



Universidade Federal do Piauí
Centro de Ciências da Natureza
Programa de Pós-Graduação em Ciência da Computação

Método de Identificação de Cédulas Monetárias para Deficientes Visuais

Leonardo Pereira de Sousa

Teresina-PI, 29 de Março de 2019

Leonardo Pereira de Sousa

Método de Identificação de Cédulas Monetárias para Deficientes Visuais

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Sistemas de Computação), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ciência da Computação.

Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação

Orientador: Prof. Dr. Rodrigo de Melo Souza Veras

Coorientador: Prof. Dr. Laurindo de Sousa Britto Neto

Teresina-PI

29 de Março de 2019

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca Setorial do CCN

S725m Sousa, Leonardo Pereira de.
Método de identificação de cédulas momentárias para deficientes visuais. / Leonardo Pereira de Sousa. – Teresina, 2019.
44f.: il.:color.

Dissertação (Mestrado) – Universidade Federal do Piauí, Centro de Ciências da Natureza, Pós-Graduação em Ciência da Computação, 2019.

Orientador: Prof. Dr. Rodrigo de Melo Souza Veras.
Coorientador: Prof. Dr. Laurindo de Sousa Britto Neto

1. Sistemas de Computação. 2. Acessibilidade – Deficientes Visuais. 3. Reconhecimento de Cédulas. 4. Tecnologias Assitivas. I. Título.

CDD 004.21

Bibliotecária: Caryne Maria da Silva Gomes. CRB/3 1461

“Método de Identificação de Cédulas Monetárias para Deficientes Visuais”

LEONARDO PEREIRA DE SOUSA

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Natureza da Universidade Federal do Piauí, como parte integrante dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Aprovada por:



Prof. Rodrigo de Melo Souza Veras
(Presidente da Banca Examinadora)



Prof. Fátima Nelszeuma Sombra de Medeiros
(Examinadora Externa à Instituição)



Prof. Romuere Rodrigues Veloso e Silva
(Examinador Externo ao Programa)



Prof. Laurindo de Sousa Britto Neto
(Examinador Interno)

Teresina, 29 de março de 2019

Para Laíla, Leonardo Júnior e Lavínia.

Agradecimentos

Agradeço primeiramente a Deus, por me dar forças e saúde para vencer essa etapa de minha vida. A fé no Senhor, sem dúvidas, me ajudou a lutar até o fim.

Agradeço ao meu orientador, Rodrigo de Melo Souza Veras, pela paciência, compreensão, disponibilidade e grande contribuição para realização deste trabalho.

Aos professores do Programa de Pós-Graduação em Ciência da Computação (PPGCC) pelos ensinamentos e apoio.

À minha esposa Laíla Carvalho Taveira Sousa e meus filhos Leonardo Pereira de Sousa Júnior e Lavínia Carvalho de Sousa, sempre presentes com compreensão nos momentos de ausência, apoio e estímulo durante todo o caminho percorrido.

Agradeço aos meus pais, Manoel Pereira de Sousa e Cecília Moura de Sousa, que não mediram esforços para que eu chegasse até esta etapa de minha vida.

Aos meus amigos do mestrado em especial ao José Lins, Luis Henrique Vogado, Luis Guilherme, Maíla e Nayara pela ajuda, sempre que foi necessário. Aos amigos e companheiros de viagem: Alan Rafael, Allan Jheyson, Dennis Sávio, Francisco Imperes, Ismael Leal e Ivenilton Moura, pelas aventuras vividas nas viagens semanais às aulas do mestrado.

Ao meu irmão, Leonilson Pereira de Sousa, pelo apoio, compreensão e hospedagem ao longo dessa caminhada.

Ao corpo docente do curso de Bacharelado em Sistema de Informação do Campus Senador Helvídio Nunes de Barros da UFPI, ao qual estou vinculado, pelo apoio, em especial a Alcilene Dalília, Patrícia Medyna e Patrícia Vieira, que sempre fizeram o possível para conciliar nossos horários de trabalho e qualificação.

Aos meus amigos pelo incentivo e apoio constantes.

*“A nossa maior glória não reside no fato de nunca cairmos,
mas sim em levantarmo-nos sempre depois de cada queda.”*
(Oliver Goldsmith)

Resumo

Estima-se que, atualmente, existam no mundo cerca de 285 milhões de deficientes visuais, sendo 39 milhões cegos e 246 milhões com baixa visão. O uso de Tecnologias de Informação e Comunicação contribui para a independência, qualidade de vida e inclusão dos deficientes visuais, por meio do suplemento, manutenção ou devolução de suas capacidades funcionais. Um dos problemas enfrentados é a identificação de cédulas monetárias, sendo esse o meio mais comum de pagamento e a maneira mais usada para a realização de transações pessoais por indivíduos que possuem deficiência visual. Com esse problema em mente, alguns países utilizam marcas táteis como orientação do valor correspondente a cada cédula. Porém, tal marca é considerada insuficiente por muitos deficientes visuais, devido ao desgaste na intensa movimentação da cédula. A visão computacional têm ferramentas que podem ser aplicadas para replicar a visão humana. Nesse contexto, o presente trabalho tem como objetivo principal desenvolver uma metodologia para detecção de cédulas monetárias a partir de imagens digitais, sob várias situações, como oclusão (utilizando apenas parte da imagem no reconhecimento) e variação de iluminação, com a utilização de descritores locais. Foram avaliados quatro detectores de pontos de interesse (BRISK, FAST, MSER e SURF), dois descritores (BRISK e SURF), quatro métodos de assinatura da imagem (média, moda, mediana e BoVW), quatro classificadores (*Radial Basis Function Networks*, *Random Tree*, *Multilayer Perceptron* e *Sequential Minimal Optimization*) e dois comitês de classificadores (*Random Forest* e um comitê formado pelos melhores classificadores testados: *Radial Basis Function Networks*, *Multilayer Perceptron* e *Sequential Minimal Optimization*). Nos testes foram utilizadas as cédulas de dólar, euro e notas brasileira de real. Para isso, foram construídas bases de imagens, composta por 1.008 imagens de cada valor de cédula, totalizando 6.048 imagens para cédulas de real e 7.056 imagens para as cédulas de dólar e euro. Nos resultados se obteve um índice *Kappa* excelente e valores de acurácia de 99,77%, 99,12% e 96,95% em cédulas de dólar, euro e real, respectivamente. Esses resultados demonstram que a metodologia desenvolvida pode ser utilizada como base para o desenvolvimento de aplicações para identificação de cédulas monetárias.

Palavras-chaves: Acessibilidade. Deficientes Visuais. Reconhecimento de Cédulas. Tecnologias Assistivas.

Abstract

It is estimated that there are currently around 285 million visually impaired people in the world, 39 million of whom are blind and 246 million with low vision. The use of Information and Communication Technologies contributes to the independence, quality of life and inclusion of the visually impaired, through the supplementation, maintenance or devolution of their functional capacities. One of the problems faced is the identification of banknotes, which is the most common means of payment and the most used way to carry out personal transactions by individuals who are visually impaired. With this problem in mind some countries use tactile marks as a guideline of the value corresponding to each banknote. However, such a mark is considered insufficient by many visually impaired, due to the wear and tear on the intense movement of the banknote. Computer vision has tools that can be applied to replicate human vision. In this context, the main objective of the present work is to develop a methodology for the detection of monetary banknotes from digital images, under various situations, such as occlusion (using only part of the image in the recognition) and variation of illumination, using local descriptors. A set of four detectors (BRISK, FAST, MSER and SURF), two descriptors (BRISK and SURF), four ways of generating the image signature (mean, mode, median and BoVW), four classifiers (Radial Basis Function Networks, Random Tree, Multilayer Perceptron and Sequential Minimal Optimization) and two classifier committees (Random Forest and a committee made up of the best classifiers tested: Radial Basis Function Networks, Multilayer Perceptron and Sequential Minimal Optimization). In the tests were used the dollar, euro and Brazilian notes of real. In the results we obtained an excellent Kappa index and accuracy values of 99.77%, 99.12% and 96.95% in dollar, euro and real, respectively. These results show that the developed methodology can be used as a basis for the development of applications for the identification of monetary banknotes.

Keywords: Accessibility. Visually Impaired. Banknote Recognition. Assistive Technologies.

Lista de ilustrações

Figura 1 – Abordagens utilizadas na identificação de cédulas monetárias.	2
Figura 2 – Exemplo de regiões selecionadas utilizadas no trabalho de Hasanuzzaman, Yang e Tian (2012).	6
Figura 3 – Exemplo de máscaras usadas para o cálculo dos pontos chave, utilizadas na pesquisa de Costa, Veiga e Sousa (2016).	7
Figura 4 – Diagrama de operação da metodologia proposta para o reconhecimento de cédulas monetárias.	9
Figura 5 – Amostras de imagens de cédulas monetárias, apresentadas na base de dados, utilizadas nos testes.	11
Figura 6 – Imagem mostrando o ponto de interesse em teste e os 16 pixels no círculo, utilizando o algoritmo FAST	14
Figura 7 – Filtros Gaussianos de segunda ordem nas direções yy e xy (a), aproximação por filtros de caixa (b). Fonte: Fonte: Bay et al. (2008).	15
Figura 8 – Pirâmide de escala montada pelo algoritmo BRISK. Fonte: Leutenegger, Chli e Siewwart (2011).	16
Figura 9 – Filtros <i>Haar Wavelets</i> para calcular a resposta em x (esquerda) e y (direita). Fonte: Bay et al. (2008).	17
Figura 10 – Representação do vetor com as sub-regiões e suas respectivas respostas na janela utilizada. Fonte: Bay et al. (2008).	18
Figura 11 – Janelas de descrição em diferentes escalas. Fonte: Bay et al. (2008).	18
Figura 12 – Padrão de amostragem do algoritmo BRISK. Fonte: Leutenegger, Chli e Siewwart (2011).	19
Figura 13 – Fluxograma para assinaturas utilizando técnicas de média, mediana e moda.	20
Figura 14 – Fluxograma de geração das assinaturas utilizando o método de BoVW.	20
Figura 15 – Amostras de imagens das cédulas de reais.	31
Figura 16 – Amostra de detecção de pontos.	32

Lista de tabelas

Tabela 1 – Montagem da base de imagens, para cada valor de cédula monetária. . .	12
Tabela 2 – Nível de exatidão de uma classificação conforme o valor do índice <i>Kappa</i> . . .	27
Tabela 3 – Combinações para a melhor classificação utilizando as imagens das cédulas de dólar.	30
Tabela 4 – Combinações para a melhor classificação utilizando as imagens das cédulas de euro.	30
Tabela 5 – Combinações para a melhor classificação utilizando as imagens das cédulas de real.	31
Tabela 6 – Resultado utilizando cédulas de real, obtidas com câmera de <i>smartphone</i> . . .	32
Tabela 7 – Pontos retornados pelos detectores analisados.	33
Tabela 8 – Principais resultados de acurácia na classificação das cédulas monetárias de dólar.	42
Tabela 9 – Principais resultados de acurácia na classificação das cédulas monetárias de euro.	43
Tabela 10 – Principais resultados de acurácia na classificação das cédulas monetárias de real.	44

Lista de abreviaturas e siglas

A	Acurácia
AGAST	<i>Adaptive and Generic Accelerated Segment Test</i>
BoVW	<i>Bag of Visual Words</i>
BRISK	<i>Binary Robust Invariant Scalable Keypoints</i>
CNN	<i>Competitive Neural Network</i>
FAST	<i>Features from Accelerated Segment Test</i>
FLANN	<i>Fast Library for Approximate Nearest Neighbors</i>
FN	Falso Negativo
FP	Falso Positivo
FREAK	<i>Fast Retina Keypoint</i>
INR	Rúpias Indianas
K	Índice <i>Kappa</i>
MLP	<i>Multilayer Perceptron</i>
MSER	<i>Maximally Stable External Regions</i>
PoI	<i>Points of Interest</i>
PCA	<i>Principal Component Analysis</i>
RBF	<i>Radial-Basis Function Network</i>
RNA	Redes Neurais Artificiais
RoI	<i>Region of Interest</i>
S	Sensibilidade
SIFT	<i>Scale-Invariant Feature Transform</i>
SMO	<i>Sequential Minimal Optimization</i>
SURF	<i>Speeded Up Robust Features</i>

SVM	<i>Support Vector Machines</i>
TIC	Tecnologia de Informação e Comunicação
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
WHO	<i>World Health Organization</i>

Lista de símbolos

σ	Desvio Padrão
\$	Dólar
€	Euro
Rs	Rúpias Indianas
R\$	Real Brasileiro

Sumário

1	INTRODUÇÃO	1
1.1	Motivação	1
1.2	Objetivos	2
1.2.1	Objetivo Geral	2
1.2.2	Objetivos Específicos	2
1.3	Contribuições	3
1.4	Produção Científica	3
1.5	Organização da Dissertação	4
2	TRABALHOS RELACIONADOS	5
2.1	Considerações Finais	8
3	MATERIAIS E MÉTODOS	9
3.1	Metodologia para Identificação de Cédulas Monetárias	9
3.2	Bases de Imagens	10
3.3	Detectores de Pontos de Interesse - <i>Points of Interest</i> (PoI)	12
3.3.1	<i>Maximally Stable External Regions - MSER</i>	13
3.3.2	<i>Features from Accelerated Segment Test - FAST</i>	13
3.3.3	<i>Speed-Up Robust Features - SURF</i>	15
3.3.4	<i>Binary Robust Invariant Scalable Keypoints - BRISK</i>	15
3.4	Descritores	17
3.4.1	<i>Speed-Up Robust Features - SURF</i>	17
3.4.2	<i>Binary Robust Invariant Scalable Keypoints - BRISK</i>	18
3.4.3	Assinatura da Imagem	19
3.4.3.1	<i>Bag of Visual Words - BoVW</i>	20
3.5	Classificadores	21
3.5.1	<i>Multilayer Perceptron - MLP</i>	21
3.5.2	<i>Radial Basis Function Networks - RBF</i>	22
3.5.3	<i>Sequential Minimal Optimization - (SMO)</i>	23
3.5.4	<i>Random Tree</i>	23
3.5.5	<i>Random Forest</i>	24
3.5.6	Comitê de Classificadores	24
3.6	Métricas de Avaliação dos Resultados	25
3.6.1	Acurácia (A)	26
3.6.2	Sensibilidade (S)	26
3.6.3	Índice <i>Kappa</i> (K)	26

3.7	Considerações Finais	28
4	RESULTADOS E DISCUSSÕES	29
4.1	Resultados Utilizando Imagens de Dólar	29
4.2	Resultados Utilizando Imagens de Euro	30
4.3	Resultados Utilizando Imagens de Real	31
4.4	Discussões	32
4.5	Considerações Finais	33
5	CONCLUSÕES E TRABALHOS FUTUROS	35
5.1	Conclusões	35
5.2	Trabalhos Futuros	35
	REFERÊNCIAS	36
	ANEXOS	41
	ANEXO A – RESULTADOS DA CLASSIFICAÇÃO UTILIZANDO A BASE DE IMAGENS DAS CÉDULAS DE DÓLAR.	42
	ANEXO B – RESULTADOS DA CLASSIFICAÇÃO UTILIZANDO A BASE DE IMAGENS DAS CÉDULAS DE EURO.	43
	ANEXO C – RESULTADOS DA CLASSIFICAÇÃO UTILIZANDO A BASE DE IMAGENS DAS CÉDULAS DE REAL.	44

1 Introdução

A deficiência visual compromete a visualização e identificação de objetos. Segundo levantamento de 2010 da World Health Organization (WHO) ([PASCOLINI; MARIOTTI, 2011](#)), existiam cerca de 285 milhões de pessoas com dificuldade de enxergar, sendo que 39 milhões se declararam totalmente cegos.

Com o surgimento de novas tecnologias, houve um aumento nos mecanismos para oferecer acessibilidade e inclusão a indivíduos com deficiência. As principais metas são proporcionar benefícios, como uma maior independência e qualidade de vida. O uso de recursos tecnológicos auxiliam os deficientes visuais na realização de tarefas, como o uso de computadores para ler jornais, realizar pesquisas acadêmicas, dentre outros.

O reconhecimento de cédulas é um dos problemas mais desafiadores enfrentados pelas pessoas com deficiência visual, uma vez que constitui um processo essencial no gerenciamento de transações monetárias ([GRIJALVA et al., 2010](#)). Essas pessoas possuem grande dificuldade em identificar os valores das cédulas monetárias, prejudicando a realização de tarefas cotidianas, tais como: pagamentos de contas, realização de compras, operações de serviços bancários, dentre outras. Além disso, o deficiente pode sofrer golpes financeiros.

Neste trabalho é proposta uma metodologia de identificação de cédulas monetárias com a aplicação de técnicas de processamento de imagens, visão computacional e reconhecimento de padrões. Para isso, realizamos o levantamento bibliográfico das principais técnicas aplicadas e avaliamos um conjunto de quatro detectores de pontos de interesse, dois descritores, quatro métodos de geração de assinatura da imagem, quatro classificadores e dois comitês de classificação. Os testes foram realizados em cédulas de dólar, euro e real, com o intuito de identificar os melhores algoritmos para compor a abordagem.

1.1 Motivação

Cédulas monetárias são utilizadas diariamente para pagamento de bens e serviços por toda a sociedade. Contudo, os deficientes visuais possuem dificuldades em identificar os valores das cédulas. Muitas abordagens já foram desenvolvidas para auxiliar na identificação de notas, como sistemas de dobragens, equipamentos eletrônicos, notas de tamanhos diferenciados e marcações táteis, conforme ilustrado na [Figura 7](#). Porém, a maioria das abordagens não asseguram o reconhecimento, necessitando a participação de uma segunda pessoa na validação do valor da cédula.

Um estudo realizado por [Braga et al. \(2012\)](#) teve como objetivo analisar o uso



Figura 1 – Abordagens utilizadas na identificação de cédulas monetárias.

das Tecnologias de Informação e Comunicação (TIC) pelos deficientes visuais. Um dos resultados da análise foi a verificação do interesse num aplicativo que identifica-se as cédulas monetárias. A principal vantagem de um sistema para identificação de cédulas monetárias é possuir uma forma de validar o valor do dinheiro utilizado.

A maioria dos dispositivos móveis atualmente possuem configuração de *hardware* com elevado poder de processamento. Com o avanço desses dispositivos, supõe-se que esses equipamentos suportem implementar aplicações com processamento de imagens.

1.2 Objetivos

1.2.1 Objetivo Geral

O objetivo principal deste trabalho é desenvolver uma metodologia para detecção de cédulas monetárias a partir de imagens digitais, sob várias situações, como oclusão (utilizando apenas parte da imagem no reconhecimento) e variação de iluminação.

1.2.2 Objetivos Específicos

Após a definição do objetivo geral, os seguintes objetivos específicos foram propostos:

- Estudo e teste de algoritmos de extração de atributos em imagens (detectores e descritores);
- Estudo e teste de técnicas de geração de assinaturas;
- Estudo e teste de algoritmos de classificação;
- Estudo e teste da combinação de algoritmos de classificação (comitê de classificadores);
- Análise comparativa da combinação entre detectores, descritores, assinaturas e classificadores.

1.3 Contribuições

As principais contribuições relacionadas a abordagem proposta neste trabalho são:

- Criação e disponibilização de bases de imagens para testes da abordagem;
- Comparação de quatro diferentes maneiras de geração de assinatura (média, mediana, moda e *Bag of Visual Words* (BoVW));
- Utilização de diversos descritores de pontos de interesse e classificadores na obtenção do melhor resultado de acurácia.

1.4 Produção Científica

Durante o desenvolvimento deste trabalho, foram produzidos os seguintes artigos científicos relacionados ao tema:

- SOUSA, L. P.; VERAS, R. M. S.; VOGADO, L. H. S.; BRITTO NETO, L. S. Metodologia de Identificação de Cédulas Monetárias para Deficientes Visuais. *Revista de Sistemas e Computação (RSC)*, v. 8, n. 1, 2018.
- SOUSA, L. P.; VERAS, R. M. S.; VOGADO, L. H. S.; BRITTO NETO, L. S.; SILVA, R. R. V.; ARAÚJO, F. H. D.; MEDEIROS, F. N. S. Banknote Identification Methodology for Visually Impaired People. *Multimedia Tools and Applications (MTAP)*, 2019 (Submetido - aguardando resultado).

1.5 Organização da Dissertação

Esta dissertação está organizada em 5 capítulos. Este capítulo apresentou os fatos que motivaram a realização deste trabalho e os objetivos propostos. O Capítulo 2 introduz os principais trabalhos encontrados na literatura sobre o reconhecimento de cédulas monetárias. O Capítulo 3 descreve a abordagem proposta neste trabalho, apresenta as principais ferramentas e procedimentos computacionais utilizados, tais como: bases de imagens, detectores, descritores, classificadores e as métricas de avaliação dos resultados. O Capítulo 4 ilustra os resultados obtidos. Ao final, o Capítulo 5 apresenta as conclusões referentes a abordagem desenvolvida e trabalhos futuros.

2 Trabalhos Relacionados

Neste capítulo são apresentados os trabalhos presentes na literatura que abordam o reconhecimento de cédulas monetárias, sendo analisados com o intuito de identificar os descritores e classificadores utilizados. Inúmeros trabalhos relacionados a esse assunto estão disponíveis, contudo considerou-se apenas os artigos que utilizam técnicas de visão computacional para auxiliar os deficientes visuais.

Grijalva et al. (2010) desenvolveram um sistema para o reconhecimento de cédulas de dólar para a utilização em *smartphones*. Os autores definiram o lado direito da imagem como a Região de Interesse (*Region of Interest - RoI*). A partir dessas regiões, os vetores de características são extraídos utilizando o método de Reconhecimento Facial *Eigenfaces* (TURK; PENTLAND, 1991) que funciona com base em Análise de Componentes Principais (*Principal Component Analysis - PCA*) (HOLLAND, 2008). Para identificação, cada nova cédula é comparada com um conjunto de treinamento usando a distância de Mahalanobis. Sendo que os autores reportaram uma taxa de precisão de 99,8%. Contudo, essa abordagem proposta não suporta mudança de perspectiva, existência de fundo e exposição parcial da cédula.

Hasanuzzaman, Yang e Tian (2012), propuseram uma abordagem para reconhecimento de cédulas usando o descritor *Speeded Up Robust Features* (SURF) (BAY et al., 2008). Segundo os autores o algoritmo SURF é de baixo custo computacional, o que permite a localização rápida dos pontos de interesse. Para a geração das regiões de referência foram realizados cortes manuais de vários pontos em cada cédula monetária (frente e verso), conforme ilustrado na Figura 2. Nos testes realizados foram utilizadas, para treino, 140 imagens de cédulas bancárias de dólar americano, contendo 20 imagens de cada classe (\$1, \$2, \$5, \$10, \$20, \$50 e \$100) e 579 imagens para testes, sendo essas cédulas do mesmo tamanho e regiões de referência diferenciadas para cada valor de cédula. Como resultado o algoritmo atingiu 100% de taxa de verdadeiro positivo. No entanto, nessa pesquisa, uma cédula só será reconhecida se, ao menos, duas regiões de referências forem identificadas.

Mulmule-Shirkhedkar e Dani (2015) realizaram um estudo comparativo dos descritores SURF e FREAK (ALAHY; ORTIZ; VANDERGHEYNST, 2012), para extração de atributos e classificação de cédulas indianas, usando um modelo baseado em regiões de referência. Para a classificação foi utilizada a abordagem de força bruta, comparando os descritores da imagem pesquisada com os descritores das imagens armazenadas. As cédulas de INR (Rúpias Indianas) utilizadas foram: Rs.5, Rs.10, Rs.20, Rs.50, Rs.100, Rs.500 e Rs.1000. No experimento foram utilizadas 210 imagens, sendo 30 de cada valor monetário, dividindo o conjunto de imagens da seguinte forma: 60% treino, 20% testes e



Figura 2 – Exemplo de regiões selecionadas utilizadas no trabalho de [Hasanuzzaman, Yang e Tian \(2012\)](#).

20% validação. Os autores obtiveram acurácia de 95,15% para o SURF e de 92,85% para o FREAK. A abordagem utilizada só reconhece uma cédula monetária, caso identifique alguma região de referência previamente cadastrada, sendo bem reduzido o número de regiões de referência composta na base de treinamento.

No trabalho de [Costa, Veiga e Sousa \(2016\)](#), os autores desenvolveram uma aplicação para o reconhecimento de cédulas monetárias de Euro (€). Nessa pesquisa foram utilizados 8 detectores e 6 descritores. Para melhorar a detecção das partes relevantes das notas, o algoritmo de detecção de ponto de interesse é aplicado apenas dentro das máscaras criadas manualmente (apenas as áreas brancas mostradas na Figura 3 são utilizadas para calcular os PoIs. Para classificar, os autores avaliaram o uso de busca extensiva e da abordagem heurística *Fast Library for Approximate Nearest Neighbors* (FLANN) ([MUJA; LOWE, 2009](#)). O sistema foi testado com 80 imagens, que continham notas nas condições mais comuns, com diferentes visões em perspectiva, ambientes desordenados, notas parcialmente ocultas e várias notas por imagem. O melhor resultado foi obtido utilizando o algoritmo *Scale-Invariant Feature Transform* (SIFT) ([MORTENSEN; DENG; SHAPIRO, 2005](#)) como detector e descritor e classificação por meio de busca extensiva. O sistema reconheceu com sucesso todas as 95 imagens de cédulas monetárias. Contudo, os autores não relataram qual a combinação de detectores e descritores utilizados no reconhecimento onde obtiveram uma acurácia de 100%. Apenas definiram que os algoritmos de detecção e descrição SURF são mais adequados para uso em tempo real.

Em [Teixeira, Braga e Nascimento \(2016\)](#), os autores propuseram um aplicativo para reconhecimento de cédulas, empregando o algoritmo SURF na identificação dos pontos de interesse (*Points of Interest* - PoI) e definição dos descritores. Nesse estudo, foram

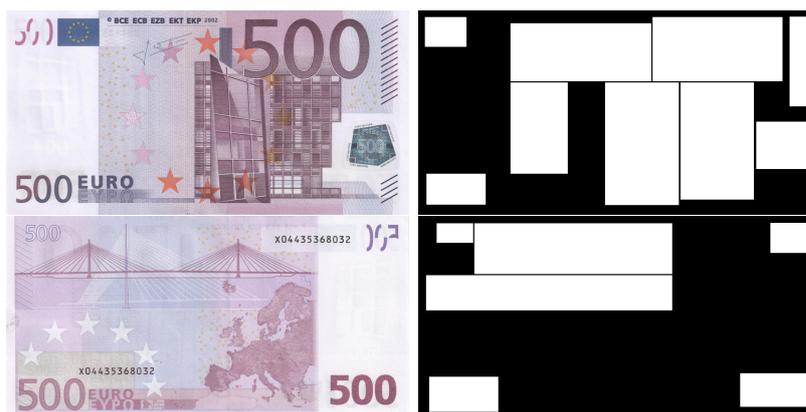


Figura 3 – Exemplo de máscaras usadas para o cálculo dos pontos chave, utilizadas na pesquisa de [Costa, Veiga e Sousa \(2016\)](#).

elaborados dois vetores, um contendo os atributos de cada PoI da imagem e outro com os seus respectivos descritores. Para a classificação foi utilizado o classificador *Sequential Minimal Optimization* (SMO) ([PLATT, 1999](#)). Nos testes foram utilizados 1440 imagens digitais, sendo 240 para cada cédula monetária (da 1ª Família de Real, com exceção da cédula de um real), capturadas em ambiente controlado com iluminação artificial. Para a verificação do experimento foi utilizado o método *n-fold cross validation* ([KOHAVI, 1995](#)). Os resultados atingiram uma taxa de 97,9% de identificação correta. Contudo, não ocorreram testes com oclusões parciais, desgaste pelo uso e variação de iluminação.

[Oyedotun e Khashman \(2017\)](#) investigaram três estruturas cognitivas hipotéticas para o reconhecimento de cédulas bancárias utilizando Redes Neurais Competitivas (*Competitive Neural Network - CNN*) ([CASTRO; ZUBEN; JR, 2003](#)). O sistema implementado reconhece as cédulas com até 75% de oclusão. Nas simulações foram utilizadas as notas de Naira (moeda oficial da Nigéria), nas seguintes denominações: 1000 Naira, 200 Naira e 50 Naira. Foram utilizadas 100 amostras para cada estrutura, totalizando 300 imagens. Os resultados da simulação baseados nas taxas globais de reconhecimento mostraram que o melhor desempenho é obtido utilizando 3 sistemas para análise das cédulas, sendo o primeiro para a identificação do lado (frente ou verso) e os dois restantes para a identificação das características da denominação do lado correspondente, alcançando uma taxa de acurácia de 86,67%.

[ABBURU et al. \(2017\)](#) propuseram um sistema para o reconhecimento automático de moedas. O método proposto é capaz de reconhecer tanto a nacionalidade quanto o valor da cédula. O funcionamento desse método se dá por meio da identificação do país de origem utilizando determinadas áreas de interesse predefinidas e, em seguida, extraindo o valor de denominação usando características como tamanho, cor ou texto na nota, dependendo de quanto as notas dentro do mesmo país sejam diferentes. Foram consideradas as 20 moedas mais negociadas, assim como suas denominações. Para fins de demonstração foi escolhida a nota canadense de 20 dólares. O sistema proposto obteve uma precisão de 93,3% na

realização dos experimentos, identificando adequadamente as nacionalidades das células apresentadas.

Mittal e Mittal (2018) desenvolveram um método baseado em aprendizado profundo para identificação de denominações de notas de rupias da moeda indiana a partir de suas imagens coloridas. A estrutura utiliza o conceito de aprendizado por transferência, no qual uma rede neural convolucional profunda (Convolutional Neural Network - CNN) (LAWRENCE et al., 1997), já treinada sobre um enorme conjunto de dados de imagens naturais, é reutilizada para o problema da classificação da denominação a partir de imagens de notas. As cédulas de INR (Rúpias Indianas) utilizadas foram: Rs.10, Rs.50, Rs.100 e Rs.500. No experimento foram utilizadas um total de 95 imagens para cada denominação, sendo adquirido sob diferentes condições, totalizando 380 imagens. Os autores obtiveram acurácia de 96,60. Contudo, não ocorreram testes com oclusões parciais.

2.1 Considerações Finais

Neste capítulo foram apresentados estudos encontrados na literatura que abordam a identificação de cédulas monetárias, sendo que a maioria dos trabalhos citados reconhecem apenas um tipo de cédula monetária. É importante salientar que nenhum dos trabalhos citados descreve a maneira como foi criada a assinatura da imagem. Esse é um passo importante visto que os classificadores recebem como entrada vetores de atributos de mesmo comprimento enquanto que os algoritmos de detecção de POIs detectam uma quantidade de pontos diferentes para cada imagem.

3 Materiais e Métodos

Neste capítulo serão apresentados os materiais e métodos utilizados para a detecção de cédulas monetárias. Será apresentada a metodologia utilizada e na sequência são descritas as bases de imagens utilizadas neste trabalho, os detectores e descritores utilizados na etapa de extração de características, com ilustração de maneiras de geração das assinaturas utilizadas. Em seguida, os classificadores utilizados na classificação são apresentados e finalmente as métricas de avaliação dos resultados utilizadas para a obtenção dos resultados.

3.1 Metodologia para Identificação de Cédulas Monetárias

Esta seção descreve a metodologia empregada na realização deste trabalho. O sistema proposto tem como objetivo a identificação automática do valor da cédula monetária a partir de imagens digitais disponibilizadas nas bases de imagens construídas.

A imagem de entrada consiste em uma cédula obtida da base de dados. Inicialmente é realizada a detecção dos PoI, buscando identificar as regiões mais significativas na imagem. No passo seguinte ocorre a descrição dos PoIs selecionados. Em seguida, utilizam-se as técnicas de geração da assinatura da imagem. Por fim, as características selecionadas são classificadas tendo como notificação o resultado da classificação. A Figura 4 apresenta o fluxograma do método proposto.

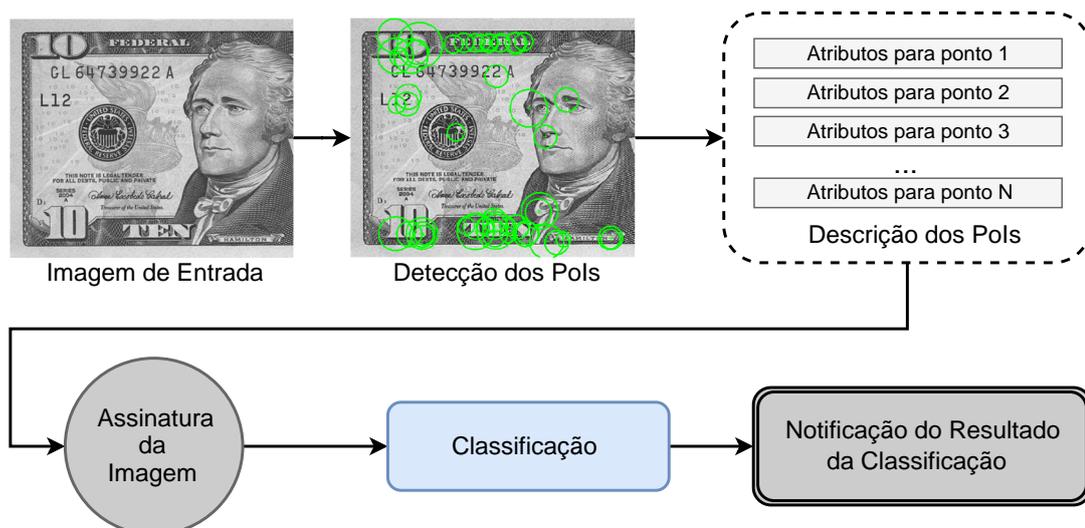


Figura 4 – Diagrama de operação da metodologia proposta para o reconhecimento de cédulas monetárias.

A detecção de PoI é a etapa inicial de reconhecimento, na qual os principais pontos de interesse para correspondência são identificados nas imagens do conjunto de

treinamento. Neste trabalho, os detectores de características MSER (MATAS et al., 2004), FAST (ROSTEN; DRUMMOND, 2006), SURF (BAY et al., 2008) e BRISK (LEUTENEGGER; CHLI; SIEGWART, 2011) foram utilizados, sendo os detectores mais utilizados na literatura para o reconhecimento de cédulas e por utilizarem formas diferentes na detecção de PoI.

Em seguida, o descritor, que fornece parâmetros que descrevem cada ponto de forma, gera um vetor de atributos que podem ser posteriormente comparados e associados aos pontos de outras assinaturas. Os descritores detectam numerosos pontos em cada imagem e geram uma matriz de atributos. Os descritores BRISK (LEUTENEGGER; CHLI; SIEGWART, 2011) e SURF (BAY et al., 2008) foram utilizados.

O BRISK é um descritor binário, cujo o reduzido tempo de processamento o torna indicado para aplicações em tempo real. No entanto, para a metodologia proposta foi escolhido o descritor SURF, devido à sua eficiência. O resultado da aplicação do descritor é um vetor de características para cada imagem.

Foi utilizada a geração de uma assinatura da imagem em que a mesma usa apenas um único vetor de atributo. A assinatura da imagem é um conjunto de características visuais detectadas que descreve uma cena específica ou um elemento da cena.

O estágio de classificação é composto pelo processo de treinamento de classificador, cujo objetivo é dividir o espaço de atributos em regiões de decisão. Dessa forma, os vetores de atributos contidos na mesma região de decisão compartilham a mesma classe. No presente trabalho, a entrada para o classificador é a assinatura da imagem, e a saída é a classe à qual a nota pertence.

3.2 Bases de Imagens

A metodologia de criação da base de imagens seguiu os seguintes passos. Inicialmente foram utilizadas imagens de frente e verso, variando a posição em horizontal, vertical, rotação em 90° a direita e 90° a esquerda e rotação a 180°, totalizando 8 imagens de cada valor de notas, para cédulas inteiras.

Na utilização em aplicações reais dificilmente se consegue analisar toda a cédula, sendo incluído na base de imagens as seguintes divisões: notas divididas ao meio e divididas em 4 partes (veja a Figura 5a). Com isso, se obtém um total de 56 imagens de cada classe.

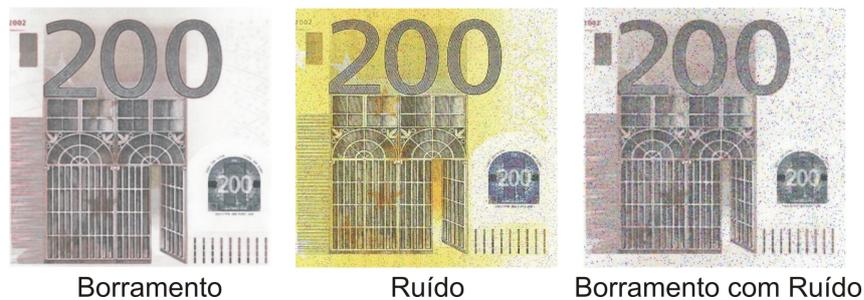
Para aumentar a robustez do método, modificou-se as imagens até aqui disponíveis para simular possíveis eventos na captura das mesmas, a saber: aplicação do filtro da mediana, com janela de tamanho $[5 \times 5]$, simulando uma imagem borrada; inserção do ruído “Salt & Pepper” na proporção de 0,05; aplicação do efeito de borramento em conjunto com o efeito de ruído (Figura 5b); diminuição da luminosidade da cédula em

30%, simulando uma imagem adquirida em ambiente com pouca luminosidade e aumento da luminosidade em 30%.

Com o objetivo de simular imagens capturadas em diversos ambientes, se inseriu imagens de fundo nas cédulas, nas cores branco, preto e cinza, nas imagens de boa qualidade, ruído, borramento e borramento com ruído (veja a Figura 5c).



(a) Exemplos de notas inteiras e as duas divisões propostas.



(b) Exemplos de imagens com modificações.



(c) Exemplos de imagens com inserção de cores de fundo.

Figura 5 – Amostras de imagens de cédulas monetárias, apresentadas na base de dados, utilizadas nos testes.

Para avaliar os resultados, criamos uma base de dados para cada uma das moedas. Geramos para cada valor de cédula 1008 imagens, conforme ilustrado na Tabela 1. No total, as bases de imagem de dólar e euro possuem 7056 imagens, enquanto que a de real possui 6048 imagens. As imagens das cédulas de dólar foram obtidas por meio de pesquisas na internet. As imagens das cédulas de euro foram retiradas do trabalho de [Costa, Veiga e Sousa \(2016\)](#) e as imagens da segunda família de real foram adquiridas na *webpage*¹ do Banco Central do Brasil.

¹ <https://www.bcb.gov.br/cedulasemoedas/cedulasemitidas>.

Tabela 1 – Montagem da base de imagens, para cada valor de cédula monetária.

Base de Imagens		Numeração
Boa Qualidade		1 a 56
Borradas		57 a 112
Ruído		113 a 168
Borramento com Ruído		169 a 224
Diminuição Iluminação		225 a 280
Aumento Iluminação		281 a 336
Fundo Branco	Normal	337 a 392
	Borradas	393 a 448
	Ruído	449 a 504
	Borramento com Ruído	505 a 560
Fundo Preto	Normal	561 a 616
	Borradas	617 a 672
	Ruído	673 a 728
	Borramento com Ruído	729 a 784
Fundo Cinza	Normal	785 a 840
	Borradas	841 a 896
	Ruído	897 a 952
	Borramento com Ruído	953 a 1008

3.3 Detectores de Pontos de Interesse - *Points of Interest* (PoI)

A detecção de pontos de interesse remonta a 1954, quando [Attneave \(1954\)](#) observou que “a informação é aglomerada ao longo dos contornos e é concentrada nos pontos em um contorno no qual sua direção muda mais rapidamente (por exemplo, nos picos de curvatura)”. Desde então, vários avanços foram feitos no esforço de interpretar imagens com características locais e a literatura sobre detecção de PoIs se tornou extensa ([LOWE, 1999](#); [MATAS et al., 2004](#); [MIKOLAJCZYK](#); [SCHMID, 2005](#); [ROSTEN](#); [DRUMMOND, 2006](#); [BAY et al., 2008](#); [TOLA](#); [LEPETIT](#); [FUA, 2010](#); [LEUTENEGGER](#); [CHLI](#); [SIEGWART, 2011](#)).

As informações mais úteis nas imagens estão localizadas em torno de zonas específicas, que normalmente correspondem aos pontos e regiões mais importantes. Em diversas aplicações de reconhecimento de objetos, o processamento local em torno desses pontos é suficiente, desde que esses pontos sejam estáveis e distintos ([SUAREZ et al., 2014](#)). Características locais, também chamadas de pontos de interesse, são caracterizadas por mudanças bruscas de intensidade na região. Esses recursos locais são geralmente classificados em cantos, bordas e manchas ([KASPERS, 2011](#)). Os pontos de interesse possuem várias informações locais e devem possuir uma alta repetitividade entre diferentes imagens. Cada ponto de interesse armazena o valor que representa uma determinada região ao redor dela, que é representado por um vetor de dados.

3.3.1 *Maximally Stable External Regions - MSER*

Algoritmo proposto por [Matas et al. \(2004\)](#) consiste em um método robusto diante de mudanças de perspectiva, originalmente utilizado para detecção de características e alinhamento de imagens estéreo. O algoritmo identifica pontos extremos na imagem buscando regiões conexas a partir da intensidade do brilho dos pixels. Esse aplica limiares de diferentes valores e detecta regiões de bordas com grande variação de intensidade.

Seu princípio algorítmico, de acordo com [Matas et al. \(2004\)](#), é baseado no limiar da imagem com um limiar de brilho variável. Sendo uma binarização de uma imagem de cena dependendo de um limiar de valor de cinza t . Todos os pixels com valor de cinza abaixo de t são definidos como zero (preto) na imagem limite. Todos os pixels com valor de cinza igual ou maior que t são definidos como um (brilhante). A partir de $t = 0$, o limiar é aumentado sucessivamente. No início, a imagem limiar é completamente brilhante. A medida que t aumenta, áreas pretas aparecerão na imagem binarizada, que crescem e finalmente se fundem. Algumas áreas pretas serão estáveis para uma grande variedade de t . Essas são as regiões MSER, revelando uma posição (por exemplo, o ponto central), bem como uma escala característica derivada do tamanho da região como dados de entrada para o cálculo do descritor da região. Todas as regiões da imagem da cena são detectadas e são significativamente mais escuras do que o ambiente ao redor. Inverter a imagem e repetir o mesmo procedimento com a imagem invertida revelam regiões características.

O detector MSER revela poucas regiões, mas sua detecção é muito estável. Além disso, o detector MSER é invariante em relação às transformações afins, o que o torna adequado para aplicações que precisam lidar com mudanças de ponto de vista ([TREIBER, 2010](#)).

3.3.2 *Features from Accelerated Segment Test - FAST*

FAST é um algoritmo proposto originalmente por [Rosten e Drummond \(2006\)](#) para reconhecer pontos de interesse em uma imagem. Um ponto de interesse pode ser determinado como um pixel em uma imagem que possui uma posição bem definida e pode ser identificada de forma robusta. Pontos de interesse têm alto conteúdo de informação local e devem ser igualmente reproduzíveis entre imagens diferentes.

A ideia básica dessa abordagem é reduzir o número de cálculos necessários em cada pixel para decidir se um ponto chave é detectado no pixel ou não. Isso é realizado colocando um círculo consistindo de 16 pixels centralizados no pixel sob investigação. Para o teste de canto, apenas as diferenças de valores de cinza entre cada um dos 16 pixels do círculo e o pixel central são avaliados, como mostrado na Figura 6 ([ROSTEN; DRUMMOND, 2006](#)).

O algoritmo FAST é descrito em 5 passos, conforme ([VISWANATHAN, 2011](#)):

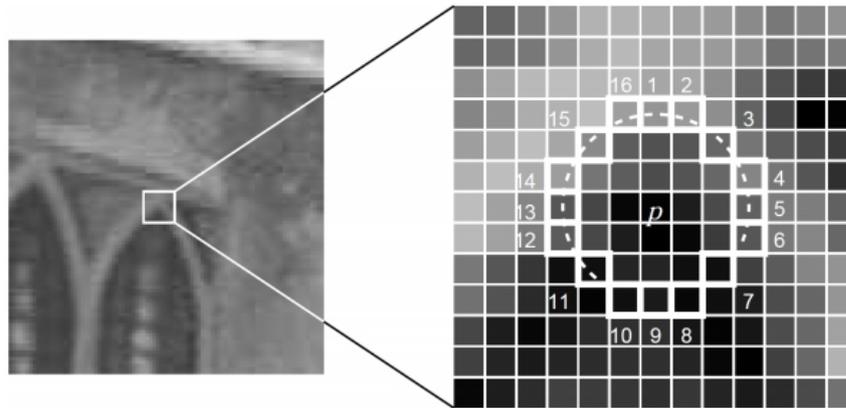


Figura 6 – Imagem mostrando o ponto de interesse em teste e os 16 pixels no círculo, utilizando o algoritmo FAST

1. Para saber se um pixel ' p ' na imagem é um ponto de interesse, será primeiramente obtida a intensidade desse pixel identificado como I_p ;
2. Definir um limiar de valores de intensidade T ;
3. Considerar um círculo de 16 pixels que cercam o pixel ' p ', conforme é ilustrado na Figura 6;
4. Comparar as intensidades dos pixels 1, 5, 9 e 13 do círculo com I_p . Pelo menos três desses quatro pixels devem estar acima ou abaixo de I_p pelo valor de T ;
5. Se o passo 4 é verdadeiro, e as intensidades de 12 pixels contíguos estão acima ou abaixo de I_p pelo valor de T , então o pixel ' p ' é um ponto de interesse;
6. Repetir o procedimento para todos os pixels da imagem.

Portanto, um PoI é detectado pela supressão não máxima em uma segunda etapa. Para esse fim, calcula-se uma função de pontuação V para cada um dos pontos detectados. A função de pontuação é definida como: “a soma da diferença absoluta entre os pixels no arco contíguo e o pixel central”. Em seguida, considere dois pontos de interesse adjacentes e compare seus valores de V . Descarte aquele com o valor V inferior. Todo o processo pode ser resumido matematicamente conforme a Equação 3.1.

$$V = \max \begin{cases} \sum(\text{valor do pixel} - p) \text{ se } (\text{valor} - p) > t \\ \sum(p - \text{valor do pixel}) \text{ se } (p - \text{valor}) > t \end{cases} \quad (3.1)$$

O algoritmo FAST é mais sensível ao ruído (que decorre do fato de que por razões de velocidade, o número de pixels avaliados em cada posição é reduzido) e não fornece informações de escala nem rotação para o cálculo do descritor (TREIBER, 2010). O algoritmo propõe uma melhoria de desempenho utilizando aprendizado de máquina e

criação de uma árvore de decisão. É um método ideal para a extração de características em ambientes de recursos limitados (JEONG; MOON, 2011).

3.3.3 Speed-Up Robust Features - SURF

O algoritmo proposto por Bay et al. (2008) é inspirado no SIFT (LOWE, 1999), porém com o objetivo de ser rápido na extração e comparação de características. Foi projetado para detecção de pontos de interesse independentemente de escala, rotação, mudanças de iluminação e pontos de vista. O SURF utiliza somente características de imagens de tons de cinza. O algoritmo SURF consiste em duas etapas, a primeira de localização dos pontos de interesse e a segunda de descrição desses pontos de interesse.

A detecção de pontos de interesse para o descritor utiliza filtros caixa (Figura 7a). Para determinar a localização e dimensão, a determinante da Hessiana é aplicada. Dado um ponto $x = (x, y)$ em uma imagem, a matriz Hessiana $H(x, \sigma)$ onde x na dimensão σ é definida pela Equação 3.2 (BAY et al., 2008),

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix}, \quad (3.2)$$

aproximando filtros gaussianos de segunda ordem (Figura 7b). É realizada a convolução desses filtros com a imagem em diversas escalas e 4 orientações (xx, xy, yx e yy). O uso de filtros caixa sobre uma imagem integral é realizado em tempo linear no tamanho da imagem, independentemente do tamanho do filtro. No caso de filtros gaussianos, a complexidade de tempo é na ordem do tamanho da imagem multiplicada pelo tamanho do filtro. Em seguida pontos de máximo local são selecionados, eliminando-se pontos em arestas e em regiões de pouco contraste da imagem.

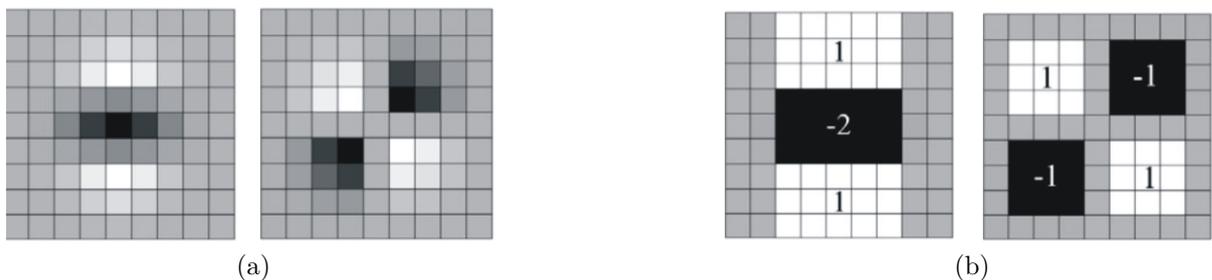


Figura 7 – Filtros Gaussianos de segunda ordem nas direções yy e xy (a), aproximação por filtros de caixa (b). Fonte: Fonte: Bay et al. (2008).

3.3.4 Binary Robust Invariant Scalable Keypoints - BRISK

O algoritmo BRISK, proposto por Leutenegger, Chli e Siegwart (2011), é um algoritmo para detecção e descrição de pontos-chave que é invariante a transformações.

Seu detector de pontos chave é fundamentado pelo método AGAST (*Adaptive and Generic Accelerated Segment Test*) (MAIR et al., 2010), o qual corresponde essencialmente ao método FAST (ROSTEN; DRUMMOND, 2006) com melhorias no tempo de processamento.

Na etapa de detecção, para manter a invariância à escala, é montada uma pirâmide de escalas da imagem. A pirâmide de escalas do algoritmo BRISK consiste em n oitavas c_i e n intraoitavas d_i , para $i = \{0, 1, \dots, n - 1\}$ em que, por padrão, os autores definem $n = 4$. As oitavas da pirâmide são formadas por meio do redimensionamento consecutivo da imagem original c_0 pela metade, realizando assim uma subamostragem. Cada intraoitava d_i é localizada entre as camadas c_i e $c_i + 1$, como ilustrado na Figura 8.

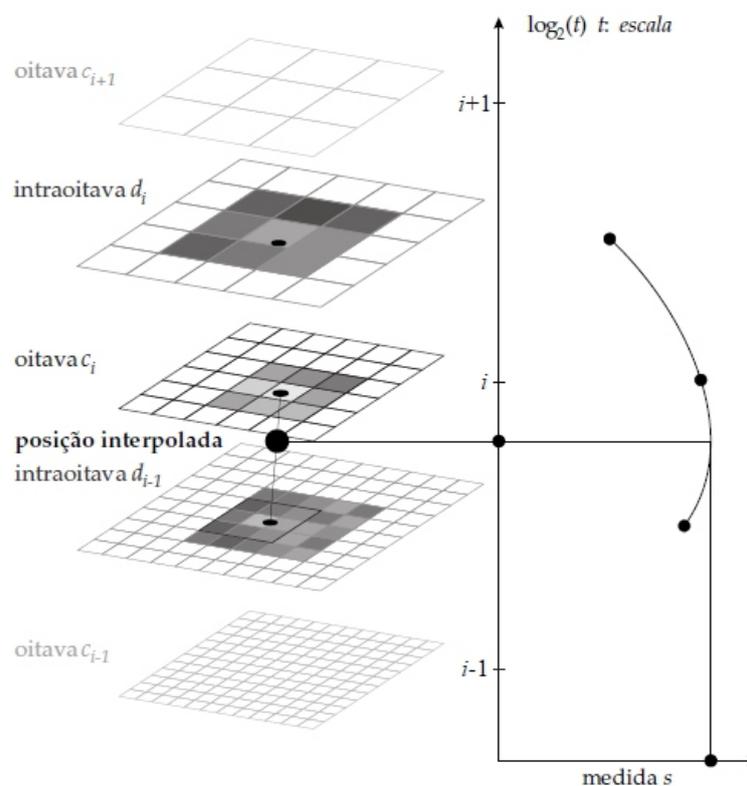


Figura 8 – Pirâmide de escala montada pelo algoritmo BRISK. Fonte: Leutenegger, Chli e Siegwart (2011).

Após a construção da pirâmide, é utilizado o detector AGAST em cada oitava e intraoitava da pirâmide para identificar potenciais regiões de interesse. Em seguida, os pontos obtidos nessas regiões são aplicados a uma supressão de não-máximos. Na realização da supressão de não-máximos os autores definem a medida s . Essa medida é determinada como sendo o limiar máximo em que um ponto é considerado canto por meio do algoritmo FAST.

No padrão utilizado em (LEUTENEGGER; CHLI; SIEGWART, 2011) os autores definem dois tipos de pares de pontos: pares longos e pares curtos. Pares curtos são pares de pontos de amostragem no qual sua distância euclidiana é menor que um determinado

limiar σ_{max} , e pares longos são aqueles no qual sua distância é maior que um certo limiar σ_{min} . Os pares longos são utilizados para determinar a orientação, e os pares curtos são utilizados na comparação de intensidade.

O algoritmo BRISK define a orientação do ponto-chave e a rotação do padrão de amostragem por meio da soma de todos os gradientes locais entre todos os pares longos, o que o torna invariante a rotação, retornando o ângulo do ponto-chave.

3.4 Descritores

A detecção dos PoIs deve ser capaz de identificar e extrair regiões que consigam ser invariantes a iluminação, rotação e translação do objeto a ser identificado. Para detectar PoIs e extrair suas características, utilizamos os algoritmos SURF (BAY et al., 2008) e BRISK (LEUTENEGGER; CHLI; SIEGWART, 2011).

3.4.1 Speed-Up Robust Features - SURF

O descritor SURF é extraído em duas etapas. Na primeira é realizada a atribuição de uma orientação dominante com base nas informações de uma região circular em torno dos pontos de interesse detectados, o que vai tornar o algoritmo invariante a rotação. A orientação computada por convolução com filtros *Haar wavelets*², nas direções x e y . A Figura 9 representa os filtros *Haar Wavelets*.



Figura 9 – Filtros *Haar Wavelets* para calcular a resposta em x (esquerda) e y (direita).
Fonte: Bay et al. (2008).

Em seguida, uma janela quadrada centrada de mesma orientação é utilizada. A Figura 10 representa essas janelas em diferentes escalas na imagem. Esta janela é segmentada em sub-regiões e respostas *Haar Wavelets* em direções dx horizontal e vertical dy são adicionadas em cada sub-região e os valores absolutos $|dx|$ e $|dy|$ são somados a fim de obter informação sobre a polaridade das alterações de intensidade. O resultado

² Wavelet é uma função capaz de decompor e representar outra função descrita no domínio do tempo ou do espaço, com o objetivo de analisar a função em diferentes escalas de frequência e de tempo.

final é um vetor de descritores de 64 valores (16 x 4) com as respectivas respostas desse processo, conforme demonstra a Figura 11.



Figura 10 – Representação do vetor com as sub-regiões e suas respectivas respostas na janela utilizada. Fonte: Bay et al. (2008).

Assim para cada ponto de interesse haverá um vetor de 64 posições descrevendo a forma como a imagem varia nesse ponto. A correspondência entre os descritores obtidos das imagens de referência e uma imagem de consulta é realizada por meio da medição de distância Euclidiana.

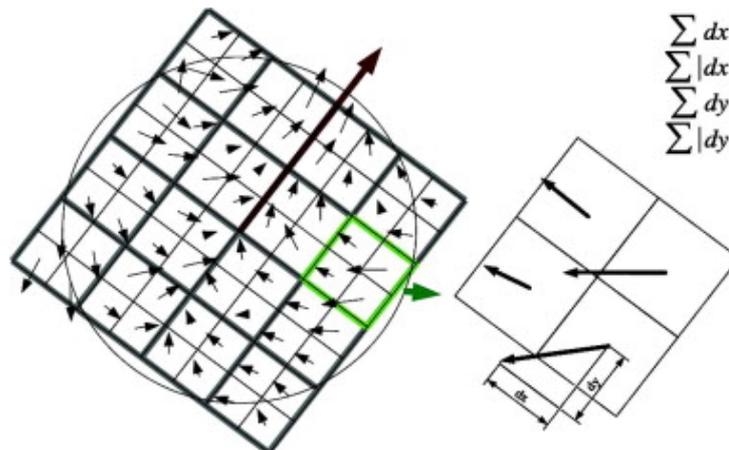


Figura 11 – Janelas de descrição em diferentes escalas. Fonte: Bay et al. (2008).

3.4.2 Binary Robust Invariant Scalable Keypoints - BRISK

O descritor BRISK é composto por uma cadeia binária concatenada, resultante de simples comparações de intensidade. O descritor BRISK faz uso de um padrão utilizado para a amostragem da vizinhança do ponto-chave detectado, ilustrado na Figura 12. Esse,

consiste em um padrão simétrico que possui pontos de amostragem, posicionados em círculos concêntricos em torno do ponto-chave. Cada círculo concêntrico representa uma suavização gaussiana com desvio padrão proporcional ao tamanho do círculo, para redução de ruídos.

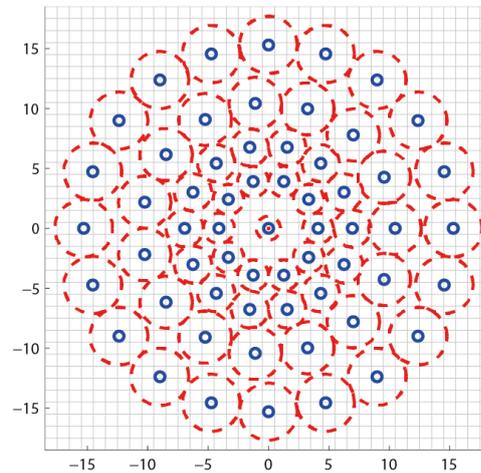


Figura 12 – Padrão de amostragem do algoritmo BRISK. Fonte: [Leutenegger, Chli e Siegart \(2011\)](#).

O algoritmo BRISK estima a orientação do ponto-chave e a rotação do padrão de amostragem por meio da soma de todos os gradientes locais entre todos os pares longos, o que o torna invariante a rotação, retornando o ângulo do ponto-chave. Por meio do conjunto de comparações de intensidade entre todos os pares curtos, o descritor BRISK resulta em uma cadeia binária de tamanho 512.

3.4.3 Assinatura da Imagem

Assinatura da imagem é o conjunto de características visuais detectadas que descreve uma determinada cena ou determinado elemento da cena. Primeiramente o detector analisa a imagem e seus pontos de referência, sendo referenciados por círculos verde na Figura 13. Em seguida o descritor, que oferece parâmetros que descrevem cada ponto, gera um vetor de atributos que possam ser posteriormente comparados e associados a pontos de outras assinaturas.

Neste trabalho utilizamos quatro técnicas de geração de assinatura, a saber: média, mediana, moda e *Bag of Words*. Para cada imagem em um banco de dados, os algoritmos de detecção PoI retornam diferentes números de pontos (N). Cada ponto gera um vetor de características, formando assim um conjunto de características. No entanto, a entrada para os classificadores deve ser uma matriz unidimensional (1×64 nesse caso, uma vez que 64 é o número de características geradas pelo descritor). Assim, devemos usar técnicas

adequadas para representar as informações na matriz como um vetor de características, chamado assinatura da imagem, conforme ilustra a Figura 13.



Figura 13 – Fluxograma para assinaturas utilizando técnicas de média, mediana e moda.

3.4.3.1 Bag of Visual Words - BoVW

O algoritmo BoVW (YANG et al., 2007) permite representar uma imagem como um único vetor (assinatura da imagem) permitindo assim que possamos treinar qualquer tipo de classificador. Neste trabalho, utiliza-se essa metodologia para alterar a representação dos atributos gerados por descritores de pontos de interesse. Esses são capazes de gerar atributos locais para cada ponto. Os atributos obtidos são rearranjados em vetores de mesmo tamanho. O fluxograma apresentado na Figura 14 detalha a metodologia para a geração da assinatura das imagens, utilizando a abordagem de BoVW.

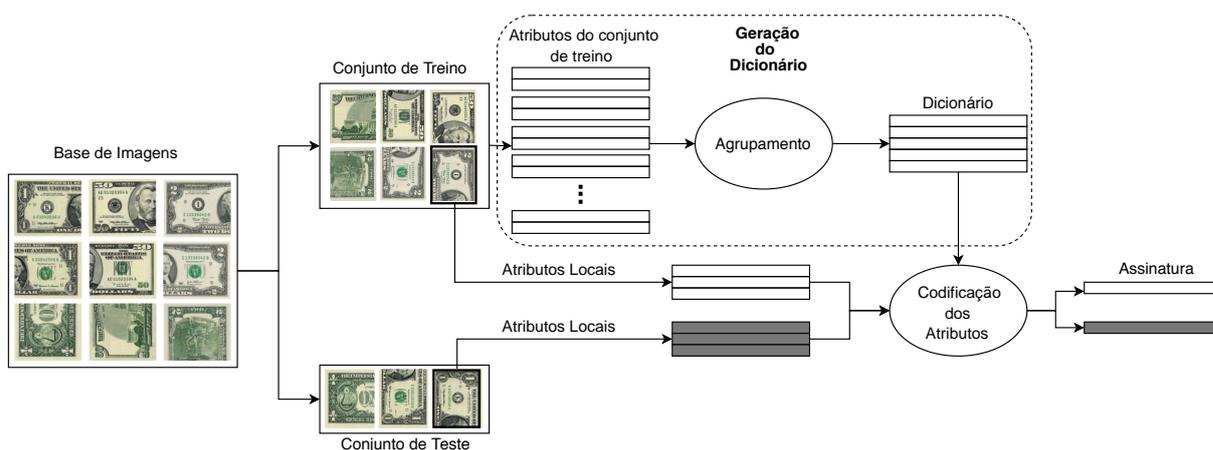


Figura 14 – Fluxograma de geração das assinaturas utilizando o método de BoVW.

Inicialmente, a base de dados é dividida em conjuntos de treino e teste. Em seguida é necessário saber quais podem ser as palavras visuais mais representativas das imagens que queremos classificar, isso significa que será construído um vocabulário de palavras visuais. Para isso, o primeiro passo a ser realizado é selecionar o conjunto de imagens de treino e extrair os atributos locais que serão utilizados para a classificação. Usando

a descrição dos atributos, cada uma delas será um ponto de interesse do descritor. Essa distribuição de pontos (características) será dividida em grupos de características similares.

Com as palavras visuais identificadas, é possível construir o histograma de palavras visuais, como representação final da imagem. Primeiro é definido cada atributo local da imagem à palavra visual mais parecida e armazena quantas vezes aparece cada palavra. Assim cada imagem que queremos classificar será representada por um vetor com a mesma dimensão. A abordagem de BoVW é representada, basicamente em 4 passos:

1. Extração dos atributos locais do conjunto de imagens de treinamento usando algum algoritmo detector de pontos de interesse;
2. Geração do dicionário de palavras visuais usando N amostras aleatórias (pontos de interesse obtidos pelo detector);
3. Quantização dos vetores, para isso é necessário obter a descrição dos atributos obtidos no passo 1, com algum algoritmo de descrição e agrupar todas as características similares entre elas para poder encontrar as palavras visuais que sejam os representantes de cada grupo de atributos similares (por exemplo, o algoritmo de agrupamento *K-Means*). Uma vez construído o vocabulário, já podemos realizar a representação e classificação de imagens;
4. Após a obtenção das palavras, por meio da quantização dos vetores, cada vetor de cada uma das imagens é rotulado com a palavra mais próxima. Na codificação dos atributos é realizado o histograma com a contagem de quantas palavras possui é gerado e normalizado. Ao término dessa etapa, cada imagem passa a ser representada por um vetor (assinatura da imagem) contendo a quantidade de cada uma das palavras que a imagem possui.

3.5 Classificadores

O objetivo de um classificador é dividir o espaço de atributos em regiões de decisão. Dessa forma, os vetores de atributos que estiverem contidos na mesma região de decisão compartilham a mesma classe. Neste trabalho, a entrada do classificador é a assinatura da imagem e a saída é a classe a qual a cédula pertence. Foram testados os seguintes classificadores: *Multilayer Perceptron*, *Radial Basis Function Networks*, *Sequential Minimal Optimization*, *Random Tree* e *Random Forest*.

3.5.1 *Multilayer Perceptron* - MLP

As Redes *Multilayer Perceptron* (MLP) são compostas por conjuntos de neurônios artificiais organizados em uma ou mais camadas intermediárias. Redes com dois ou mais

neurônios artificiais na camada intermediária podem ser utilizadas para a classificação de problemas não lineares (BRAGA; FERREIRA; LUDERMIR, 2007).

A MLP foi concebida para solucionar problemas que não poderiam ser resolvidos pelo modelo de neurônio básico. Para isso, são necessárias mais conexões, que só existem em uma rede de *perceptrons* dispostos em camadas. Os principais componentes são a camada de entrada, camada oculta ou intermediária e camada de saída.

A camada de entrada é responsável por receber os dados, não acontecendo nenhum processamento, sendo os dados apenas repassados para a camada intermediária. A camada intermediária é composta por uma ou mais camadas com um ou mais nós (neurônios artificiais). Nessa camada é aplicada uma função de ativação não linear, a qual determina se o neurônio irá propagar a informação ou não com base nos estímulos recebidos da camada anterior multiplicado pelos pesos obtidos durante a etapa de treinamento. A última camada, chamada de camada de saída, recebe as informações das camadas intermediárias e produz a resposta.

O método mais popular para aprendizado em MLP é chamado de retro-propagação (*backpropagation*). Esse algoritmo é baseado numa regra de aprendizagem que “corrige” o erro durante o treinamento (HAYKIN, 1998). A aprendizagem por retro-propagação do erro consiste em dois passos através das diferentes camadas da rede. O primeiro passo é chamado de propagação. Nesse passo um padrão de atividades é aplicado aos nós sensoriais e é produzida a resposta real da rede. O segundo passo é a retro-propagação. Nesse passo os pesos sinápticos são ajustados de acordo com um sinal de erro. Esse sinal é então propagado na direção contrária das conexões sinápticas. O objetivo é minimizar o erro médio global, ou seja, fazer com que a resposta real da rede se aproxime da resposta desejada.

3.5.2 Radial Basis Function Networks - RBF

A Rede RBF é uma das Redes Neurais Artificiais (RNAs) que tem sido aplicada com sucesso para tarefas de classificação (SUG, 2009) e é constituída de três camadas. A camada de entrada é formada por nós de fonte que interligam a rede ao seu ambiente. A segunda camada executa uma transformação não linear do espaço de entrada para o espaço oculto. Por fim, a camada de saída é sempre linear.

Uma RBF executa uma tarefa complexa de classificação de padrões. O problema de classificação é resolvido basicamente pela transformação do espaço de entrada para um espaço de dimensionalidade maior. Tal transformação é feita de forma não-linear de acordo com o Teorema de Cover (COVER, 1965). De acordo com esse teorema, um problema complexo de classificação de padrões, que esteja disposto não-linearmente em um espaço de alta dimensão, tem uma probabilidade maior de ser linearmente separável do que em

um espaço de dimensionalidade baixa.

A RBF apresenta três diferenças principais em relação à *Multilayer Perceptron* (MLP). A primeira é que ela sempre apresenta uma única camada intermediária, ou seja, uma camada oculta. Outra importante característica é que os neurônios da camada intermediária têm apenas funções de base radial como função de ativação, ao invés de função sigmoideal ou tangente hiperbólica. Por fim, em uma RBF os neurônios da camada de saída são sempre lineares.

3.5.3 *Sequential Minimal Optimization* - (SMO)

A *Support Vector Machine* - SVM é uma técnica de aprendizagem supervisionada proposta por Cortes e Vapnik (1995). Na fase de treinamento, atributos e saídas desejadas são disponibilizadas a fim de projetar os vetores de suporte. Os vetores de suporte são utilizados para obter um classificador. Na fase de teste, o classificador é utilizado para encontrar uma saída a partir de um vetor de entrada.

Ao utilizar SVM para reconhecimento de padrões, se faz necessário encontrar uma transformação não linear para manipular dados em um espaço que define o hiperplano de separação em uma função linearmente separável. Para isso, é necessário que seja aumentada a dimensionalidade do problema (COVER, 1965). As funções que aumentam a dimensionalidade do espaço de entrada são chamadas Funções de *Kernel*.

O modelo de otimização gerado pelo SVM envolve em sua formulação um problema quadrático. Para a solução desse problema do SVM, utilizou-se o algoritmo *Sequential Minimal Optimization* - SMO como *kernel*, desenvolvido por Platt (1999), que escolhe dois pontos por iteração para fazer a otimização.

O algoritmo SMO (PLATT, 1999) é reconhecido como um dos mais rápidos e o mais fácil para implementar em software para o caso geral de SVM, o não linear e não separável. O algoritmo é iterativo e adota uma solução analítica para a otimização de um par de Multiplicadores de Lagrange³ em cada iteração, evitando o armazenamento de matrizes de grandes dimensões em memória (HERNÁNDEZ, 2009).

3.5.4 *Random Tree*

O algoritmo *Random Tree* (BREIMAN, 2001) constrói diversas árvores de decisão de forma aleatória. A classificação funciona da seguinte maneira: o classificador de árvores aleatórias recebe o vetor de características, como entrada, classifica-o em cada árvore da floresta e exhibe o rótulo da classe que recebeu a maioria dos votos.

³ Multiplicadores de Lagrange é um procedimento para obtermos os pontos de máximo local ou mínimo local de uma função diferenciável quando suas variáveis estão sujeitas a restrições.

A representação dos dados como árvore tem a vantagem em comparação com outras abordagens de ser significativa e fácil de interpretar. O objetivo é criar um modelo de classificação que prevê o valor do rótulo com base em vários atributos de entrada. Cada nó interior da árvore corresponde a um dos atributos de entrada. O número de arestas de um nó interior é igual ao número de valores possíveis do atributo de entrada correspondente. Cada nó folha representa um valor do rótulo, dados os valores dos atributos de entrada representados pelo caminho da raiz até a folha. Uma árvore pára de crescer quando as seguintes condições são alcançadas: (a) um nó se torna vazio ou não haja mais exemplos para dividir no nó atual; ou (b) a profundidade da árvore exceder certos limites (HALL et al., 2009).

A poda é uma técnica na qual os nós de folhas que não contribuem para o poder discriminativo da árvore de decisão são removidos. Isso é feito para converter uma árvore super específica ou ajustada em excesso para uma forma mais geral, a fim de aumentar seu poder preditivo em conjuntos de dados não vistos.

3.5.5 *Random Forest*

O algoritmo *Random Forest* (BREIMAN, 2001) é uma combinação de predições de diversas árvores em que cada árvore depende dos valores de um vetor independente, amostrados aleatoriamente e com a mesma distribuição para todas as árvores da floresta. Aqui, floresta é o que se denomina para uma série de árvores de decisão. Após a geração de um grande número de árvores, as classes com maior número de votos são eleitas.

Em um *Random Forest*, cada nó é dividido usando o melhor dentre um subconjunto de indicadores escolhidos aleatoriamente naquele nó. Essa estratégia um tanto contraditória acaba por funcionar muito bem em comparação com muitos outros classificadores, além de ser robusto a superajuste nos parâmetros (BREIMAN, 2001). Além disso, é de fácil utilização, no sentido que ele tem apenas dois parâmetros (o número de variáveis no subconjunto aleatório em cada nó e o número de árvores da floresta) e, normalmente, não é muito sensível aos seus valores.

A partir de um vetor de atributos, são gerados outros vetores de atributos, que são embaralhados em relação ao vetor original. É gerado um vetor para cada árvore do *Random Forest*. Em seguida, os vetores de atributos são passados como parâmetro para as árvores de decisão. Cada árvore irá gerar um resultado para a classificação e, os resultados são combinados obtendo uma saída unificada.

3.5.6 Comitê de Classificadores

Comitê de classificadores é uma técnica que utiliza o resultado de diversos classificadores para obter a resposta final do problema, seu uso no aprendizado de máquinas

é crescente há alguns anos. Ao utilizar múltiplos classificadores e tomar a combinação de seus resultados, evita-se que seja utilizado apenas um classificador fraco (PONTI JR., 2011). O objetivo de combinar é aumentar a taxa de acerto em relação a classificadores utilizados individualmente, proporcionando uma maior confiabilidade nos dados. Em um comitê, cada classificador é um membro com poder de voto e a decisão final depende dos votos de todos os membros, podendo cada um ter voto com peso diferente. Trabalhos apresentados na literatura ((KUMAR et al., 2017; MOHAPATRA; PATRA; SATPATHY, 2014)) difundem o uso dessa técnica em sistemas de auxílio por computador.

A escolha de diferentes classificadores, com propriedades distintas e sua combinação, verificam a eficiência do uso de comitê. No presente trabalho, foram selecionados três classificadores, baseados em testes e em trabalhos da literatura, são eles: *Multilayer Perceptron* (MLP), *Radial Basis Function Networks* (RBF) e *Sequential Minimal Optimization* (SMO).

Existem inúmeras formas de se combinar classificadores, normalmente se utiliza a saída individual de cada um nesse processo, isso minimiza a ocorrência de decisões errôneas. Na combinação dos classificadores se utilizou o voto majoritário. Ele avalia a saída de cada classificador, e se a maioria classificar a imagem como pertencente a uma das classes, então essa será a sua classificação final. Diversos estudos são dedicados ao uso do voto majoritário para a combinação de classificadores (LIN et al., 2003; RUTA; GABRYS, 2002; LAM; SUEN, 1994). Além do voto majoritário, existem outras regras como o voto majoritário ponderado, contagem de borda, média e produto das probabilidades e a mediana (PONTI JR., 2011).

3.6 Métricas de Avaliação dos Resultados

A maioria dos critérios de análise dos resultados provém de uma matriz de confusão, que indica a quantidade de respostas corretas e incorretas para cada um dos grupos (CHIMIESKI; FAGUNDES, 2013).

A matriz de confusão é uma tabela que mostra o resultado da classificação, comparando-o com o resultado real. A matriz é composta por quatro valores: Verdadeiro Positivo (VP), Falso Positivo (FP), Falso Negativo (FN) e Verdadeiro Negativo (VN). Tais valores são obtidos a partir do resultado da classificação. Supondo a existência de duas classes hipotéticas: classe X e classe Y, são computados os seguintes valores:

- VP: número de objetos da classe X classificados como da classe X;
- FP: número de objetos da classe Y classificados como X;
- FN: número de objetos da classe X classificados como Y;

- VN: número de objetos da classe Y classificados como sendo da classe Y.

Os resultados do método proposto neste trabalho foram analisados por meio de três métricas existentes na literatura, são elas: Acurácia (A) (STORY; CONGALTON, 1986), índice *Kappa* (K) (CHIMIESKI; FAGUNDES, 2013) e Sensibilidade (S) (HONEST; KHAN, 2002)

3.6.1 Acurácia (A)

A acurácia de um classificador é a porcentagem de casos corretamente classificados em um conjunto de teste. Essa medida mede o quão bem um classificador reconhece instâncias de diversas classes. Ela pode ser calculada a partir da Equação 3.3.

$$A = \frac{VP + VN}{VP + FP + FN + VN}. \quad (3.3)$$

3.6.2 Sensibilidade (S)

Reflete a proporção de verdadeiros positivos em relação a suas predições positivas e as suas incorretas predições negativas. Essa medida mostra o comportamento dos objetos classificados como X, ou seja, de todos os objetos da classe X, quantos foram classificados como X. Ela pode ser calculada a partir da Equação 3.4.

$$S = \frac{VP}{VP + FN}. \quad (3.4)$$

3.6.3 Índice *Kappa* (K)

O índice *Kappa* é utilizado como uma medida apropriada da exatidão por representar inteiramente a matriz de confusão. Esse índice toma todos os elementos da matriz em consideração, ao invés de apenas aqueles que retratam a quantidade de classificações verdadeiras, o que ocorre quando se calcula a exatidão global da classificação (ROSENFELD; FITZPATRICK-LINS, 1986).

O índice *Kappa* é um coeficiente de concordância para escalas nominais que mede o relacionamento entre a concordância, além da casualidade, e a discordância esperada (ROSENFELD; FITZPATRICK-LINS, 1986). O índice *Kappa* pode ser calculado com base na Equação 3.5.

$$K = \frac{(\text{observado} - \text{esperado})}{1 - \text{esperado}} \times 100. \quad (3.5)$$

Nesse caso, “observado” é o valor global para a porcentagem correta, ou seja, o somatório da diagonal principal da matriz dividido pela quantidade de elementos e

“esperado” são os valores calculados usando-se os totais de cada linha e cada coluna da matriz de confusão. A qualidade dos resultados podem ser avaliados pelo valor do índice *Kappa* observado na Tabela 2, conforme definido por Landis e Koch (1977).

Tabela 2 – Nível de exatidão de uma classificação conforme o valor do índice *Kappa*.

Índice <i>Kappa</i> (K)	Qualidade
$K \leq 20\%$	Ruim
$20\% < K \leq 40\%$	Razoável
$40\% < K \leq 60\%$	Bom
$60\% < K \leq 80\%$	Muito Bom
$K > 80\%$	Excelente

Para testar a significância estatística da diferença entre dois índices *Kappa*, realizou-se o teste de hipótese conhecido por Teste Z e proposto por Congalton e Green (2008), descrito a seguir:

$$Z = \frac{K_2 - K_1}{\sqrt{\sigma_{K_2}^2 + \sigma_{K_1}^2}}, \quad (3.6)$$

no qual K_1 = índice *Kappa* do classificador 1; K_2 = índice *Kappa* do classificador 2 e σ^2 = variância do índice *Kappa*, que pode ser calculado por:

$$\sigma^2 = \frac{\theta_1(1 - \theta_1)}{N(1 - \theta_2)^2}, \quad (3.7)$$

em que N é o número total de elementos na matriz de confusão. Vale ressaltar que neste trabalho o Teste Z foi o escolhido para a comparação dos índices *Kappa*, pois, de acordo com a análise da literatura, esse teste é mais utilizado para essa finalidade (CONGALTON; GREEN, 2008).

Com as classificações, comparadas por meio do Teste Z, testou-se a hipótese de nulidade (H_0), que ambos os classificadores apresentaram classificações equivalentes, em contraposição a H_1 em que as classificações são diferentes, com nível de significância de 5%.

Por fim, para avaliação dos resultados obtidos foram utilizadas as métricas mostradas na Seção 3.6. Os parâmetros utilizados para a classificação foram os padrões de cada classificador no WEKA (HALL et al., 2009) e, o método de validação utilizado foi subamostragem aleatória, variação do método *Holdout* (HAN; PEI; KAMBER, 2011), sendo os dados divididos em dois conjuntos distintos, denominados de conjunto de treinamento e conjunto de teste, repetindo a execução de forma aleatória N vezes. Para os experimentos, foi utilizado o valor de $N = 10$, com 90% da base de imagens para treino e 10% para testes. Para a análise dos resultados, apenas a média das execuções são apresentadas. Os

resultados são ilustrados utilizando a assinatura da imagem que obteve o melhor resultado na classificação.

3.7 Considerações Finais

Neste capítulo foram apresentados os principais materiais e métodos utilizados no desenvolvimento deste trabalho.

Descritores de imagens são algoritmos de visão computacional responsáveis por extrair atributos significativos das imagens. Algoritmos de classificação são utilizados em diversas áreas da computação, visto que na literatura, existe inúmeros algoritmos com esse propósito. Neste trabalho, aplica-se os classificadores mais utilizados.

4 Resultados e Discussões

Neste capítulo são apresentados e discutidos os resultados encontrados nos testes realizados com o objetivo de analisar o desempenho dos algoritmos e método proposto para a classificação de cédulas monetárias.

Para a obtenção dos resultados da abordagem proposta foram testadas diferentes combinações para cada cédula, variando os algoritmos de detecção de PoI, descrição de PoI, geração de assinaturas e classificadores, sendo os principais resultados ilustrados nos anexos [A](#), [B](#) e [C](#). Todos os algoritmos foram implementados utilizando a linguagem de programação *Python* (ROSSUM, 1995) em conjunto com os algoritmos de mineração de dados disponíveis no aplicativo *Waikato Environment for Knowledge Analysis* (WEKA) (HALL et al., 2009).

Conforme apresentado na Seção 3.2, foram utilizadas três bases de imagens de cédulas monetárias para a realização dos testes:

- Cédulas monetárias de dólar - 7 classes e 7056 imagens;
- Cédulas monetárias de euro - 7 classes e 7056 imagens;
- Cédulas monetárias de real (segunda família de cédulas) - 6 classes e 6048 imagens.

Devido ao baixo poder computacional disponível para a realização dos testes, foi utilizado inicialmente a base de imagens das cédulas de real, por conter uma menor quantidade de classes e imagens, para analisar o desempenho das combinações de algoritmos propostas. A partir dos resultados obtidos nessa base, as melhores combinações de algoritmos foram aplicadas nas bases de imagens de dólar e euro.

Para cada cédula monetária, os melhores resultados foram obtidos por diferentes combinações de detector, descritor, assinatura e classificador. Foi realizado o Teste Z (CONGALTON; GREEN, 2008) para comparar estatisticamente os resultados em um nível de significância de 5% para avaliar se uma configuração era significativamente diferente de outra.

4.1 Resultados Utilizando Imagens de Dólar

A Tabela 3 apresenta os melhores resultados alcançados, para os classificadores testados, para as imagens das cédulas de dólar, bem como a combinação de detector, descritor e assinatura da imagem que originou esse resultado.

Tabela 3 – Combinações para a melhor classificação utilizando as imagens das cédulas de dólar.

Classificador	A (%)	S (%)	K (%)	Detector	Descritor	Assinatura
MLP	98,40 ±0,62	98,40	98,10	MSER	SURF	Média
RBF	92,64 ±0,77	92,60	91,42	MSER	SURF	BoVW (500)
Random Tree	78,37 ±3,78	75,20	74,70	MSER	SURF	BoVW (400)
SMO	99,77 ±0,30	99,60	99,60	SURF	SURF	BoVW (500)
SMO+MLP+RBF	94,95 ±1,20	94,50	94,10	MSER	SURF	Média
Random Forest	99,15 ±0,36	99,20	99,01	MSER	SURF	BoVW (500)

A abordagem BoVW com um tamanho de palavra de 500 foi usada como assinatura da imagem. O detector SURF combinado com o descritor SURF e o classificador SMO obteve o melhor resultado, com uma taxa de acurácia de 99,77% e uma taxa de sensibilidade de 99,60%. De acordo com a Tabela 3, o índice *Kappa* classificou o resultado como excelente, com uma taxa de 99,60%.

O Teste Z mostrou que a configuração usando um comprimento de palavra de tamanho 400 e a mesma combinação de detector e descritor de PoI e mesmo classificador exibiram desempenho equivalente.

4.2 Resultados Utilizando Imagens de Euro

A Tabela 4 mostra a combinação de detector, descritor e tipo de assinatura da imagem que proporcionou o melhor resultado para cada classificador testado. Os melhores resultados para a identificação das imagens em euro foram dados quando a assinatura da imagem foi gerada por BoVW com palavras de tamanho 300. O classificador *Random Forest* usado em conjunto com o detector MSER e o descritor SURF obteve a melhor acurácia com 99,12%, uma sensibilidade de 99,00% e índice *Kappa* de 99,00%, o que é considerado excelente.

Um Teste Z mostrou que uma configuração com comprimentos de palavras de 200, 400 e 500 com a combinação do detector MSER, descritor SURF e classificador *Random Forest* exibiram desempenhos equivalentes.

Tabela 4 – Combinações para a melhor classificação utilizando as imagens das cédulas de euro.

Classificador	A (%)	S (%)	K (%)	Detector	Descritor	Assinatura
MLP	97,99 ±0,75	98,00	97,60	MSER	SURF	Média
RBF	89,94 ±1,64	88,70	88,20	MSER	SURF	BoVW (400)
Random Tree	79,30 ±2,42	76,00	75,80	MSER	BRISK	BoVW (400)
SMO	97,80 ±0,31	97,40	97,40	MSER	SURF	BoVW (400)
SMO+MLP+RBF	94,36 ±1,26	94,40	93,40	MSER	SURF	Mean
Random Forest	99,12±0,14	99,00	99,00	MSER	SURF	BoVW (300)

4.3 Resultados Utilizando Imagens de Real

A Tabela 5 mostra a combinação de detector, descritor e tipo de assinatura que obteve o melhor resultado para cada classificador testado. A abordagem BoVW foi usada como assinatura de imagem de tamanho de palavra igual a 400. O detector MSER combinado com o descritor SURF e classificador *Random Forest* obteve a melhor classificação para as notas de real, com uma taxa de acurácia de 96,95%, uma sensibilidade de 97,30% e um índice *Kappa* de 96,30%, o que representa um excelente resultado.

Tabela 5 – Combinações para a melhor classificação utilizando as imagens das cédulas de real.

Classificador	A (%)	S (%)	K (%)	Detector	Descritor	Assinatura
MLP	94,91 ±1,21	94,90	93,80	MSER	SURF	Média
RBF	84,36 ±2,86	84,30	81,20	MSER	SURF	BoVW (400)
Random Tree	76,90 ±8,70	76,80	72,20	MSER	SURF	BoVW (400)
SMO	92,70 ±5,55	92,60	91,20	MSER	SURF	BoVW (400)
SMO+MLP+RBF	91,55 ±1,67	91,30	89,80	MSER	BRISK	Média
Random Forest	96,95 ±0,72	97,30	96,30	MSER	SURF	BoVW (400)

Um Teste Z mostrou que configurações com comprimentos de palavras de 300 e 500 com a mesma combinação de detector de PoI, descritor de PoI e classificador exibiram desempenhos equivalentes. Desempenho equivalente também foi alcançado com a combinação de SURF como detector e descritor, assinatura da imagem gerada pela abordagem de BoVW com tamanho de palavra igual a 300 e classificador *Random Forest*.

Em cédulas monetárias de real realizamos testes com imagens utilizadas no cotidiano, obtidas com uma câmera de *smartphone*, conforme ilustrado na Figura 15, com objetivo de validar a nossa proposta e verificar o reconhecimento utilizando cédulas diferentes das contidas na base de imagens gerada.



Figura 15 – Amostras de imagens das cédulas de reais.

Nos testes realizados nas cédulas de real foi constatado o detector MSER obtendo a melhor detecção de PoIs, a Figura 16 ilustra a detecção de pontos, realizando um comparativo entre imagens da base gerada com imagens capturadas com o uso de um *smartphone*. Os pontos detectados nas imagens possuem áreas equivalentes, validando o uso do detector MSER.

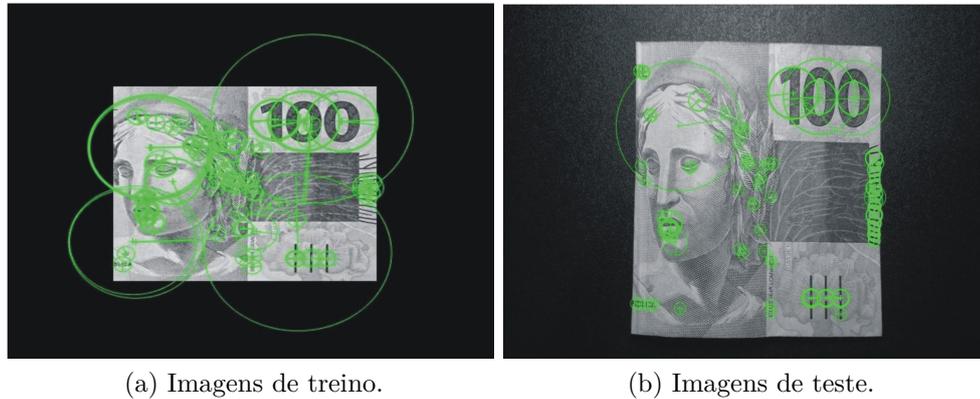


Figura 16 – Amostra de detecção de pontos.

A Tabela 6 ilustra o resultado da classificação utilizando imagens de cédulas obtidas por câmera de *smartphone*. Foram testadas imagens inteiras e dobradas ao meio, com fundos na cor branca e preta. Foi utilizada a combinação que obteve o melhor resultado nos testes iniciais, sendo a abordagem BoVW com assinatura de imagem de tamanho de palavra igual a 400, detector MSER combinado com o descritor SURF e classificador *Random Forest*.

Nas cédulas inteiras foi obtido uma taxa de acurácia de 91,66% e um índice *Kappa* de 90,00%, o que representa um excelente resultado. Nas cédulas dobradas foi obtido uma taxa de acurácia de 80,55% e um índice *Kappa* de 76,67%, o que representa um resultado muito bom.

Tabela 6 – Resultado utilizando cédulas de real, obtidas com câmera de *smartphone*.

Cédulas	Fundo	Quantidade	A (%)	S (%)	K (%)
Inteiras	Preto	12	91,66	91,70	90,00
Inteiras e Metade	Preto e Branco	72	80,55	80,60	76,67

4.4 Discussões

A Tabela 7 ilustra a média dos PoIs detectados nas imagens das cédulas utilizadas. Os resultados apresentados demonstram que os algoritmos MSER e SURF são mais robustos a variações de iluminação, escala, transformação, rotação e ruído durante a aquisição de imagens. Como eles extraem recursos com menos pontos que os outros, eles também são computacionalmente mais rápidos.

Tabela 7 – Pontos retornados pelos detectores analisados.

Dólar				
	BRISK	FAST	MSER	SURF
Originais	1.599	1.081	678	596
Imagens com filtro	3.172	2.566	511	582
Com alteração de iluminação	393	290	228	169
Com inserção de fundo	2.882	2.144	505	692
Média Geral	2.012	1.520	481	510
Euro				
Originais	857	618	395	266
Imagens com filtro	3.534	2.779	274	335
Com alteração de iluminação	159	113	120	47
Com inserção de fundo	2.622	2.255	290	446
Média Geral	1.793	1.441	270	274
Real				
Originais	484	292	381	191
Imagens com filtro	1.716	1.434	272	209
Com alteração de iluminação	180	110	188	72
Com inserção de fundo	1.400	1.136	271	289
Média Geral	945	743	278	190

De acordo com (LEUTENEGGER; CHLI; SIEGWART, 2011), BRISK oferece uma alternativa marcadamente mais rápida e um desempenho que é comparável ao de SIFT e SURF. Contudo, nos testes realizados, o descritor SURF foi superior ao BRISK. Para as notas de dólar, não houve resultados para o descritor BRISK entre os melhores resultados. Para as notas euro e real, houve apenas um exemplo, como pode ser observado nas Tabelas 3, 4 e 5.

Para as notas de euro e real, houve combinações que eram estatisticamente iguais, ou seja, o BoVW com tamanho de palavras de 300, 400 e 500 com o detector MSER, o descritor SURF e o classificador *Random Forest*. Como o resultado obtido não possui diferenças estatísticas significantes, aconselha-se a utilização do tamanho da palavra igual a 300, diminuindo a quantidade de atributos e aumentando o desempenho do processamento. Para as notas de dólar, não houve combinação entre os melhores resultados em que as notas de euro e real eram estatisticamente as mesmas.

Utilizou-se a combinação com os melhores classificadores avaliados a saber: MLP, RBF e SMO. Contudo, os resultados obtidos não foram superiores ao comitê *Random Forest*, utilizando um tempo de execução muito inferior à combinação dos classificadores criado.

4.5 Considerações Finais

Este capítulo apresentou os experimentos da metodologia proposta, empregando bases de imagens criadas, utilizando aumento de dados na obtenção de uma melhor perfor-

mance e confiabilidade nos resultados. Com base nos resultados obtidos pela abordagem, pode-se perceber que a proporção da quantidade de imagens utilizadas pelos resultados obtidos é possível validar a sua robustez, atingindo em todos os tipos de cédulas monetárias um índice *Kappa* considerado “Excelente”.

5 Conclusões e Trabalhos Futuros

5.1 Conclusões

A visão computacional está evoluindo constantemente, tornando-se uma ferramenta que pode ser utilizado em diferentes áreas, bem como no auxílio as pessoas com deficiência, auxiliado com o uso das tecnologias assistivas.

De acordo com os testes realizados neste estudo, os melhores resultados foram obtidos utilizando uma assinatura da imagem gerada com o BoVW. Para a classificação das notas de dólar, o melhor resultado foi obtido utilizando o detector SURF, descritor SURF e classificador SMO; e para as notas de euro e real, os melhores resultados foram obtidos com o detector MSER, descritor SURF e classificador *Random Forest*. Essa configuração foi a segunda melhor combinação para a classificação nas cédulas de dólar, obtendo um índice *kappa* excelente, conforme ilustrado na Tabela 8.

Em um sistema automatizado no qual o tipo de nota testada é desconhecido, o detector MSER, os descritor SURF, a abordagem BoVW e o comitê *Random Forest* devem ser usados, pois eles alcançaram os melhores resultados para duas das três bases de dados testadas e um índice *kappa* excelente para todas as bases de dados.

Este estudo se destacou dos demais trabalhos relacionados, pois busca identificar o melhor detector, descritor e classificador para as cédulas monetárias mais utilizadas mundialmente e a de uso nacional. Diferenciou-se também pela forma de análise das cédulas monetárias, buscando identificar cédulas em diferentes perspectivas, orientação e pedaços da imagem, aumentando o poder de reconhecimento das notas em diferentes maneiras.

5.2 Trabalhos Futuros

Em trabalhos futuros, pretendemos utilizar os detectores, descritores e classificadores analisados neste estudo como base para a criação de aplicativos para dispositivos móveis que possam auxiliar deficientes visuais na identificação de denominações de cédulas monetárias. Devem ser realizados testes utilizando outros descritores em conjunto com classificadores, em uma tentativa de obter melhores resultados, pretendemos investigar o efeito dessa operação no resultado final da classificação. Outro ponto crítico é o passo de classificação, já que nosso objetivo principal era uma descrição das imagens. Os classificadores avaliados não estavam sintonizados e acreditamos que os resultados poderiam ser melhorados com o ajuste fino dos parâmetros.

Referências

- ABBURU, V.; GUPTA, S.; RIMITHA, S. R.; MULIMANI, M.; KOOLAGUDI, S. G. Currency recognition system using image processing. In: *2017 Tenth International Conference on Contemporary Computing (IC3)*. [S.l.: s.n.], 2017. p. 1–6. Citado na página [7](#).
- ALAHY, A.; ORTIZ, R.; VANDERGHEYNST, P. Freak: Fast retina keypoint. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2012. p. 510–517. Citado na página [5](#).
- ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, p. 183–193, 1954. Citado na página [12](#).
- BAY, H.; ESS, A.; TUYTELAARS, T.; GOOL, L. V. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, v. 110, n. 3, p. 346 – 359, 2008. Similarity Matching in Computer Vision and Multimedia. Citado 7 vezes nas páginas [9](#), [5](#), [10](#), [12](#), [15](#), [17](#) e [18](#).
- BRAGA, A. de P.; FERREIRA, A. C. P. de L.; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. [S.l.]: LTC Editora, 2007. Citado na página [22](#).
- BRAGA, J. C.; JR, A. C. C. C.; GRADUANDO, R. J. P. D.; ALBERNAZ, N. H. da C. Estudo e relato sobre a utilização da tecnologia pelos deficientes visuais. In: BRAZILIAN COMPUTER SOCIETY. *Proceedings of the 11th Brazilian Symposium on Human Factors in Computing Systems*. [S.l.], 2012. p. 37–46. Citado na página [1](#).
- BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, Oct 2001. Citado 2 vezes nas páginas [23](#) e [24](#).
- CASTRO, L. N. D.; ZUBEN, F. J. V.; JR, G. A. de D. The construction of a boolean competitive neural network using ideas from immunology. *Neurocomputing*, Elsevier, v. 50, p. 51–85, 2003. Citado na página [7](#).
- CHIMIESKI, B. F.; FAGUNDES, R. D. R. Association and classification data mining algorithms comparison over medical datasets. *Journal of Health Informatics (JHI)*, p. 44–51, 2013. Citado 2 vezes nas páginas [25](#) e [26](#).
- CONGALTON, R. G.; GREEN, K. *Assessing the accuracy of remotely sensed data: principles and practices*. [S.l.]: CRC press, 2008. Citado 2 vezes nas páginas [27](#) e [29](#).
- CORTES, C.; VAPNIK, V. Support-vector networks. *Machine Learning*, v. 20, n. 3, p. 273–297, Sep 1995. Citado na página [23](#).
- COSTA, C. M.; VEIGA, G.; SOUSA, A. Recognition of banknotes in multiple perspectives using selective feature matching and shape analysis. In: IEEE. *2016 International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. [S.l.], 2016. p. 235–240. Citado 4 vezes nas páginas [9](#), [6](#), [7](#) e [11](#).

- COVER, T. M. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, EC-14, n. 3, p. 326–334, June 1965. Citado 2 vezes nas páginas 22 e 23.
- GRIJALVA, F.; RODRÍGUEZ, J. C.; LARCO, J.; OROZCO, L. Smartphone recognition of the u.s. banknotes' denomination, for visually impaired people. In: *2010 IEEE ANDESCON*. [S.l.: s.n.], 2010. p. 1–6. Citado 2 vezes nas páginas 1 e 5.
- HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The Weka data mining software: An update. *SIGKDD Explor. Newsl.*, ACM, v. 11, n. 1, p. 10–18, nov. 2009. Citado 3 vezes nas páginas 24, 27 e 29.
- HAN, J.; PEI, J.; KAMBER, M. *Data mining: concepts and techniques*. [S.l.]: Elsevier, 2011. Citado na página 27.
- HASANUZZAMAN, F. M.; YANG, X.; TIAN, Y. Robust and effective component-based banknote recognition for the blind. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, v. 42, n. 6, p. 1021–1030, Nov 2012. Citado 3 vezes nas páginas 9, 5 e 6.
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2nd. ed. [S.l.]: Prentice Hall PTR, 1998. Citado na página 22.
- HERNÁNDEZ, R. A. *MP-SMO: um algoritmo para a implementação VLSI do treinamento de máquinas de vetores de suporte*. Dissertação (Mestrado) — Universidade de São Paulo, 2009. Citado na página 23.
- HOLLAND, S. M. Principal components analysis (pca). *Department of Geology, University of Georgia, Athens, GA*, p. 30602–2501, 2008. Citado na página 5.
- HONEST, H.; KHAN, K. S. Reporting of measures of accuracy in systematic reviews of diagnostic literature. *BMC Health Services Research*, v. 2, n. 1, p. 4, Mar 2002. Citado na página 26.
- JEONG, K.; MOON, H. Object detection using fast corner detector based on smartphone platforms. In: IEEE. *Computers, Networks, Systems and Industrial Engineering (CNSI), 2011 First ACIS/JNU International Conference on*. [S.l.], 2011. p. 111–115. Citado na página 15.
- KASPERS, A. *Blob detection*. Dissertação (Mestrado) — Image Sciences Institute - UMC Utrecht, 2011. Citado na página 12.
- KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI) - Volume 2*. [S.l.]: Morgan Kaufmann Publishers Inc., 1995. p. 1137–1143. Citado na página 7.
- KUMAR, A.; KIM, J.; LYNDON, D.; FULHAM, M.; FENG, D. An ensemble of fine-tuned convolutional neural networks for medical image classification. *IEEE Journal of Biomedical and Health Informatics*, v. 21, n. 1, p. 31–40, Jan 2017. Citado na página 25.

- LAM, L.; SUEN, C. Y. A theoretical analysis of the application of majority voting to pattern recognition. In: *Proceedings of the 12th International Conference on Pattern Recognition (IAPR), Vol. 3 - Conference C: Signal Processing*. [S.l.: s.n.], 1994. v. 2, p. 418–420 vol.2. Citado na página 25.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, [Wiley, International Biometric Society], v. 33, n. 1, p. 159–174, 1977. Citado na página 27.
- LAWRENCE, S.; GILES, C. L.; TSOI, A. C.; BACK, A. D. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*, IEEE, v. 8, n. 1, p. 98–113, 1997. Citado na página 8.
- LEUTENEGGER, S.; CHLI, M.; SIEGWART, R. Y. Brisk: Binary robust invariant scalable keypoints. In: *2011 International Conference on Computer Vision*. [S.l.: s.n.], 2011. p. 2548–2555. Citado 8 vezes nas páginas 9, 10, 12, 15, 16, 17, 19 e 33.
- LIN, X.; YACOUB, S.; BURNS, J.; SIMSKE, S. Performance analysis of pattern classifier combination by plurality voting. *Pattern Recogn. Lett.*, Elsevier Science Inc., v. 24, n. 12, p. 1959–1969, ago. 2003. Citado na página 25.
- LOWE, D. G. Object recognition from local scale-invariant features. In: *Proceedings of the Seventh IEEE International Conference on Computer Vision*. [S.l.: s.n.], 1999. v. 2, p. 1150–1157 vol.2. Citado 2 vezes nas páginas 12 e 15.
- MAIR, E.; HAGER, G. D.; BURSCHKA, D.; SUPPA, M.; HIRZINGER, G. Adaptive and generic corner detection based on the accelerated segment test. In: *Computer Vision – ECCV 2010*. [S.l.]: Springer Berlin Heidelberg, 2010. p. 183–196. Citado na página 16.
- MATAS, J.; CHUM, O.; URBAN, M.; PAJDLA, T. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, v. 22, n. 10, p. 761 – 767, 2004. Citado 3 vezes nas páginas 10, 12 e 13.
- MIKOLAJCZYK, K.; SCHMID, C. A performance evaluation of local descriptors. *IEEE transactions on pattern analysis and machine intelligence*, v. 27, n. 10, p. 1615–1630, 2005. Citado na página 12.
- MITTAL, S.; MITTAL, S. Indian banknote recognition using convolutional neural network. In: IEEE. *2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU)*. [S.l.], 2018. p. 1–6. Citado na página 8.
- MOHAPATRA, S.; PATRA, D.; SATPATHY, S. An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images. *Neural Computing and Applications*, Springer, v. 24, n. 7-8, p. 1887–1904, 2014. Citado na página 25.
- MORTENSEN, E. N.; DENG, H.; SHAPIRO, L. A sift descriptor with global context. In: IEEE. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. [S.l.], 2005. v. 1, p. 184–190. Citado na página 6.
- MUJA, M.; LOWE, D. G. Fast approximate nearest neighbors with automatic algorithm configuration. *VISAPP (1)*, v. 2, n. 331-340, p. 2, 2009. Citado na página 6.

- MULMULE-SHIRKHEDKAR, D.; DANI, A. R. Comparative study of surf and freak descriptor on indian rupee currency notes. In: *2015 International Conference on Information Processing (ICIP)*. [S.l.: s.n.], 2015. p. 784–789. Citado na página 5.
- OYEDOTUN, O. K.; KHASHMAN, A. Banknote recognition: investigating processing and cognition framework using competitive neural network. *Cognitive Neurodynamics*, Springer, v. 11, n. 1, p. 67–79, 2017. Citado na página 7.
- PASCOLINI, D.; MARIOTTI, S. P. Global estimates of visual impairment: 2010. *British Journal of Ophthalmology*, BMJ Publishing Group Ltd., p. bjophthalmol–2011, 2011. Citado na página 1.
- PLATT, J. C. 12 fast training of support vector machines using sequential minimal optimization. *Advances in Kernel Methods*, p. 185–208, 1999. Citado 2 vezes nas páginas 7 e 23.
- PONTI JR., M. P. Combining classifiers: From the creation of ensembles to the decision fusion. In: *2011 24th SIBGRAPI Conference on Graphics, Patterns, and Images Tutoriais*. [S.l.: s.n.], 2011. p. 1–10. Citado na página 25.
- ROSENFELD, G. H.; FITZPATRICK-LINS, K. A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing*, v. 52, n. 2, p. 223–227, 1986. Citado na página 26.
- ROSSUM, G. Python reference manual. CWI (Centre for Mathematics and Computer Science), 1995. Citado na página 29.
- ROSTEN, E.; DRUMMOND, T. Machine learning for high speed corner detection. In: SPRINGER. *European Conference on Computer Vision*. [S.l.], 2006. p. 430–443. Citado 4 vezes nas páginas 10, 12, 13 e 16.
- RUTA, D.; GABRYS, B. A theoretical analysis of the limits of majority voting errors for multiple classifier systems. *Pattern Analysis and Applications*, Springer, v. 5, n. 4, p. 333–350, 2002. Citado na página 25.
- STORY, M.; CONGALTON, R. G. Accuracy assessment: a user’s perspective. *Photogrammetric Engineering and Remote Sensing*, Falls Church, v. 52, n. 3, p. 397–399, 1986. Citado na página 26.
- SUAREZ, O. D.; CARROBLES, M. d. M. F.; ENANO, N. V.; GARCÍA, G. B.; GRACIA, I. S. *OpenCV Essentials*. [S.l.]: Packt Publishing, 2014. ISBN 1783984244, 9781783984244. Citado na página 12.
- SUG, H. Performance comparison of rbf networks and mlps for classification. In: *Proceedings of the 9th WSEAS International Conference on Applied Informatics and Communications*. [S.l.]: World Scientific and Engineering Academy and Society (WSEAS), 2009. p. 450–454. Citado na página 22.
- TEIXEIRA, V. V.; BRAGA, J. C.; NASCIMENTO, M. Z. do. Aplicativo identificador de cédulas para deficientes visuais. *Revista de Informática Aplicada*, v. 11, n. 1, 2016. Citado na página 6.

- TOLA, E.; LEPETIT, V.; FUA, P. Daisy: An efficient dense descriptor applied to wide-baseline stereo. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 32, n. 5, p. 815–830, 2010. Citado na página 12.
- TREIBER, M. A. *An introduction to object recognition: selected algorithms for a wide variety of applications*. [S.l.]: Springer Science & Business Media, 2010. Citado 2 vezes nas páginas 13 e 14.
- TURK, M.; PENTLAND, A. Eigenfaces for recognition. *Journal of cognitive neuroscience*, MIT Press, v. 3, n. 1, p. 71–86, 1991. Citado na página 5.
- VISWANATHAN, D. Features from accelerated segment test (fast). In: . [S.l.: s.n.], 2011. Citado na página 13.
- YANG, J.; JIANG, Y.-G.; HAUPTMANN, A. G.; NGO, C.-W. Evaluating bag-of-visual-words representations in scene classification. In: *Proceedings of the International Workshop on Workshop on Multimedia Information Retrieval*. [S.l.]: ACM, 2007. p. 197–206. Citado na página 20.

Anexos

ANEXO A – Resultados da Classificação Utilizando a Base de Imagens das Cédulas de Dólar.

Tabela 8 – Principais resultados de acurácia na classificação das cédulas monetárias de dólar.

Classificadores Individuais				
<i>Multilayer Perceptron</i>				
Assinatura	MSER		SURF	
	BRISK	SURF	BRISK	SURF
Média	96,272	98,398	88,633	96,159
Mediana	91,199	97,463	77,820	91,737
Moda	72,250	44,756	49,206	37,514
BoW - 500	68,698	14,497	51,619	36,753
<i>Radial Basic Fuction Network</i>				
Média	65,504	67,261	41,454	47,831
Mediana	62,528	60,756	42,276	44,118
Moda	60,600	29,209	34,750	19,019
BoW - 500	90,730	92,618	53,619	92,361
<i>Random Tree</i>				
Média	68,707	74,872	60,870	68,353
Mediana	68,948	70,039	58,999	65,121
Moda	63,265	48,823	48,214	38,945
BoW - 500	67,492	61,350	38,000	63,988
<i>Sequential Minimal Optimization</i>				
Média	86,408	87,896	69,316	77,763
Mediana	81,235	85,034	64,427	73,937
Moda	68,990	36,763	46,882	29,365
BoW - 500	97,555	98,557	97,015	99,773
Comitê de Classificadores				
MLP + RBF + SMO				
Média	94,416	94,954	82,894	88,733
Mediana	89,795	93,693	74,759	86,281
Moda	73,356	42,063	50,283	35,856
BoW - 500	94,030	93,522	71,491	94,688
<i>Random Forest</i>				
Média	95,223	96,612	91,468	93,891
Mediana	94,331	96,230	90,178	92,800
Moda	88,931	71,343	76,615	67,715
BoW - 500	97,365	98,550	92,095	99,136

ANEXO B – Resultados da Classificação Utilizando a Base de Imagens das Cédulas de Euro.

Tabela 9 – Principais resultados de acurácia na classificação das cédulas monetárias de euro.

Classificadores Individuais				
<i>Multilayer Perceptron</i>				
Assinaturas	MSER		SURF	
	BRISK	SURF	BRISK	SURF
Média	96,343	97,987	75,382	84,963
Mediana	91,964	95,734	60,303	72,505
Moda	72,392	52,097	36,663	35,459
BoW - 300	54,349	26,548	34,698	34,087
<i>Radial Basic Fuction Network</i>				
Média	61,238	61,706	25,977	29,450
Mediana	63,917	55,966	27,480	27,735
Moda	56,760	28,784	26,658	19,345
BoW - 300	85,841	86,647	41,968	75,106
<i>Random Tree</i>				
Média	75,552	73,469	48,256	53,769
Mediana	74,560	72,718	44,543	50,807
Moda	69,331	59,949	38,407	39,044
BoW - 300	79,301	79,123	31,238	60,792
<i>Sequential Minimal Optimization</i>				
Média	79,988	86,210	53,202	63,081
Mediana	77,253	82,681	47,491	51,828
Moda	65,759	39,384	33,404	31,901
BoW - 300	94,222	97,312	86,063	95,318
Comitê de Classificadores				
MLP + RBF + SMO				
Média	92,885	94,359	69,883	49,263
Mediana	89,271	91,751	57,341	67,531
Moda	73,072	48,001	37,386	35,104
BoW - 300	88,507	88,542	54,793	79,815
<i>Random Forest</i>				
Média	96,697	97,335	82,398	89,200
Mediana	96,712	97,094	79,761	85,657
Moda	93,976	81,108	82,894	66,496
BoW - 300	95,523	99,123	74,759	96,676

ANEXO C – Resultados da Classificação Utilizando a Base de Imagens das Cédulas de Real.

Tabela 10 – Principais resultados de acurácia na classificação das cédulas monetárias de real.

Classificadores Individuais								
Multilayer Perceptron								
Assinatura	BRISK		FAST		MSER		SURF	
	BRISK	SURF	BRISK	SURF	BRISK	SURF	BRISK	SURF
Média	67,890	77,314	21,808	75,000	94,808	94,907	78,125	84,060
Mediana	53,571	73,644	21,808	62,797	89,682	90,476	68,468	74,619
Moda	34,391	26,570	45,618	38,839	74,272	51,851	45,932	43,220
BoW - 400	22,293	30,918	23,894	18,234	55,333	43,052	22,656	37,772
Radial Basic Fuction Network								
Média	29,332	28,108	23,743	29,761	65,162	52,827	26,653	34,093
Mediana	28,687	29,761	23,743	27,050	63,425	50,694	28,802	31,233
Moda	25,611	19,229	33,019	25,297	58,945	29,761	33,482	21,726
BoW - 400	59,026	34,366	67,310	51,930	78,952	69,306	70,429	67,508
Random Tree								
Média	44,758	52,595	24,917	49,404	74,107	71,015	52,662	53,290
Mediana	43,518	52,430	24,917	48,974	72,933	69,824	51,868	51,008
Moda	33,267	28,422	44,659	39,037	68,535	59,093	43,898	40,988
BoW - 400	36,419	28,014	47,640	44,405	67,714	70,462	50,181	68,729
Sequential Minimal Optimization								
Média	45,287	46,974	20,634	51,785	84,705	81,944	60,367	65,261
Mediana	40,691	46,362	20,634	48,131	79,067	76,074	56,216	54,431
Moda	33,614	24,239	42,476	34,804	66,501	39,632	43,634	65,029
BoW - 400	77,970	76,407	85,940	77,937	85,238	86,518	88,448	85,132
Comitê de Classificadores								
MLP + RBF + SMO								
Média	59,821	66,435	23,263	69,791	91,550	90,244	74,437	78,687
Mediana	49,586	62,681	23,263	57,622	87,202	86,326	65,740	70,899
Moda	35,416	26,504	45,916	38,789	73,743	48,363	46,031	41,931
BoW - 400	61,847	56,878	76,874	54,058	80,285	75,098	78,626	73,861
Random Forest								
Média	73,131	86,061	30,985	78,091	94,246	93,204	85,003	86,739
Mediana	70,122	84,639	30,985	77,546	94,494	92,840	82,225	82,787
Moda	55,985	49,404	66,650	62,516	91,236	80,638	72,387	69,725
BoW - 400	77,392	58,516	88,052	84,504	89,333	96,947	89,603	95,709