre			

Fonte: do próprio autor

Como na tabela Tabela 5.5, imagens classificadas corretamente são apresentadas na linha diagonal. Pela análise visual, é possível observar que amostras de classe “lacre rompido” ou “sem lacre” foram rotuladas incorretamente como “lacre intacto” (sem fraude) quando há um predomínio da cor azul, seja pela ocorrência de uma ruptura de pequena abertura ou pela inclusão parcial do corpo do hidrômetro, que possui mesma cor do lacre. As amostras da classe “lacre intacto” foram classificadas incorretamente usualmente devido a alterações de cor por variação de iluminação.

Estes resultados sugerem que a combinação de dois ou mais descritores que extraiam informação de características distintas (por exemplo, um descritor de cor combinado a um de textura) pode potencialmente favorecer a aplicação.

### 5.3.3 Combinação de descritores

Segundo Mansano et al. (2012) a combinação de descritores é uma abordagem comumente empregada a fim de melhorar o desempenho da classificação alcançada individualmente por descritores que possuem taxas de acerto razoáveis, uma vez que podem fornecer informações diferentes e complementares sobre a amostra avaliada.

Considerando isto, no quarto experimento foi realizada a classificação dos lacares utilizando 4 diferentes combinações entre os descritores de cor e textura, sendo estas apresentadas na Tabela 5.6 em notação polonesa (operadores prefixados). As combinações foram obtidas através da aplicação do *framework* de programação genética proposto por Torres et al. (2009), tendo como entrada as matrizes de distância entre as 740 amostras, considerando os doze descritores e as medidas de distância utilizadas no experimento 3. Utilizou-se a função de avaliação FFP4 (TORRES et al., 2009).

Novas matrizes de distâncias foram calculadas originadas pela aplicação das fórmulas de cada combinação apresentada na Tabela 5.6. As mesmas foram submetidas à classificação OPF, utilizando validação cruzada *hold-out* na proporção de 75% para treinamento e 25% para teste, sendo a divisão da base realizada de forma aleatória, mas mantendo a proporção das classes. A Tabela 5.7 mostra a taxa de acerto média de 30 execuções.

Observa-se que mesmo o melhor resultado das combinações (indivíduo 4) não atingiu o resultado alcançado no experimento 3 pelo melhor descritor individualmente (ACC). Porém, a combinação representada pelo indivíduo 4 é maior do que a média dos 12 descritores. Apesar disto,

Tabela 5.6: Experimento 4 - Combinação de descritores

<b>Indivíduo 1</b>	(+ (+ (+ (% CSD (% SMS JAC)) (+ (+ HTD ACC) (+ HTD SID))) CSD) (+ (+ JAC (* 0.9333663937985948 ACC)) (+ (% (+ HTD HTD) (* 0.9333663937985948 LBP)) (+ HTD SID))))
<b>Indivíduo 2</b>	(+ (+ (* HTD LAS) (+ (+ (* LAS SID) (% SID 0.7586095168194879)) (+ (+ HTD 0.326088518823227) (* CSD 0.5378393631149826)))) (% (+ (+ (* LAS SID) (% SID 0.6530895122504692)) (+ (+ HTD 0.3281778545177495) (* CSD 0.5378393631149826))) (% SMS (* 0.4510786475402585 (+ ACC 0.5842577744521885))))
<b>Indivíduo 3</b>	(+ (+ (* CSD 0.028111079075687107) (% GCH (+ (+ 0.7780880053772969 0.829500619014981) LBP))) (* (% JAC (+ SMS 0.28211717977323314)) (* (+ 0.17962934415632759 SID) LAS) (* 0.731858001493336 0.8055029076598649))))
<b>Indivíduo 4</b>	(* (* (* CSD JAC) (% ACC BIC)) (+ (+ HTD SASI) (* LAS 0.9787088921130868))

Tabela 5.7: Experimento 4 - Taxa de acerto de combinação de descritores

Combinação	Indivíduo 1	Indivíduo 2	Indivíduo 3	Indivíduo 4
<b>Taxa de acerto (%)</b>	67,32	68,18	69,72	<b>72,76</b>

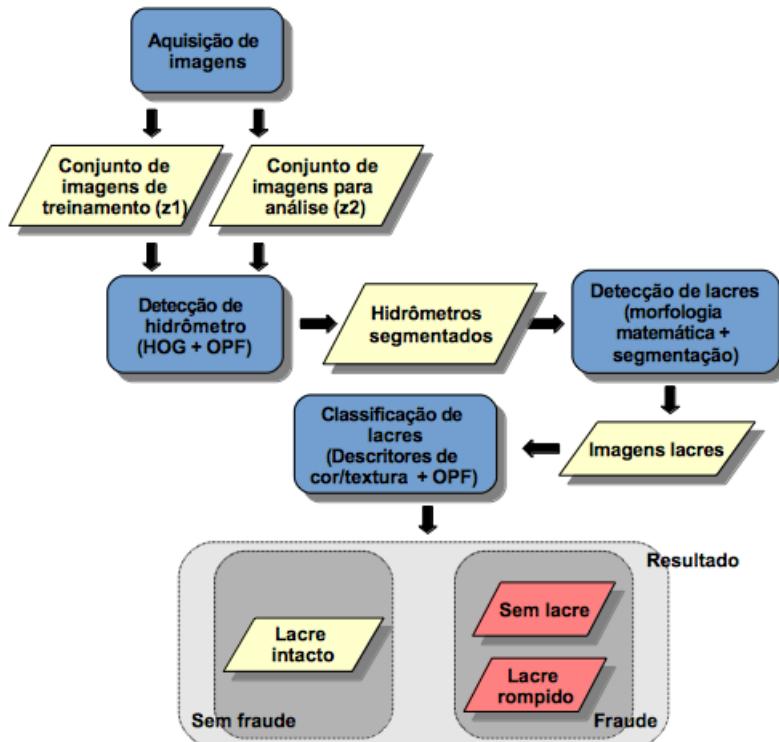
a abordagem mostra-se como uma alternativa promissora, uma vez que a função de avaliação adotada é usualmente utilizada para problemas de *ranking* e CBIR. Uma pesquisa com o intuito de encontrar uma função de *fitness* mais adequada na utilização do *framework* de programação genética na busca de melhores combinações de descritores de imagem para o problema de classificação OPF merece mais estudos.

## 5.4 Considerações

Foram discutidos neste capítulo os resultados alcançados pelo método proposto. A Figura 5.5 mostra a visão macro do esquema da solução proposta, semelhante à apresentada na Figura 4.3, destacando às principais técnicas utilizadas em cada etapa.

Na etapa de detecção de hidrômetros, comparou-se a taxa de

Figura 5.5: Visão macro da solução - técnicas aplicadas



Fonte: do próprio autor

acertos da proposta utilizando o classificador OPF aos métodos SVM linear e RBF. Observou-se que a abordagem HOG+OPF apresenta-se como uma solução com estrutura mais simples, beneficiando-se do caráter multiclasses do método de classificação.

Em relação à etapa de classificação dos lacres, realizou-se uma análise comparativa de doze descritores de cor e textura. Diversos experimentos foram realizados utilizando diferentes métodos de validação cruzada, medidas de distância, além de combinação de descritores através de programação genética. Nesta etapa, observou-se uma melhor taxa de

acerto alcançada pelos descritores de cor, indicando uma tendência desta característica ser mais relevante na classificação de condição dos lacres.

Desta forma, considerando o melhor resultado de cada uma das três etapas, alcança-se para o ciclo completo a eficiência de 65,93%, resultado da detecção de hidrômetros por HOG+OPF (89,03%), detecção de lacres por processamento de imagens (91,0876%) e a classificação da condição dos lacres por OPF associado ao descritor ACC, utilizando distância Manhattan como função de similaridade (81,2957%). A execução do método aplicado a uma imagem levou em média 1,4 segundos.



## 6 CONCLUSÃO

O trabalho desenvolvido apresentou uma pesquisa sobre a aplicação de técnicas de reconhecimento de padrões para o problema de detecção de fraudes visíveis em hidrômetros através de análise de imagens. A proposta é composta de três etapas principais: detecção de hidrômetros, detecção da área estimada dos lacres e classificação dos lacres/suspeitas de fraude. Com isto, procurou-se empregar uma abordagem de exploração por contexto, localizando primeiramente o objeto mais distingível, o hidrômetro, para em seguida identificar os lacres associados a ele, sendo estes os alvos das fraudes analisadas no escopo desta pesquisa.

Destacou-se a importância de um mecanismo automatizado para auxiliar na detecção de suspeitas de fraudes, uma vez que estas são uma das principais causas de vazamentos ocultos e aparentes, consumo irregular de água e, consequentemente, do desperdício do recurso hídrico.

Foi realizada uma pesquisa exploratória através da qual foram encontrados estudos que tratam do uso de técnicas de reconhecimento de padrões para o problema de detecção de fraudes em hidrômetros ou relacionados à fraudes em consumo de água. Poucos trabalhos foram identificados que tratam-se da detecção de fraudes por análise de imagens.

Neste contexto, o método de classificação OPF foi utilizado junto ao *framework* de detecção HOG na detecção de hidrômetros. Foi mostrado que o uso destes métodos combinados é eficiente e possibilita a construção de um modelo de estrutura mais simples ao ser aplicado a problemas multiclasses. A característica nativamente multiclasses do classificador e a independência de parâmetros mostram-se como fatores vantajosos e que favorecem sua adoção.

Desta forma, os resultados observados, para o conjunto de da-

dos utilizado nesta pesquisa, demonstram que a combinação do descritor HOG ao classificador OPF apresenta-se como uma solução promissora para aplicações de detecção de objetos, especialmente quando trata-se de um problema com múltiplas classes. Porém, observou-se que esta solução beneficia-se principalmente quando o objeto de interesse pode ser representando, em todas as classes, por um mesmo tamanho de janela de detecção, sem que isto leve em uma distorção da forma do objeto. Esta restrição está relacionada ao mecanismo de funcionamento do método HOG. Dito isto, caso as classes necessitem de janelas de dimensões diferentes, faz-se necessário o uso de mais de um classificador OPF aninhado, de forma semelhante à adotada usualmente para o classificador SVM.

O método de detecção de hidrômetro HOG+OPF proposto, correspondente à primeira etapa, foi comparado à alternativa HOG+SVM linear e RBF, sendo o primeiro, o método de classificação utilizado na forma canônica do *framework* de detecção HOG. A abordagem proposta apresentou maior percentual de FP, detectando corretamente o hidrômetro em 89,03% dos casos, em comparação aos 82,4% e 71,53% alcançados pelo SVM linear e RBF, respectivamente.

Observa-se que, por ser não paramétrico, o classificador OPF dispensa a necessidade de uma etapa de controle de parâmetros, como ocorre com os classificadores SVM linear e RBF, para os quais é preciso encontrar a melhor configuração das variáveis  $C$  e  $\gamma$ , para SVM RBF, e  $C$ , para SVM linear, que irão retornar um resultado satisfatório. Mudanças no conjunto de treinamento podem implicar na necessidade de reajustar os parâmetros. Visto isto, a independência de parâmetros favorece a obtenção de resultados mais acertivos já nas primeiras execuções do OPF, uma vez que a definição dos parâmetros mais adequados ao problema em questão nem sempre é uma tarefa trivial. Ressalta-se a importância da qualidade da formação do conjunto de dados para o sucesso da aplicação.

Métodos de processamento de imagens, como morfologia mate-

mática e segmentação, foram aplicados para a detecção da localização estimada da região dos lacres.

Diversos experimentos foram realizados na análise comparativa de doze descritores de cor e textura associados ao classificador OPF e aplicados à terceira etapa da solução proposta, ou seja, na classificação da condição dos lacres. Estes experimentos envolveram diferentes técnicas de validação cruzada, medidas de distância e combinação de descritores por programação genética. Como resultado desta análise, pode-se observar uma tendência dos descritores de cor alcançarem melhor eficiência do que os descritores de textura. Isto pode indicar que a característica de cor é mais relevante ao problema estudado.

Em relação à área de aplicação, a utilização de técnicas de reconhecimento de padrões na construção de um sistema automatizado de detecção de fraudes em hidrômetros apresenta-se como uma solução viável que traz benefícios na agilidade de identificação de casos suspeitos, suportando a tomada de decisões.

Nota-se que ações simples podem ser adotadas na tarefa de aquisição de imagens para garantir um maior controle e padronização de condições como enquadramento e qualidade das imagens, podem influenciar positivamente a aplicação, simplificando o modelo e, potencialmente, alcançando níveis de acerto ainda mais altos.

## 6.1 Trabalhos futuros

Como trabalhos futuros sugere-se ampliar os experimentos de combinação de descritores e um estudo para identificar a função de avaliação mais adequada ao problema de classificação, com o intuito de encontrar uma combinação de descritores que possibilite uma maior taxa de acerto na classificação da condição dos lacres.

Propõe-se um estudo para avaliar o uso de outros tipos de descritores ao problema em questão, além dos presentes no escopo desta pesquisa. Também faz-se necessário pesquisar a aplicação de métodos

para identificar e salientar as características mais relevantes das fraudes em lacres.

Também é relevante, ampliar o conjunto de dados de treinamento, incluindo novas amostras de hidrômetros, aumentando assim a variabilidade de informação *a priori* acerca dos objetos procurados.

Além disto, sugere-se abranger o método para detectar outros tipos de fraudes em hidrômetros. No escopo desta pesquisa, foram consideradas as violações ocorridas nos lacres. Poderão ser tratadas outras irregularidades visíveis como, por exemplo, remoção da cúpula do hidrômetro, inversão do medidor, obstruções, entre outros.

Partindo do projeto realizado, diversas direções para novas pesquisas podem ser seguidas como, por exemplo, estudar a aplicação de métodos de apredizado semi-supervisionado, possibilitando o uso de um conjunto maior de imagens sem a necessidade destas estarem previamente rotuladas, uma vez que a rotulação é frequentemente feita de forma manual, o que torna a atividade mais demorada. A partir da abordagem semi-supervisionada é possível obter informações através análise da estrutura das amostras no espaço de características.

Além disto, novos estudos poderão propor a inclusão da interação de usuários especialistas no aprimoramento do treinamento dos classificadores, possibilitando o usuário avaliar os resultados retornados pelo sistema a cada iteração. Isto permite unir seus pontos fortes, a *expertise* do usuário especialista e a capacidade do sistema em processar grandes volumes de informação, na construção e evolução do modelo.

## REFERÊNCIAS

- ANDREOPoulos, A.; TSOTSOS, J. K. 50 years of object recognition: Directions forward. **Computer Vision and Image Understanding**, Elsevier, v. 117, n. 8, p. 827–891, 2013.
- BEN-HUR, A.; WESTON, J. A user's guide to support vector machines. In: **Data Mining Techniques for the Life Sciences**. [S.I.]: Springer, 2010. p. 223–239.
- BERGBOER, N.; POSTMA, E. O.; HERIK, H. J. van den. Context-based object detection in still images. **Image and Vision Computing**, Elsevier, v. 24, n. 9, p. 987–1000, 2006.
- BERNAL, J.; VILARIÑO, F.; SÁNCHEZ, J. **Feature detectors and feature descriptors: Where we are now**. [S.I.], 2010.
- BERTOZZI, M.; BROGGI, A.; ROSE, M. D.; FELISA, M.; RAKOTOMAMONJY, A.; SUARD, F. A pedestrian detector using histograms of oriented gradients and a support vector machine classifier. In: IEEE. **Intelligent Transportation Systems Conference, 2007. ITSC 2007**. IEEE. [S.I.], 2007. p. 143–148.
- BEUCHER, S.; LANTUÉJOUL, C. Use of watersheds in contour detection. In: **International Workshop on Image Processing, Real-time Edge and Motion Detection**. [S.I.]: s.n.], 1979.
- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: ACM. **Proceedings of the Fifth Annual Workshop on Computational Learning Theory**. [S.I.], 1992. p. 144–152.
- BOYKOV, Y.; KOLMOGOROV, V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. In: **Energy Minimization Methods in Computer Vision and Pattern Recognition**. [S.I.]: s.n.], 2001. p. 359–374.
- BOYKOV, Y. Y.; JOLLY, M.-P. Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In: IEEE. **Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on**. [S.I.], 2001. v. 1, p. 105–112.
- BRASIL. **Sistema nacional de informações sobre saneamento**: diagnóstico dos serviços de água e esgoto. Brasília, 2015. 15 p.

- ÇARKACIOGLU, A.; YARMAN-VURAL, F. Sasi: a generic texture descriptor for image retrieval. **Pattern Recognition**, Elsevier, v. 36, n. 11, p. 2615–2633, 2003.
- CASELLES, V.; KIMMEL, R.; SAPIRO, G. Geodesic active contours. **International Journal of Computer Vision**, Springer, v. 22, n. 1, p. 61–79, 1997.
- CHANDOLA, V.; BANERJEE, A.; KUMAR, V. Anomaly detection: A survey. **ACM Computing Surveys (CSUR)**, ACM, v. 41, n. 3, p. 15, 2009.
- CHEN, Z.; ELLIS, T. Multi-shape descriptor vehicle classification for urban traffic. In: IEEE. **Digital Image Computing Techniques and Applications (DICTA), 2011 International Conference on**. [S.I.], 2011. p. 456–461.
- CHIACHIA, G.; MARANA, A. N.; PAPA, J. P.; FALCÃO, A. X. Infrared face recognition by optimum-path forest. In: IEEE. **Systems, Signals and Image Processing, 2009. IWSSIP 2009. 16th International Conference on**. [S.I.], 2009. p. 1–4.
- CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, Springer, v. 20, n. 3, p. 273–297, 1995.
- DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: IEEE. **Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on**. [S.I.], 2005. v. 1, p. 886–893.
- DOUGHERTY, E. R.; LOTUFO, R. A. **Hands-on morphological image processing**. [S.I.]: SPIE press Bellingham, 2003.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification (2a Edição)**. [S.I.]: Wiley-Interscience, 2000. ISBN 0471056693.
- ENZWEILER, M.; GAVRILA, D. M. Monocular pedestrian detection: Survey and experiments. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 31, n. 12, p. 2179–2195, 2009.
- FELZENSZWALB, P. F. Learning models for object recognition. In: IEEE. **Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on**. [S.I.], 2001. v. 1, p. I–1056.

- FELZENZWALB, P. F.; GIRSHICK, R. B.; MCALLESTER, D. Cascade object detection with deformable part models. In: IEEE. **Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on**. [S.I.], 2010. p. 2241–2248.
- FRIEDRICH, F.; IGEL, C. Evolutionary tuning of multiple svm parameters. **Neurocomputing**, Elsevier, v. 64, p. 107–117, 2005.
- GHOSH, S.; REILLY, D. L. Credit card fraud detection with a neural-network. In: IEEE. **System Sciences, 1994. Proceedings of the Twenty-Seventh Hawaii International Conference on**. [S.I.], 1994. v. 3, p. 621–630.
- GONÇALVES, V. M.; NUNES, F. L.; DELAMARO, M. E. Avaliação de funções de similaridade em sistemas de cbir: uma revisão sistemática. 2011.
- GONZALEZ, R. C.; WOODS, R. E. Digital image processing (2a edição). **SL: Prentice Hall**, v. 2, 2002.
- GUYON, I.; ELISSEFF, A. An introduction to feature extraction. In: **Feature Extraction**. [S.I.]: Springer, 2006. p. 1–25.
- HARALICK, R. M.; STERNBERG, S. R.; ZHUANG, X. Image analysis using mathematical morphology. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, n. 4, p. 532–550, 1987.
- HILAS, C. S.; MASTOROCOSTAS, P. A. An application of supervised and unsupervised learning approaches to telecommunications fraud detection. **Knowledge-Based Systems**, Elsevier, v. 21, n. 7, p. 721–726, 2008.
- HOANG, V.-D.; LE, M.-H.; JO, K.-H. Hybrid cascade boosting machine using variant scale blocks based hog features for pedestrian detection. **Neurocomputing**, Elsevier, v. 135, p. 357–366, 2014.
- HUANG, J.; KUMAR, S. R.; MITRA, M.; ZHU, W.-J.; ZABIH, R. Image indexing using color correlograms. In: IEEE. **Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on**. [S.I.], 1997. p. 762–768.
- JAFAR, O. M.; SIVAKUMAR, R. Ant-based clustering algorithms: a brief survey. **International Journal of Computer Theory and Engineering**, v. 2, n. 5, p. 1793–8201, 2010.
- JAIN, A. K.; RATHA, N. K.; LAKSHMANAN, S. Object detection using gabor filters. **Pattern Recognition**, Elsevier, v. 30, n. 2, p. 295–309, 1997.

- JOACHIMS, T. **Making large-scale SVM learning practical**. [S.I.], 1998.
- KASS, M.; WITKIN, A.; TERZOPOULOS, D. Snakes: Active contour models. **International Journal of Computer Vision**, Springer, v. 1, n. 4, p. 321–331, 1988.
- KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: **Ijcai**. [S.I.: s.n.], 1995. v. 14, n. 2, p. 1137–1145.
- KOU, Y.; LU, C.-T.; SIRWONGWATTANA, S.; HUANG, Y.-P. Survey of fraud detection techniques. In: **IEEE. Networking, Sensing and Control, 2004 IEEE International Conference on**. [S.I.], 2004. v. 2, p. 749–754.
- KUNCHEVA, L. I. **Combining pattern classifiers: methods and algorithms**. [S.I.]: John Wiley & Sons, 2004.
- LEE, S. E.; MIN, K.; SUH, T. Accelerating histograms of oriented gradients descriptor extraction for pedestrian recognition. **Computers & Electrical Engineering**, Elsevier, v. 39, n. 4, p. 1043–1048, 2013.
- LEIBE, B.; LEONARDIS, A.; SCHIELE, B. Robust object detection with interleaved categorization and segmentation. **International Journal of Computer Vision**, Springer, v. 77, n. 1-3, p. 259–289, 2008.
- LIU, Y.-F.; GUO, J.-M.; CHANG, C.-H. Low resolution pedestrian detection using light robust features and hierarchical system. **Pattern Recognition**, Elsevier, v. 47, n. 4, p. 1616–1625, 2014.
- LLORCA, D.; VILARINO, F.; ZHOU, Z.; LACEY, G. A multi-class svm classifier ensemble for automatic hand washing quality assessment. In: **BMVC Proc. Brit Mach Vision Conference, Warwick, UK**. [S.I.: s.n.], 2007. p. 213–223.
- LU, T.-C.; CHANG, C.-C. Color image retrieval technique based on color features and image bitmap. **Information Processing & Management**, Elsevier, v. 43, n. 2, p. 461–472, 2007.
- LUGOSI, G. Pattern classification and learning theory. In: **Principles of Nonparametric Learning**. [S.I.]: Springer, 2002. p. 1–56.
- LUZ, E. J. d. S.; NUNES, T. M.; ALBUQUERQUE, V. H. C. D.; PAPA, J. P.; MENOTTI, D. Ecg arrhythmia classification based on optimum-path forest. **Expert Systems with Applications**, Elsevier, v. 40, n. 9, p. 3561–3573, 2013.

- MADHOGARIA, S.; BAGGENSTOSS, P. M.; SCHIKORA, M.; KOCH, W.; CREMERS, D. Car detection by fusion of hog and causal mrf. **Aerospace and Electronic Systems, IEEE Transactions on**, IEEE, v. 51, n. 1, p. 575–590, 2015.
- MANJUNATH, B. S.; OHM, J.-R.; VASUDEVAN, V. V.; YAMADA, A. Color and texture descriptors. **Circuits and Systems for Video Technology, IEEE Transactions on**, IEEE, v. 11, n. 6, p. 703–715, 2001.
- MANSANO, A.; MATSUOKA, J.; AFONSO, L. C.; PAPA, J. P.; FARIA, F.; TORRES, R. d. S. Improving image classification through descriptor combination. In: IEEE. **Graphics, Patterns and Images (SIBGRAPI), 2012 25th SIBGRAPI Conference on**. [S.I.], 2012. p. 324–329.
- MILGRAM, J.; CHERIET, M.; SABOURIN, R. “one against one” or “one against all”: which one is better for handwriting recognition with svms? In: SUVISOFT. **Tenth International Workshop on Frontiers in Handwriting Recognition**. [S.I.], 2006.
- MISSAOUI, R.; SARIFUDDIN, M.; VAILLANCOURT, J. Similarity measures for efficient content-based image retrieval. **IEE Proceedings-Vision, Image and Signal Processing**, IET, v. 152, n. 6, p. 875–887, 2005.
- MITCHELL, T. M. **Machine learning**. WCB. [S.I.]: McGraw-Hill Boston, MA;, 1997.
- MOHAN, A.; PAPAGEORGIOU, C.; POGGIO, T. Example-based object detection in images by components. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 23, n. 4, p. 349–361, 2001.
- MONEDERO, I.; BISCARRI, F.; LEÓN, C.; GUERRERO, J. I.; BISCARRI, J.; MILLÁN, R. Detection of frauds and other non-technical losses in a power utility using pearson coefficient, bayesian networks and decision trees. **International Journal of Electrical Power & Energy Systems**, Elsevier, v. 34, n. 1, p. 90–98, 2012.
- MONTOYA-ZEGARRA, J.; LEITE, N. J.; TORRES, R. da S. et al. Rotation-invariant and scale-invariant steerable pyramid decomposition for texture image retrieval. In: IEEE. **Computer Graphics and Image Processing, 2007. SIBGRAPI 2007. XX Brazilian Symposium on**. [S.I.], 2007. p. 121–128.
- MORANDUZZO, T.; MELGANI, F. Detecting cars in uav images with a catalog-based approach. **Geoscience and Remote Sensing, IEEE Transactions on**, IEEE, v. 52, n. 10, p. 6356–6367, 2014.

- NAGI, J.; MOHAMMAD, A.; YAP, K.; TIONG, S.; AHMED, S. Non-technical loss analysis for detection of electricity theft using support vector machines. In: **IEEE. Power and Energy Conference, 2008. PECon 2008. IEEE 2nd International.** [S.I.], 2008. p. 907–912.
- OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 24, n. 7, p. 971–987, 2002.
- OLIVA, A.; TORRALBA, A. The role of context in object recognition. **Trends in Cognitive Sciences**, Elsevier, v. 11, n. 12, p. 520–527, 2007.
- OPELT, A.; PINZ, A.; FUSSENEGGER, M.; AUER, P. Generic object recognition with boosting. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 28, n. 3, p. 416–431, 2006.
- OREN, M.; PAPAGEORGIOU, C.; SINHA, P.; OSUNA, E.; POGGIO, T. Pedestrian detection using wavelet templates. In: **IEEE. Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on.** [S.I.], 1997. p. 193–199.
- OTSU, N. A threshold selection method from gray-level histograms. **Automatica**, v. 11, n. 285-296, p. 23–27, 1975.
- PAGET, R.; LONGSTAFF, I. D. Texture synthesis via a noncausal nonparametric multiscale markov random field. **IEEE Transactions on Image Processing**, v. 7, n. 6, p. 925–931, 1998. Disponível em: <<http://dx.doi.org/10.1109/83.679446>>.
- PAL, N. R.; PAL, S. K. A review on image segmentation techniques. **Pattern Recognition**, Elsevier, v. 26, n. 9, p. 1277–1294, 1993.
- PAPA, J. P.; FALCÃO, A. X.; ALBUQUERQUE, V. H. C. D.; TAVARES, J. M. R. Efficient supervised optimum-path forest classification for large datasets. **Pattern Recognition**, Elsevier, v. 45, n. 1, p. 512–520, 2012.
- PAPA, J. P.; FALCÃO, A. X.; SUZUKI, C. T. Supervised pattern classification based on optimum-path forest. **International Journal of Imaging Systems and Technology**, Wiley Online Library, v. 19, n. 2, p. 120–131, 2009.
- PAPAGEORGIOU, C.; POGGIO, T. A trainable system for object detection. **International Journal of Computer Vision**, Springer, v. 38, n. 1, p. 15–33, 2000.

- PAPAGEORGIOU, C. P.; OREN, M.; POGGIO, T. A general framework for object detection. In: IEEE. **Computer vision, 1998. Sixth International Conference on**. [S.I.], 1998. p. 555–562.
- PENATTI, O. A.; VALLE, E.; TORRES, R. d. S. Comparative study of global color and texture descriptors for web image retrieval. **Journal of Visual Communication and Image Representation**, Elsevier, v. 23, n. 2, p. 359–380, 2012.
- PEREIRA, L.; PAPA, J. P.; ALMEIDA, J.; TORRES, R. d. S.; AMORIM, W. P. et al. A multiple labeling-based optimum-path forest for video content classification. In: IEEE. **Graphics, Patterns and Images (SIBGRAPI), 2013 26th SIBGRAPI-Conference on**. [S.I.], 2013. p. 334–340.
- RAMOS, C. C. O.; SOUSA, A. N. de; PAPA, J. P.; FALCÃO, A. X. A new approach for nontechnical losses detection based on optimum-path forest. **Power Systems, IEEE Transactions on**, IEEE, v. 26, n. 1, p. 181–189, 2011.
- RIBEIRO, P. B.; COSTA, K. D.; PAPA, J. P.; ROMERO, R. et al. Optimum-path forest applied for breast masses classification. In: IEEE. **Computer-Based Medical Systems (CBMS), 2014 IEEE 27th International Symposium on**. [S.I.], 2014. p. 52–55.
- RIDLER, T.; CALVARD, S. Picture thresholding using an iterative selection method. **IEEE Transactions on Systems, Man and Cybernetics**, v. 8, n. 8, p. 630–632, 1978.
- ROCHA, L. M.; CAPPABIANCO, F. A.; FALCÃO, A. X. Data clustering as an optimum-path forest problem with applications in image analysis. **International Journal of Imaging Systems and Technology**, Wiley Online Library, v. 19, n. 2, p. 50–68, 2009.
- ROTH, P. M.; WINTER, M. Survey of appearance-based methods for object recognition. **Inst. for Computer Graphics and Vision, Graz University of Technology, Austria, Technical Report ICGTR0108 (ICG-TR-01/08)**, 2008.
- ROTHÉR, C.; KOLMOGOROV, V.; BLAKE, A. Grabcut: Interactive foreground extraction using iterated graph cuts. **ACM Transactions on Graphics (TOG)**, ACM, v. 23, n. 3, p. 309–314, 2004.
- SAITO, P. T.; SUZUKI, C. T.; GOMES, J. F.; REZENDE, P. J. de; FALCÃO, A. X. Robust active learning for the diagnosis of parasites. **Pattern Recognition**, Elsevier, 2015.

SILVA, A. T. D.; FALCÃO, A. X.; MAGALHÃES, L. P. Active learning paradigms for cbir systems based on optimum-path forest classification. **Pattern Recognition**, Elsevier, v. 44, n. 12, p. 2971–2978, 2011.

SILVA, A. T. da. **Recuperação de imagens por conteúdo baseada em realimentação de relevância e classificador por floresta de caminhos ótimos**. Tese (Doutorado) — UNIVERSIDADE ESTADUAL DE CAMPINAS, 2011.

SONG, Z.; CHEN, Q.; HUANG, Z.; HUA, Y.; YAN, S. Contextualizing object detection and classification. In: IEEE. **Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on**. [S.I.], 2011. p. 1585–1592.

STEHLING, R. O.; NASCIMENTO, M. A.; FALCÃO, A. X. A compact and efficient image retrieval approach based on border/interior pixel classification. In: ACM. **Proceedings of the Eleventh International Conference on Information and Knowledge Management**. [S.I.], 2002. p. 102–109.

SUZUKI, S.; ABE, K. Topological structural analysis of digitized binary images by border following. **Computer Vision, Graphics, and Image Processing**, Elsevier, v. 30, n. 1, p. 32–46, 1985.

SWAIN, M. J.; BALLARD, D. H. Color indexing. **International Journal of Computer Vision**, Springer, v. 7, n. 1, p. 11–32, 1991.

TAO, B.; DICKINSON, B. W. Texture recognition and image retrieval using gradient indexing. **Journal of Visual Communication and Image Representation**, Elsevier, v. 11, n. 3, p. 327–342, 2000.

TORRALBA, A.; MURPHY, K. P.; FREEMAN, W. T. Using the forest to see the trees: exploiting context for visual object detection and localization. **Communications of the ACM**, ACM, v. 53, n. 3, p. 107–114, 2010.

TORRES, R. S.; FALCÃO, A. X.; GONÇALVES, M. A.; PAPA, J. P.; ZHANG, B.; FAN, W.; FOX, E. A. A genetic programming framework for content-based image retrieval. **Pattern Recognition**, Elsevier, v. 42, n. 2, p. 283–292, 2009.

TRUCCO, E.; VERRI, A. **Introductory techniques for 3-D computer vision**. [S.I.]: Prentice Hall Englewood Cliffs, 1998.

VIOLA, P.; JONES, M. J. Robust real-time face detection. **International Journal of Computer Vision**, Springer, v. 57, n. 2, p. 137–154, 2004.

- WILLIAMS, A.; YOON, P. Content-based image retrieval using joint correlograms. **Multimedia Tools and Applications**, Springer, v. 34, n. 2, p. 239–248, 2007.
- WU, S.; LAGANIÈRE, R.; PAYEUR, P. Improving pedestrian detection with selective gradient self-similarity feature. **Pattern Recognition**, Elsevier, v. 48, n. 8, p. 2364–2376, 2015.
- YANG, M.; KPALMA, K.; RONSIN, J. et al. A survey of shape feature extraction techniques. **Pattern Recognition**, p. 43–90, 2008.
- YANG, S.; WU, J.; ZHANG, L. A fast pedestrian detection method based on simplified hog descriptor. **International Journal of Digital Content Technology & its Applications**, v. 6, n. 4, 2012.
- YILMAZ, A.; JAVED, O.; SHAH, M. Object tracking: A survey. **Acm Computing Surveys (CSUR)**, ACM, v. 38, n. 4, p. 13, 2006.
- ZAKLOUTA, F.; STANCIULESCU, B. Real-time traffic sign recognition in three stages. **Robotics and Autonomous Systems**, Elsevier, v. 62, n. 1, p. 16–24, 2014.
- ZHANG, D.; LU, G. Review of shape representation and description techniques. **Pattern Recognition**, Elsevier, v. 37, n. 1, p. 1–19, 2004.