



**DETECÇÃO E RECONHECIMENTO DE NÚMEROS SERIAIS
EM CÉDULAS DA SEGUNDA FAMÍLIA DO REAL**

IAN ALVES ULIAN

**DISSERTAÇÃO DE MESTRADO PROFISSIONAL EM INFORMÁTICA
FORENSE E SEGURANÇA DA INFORMAÇÃO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

FACULDADE DE TECNOLOGIA

UNIVERSIDADE DE BRASÍLIA

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**DETECÇÃO E RECONHECIMENTO DE NÚMEROS SERIAIS
EM CÉDULAS DA SEGUNDA FAMÍLIA DO REAL**

IAN ALVES ULIAN

Orientador: PROF. DR. FLÁVIO DE BARROS VIDAL, CIC/UNB

**DISSERTAÇÃO DE MESTRADO PROFISSIONAL EM INFORMÁTICA
FORENSE E SEGURANÇA DA INFORMAÇÃO**

**PUBLICAÇÃO PPGENE.DM - 633/2016
BRASÍLIA-DF, 13 DE DEZEMBRO DE 2016.**

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

DETECÇÃO E RECONHECIMENTO DE NÚMEROS SERIAIS EM
CÉDULAS DA SEGUNDA FAMÍLIA DO REAL

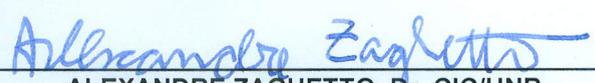
IAN ALVES ULIAN

DISSERTAÇÃO DE MESTRADO PROFISSIONAL SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.

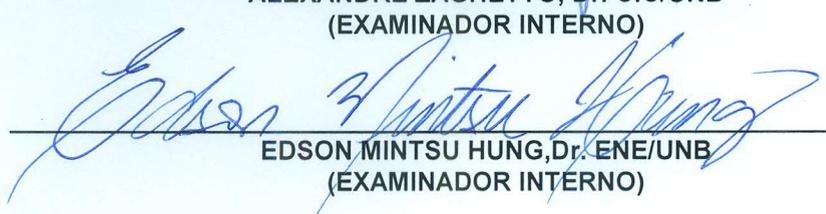
APROVADA POR:



FLÁVIO DE BARROS VIDAL, Dr. CIC/UNB
(ORIENTADOR)



ALEXANDRE ZAGHETTO, Dr. CIC/UNB
(EXAMINADOR INTERNO)



EDSON MINTSU HUNG, Dr. ENE/UNB
(EXAMINADOR INTERNO)

Brasília, 13 de Dezembro de 2016.

FICHA CATALOGRÁFICA

IAN ALVES ULIAN

**Detecção e reconhecimento de números seriais em cédulas da segunda família do Real
2016xv, 61p., 201x297 mm**

(ENE/FT/UnB, Mestre, INFORMÁTICA FORENSE E SEGURANÇA DA INFORMAÇÃO,
2016)

Dissertação de Mestrado - Universidade de Brasília

Faculdade de Tecnologia - Departamento de Engenharia Elétrica

REFERÊNCIA BIBLIOGRÁFICA

IAN ALVES ULIAN (2016) Detecção e reconhecimento de números seriais em cédulas da segunda família do Real. Dissertação de Mestrado em INFORMÁTICA FORENSE E SEGURANÇA DA INFORMAÇÃO, Publicação 633/2016, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 61p.

CESSÃO DE DIREITOS

AUTOR: IAN ALVES ULIAN

TÍTULO: Detecção e reconhecimento de números seriais em cédulas da segunda família do Real.

GRAU: Mestre ANO: 2016

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de Mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. Do mesmo modo, a Universidade de Brasília tem permissão para divulgar este documento em biblioteca virtual, em formato que permita o acesso via redes de comunicação e a reprodução de cópias, desde que protegida a integridade do conteúdo dessas cópias e proibido o acesso a partes isoladas desse conteúdo. O autor se reserva a outros direitos de publicação e nenhuma parte desta dissertação de Mestrado pode ser reproduzida sem a autorização por escrito do autor.

IAN ALVES ULIAN

SQS 310 BLOCO J APARTAMENTO 101, ASA SUL, BRASÍLIA, DF

Agradecimentos

Em primeiro lugar, um agradecimento à Deus pela minha saúde, que me faz capaz de alcançar perseverança, dedicação, serenidade, paciência e sabedoria.

Gratidão também não pode faltar à minha família, capitaneada pelo meu pai, o Paulo Show, o homem mais louco, focado, perseverante e dedicado que já conheci; e pela minha mãe, Simone, mulher que abriu mão de seus inúmeros sonhos, interesses e desejos em prol dos seus três filhos. Aos meus irmãos, Samanta e Ravi, o meu abraço mais apertado, vocês são a companhia mais agradável que posso ter. Ao meu sobrinho Lorenzo, desejo-te o mundo, você é a maior alegria da nossa família.

Um agradecimento especial ao meu avô, Sebastião Alves Pereira, o meu maior exemplo de simplicidade, bom humor, coração bom e trabalho. Apesar da sua ausência desde que iniciei o meu projeto de ser Perito Criminal Federal, os bons momentos vividos ao seu lado me trazem extrema felicidade quando lembrados.

Ao Flávio de Barros Vidal, meu orientador, um muito obrigado. Obrigado pelo profissionalismo, pela prontidão em ajudar, pelo entusiasmo nas discussões, pela disponibilidade frequente, pelas respostas rápidas e pelo caminho bem orientado.

Agradeço também à Polícia Federal e à Universidade de Brasília pela oportunidade de ingressar no programa de Mestrado que deu origem a este trabalho.

Por fim, agradeço pela companhia dos colegas de turma ao longo das aulas, momento em que aprendi muito com a presença de alguns pontos fora da curva.

RESUMO

Um sistema de monitoramento de cédulas de Real se mostra de extrema importância para questões criminais e econômicas. Nesse caso, faz-se necessário um eficiente método de detecção e reconhecimento de número serial de cédulas. Este trabalho apresenta uma proposta robusta para a utilização de um sistema de reconhecimento óptico de caracteres (OCR) em situações complexas de reconhecimento automático de números seriais da Segunda Família do Real. Todos os resultados são promissores, o que mostra ser factível a utilização da abordagem proposta em aplicações reais.

ABSTRACT

A tracking system for Real banknote shows extremely importance for criminal and economic purposes. In this case an efficient method for detecting and recognizing the banknote serial number is required. This work presents a robust proposal to use an Optical Character Recognition (OCR) system in complex situations in an automatic serial number recognition of the Second Series of Real. All shown results are promising, showing to be feasible the use of the proposed approach in real applications.

SUMÁRIO

| | | |
|----------|--|-----------|
| 1 | INTRODUÇÃO | 1 |
| 1.1 | MOTIVAÇÃO | 2 |
| 1.2 | OBJETIVOS..... | 4 |
| 1.3 | ESTRUTURA DO TRABALHO | 5 |
| 2 | REFERENCIAL TEÓRICO | 6 |
| 2.1 | NORMATIVOS REFERENTES A CÉDULAS DE REAL | 7 |
| 2.2 | SEGUNDA FAMÍLIA DO REAL | 8 |
| 2.3 | SISTEMA DE MONITORAMENTO DE CÉDULAS | 13 |
| 2.4 | RECONHECIMENTO DE CARACTERES EM IMAGENS..... | 15 |
| 2.4.1 | PROBLEMAS EM LOCALIZAR E RECONHECER CARACTERES EM IMAGENS..... | 15 |
| 2.4.2 | PRINCIPAIS TÉCNICAS UTILIZADAS | 18 |
| 2.4.3 | ABORDAGENS PARA A LOCALIZAÇÃO | 29 |
| 2.5 | RECONHECIMENTO ÓTICO DE CARACTERES - <i>Tesseract</i> | 30 |
| 2.6 | TRABALHOS CORRELATOS | 32 |
| 3 | METODOLOGIA PROPOSTA..... | 34 |
| 3.1 | ELABORAÇÃO DO BANCO DE IMAGENS | 35 |
| 3.2 | CENÁRIOS INICIAIS | 36 |
| 3.3 | DETECÇÃO DOS NÚMEROS SERIAIS..... | 39 |
| 3.3.1 | DETERMINAÇÃO DA INCLINAÇÃO | 39 |
| 3.3.2 | PROVÁVEIS REGIÕES COM NÚMERO SERIAL..... | 41 |
| 3.3.3 | VALIDAÇÃO DAS REGIÕES COM CARACTERES DO NÚMERO SERIAL.... | 44 |
| 3.4 | RECONHECIMENTO DOS CARACTERES | 45 |
| 4 | RESULTADOS EXPERIMENTAIS | 48 |
| 4.1 | METODOLOGIA DE AVALIAÇÃO..... | 49 |
| 4.2 | DETERMINAÇÃO DA INCLINAÇÃO | 49 |
| 4.3 | LOCALIZAÇÃO DOS NÚMEROS SERIAIS | 50 |
| 4.4 | AVALIAÇÃO GLOBAL DO SISTEMA | 51 |
| 5 | CONCLUSÕES E TRABALHOS FUTUROS..... | 55 |

REFERÊNCIAS BIBLIOGRÁFICAS..... 57

LISTA DE FIGURAS

| | | |
|------|--|----|
| 1.1 | Localização de alguns elementos de segurança da cédula de R\$ 100 da Segunda Família do Real. Imagens retiradas de https://www.bcb.gov.br/novasnotas/nota-100-reais.html | 4 |
| 2.1 | Elementos visuais de segurança das cédulas da Segunda Família do Real, imagens retiradas de [1]. | 10 |
| 2.2 | Elementos táteis de segurança das cédulas da Segunda Família do Real. Imagens retiradas de [1]. | 11 |
| 2.3 | Elementos de segurança ocultos das cédulas da Segunda Família do Real. | 12 |
| 2.4 | Figura em que se pode ver ambas as numerações da cédula | 13 |
| 2.5 | Imagens ilustrativas dos contratemplos enfrentados na tarefa de localização e reconhecimento de caracteres em imagens. | 16 |
| 2.6 | Imagem que ilustra os pontos característicos robustos obtidos com o algoritmo SURF. | 17 |
| 2.7 | Imagens de cédulas que ilustram a variação de escala e de rotação de caracteres. | 17 |
| 2.8 | Destaque para o complexo fundo de uma cédula de Real. | 18 |
| 2.9 | Ilustração da conversão de uma imagem colorida para uma em escala de cinza. | 20 |
| 2.10 | Ilustração da conversão de uma imagem colorida para uma em escala de cinza. | 22 |
| 2.11 | | 24 |
| 2.12 | Ilustração da aplicação do detector de bordas de Canny a uma imagem. | 26 |
| 2.13 | Os quatro pontos colineares em (a) são mapeados em quatro retas que se cruzam no mesmo ponto no espaço de parâmetros em (b). | 27 |
| 2.14 | Ilustração de uma imagem no espaço de Hough. | 28 |
| 2.15 | Fluxograma de funcionamento do <i>Tesseract</i> , adaptado de [2]. | 31 |
| 2.16 | Tela de comando com as opções do <i>Tesseract</i> | 32 |
| 3.1 | Fluxograma da metodologia proposta do trabalho. | 35 |
| 3.2 | Amostra do banco de imagens criado. | 37 |
| 3.3 | Exemplos de imagens utilizadas no cenários iniciais | 38 |
| 3.4 | Ilustração dos passos para determinar as retas. | 40 |
| 3.5 | Ilustração dos passos para determinação do ângulo. | 41 |
| 3.6 | Fluxograma para a localização prováveis regiões com número serial. | 42 |
| 3.7 | Ilustração de imagens obtidas com o fluxograma ilustrado pela Figura 3.6. | 43 |
| 3.8 | Ilustração dos passos para a validação das 11 regiões. | 46 |

| | | |
|-----|---|----|
| 3.9 | Exemplos de imagens de entrada do <i>Tesseract</i> | 46 |
| 4.1 | Exemplo de erro obtido na estimativa do ângulo de inclinação da cédula na imagem. | 50 |
| 4.2 | Ilustração de imagem em que não se consegue localizar um número serial, por conta da sombra de uma cédula na outra. | 51 |
| 4.3 | Exemplos de entrada e saída para o Cenário de Referência. | 53 |
| 4.4 | A cadeia reconhecida posta à direita do número serial difere por apenas um caractere dele, há um \ no lugar de um 8. | 54 |

LISTA DE TABELAS

| | | |
|-----|---|----|
| 2.1 | Dimensões das cédulas da Segunda Família do Real | 9 |
| 3.1 | Configurações da câmera utilizada para construção do banco de imagens. | 36 |
| 4.1 | Resultados da etapa de determinação da inclinação das cédulas..... | 49 |
| 4.2 | Resultados da etapa de localização dos números seriais | 51 |
| 4.3 | Resultados para mensuração por cadeia | 53 |
| 4.4 | Resultados para mensuração por caractere | 53 |

LISTA DE TERMOS E SIGLAS

| | |
|--------|--|
| BACEN | Banco Central do Brasil |
| IPED | Indexador e Processador de Evidências Digitais |
| MPF | Ministério Público Federal |
| MSER | Maximally Stable Extremal Regions |
| OCR | Optical Character Recognition - Reconhecimento Ótico de Caracteres |
| PF | Polícia Federal |
| SENASP | Secretaria Nacional de Segurança Pública |
| SIFT | Scale Invariant Feature Transform |
| SURF | Speed Up Robust Features |

Capítulo 1

Introdução

Ao se analisar a história, percebe-se que um dos maiores impulsionadores de avanços tecnológicos foram as diversas guerras existentes mundo afora. Como exemplos, podem ser citadas as bombas atômicas da Segunda Guerra Mundial e a máquina Enigma, máquina criptográfica utilizada para transmissão de informações sensíveis durante guerras. Fazendo uma extensão desse pensamento, pode-se dizer também que o combate à criminalidade é um combustível para o desenvolvimento da tecnologia.

Nesse contexto, estão inseridos os órgãos de segurança pública, que, por serem responsáveis por combater o crime, devem instituir ações para tal fim. Uma dessas ações se traduz na capacitação do seu corpo funcional, tanto é que a Secretaria Nacional de Segurança Pública (SENASP) possui uma Rede Nacional de Ensino a Distância - Rede EAD - SENASP [3].

Na esteira do que foi dito e frisando que este trabalho é uma consequência de uma política de capacitação de um órgão de segurança pública ¹, as suas motivações incluem, dentre outras, a utilização de tecnologias relacionadas à área de visão computacional na área forense com a finalidade de se otimizar e tornar mais eficiente os trabalhos de investigação e combate ao crime organizado, bem como auxiliar em alguns aspectos atrelados à economia.

Mais especificamente, a principal motivação é a necessidade de determinação da rota de cédulas de Real no território Brasileiro e, como consequência, as implicações que isso teria para a persecução criminal e para a economia.

Partindo disso, o objetivo do trabalho é propor uma primeira abordagem de aplicação de visão computacional na leitura de números seriais de cédulas de Real. Para a sua consecução, propõe-se um conjunto de ações com a finalidade de que números seriais de cédulas de Real da Segunda Família sejam reconhecidos de forma automática a partir de imagens de um banco de imagens criado para tal fim. Deve-se observar que a precisão na execução dessa tarefa sofre influência de alguns alguns contratemplos, como, por exemplo, variação de

¹Termo de Cooperação 003/2013, estabelecido entre a SENASP e a Polícia Federal (PF), publicado no Diário Oficial da União em 13/01/2014, cujo objeto é: Curso de mestrado profissional em engenharia elétrica com ênfase em informática forense e segurança da informação para servidores da Polícia Federal e das Polícias Civis dos Estados

orientação e de escala. Sendo assim, é um objetivo também a mensuração da precisão da proposta diante dos cenários propostos.

Este capítulo traz, na ordem que segue, as motivações e objetivos do trabalho, bem como a forma como este texto está organizado.

1.1 Motivação

O chamado crime do colarinho branco, que pode ser entendido como um crime cometido por uma pessoa respeitável e de alta posição social de Estado, no exercício de suas ocupações [4], tem tido destaque nos últimos anos devido a uma série de casos descobertos. Essas descobertas são fruto, dentre outros detalhes, da evolução dos métodos de investigação, o que engloba a tecnologia utilizada.

Tendo isso em mente, o desenvolvimento de ferramentas de computação para o combate ao crime é uma atividade motivadora para os afetos à área computacional e criminal.

No âmbito do Departamento de Polícia Federal, podem ser citados alguns trabalhos desenvolvidos por Peritos Criminais Federais que se utilizam da tecnologia para otimizar procedimentos periciais e investigativos, como, por exemplo, NuDetective [5] e IPED². Estes foram desenvolvidos para atender a demandas pontuais, mas, por sua relevância prática, acabaram transcendendo as fronteiras dos estados.

Nesse sentido, a visão computacional, por conta do seu avanço tecnológico, tanto a nível de hardware bem como de software, também pode ser utilizada com o intuito de se criar um produto útil tanto para a justiça quanto para a economia.

Dando maior enfoque ao contexto da justiça, nos últimos anos, uma das maiores operações de combate à corrupção e lavagem de dinheiro tem tido desdobramentos no Brasil, trata-se da denominada Operação Lava Jato. É impressionante os números de tal atividade, como, por exemplo, 1.323 procedimentos instaurados, 666 buscas e apreensões, 181 conduções coercitivas, 76 prisões preventivas, 92 prisões temporárias, 6 prisões em flagrante [6]. Ainda segundo a mesma fonte, "Os crimes já denunciados envolvem pagamento de propina de cerca de R\$ 6,4 bilhões".

Por diversas vezes, o pagamento de propina é realizado mediante cédulas com o intuito de não se deixar nenhum rastro como ocorre em uma transação bancária. Mais do que isso, nesses números estão incluídas as cédulas apreendidas em buscas e apreensões da Polícia Federal. Observa-se, todavia, que não se consegue determinar qual a origem dessas cédulas, impossibilitando, por exemplo, a determinação do vínculo entre alguns investigados.

Essa constatação é fortalecida por um trecho da denúncia do Ministério Público Federal

²Indexador e Processador de Evidências Digitais (IPED) : é um sistema de indexação e processamento das evidências contidas em mídias digitais, que torna possível a identificação de arquivos definidos por palavras-chave de pesquisa

(MPF) relativa à Operação Saqueador [7]: *"Estes núcleos da organização seguiam o mesmo modus operandi: recebiam pagamento por serviços fictícios de "sociedades laranjas", dissimulando, além de desvios de recursos para obras públicas, o pagamento de propina a agentes estatais em espécie, interrompendo o rastreamento das verbas". Essa sistemática tem como objetivo principal justamente **apagar o rastro do dinheiro sujo**, uma vez que as "empresas laranjas receptoras" são criadas com uma finalidade precípua: propiciar o saque em espécie dos recursos desviados, de maneira a não só dissociá-los de sua origem ilícita e pública, mas também ocultar os seus reais beneficiários – notadamente agentes políticos e públicos -, que recebem o numerário em cash, à margem do sistema bancário oficial. Tal modalidade de lavagem de dinheiro já era conhecida, tendo se tornado ainda mais notória com a deflagração da denominada "Operação Lava-jato". (grifo meu)*

Mais do que isso, em se tratando de economia, pode-se pensar na determinação do fluxo de cédulas no território nacional e seus desdobramentos, como, por exemplo, os motivos do fluxo de cédulas de um estado para outro e, ainda, os motivos de uma inversão desse fluxo em determinada época. Pode-se também mensurar a durabilidade de uma cédula colocada em circulação, o que é um dado importante tanto para o planejamento de produção de cédulas como para o planejamento de distribuição delas no território nacional, evitando, por exemplo, a falta de cédulas em determinado local.

Outro ponto a ser observado é que, dentre os diversos alvos das forças policiais e de órgãos de controle da economia, a fraude de cédulas de Real tem grande importância, não só criminalmente, mas também economicamente. De acordo com [8], no ano de 2015 foram retidas pelo Banco Central do Brasil mais de 400.000 notas falsas, sendo mais de 150.000 somente no estado de São Paulo.

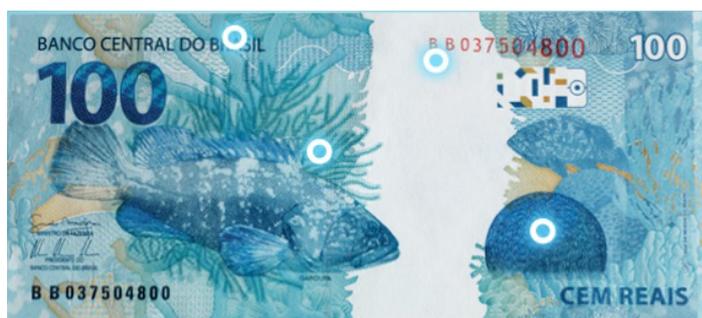
Por força de lei ³, a cédula de Real deve ser posta em circulação no Brasil e, com a finalidade de se prevenir sua falsificação, esta é confeccionada com diversos elementos de segurança, tais como: número serial, marca d'água, registro coincidente, imagem latente, entre outros [1]. A Figura 1.1 mostra a localização de alguns dos elementos de segurança existentes em uma cédula de cem reais da Segunda Família do Real.

Segundo [1], outra característica dela é a unicidade do número serial: não são postas em circulação pela Casa da Moeda duas cédulas com o mesmo número serial. Essa é uma peculiaridade que pode ser explorada para alguns fins além de prevenção e de detecção de falsificação. A partir disso, um sistema de monitoramento de cédulas baseado na detecção e reconhecimento do número serial, podendo ser utilizado, como já citado, tanto para fins econômicos quanto criminais, faz-se de fundamental importância.

³Lei nº 4.595, de 31 de dezembro 1964.



(a) Anverso da cédula de R\$ 100 da Segunda Família.



(b) Reverso da cédula de R\$ 100 da Segunda Família.

Figura 1.1: Localização de alguns elementos de segurança da cédula de R\$ 100 da Segunda Família do Real. Imagens retiradas de <https://www.bcb.gov.br/novasnotas/nota-100-reais.html>.

1.2 Objetivos

Para permitir essa concepção do sistema e a viabilidade de sua implementação em um futuro próximo, porém, o número serial deve ser alvo de um eficiente método de detecção e reconhecimento óptico de caracteres (OCR) para permitir seu rastreamento.

Dessa necessidade, extrai-se o objetivo central deste trabalho: a proposição de um eficiente método de detecção e para auxiliar no reconhecimento óptico de caracteres (OCR) para o número serial de cédulas da Segunda Família do Real.

O objetivo também pode ser encarado como uma primeira proposição para as cédulas de Real, visto que na literatura só foram encontrados trabalhos para outras moedas, como, por exemplo, Dólar [9, 10], Renminbi [9, 11, 12, 13] e Euro [9].

Para a consecução desse objetivo central, outros devem ser alcançados. O primeiro deles é a criação de um banco de imagem de cédulas da Segunda Família do Real, insumo necessário para a aplicação, análise e avaliação das técnicas propostas.

Mais do que isso, uma vez criado, o banco fica disponível para testes e trabalhos futuros, tornando-se uma contribuição relevante para a evolução das propostas.

Além disso, técnicas de OCR atingiram o grau de maturidade, sendo possível serem encontradas, atualmente, em vários sistemas industriais e comerciais. Mais do que isso, diversas abordagens promissoras [14, 15, 16] estão em constante desenvolvimento. Mesmo

com essa evolução, existem características que ainda são desafiadoras para a comunidade científica, como, por exemplo, efeitos causados pelo não-alinhamento devido à rotação, alteração de escala e influência do *Background*, que em muito comprometem a qualidade e desempenho desse tipo de sistema [17].

Sendo assim, este trabalho também tem como meta uma proposta de solução robusta que permite a utilização de pré-processamento nas imagens para auxiliar dispositivos OCRs em situações com rotação e variação de escala para a detecção e o reconhecimento automático de caracteres do número serial de cédulas de Real, usando a nova família. Ressalta-se que o alvo é o número serial localizado no canto inferior esquerdo do anverso de uma cédula da Segunda Família do Real 3.2.

Diferentemente de alguns trabalhos em que o foco está no desenvolvimento de técnicas de OCR [18, 19, 20, 21], ressalta-se que o objetivo aqui não inclui esta proposição. Em consonância com as propostas do estado da arte [14], utiliza-se uma ferramenta pronta, a amplamente difundida *Tesseract* [22, 23]. Esta é atualmente a mais poderosa e disponível implementação gratuita para aplicação de técnicas OCR. Mesmo sendo uma das melhores e mais utilizadas, ainda possui suas limitações, principalmente em imagens com *Background* que não se difere do texto e com textos desalinhados, por exemplo.

1.3 Estrutura do Trabalho

Por fim, a estrutura deste trabalho está disposta da seguinte forma: o referencial teórico a respeito do trabalho é apresentado no Capítulo 2. No Capítulo 3, a metodologia utilizada é detalhada, enquanto que, no Capítulo 4, são apresentados os resultados experimentais obtidos e as métricas de validação utilizadas. Por fim, no Capítulo 5, são apresentados as conclusões e os trabalhos futuros.

Capítulo 2

Referencial Teórico

Por mais que outros meios de utilização de dinheiro tenham se popularizado nas últimas décadas, como, por exemplo, cartão de crédito e moeda eletrônica [24], a utilização das cédulas ainda encontra espaço no cotidiano das sociedades. Mais do que isso, a circulação de papel moeda no Brasil é forçada por Lei [25].

Ainda, segundo [1], o dinheiro em espécie é um meio de pagamento conveniente, seguro e imediato. Contudo, sua utilização requer a verificação dos seus elementos de segurança.

O Dólar, por exemplo, por ser a moeda mais falsificada do mundo, prescinde de elementos de segurança bem confeccionados. Por isso, possuem fibras coloridas, impressão calcográfica, impressão tipográfica nos selos, numerações, entre outros [26].

O Euro, por ser a moeda de inúmeros países, também merece atenção especial quanto às técnicas de prevenção de contrafação. Em comum todas as cédulas de Euro apresentam papel de segurança, marca d'água, fio de segurança, impressão em calcografia, entre outros [27].

Ainda de acordo com [1], as cédulas da Segunda Família do Real, quando comparadas às da Primeira Família, contam com recursos gráficos mais sofisticados, capazes de impor obstáculos mais sólidos às tentativas de falsificação. Seus elementos de segurança ficaram mais seguros e mais fáceis de verificar.

Um desses elementos é a chave para a proposta deste trabalho, o número serial. Este é único para cada cédula e vice-versa, o que a identifica de forma unívoca.

Partindo dessa premissa, entende-se ser possível o monitoramento e rastreamento de cédulas baseado nos números seriais. A seguir, mais especificamente na seção 2.3, é dada uma sugestão para um sistema de monitoramento de cédulas, que possui aplicações tanto para a justiça como para a economia.

De antemão, ressalta-se que uma das principais funcionalidades desse sistema é o modo como o número serial é passado e identificado por ele, então propõe-se a utilização de reconhecimento ótico de caracteres (OCR). Pontua-se que a utilização de OCR aqui é em sentido

amplo, isto é, engloba, além da tarefa de atribuir a um elemento de uma imagem um caractere, a localização de tal elemento na imagem, ou seja, realizar a tarefa de detecção e reconhecimento da cadeia de caracteres que formam o número serial.

Como em todas aplicações de recursos tecnológicos, há desafios a serem superados. Nesse sentido, a aplicação de um OCR a imagens de cédulas de Real traz alguns desafios, como, por exemplo, a localização do número serial na imagem; a forma como lidar com o fundo das imagens, visto que as cédulas possuem vários detalhes e artefatos de segurança, entre outros elementos, que dificultam todo o processo a ser realizado pelo OCR.

Para abordar de forma mais aprofundada os assuntos tratados acima, este capítulo está organizado da seguinte forma. A Seção 2.1 traz alguns normativos relacionados às cédulas de Real; a Seção 2.2 aborda as principais características das cédulas da Segunda Família do Real; a Seção 2.3 descreve uma proposta para um sistema de monitoramento de cédulas; a Seção 2.4 trata de OCR, sobretudo de seus desafios. Já a Seção 2.5 traz alguns detalhes da ferramenta utilizada no trabalho para o reconhecimento de caracteres propriamente dito. Por fim, na Seção 2.6 são citados alguns trabalhos correlatos ao tema abordado aqui.

2.1 Normativos Referentes a cédulas de Real

Como já comentado, a circulação de cédulas no Brasil é forçada por lei. A Lei nº 4.595 [25], de 31 de dezembro 1964, preconiza, em seu artigo 4º, inciso II, que:

"Art. 4º Compete ao Conselho Monetário Nacional, segundo diretrizes estabelecidas pelo Presidente da República:

*II - Estabelecer condições para que o Banco Central da República do Brasil **emita moeda-papel de curso forçado**, nos termos e limites decorrentes desta Lei, bem como as normas reguladoras do meio circulante." (grifo meu).*

Nota-se, portanto, a exigência explícita da circulação de papel moeda em território nacional. Entende-se que isso também é uma entrave à sua extinção.

Mais do que isso, consoante a Lei das Contravenções Penais, Decreto-Lei nº 3.688, de 3 de outubro de 1941, a recusa em receber moeda corrente do país é uma contravenção penal: *"Art. 43 - Recusar-se a receber, pelo seu valor, moeda de curso legal no país; Pena: multa."*

Indo mais a fundo na esfera penal, conforme o Código Penal, Decreto Lei nº 2.848, de 7 de dezembro de 1940, tem-se as tipificações dos atos relacionados às falsificações de moedas.

"Art. 289 – Falsificar, fabricando-a ou alterando-a, moeda metálica ou papel moeda de curso legal no País ou no estrangeiro:

Pena: reclusão, de três a doze anos e multa.

§1º – Nas mesmas penas incorre quem, por conta própria ou alheia, importa ou exporta,

adquire, vende, troca, cede, empresta, guarda ou introduz na circulação moeda falsa.

§2º – Quem, tendo recebido de boa-fé, como verdadeira, moeda falsa ou alterada, a restitui à circulação, depois de conhecer a falsidade, é punido com detenção, de seis meses a dois anos, e multa.

Art. 290 – Formar cédula, nota ou bilhete representativo de moeda com fragmentos de cédulas, notas ou bilhetes verdadeiros; suprimir, em nota, cédula ou bilhete recolhidos, para o fim de restituí-los à circulação, sinal indicativo de sua inutilização; restituir à circulação cédula, nota ou bilhete em tais condições, ou já recolhidos para fim de inutilização:

Pena: reclusão, de dois a oito anos, e multa.

Parágrafo único - O máximo da reclusão é elevado a doze anos e multa, se o crime é cometido por funcionário que trabalha na repartição onde o dinheiro se achava recolhido, ou nela tem fácil ingresso, em razão do cargo.

Art. 291 – Fabricar, adquirir, fornecer, a título oneroso ou gratuito, possuir ou guardar maquinismo, aparelho, instrumento ou qualquer objeto especialmente destinado à falsificação de moeda:

Pena: reclusão, de dois a seis anos, e multa.

Art. 292 – Emitir, sem permissão legal, nota, bilhete, ficha, vale ou título que contenha promessa de pagamento em dinheiro ao portador ou a que falte indicação do nome da pessoa a quem deva ser pago:

Pena: detenção, de um a seis meses, ou multa.

Parágrafo único: Quem recebe ou utiliza como dinheiro qualquer dos documentos referidos neste artigo incorre na pena de detenção, de quinze dias a três meses, ou multa.

Apesar de, como abordado acima, a falsificação de cédulas ser crime, é permitido que aquelas detectadas como falsas sejam devolvidas aos bancos, que as encaminharão ao Banco Central.

Contudo, as que são encaminhadas às polícias devem passar pelo crivo de Peritos Criminais Federais, tornando-se alvo de um inquérito caso sejam apontadas como falsas.

2.2 Segunda Família do Real

O Real passou a ser a moeda corrente no Brasil a partir de 1994, com a edição do Plano Real [28]. O órgão responsável por sua emissão é o Banco Central do Brasil (BACEN), sendo que atualmente todas as cédulas e moedas metálicas de Real são produzidas pela Casa da Moeda do Brasil (CMB).

De acordo com [1], as denominações das cédulas de Real em circulação são: R\$ 1, R\$ 2 (a partir de outubro de 2001), R\$ 5, R\$ 10, R\$ 20 (a partir de junho de 2002), R\$ 50 e R\$

100.

Ressalta-se que a cédula de R\$ 1 já não é mais posta em circulação, visto que pertence à Primeira Família do Real, que, apesar de ainda ter valor, não é mais confeccionada pelos órgãos competentes.

Inclusive, as cédulas da Primeira Família já incluíam alguns elementos de segurança, tais como, fibras coloridas, marca d'água, fio de segurança, imagem latente, marca tátil, microimpressões, entre outros. Contudo, mesmo já tendo alguns elementos de segurança, por conta do avanço das tecnologias digitais, a facilidade de contrafação aumentou.

Por conta disso, em 2010, o BACEN colocou em circulação as primeiras cédulas da Segunda Família do Real, nos valores de R\$ 50 e R\$ 100 [29]. Em 2012, foram lançadas as cédulas de R\$ 20 e R\$ 10. Por fim, em 2013, foram inseridas no mercado as cédulas de R\$ 2 e R\$ 5.

De forma bem evidente e já com a intenção de melhora nas suas funcionalidades, uma das características que diferencia a Segunda da Primeira família é que as cédulas daquela possuem dimensões de acordo com a denominação, enquanto que as desta possuem dimensões iguais, de 140 x 65 mm. A Tabela 2.1 lista as dimensões das cédulas da Segunda Família do Real, bem como outras aspectos que as diferenciam, tais como, a cor predominante e o tema.

Tabela 2.1: Dimensões das cédulas da Segunda Família do Real

| Denominação (R\$) | Dimensões (mm) | Cor predominante | Tema |
|-------------------|----------------|------------------|-------------------|
| 2 | 121 x 65 | Azul marinho | Tartaruga marinha |
| 5 | 128 x 65 | Lilás | Garça |
| 10 | 135 x 65 | Vermelho | Arara vermelha |
| 20 | 142 x 65 | Amarelo | Mico leão dourado |
| 50 | 149 x 70 | Bege | Onça pintada |
| 100 | 156 x 70 | Azul turquesa | Garoupa |

Além disso, todas as cédulas da Segunda Família possuem papel de segurança 100% algodão; são impressas em calcografia, ofsete e tipografia; apresentam microimpressões, marca d'água, registro coincidente, imagem latente e elementos fluorescentes, tais como, fibras invisíveis e fluorescência na numeração do reverso e no número indicativo do valor no anverso da cédula).

Além desses elementos de segurança, que são comuns a todas as denominações, foram criados elementos para diferenciação das baixas, médias e altas denominações. Abaixo seguem as diferenças [1]

- R\$ 2 e R\$5: não há fio de segurança e há impressão calcográfica somente no anverso. Por conta de serem as de menor valor, são mais manuseadas e possuem maior circulação, possuem revestimento especial de uma camada protetora de verniz, transparente e imperceptível ao tato, nas duas faces;

- R\$10 e R\$20: fio de segurança, impressão com tinta de variação óptica. A cédula de R\$10 possui impressão calcográfica somente no anverso;
- R\$50 e R\$100: fio de segurança, faixa holográfica, impressão calcográfica no anverso e reverso.

A título de ilustração, a seguir, seguem imagens de alguns elementos de segurança. Para ser mais didático, as ilustrações serão divididas em elementos visuais, táteis e ocultos.

A Figura 2.1 mostra os elementos visuais, são eles: marca d'água, fio de segurança e registro coincidente. Sugere-se que se observe a cédula contra a luz para maior facilidade de visualização de tais elementos.



(a) Marca d'água da cédula de R\$2, tartaruga marinha e número 2.
 (b) Fio de segurança da cédula de R\$50.



(c) Registro coincidente da cédula de R\$5.

Figura 2.1: Elementos visuais de segurança das cédulas da Segunda Família do Real, imagens retiradas de [1].

A Figura 2.2 ilustra alguns elementos de segurança táteis, ou seja, partes das cédulas que são confeccionadas em alto relevo. Destaca-se que todas as cédulas possuem relevo no anverso, mas apenas as de 20, 50 e 100 apresentam alto relevo também no reverso.

Para finalizar a ilustração dos elementos de segurança, a Figura 2.3 evidencia os elementos de segurança ocultos, são eles: faixa holográfica (somente nas denominações de 50 e 100), número que muda de cor (apenas nas denominações de 10 e 20), imagem latente, microimpressões e imagem latente.

Destaca-se que, para a visualização da faixa holográfica e do número que muda de cor, deve-se movimentar as cédulas; para a percepção da imagem latente, deve-se colocar a cé-



(a) Alto relevo na legenda "BANCO CENTRAL DO BRASIL" da cédula de R\$20.



(b) Alto relevo na legenda "REPÚBLICA FEDERATIVA DO BRASIL" da cédula de R\$50.



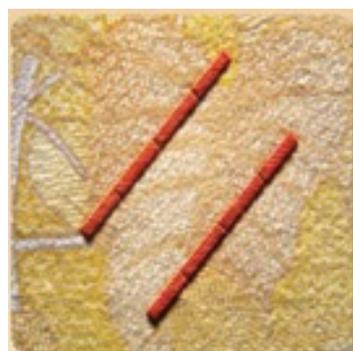
(c) Alto relevo na figura do animal da cédula de R\$100.



(d) Alto relevo na efigie da cédula de R\$10.



(e) Alto relevo nas extremidades laterais da cédula R\$50.



(f) Alto relevo na marca tátil da cédula de R\$20.



(g) Alto relevo nos numerais que indicam o valor da nota de R\$100.



(h) Alto relevo nos numerais que indicam o valor da nota de R\$50.

Figura 2.2: Elementos táteis de segurança das cédulas da Segunda Família do Real. Imagens retiradas de [1].

dula na horizontal na altura dos olhos; para se notar as microimpressões, é prudente o uso de uma lente de aumento; para a visualização dos elementos fluorescentes, faz-se necessário luz ultravioleta.



(a) Indicação da região que possui a faixa holográfica na cédula de R\$50, na de R\$100, o local é o mesmo.



(b) Ilustração das mudanças de cor no número em uma cédula de R\$10.



(c) Imagem latente em uma cédula de R\$5



(d) Ilustração de uma das regiões com microimpressões em uma cédula de R\$50.



(e) Ilustração dos elementos fluorescentes em uma cédula de R\$2, número indicativo do valor, numeração de série e pequenos fios espalhados pela nota.

Figura 2.3: Elementos de segurança ocultos das cédulas da Segunda Família do Real.

É interessante notar que, para difundir o conhecimento dos elementos de segurança das cédulas de Real, o BACEN criou um aplicativo para *smartphone*. Este, ao ter uma nota sob foco, identifica a denominação da cédula e mostra os elementos de segurança que devem ser analisados. O aplicativo é gratuito e mais informações sobre ele podem ser encontradas em [30].

Por fim, atenção será dada à numeração das cédulas, que também pode ser considerada como um elemento de segurança. Na primeira geração do Real, as cédulas contavam com a impressão da numeração apenas um única vez, já na segunda, há dois locais em que o número é impresso, ambos no reverso. A primeira impressão está em azul, no canto inferior esquerdo; a segunda, em vermelho, na parte superior direita, com caracteres crescentes. A Figura 2.4 ilustra o reverso de uma cédula de R\$2, em que é possível visualizar os seus números seriais.



Figura 2.4: Figura em que se pode ver ambas as numerações da cédula

Pontua-se que, para este trabalho, a impressão de interesse é aquela em azul, localizada no canto inferior esquerdo do reverso das cédulas.

2.3 Sistema de Monitoramento de Cédulas

Seja a seguinte situação hipotética: Um delator de um esquema de corrupção relata que irá pagar propina a um político de renome. Para que isso seja comprovado, surge a ideia de se registrar o número serial de cada uma das cédulas que será utilizada para o tal pagamento. Feito isso, após a realização do pagamento, é feita uma busca e apreensão na residência do político e as cédulas cujos números foram anotados são encontradas. Logo, sabendo que o número serial identifica uma cédula de forma unívoca, está comprovado o repasse de propina.

Essa situação, ainda que hipotética, não está distante do que pode ocorrer na realidade. Além disso, revela uma das vantagens de se ter controle sobre os números seriais de algumas cédulas específicas.

Tomando essa situação como motivação, sugere-se um Sistema de Monitoramento de Cédulas da Segunda Família do Real. O sistema seria composto por Pontos de Leitura e por uma Central. Os Pontos de Leitura seriam dispositivos de leitura automática do número serial de cédulas de Real. Eles poderiam estar localizados, por exemplo, em:

- caixas de supermercados (ou estabelecimentos com grande circulação de cédulas);
- caixas eletrônicos de bancos;
- delegacias de polícia;
- aplicativo de celular;
- transportadora de valores.

Além de ler e anotar um número serial, os Pontos de Leitura o remeteriam, juntamente com data e localização, à Central, responsável por armazenar esses dados, ou seja, um banco de dados. As informações enviadas por um Ponto de Leitura localizado em um caixa eletrônico ainda poderiam incluir dados sobre o indivíduo que realizou a operação.

Pontua-se que a ideia não é monitorar toda e qualquer cédula de Real existente, o que, pela quantidade de dados a serem armazenados, poderia inviabilizar tal sistema, mas somente aquelas de interesse, tais como, as advindas de saques de grande vulto em um banco, ou ainda, cédulas que foram roubadas de uma transportadora de valores.

Abaixo seguem algumas das aplicações que são vislumbradas, possuindo importância tanto para a justiça quanto para a economia.

1. Noção do fluxo de cédulas em território brasileiro: até onde se consegue pesquisar, no Brasil, ainda não se possui informação a respeito do fluxo de cédulas em seu território. Ou seja, não se sabe qual o destino das cédulas que são colocadas em circulação, por exemplo, em Brasília;
2. Detecção de contrafação: caso duas cédulas com mesmo número serial sejam identificadas em tempos próximos e locais distantes, há um indicativo de que uma delas é falsa;
3. Pagamento de propina (ou de resgate) e identificação de roubos, furtos: situação já abordada na situação hipotética levantada acima, anota-se os números das cédulas antes do evento e, após o evento, faz-se a conferência;
4. Determinação de transferência de cédulas entre indivíduos: em uma busca e apreensão na residência do indivíduo A, apreende-se determinada quantidade de cédulas e, ao se consultar a Central, nota-se que grande parte dessas cédulas foram objeto de um saque em caixa eletrônico efetuado pelo indivíduo B. Essa informação pode ser de grande valia para a investigação.

Até onde foi possível se pesquisar, há sistemas análogos de governos na Europa e nos Estados Unidos da América (EUA). Contudo, como a informação a respeito de possíveis sistemas não está acessível de forma simples, não se sabe como eles funcionam.

De qualquer forma, de acordo com [31], o banco holandês DNB (*De Nederlandsche Bank*) lê números seriais de cédulas desde 1973 com os seguintes objetivos:

- detecção de contrafação;
- registro de cédulas utilizadas para pagamento de resgate de sequestro;
- garantia de qualidade;
- fraudes internas e externas;
- testes de circulação.

Além disso, para as seguintes moedas, existem algumas iniciativas isoladas e independentes que fazem algo no sentido da proposta apresentada aqui: Euro [32], Dólar [33], moeda canadense [34, 35].

2.4 Reconhecimento de Caracteres em Imagens

Tanto por parte da visão computacional como da comunidade de análise de documentos, o problema de reconhecimento de caracteres em imagens tem sido alvo de inúmeras pesquisas [14]. Há, inclusive, uma competição internacional que estimula o desenvolvimento de técnicas na área, ICDAR (*International Conference on Document Analysis and Recognition*) *Robust Reading Competition*.

Baseando-se no relatório da última edição da competição [15], pode-se dividir esse problema em duas ações: localizar e reconhecer os caracteres. Por conta disso, essas duas ações serão norteadoras dos problemas e soluções trazidos a seguir.

Na tentativa de expor as ideias de forma mais clara, esta seção será subdividida em três partes. Na primeira, são abordados alguns dos problemas encontrados na localização e no reconhecimento de caracteres em imagens, sobretudo aquelas que contém cédulas de Real. Na segunda, são citados alguns algoritmos de processamento de imagens que auxiliam no tratamento dos problemas. Na última, são evidenciadas algumas formas de solução para o problema da localização dos caracteres em uma imagem.

2.4.1 Problemas em Localizar e Reconhecer Caracteres em Imagens

No contexto de localização e reconhecimento de caracteres, há uma tarefa que atrai grande interesse dos pesquisadores, que é a leitura de placas de carro, visto que possui aplicações tanto para fins de segurança quanto de controle de tráfico. Para essa tarefa, alguns problemas são enfrentados, como, por exemplo, diversidade de padrões de caracteres (diferenças de fontes, tamanhos e cores), distorção de caracteres por conta do ângulo de captura da imagem e baixa qualidade da imagem, causada por, dentre outros aspectos, iluminação irregular, oclusão e borramento. O fundo também é um problema intrigante [36].

De um modo geral, os principais problemas a serem superados na tarefa em questão são fundo da imagem complexo, variação de iluminação, variação da fonte, variação da cor, alinhamento do texto, borramento [37]. Evidentemente, essa tarefa, quando utilizada para o reconhecimento de números seriais, também enfrenta os mesmos problemas [11].

A Figura 2.5 ilustra imagens utilizadas na ICDAR 2015 *Robust Reading Competition*, em que é possível observar algumas das dificuldades enfrentadas nas tarefas de localização e reconhecimento de caracteres em imagens.

Para este trabalho, faz-se importante destacar dois dos contrastes citados acima, a variação de escala (mudança no tamanho dos caracteres) e a rotação (inclinação dos caracteres em relação a uma linha horizontal de referência).

Esses dois contrastes são desafiadores não só para reconhecimento caracteres, mas também para o reconhecimento de padrões em imagens em geral, como, por exemplo, reconhecimento facial e reconhecimento de objetos.



(a) Imagem com caracteres em diferentes ângulos de captura, diferentes tamanhos e cores. (b) Imagem borrada com caracteres na vertical.



(c) Imagem borrada, com iluminação irregular e caracteres distorcidos. (d) Imagem com iluminação irregular e caracteres distorcidos.

Figura 2.5: Imagens ilustrativas dos contratemplos enfrentados na tarefa de localização e reconhecimento de caracteres em imagens.

São desafiadores, pois, geralmente, a imagem do padrão a ser reconhecido possui uma escala e uma orientação fixa, como ocorre, por exemplo, em bancos de dados usuais de impressões digitais. Assim, em uma imagem na qual o padrão esteja em outra escala e orientação, uma abordagem de *template matching* pura talvez não funcione. Seria necessário, por exemplo, a procura em várias escalas e orientações, o que aumentaria a complexidade computacional.

É por conta disso, inclusive, que alguns algoritmos, como, por exemplo, SIFT (*Scale Invariant Feature Transform*) [38] e SURF (*Speed Up Robust Features*) [39], buscam o reconhecimento de uma forma robusta a essas variações. Fazem isso ao determinar pontos característicos na imagem que são tão discriminantes a ponto de serem identificados mesmo com as variações abordadas aqui.

A Figura 2.6 mostra uma imagem que focaliza uma cédula de Real após a utilização do algoritmo SURF. Nota-se, em vermelho, os pontos característicos robustos encontrados na imagem.

Tomando como padrão a ser reconhecido uma cadeia de caracteres, a Figura 2.7 ilustra situações de variação de escala e de rotação de números seriais de cédulas da Segunda Família do Real.

Considerando ainda as imagens de cédulas de Real e o número serial, deve-se levar em consideração o fundo complexo da nota, com diferentes cores, texturas, linhas e desenhos.



Figura 2.6: Imagem que ilustra os pontos característicos robustos obtidos com o algoritmo SURF.



(a) Imagem para ilustrar a variação de escala em relação à imagem c)



(b) Imagem para ilustrar a variação de orientação em relação à imagem d)



(c) Imagem para ilustrar a variação de escala em relação à imagem a)



(d) Imagem para ilustrar a variação de orientação em relação à imagem b)

Figura 2.7: Imagens de cédulas que ilustram a variação de escala e de rotação de caracteres.

Isso é feito para dificultar a sua contrafação, dificultando também a tarefa de reconhecimento dos caracteres.

A Figura 2.8 mostra a região que engloba o número serial de uma cédula, em que é possível perceber a complexidade do fundo próximo a ele.



Figura 2.8: Destaque para o complexo fundo de uma cédula de Real.

2.4.2 Principais Técnicas Utilizadas

As técnicas expostas aqui são utilizadas no desenvolvimento da metodologia proposta neste trabalho. Contudo, antes de tratar propriamente delas, algumas definições, sobretudo das notações a serem utilizadas, são necessárias.

Primeiramente, apesar de a imagem vista pelo olho humano ser um sinal analógico no espaço, aqui, por ser um trabalho que envolve processamento computacional, as imagens são entendidas como digitais. Ou seja, são amostradas e quantizadas.

Sendo assim, uma imagem $I(m, n)$ é entendida como uma matriz bidimensional de valores discretos, em que m e n são as coordenadas discretas de um ponto da imagem. Estipula-se que $0 \leq m \leq M - 1$ e $0 \leq n \leq N - 1$, em que M e N são, respectivamente, o número de linhas e colunas da representação, ou seja, referem-se às dimensões da matriz.

A equação (2.1) mostra a forma genérica de uma matriz $I(m, n)$ que representa uma imagem.

$$I(m, n) = \begin{bmatrix} I(0, 0) & I(1, 0) & \cdots & I(M - 1, 0) \\ I(0, 1) & I(1, 1) & \cdots & I(M - 1, 1) \\ \vdots & \vdots & \ddots & \vdots \\ I(0, N - 1) & I(1, N - 1) & \cdots & I(M - 1, N - 1) \end{bmatrix} \quad (2.1)$$

Ressalta-se que essa representação pode ser interpretada espacialmente, ou seja, o elemento $I(m, n)$ é o elemento localizado nas coordenadas (m, n) . Além disso, cada elemento dessa matriz é denominado Pixel, acrônimo do inglês *Pixel Element*, em que sua intensidade em uma imagem monocromática é denominada de nível de cinza naquele ponto, podendo também ser chamada de imagem em tons de cinza.

Uma convenção adotada é que a cor preta corresponde ao nível de cinza mais escuro,

possuindo o valor zero, e a cor branca, ao nível de cinza mais claro, possuindo o valor 255, por exemplo, em uma imagem com 8 unidades de profundidade.

Há também as imagens coloridas, que podem seguir a mesma representação dada acima, mas para cada pixel da imagem, no espaço de cores RGB, haverá um vetor com três posições: (R, G, B), em que a posição R denota o peso dado à cor vermelha (*Red*); a G, dado à cor verde (*Green*); a B, dado à cor azul (*Blue*). Assim, pode-se representar cada pixel da imagem como um combinação linear dessas três cores bases.

Tendo estipulado algumas definições, pode-se seguir para as técnicas a serem abordadas.

2.4.2.1 Conversão da imagem para tons de cinza

A conversão de uma imagem colorida para tons de cinza é o primeiro passo de inúmeros algoritmos de análise de imagens, já que reduz a quantidade de informação de uma imagem [40]. Embora haja redução, a maioria das informações relacionadas às características da imagem se preservam, tais como, bordas, regiões, junções, entre outras [41].

Uma imagem no espaço de cores RGB, I_{RGB} , é convertida para uma em tons de cinza, I_{cinza} , por meio da transformação mostrada pela Equação (2.2):

$$I_{cinza}(m, n) = \alpha I_{RGB}(m, n, r) + \beta I_{RGB}(m, n, g) + \gamma I_{RGB}(m, n, b), \quad (2.2)$$

em que (m, n) são os índices de um pixel na imagem em escala de cinza, (m, n, c) diz respeito ao pixel com localização (m, n) no canal c da imagem colorida, sendo que c assume as variáveis r , g e b , que se referem, respectivamente, aos canais de cores vermelho, verde e azul.

Nota-se, portanto, que uma imagem em escala de cinza é uma combinação linear dos canais de cores de uma imagem no espaço de cores RGB. O peso dos coeficientes (α , β e γ) são atribuídos com a finalidade de que o olho humano perceba a imagem em tons de cinza da mesma forma que a imagem colorida. Logo, sabendo que esse órgão possui mais sensibilidade ao comprimento de luz na faixa da cor verde, o peso atribuído à β tende a ser maior que os outros. No padrão de televisão NTSC (*National Television System Committee*), por exemplo, $\alpha = 0,2989$, $\beta = 0,5870$ e $\gamma = 0,1140$. Já no padrão utilizado para a HDTV (*High Definition Television*), $\alpha = 0,2126$, $\beta = 0,7152$ e $\gamma = 0,0722$.

A Figura 2.9 ilustra o resultado da conversão de uma imagem colorida para uma em escala de cinza.

Por fim, faz-se importante salientar que a conversão de uma imagem em RGB para uma em escala de cinza é não inversível, ou seja, a informação perdida na transformação não pode ser recuperada.



(a) Imagem colorida.

(b) Imagem em escala de cinza.)

Figura 2.9: Ilustração da conversão de uma imagem colorida para uma em escala de cinza.

2.4.2.2 Binarização de Otsu

Podendo ser entendida como uma técnica de segmentação de imagens, a binarização divide os pixels de uma imagem qualquer em dois grupos, aqueles que serão pretos e aqueles que serão brancos, ou ainda, respectivamente, aqueles pixels que terão valor 0 e aqueles que terão valor 1 [40]. Dessa última atribuição que deriva o nome da técnica, ou seja, a imagem se torna uma imagem com números binários.

Para que isso seja realizado, é necessário a escolha de um limiar t (*threshold*), que será responsável por dividir a imagem nos dois grupos.

Portanto, partindo de uma imagem em escala de cinza, $I_{cinza}(m, n)$, chega-se em uma imagem em preto e branco, $I_{P\&B}(m, n)$, ou seja, binária, pela equação (2.3).

$$I_{P\&B}(m, n) = \begin{cases} 0, & \text{if } I_{cinza}(m, n) < t \\ 1, & \text{caso contrário} \end{cases} \quad (2.3)$$

Nota-se, portanto, que a escolha de t é um passo importante dessa técnica. Para isso, há alguns métodos de cálculo de forma a otimizar tal escolha. O método de Otsu [42], por exemplo, é um deles.

Ele visa dividir uma imagem que possui L níveis de cinza em duas classes C_0 e C_1 , uma que contará com os pixels pretos, a outra, com os brancos. Admitindo-se que essa divisão será estabelecida no nível de cinza t , tem-se que as duas classes são formadas pelos seguintes níveis de cinza:

$$\begin{aligned} C_0 &= \{0, 1, 2, \dots, t\}; \\ C_1 &= \{t + 1, t + 2, \dots, L\}. \end{aligned} \quad (2.4)$$

Com o intuito de se avaliar se a escolha pelo limiar é boa, são introduzidos os seguintes

critérios:

$$\zeta = \frac{\sigma_B^2}{\sigma_W^2}; \quad \eta = \frac{\sigma_B^2}{\sigma_T^2}; \quad \lambda = \frac{\sigma_T^2}{\sigma_W^2} \quad (2.5)$$

Sabendo que η é a medida mais simples para a otimização, atenção será dispensada somente a ela.

Sendo assim, faz-se necessário detalhar:

$$\begin{aligned} \sigma_B &= \omega_0 (\mu_0 - \mu_T)^2 + \omega_1 (\mu_1 - \mu_T)^2 = \omega_0 \omega_1 (\mu_1 - \mu_0); \\ \sigma_T^2 &= \sum_1^L (i - \mu_T)^2 p_i, \end{aligned} \quad (2.6)$$

em que ω se refere ao momento cumulativo de ordem zero das classes 0 e 1; μ , ao momento cumulativo de segunda ordem das classes 0 e 1 e da imagem como um todo (subíndice T); p_i é a probabilidade de pixel qualquer da imagem pertencer ao nível de cinza i . Essa probabilidade pode ser entendida como uma das componentes do histograma normalizado da imagem. É por isso que esse método também é conhecido como um método que seleciona o limiar automaticamente do histograma de níveis de cinza.

Em suma, a métrica η é calculada para todos os níveis possíveis da imagem, sendo que o t escolhido será aquele que a maximiza. Equivalentemente, é aquele que maximiza o valor de σ_B^2 , que consiste da variância entre as duas classes.

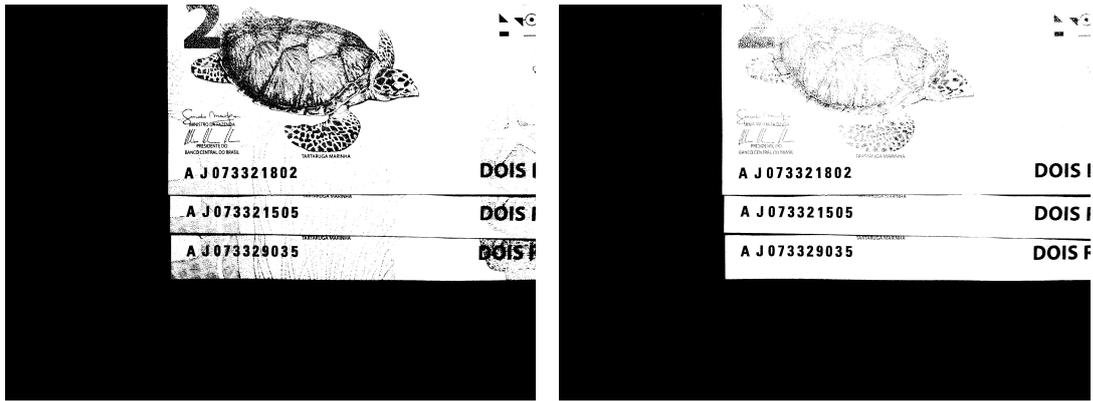
Apesar de envolver alguns cálculos, o método é simples, visto que são utilizados apenas os momentos cumulativos zero e de primeira ordem do histograma de níveis de cinza.

A Figura 2.10 mostra a imagem 2.9(b), em escala de cinza, binarizada com dois limiares, um escolhido arbitrariamente, outro calculado via método de Otsu. Nota-se que aquela imagem em que se utilizou o método ficou com os caracteres bem mais destacados enquanto que a região que contem a tartaruga perdeu destaque.

2.4.2.3 Filtragem

Dentro das técnicas de processamento de imagens, há aquelas que são denominadas técnicas de melhoramento ou realce de imagens [40]. Utilizando o termo melhoramento, pode-se dizer que se trata de algo atrelado ao aspecto visual, o que torna a sua avaliação um tanto quanto subjetiva.

No entanto, pela palavra realce, pode-se entender que as técnicas buscam acentuar a aparência de determinadas características da imagem, com o intuito de torná-la mais adequada à aplicação desejada. Sendo assim, deixa-se de lado um pouco da subjetividade. Um exemplo disso é a situação em que se deseja eliminar da imagem os pequenos detalhes de um desenho, ou melhor, quando se deseja, em uma imagem como a ilustrada pela Figura 2.7(b), eliminar



(a) Imagem binarizada com limiar 0,5.

(b) Imagem binarizada com limiar 0,3569, obtido com o método de Otsu.)

Figura 2.10: Ilustração da conversão de uma imagem colorida para uma em escala de cinza.

os detalhes do desenho da garça e, ao mesmo tempo, não eliminar as informações referentes aos caracteres.

Em sentido mais amplo, o realce pode ser necessário para imagens que passaram por um processo de degradação em virtude de introdução de ruído, perda de contraste, borrachamento, condição inadequada de iluminação e distorção oriunda do equipamento de captura.

As principais abordagens para tratar esses tipos de problemas envolvem métodos no domínio espacial e no domínio da frequência. Aqui, atenção será dada apenas ao domínio espacial, situação em que o processamento se dá diretamente no valor dos pixels.

Dentro da filtragem espacial, ainda existem as abordagens linear e não-linear. Aquela se caracteriza por atribuir um novo valor a cada pixel por meio da combinação linear dos seus pixels vizinhos; esta já requer uma combinação não linear.

Em relação à abordagem linear, a forma como ela se dá é determinada pela máscara do filtro, que consiste de uma matriz de dimensões $N \times N$, cujos componentes são os pesos da combinação linear a ser feita. Por questões de simetria, geralmente, são utilizadas janelas quadradas com N ímpar. Além disso, normalmente, para maior eficiência computacional, os valores de N são pequenos.

Basicamente, o elemento central da matriz da máscara coincide com o pixel a ser modificado, sendo posicionado sucessivamente sobre cada pixel de interesse da imagem. Portanto, o processo pode ser entendido como uma máscara deslizante sobre os pontos de interesse da imagem, que assumirão um novo valor de acordo com a combinação linear dos pixels vizinhos ponderados pelos pesos w da matriz máscara.

Esse processo todo pode ser representado por uma operação de convolução da imagem $I(m, n)$ com a máscara de pesos w , resultado na imagem filtrada $I_{filt}(m, n)$. A equação

(2.7) representa o processo em sua forma matemática.

$$I_{filt}(m, n) = \sum_{i=I_{min}}^{i=I_{max}} \sum_{j=J_{min}}^{j=J_{max}} w(i, j)I(m + i, n + j), \quad (2.7)$$

situação em que os índices $i = 0$ e $j = 0$ se referem ao pixel central da máscara, que possui dimensões $(I_{max} - I_{min} + 1, J_{max} - J_{min} + 1)$.

Destaca-se que os pontos pertencentes às bordas da imagem merecem cálculo diferenciado, visto que não possuem todos os vizinhos.

A filtragem linear tem a mesma ideia de máscara deslizante, ponto central e vizinhança, mas o novo valor a ser atribuído ao pixel da vez é dependente de alguma operação não linear envolvendo os pixels vizinhos.

Em relação aos tipos, geralmente os filtros se dividem em três: passa-baixas, passa-faixa e passa-altas.

O primeiro se caracteriza por atenuar as altas frequências, que estão atreladas às informações de detalhes das imagens, provocando um efeito visual de suavização, tendendo a minimizar o efeito do ruído nas imagens. Todavia, por conta do borramento causado, detalhes finos podem ser removidos das imagens.

O filtro passa-faixa prima por selecionar um intervalo de frequências para ser realçado. Portanto, seu efeito geral depende do intervalo selecionado.

Já o filtro passa-altas realça as altas frequências com a finalidade de realce dos detalhes, o que promove um efeito de maior nitidez nas transições de regiões, mais conhecidas como bordas. Em contrapartida, esse filtro enfatiza o ruído já presente nas imagens.

Como exemplos de filtro passa-baixas, serão elencados o filtro da média e o filtro Gaussiano.

Aplicar o filtro da média em uma imagem consiste em substituir o pixel da vez pela média dos seus *pixels* vizinhos, que depende do tamanho da máscara. A equação (2.8) mostra um exemplo de máscara de um filtro da média.

$$W = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (2.8)$$

Em relação ao filtro Gaussiano, os coeficientes da máscara são obtidos a partir de uma função Gaussiana bidimensional. A equação (2.9) mostra a função Gaussiana com média zero e desvio padrão σ .

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (2.9)$$

A função Gaussiana, em suas duas dimensões, são simétricas com relação à rotação. Logo, o grau de suavização de um filtro baseado nela é o mesmo em todas as direções, ou seja, é um filtro isotrópico. Além disso, o peso atribuído aos pixels vizinhos na computação do pixel da vez decresce monotonicamente com a distância.

Destaca-se também que o grau de suavização de um filtro Gaussiano está relacionado com o parâmetro σ , quanto maior seu valor, maior o grau de suavização. A equação (2.10) expressa um exemplo de uma máscara de um filtro Gaussiano de dimensão 5x5.

$$W = \frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \quad (2.10)$$

A Figura 2.11 mostra o resultado dos filtros passa-baixas abordados aqui, pela média e filtro Gaussiano. Nas imagens filtradas, nota-se a suavização e, conseqüentemente, a perda de alguns detalhes.



(a) Imagem a ser filtrada.

(b) Imagem filtrada por um filtro pela média.

(c) Imagem filtrada por um filtro Gaussiano.

Figura 2.11

2.4.2.4 Detector de Bordas de Canny

Voltando às técnicas utilizadas para segmentação de imagens (assim como a binarização, já abordada), as abordagens para detecção de bordas possuem importante papel, como descrito em [40].

Primeiramente, entende-se borda como o limite ou a fronteira entre duas regiões com propriedades distintas de nível de cinza. Para isso, assume-se que as duas regiões separadas pela borda são suficientemente homogêneas a ponto de a transição entre elas poder ser determinada com base apenas na descontinuidade dos níveis de cinza.

De forma geral, a maioria das técnicas de detecção de bordas utiliza o cálculo de um operador diferencial. Partindo desse ponto, pode-se utilizar a magnitude da primeira derivada para a detecção de uma borda em uma imagem. Por sua vez, o sinal da derivada segunda possui um cruzamento no zero, ou seja, indica que há transição dos níveis de cinza, o que

permite a localização das bordas na imagem.

Pontua-se que a primeira derivada em qualquer ponto da imagem é obtida a partir da magnitude do gradiente naquele ponto. Já a segunda derivada é obtida da mesma forma, mas a partir do operador Laplaciano.

Feita essa breve introdução sobre detectores de bordas, ressalta-se que um dos mais utilizados é o detector de bordas de Canny [43], que foi pensado para satisfazer três critérios:

- baixa taxa de erro: detectar as bordas que, de fato, existem na imagem, com a menor ocorrência de falsos positivos;
- boa localização dos pontos de borda: distância mínima entre os pontos detectados como bordas e os pontos que, de fato, fazem parte da bordas;
- uma única resposta para uma borda.

O detalhamento matemático da técnica foge do escopo deste trabalho, mas os passos básicos de sua aplicação seguem abaixo:

1. Primeiramente, a imagem é filtrada por um filtro Gaussiano, já que operadores diferenciais são sensíveis à ruído. Quanto maior a largura da máscara, mais suavizada ficará a imagem, em contrapartida, o erro na localização das bordas aumentará;
2. encontra-se a robustez da borda por meio da utilização de operadores de Sobel na vertical e na horizontal. A robustez será determinada pela soma nas duas direções;
3. ainda por meio da utilização de operadores de Sobel, calcula-se a direção das bordas;
4. aproxima-se a direção das bordas para que seja possível as suas representações na imagem;
5. aplica-se a supressão não-máxima, que consiste em atribuir como borda apenas aqueles pontos cuja magnitude seja localmente máxima na direção do gradiente, o que reduz a espessura das bordas;
6. realiza-se a limiarização com histeres: por ainda ser possível a existência de fragmentos espúrios causados por ruído ou textura fina, utiliza-se limiares mínimos e máximos para os gradientes com o intuito de seleção das bordas. Ou seja, as bordas que estão dentro dos limiares são selecionadas.

A Figura 2.12 mostra a aplicação do detector de bordas de Canny em uma imagem. Observa-se, para o caso, o destaque dado a algumas bordas de cédulas da imagem, a detalhes do desenho da onça da cédula e aos caracteres existentes.



(a) Imagem antes da utilização do detector de bordas de Canny. (b) Imagem após a utilização do detector de bordas de Canny, com limiares mínimo e máximo, respectivamente, iguais a 0,2 e 0,4.

Figura 2.12: Ilustração da aplicação do detector de bordas de Canny a uma imagem.

2.4.2.5 Transformada Hough

Além da detecção de bordas, neste trabalho, a detecção de um conjunto de pontos em uma imagem que pertencem a um segmento de reta também é importante. Ressalta-se que, a abordagem que é feita aqui para segmento de reta pode ser estendida para uma curva específica, como, por exemplo, uma circunferência ou elipse, seguindo a proposta apresentada por [44].

Dito isso, o problema relacionado ao segmento de reta consiste basicamente em achar subconjuntos de pontos que sejam colineares. Uma solução possível, mas inviável computacionalmente para a maioria das aplicações, é encontrar todos os segmentos de retas formados entre cada par de pontos e procurar pelos conjuntos de pontos que estejam próximos desses segmentos.

Com a finalidade de tornar essa busca mais eficiente, Hough propôs um método, denominado transformada de Hough [44], que é detalhado na sequência.

Uma reta pode ser representada pela equação (2.11)

$$y = mx + b, \tag{2.11}$$

em que m é a declividade da reta e b é o ponto de sua intersecção com o eixo y de um plano cartesiano, que também pode ser entendido como o plano de uma imagem.

Para diferentes valores de m e b , há infinitas retas que passam por um ponto $p_1(x_1, y_1)$, todas elas satisfazendo a equação $y_1 = mx_1 + b$, bem como há infinitas retas que passam por um ponto $p_2(x_2, y_2)$, todas elas satisfazendo a equação $y_2 = mx_2 + b$.

Reescrevendo a equação (2.11), tem-se que

$$b = y - mx. \tag{2.12}$$

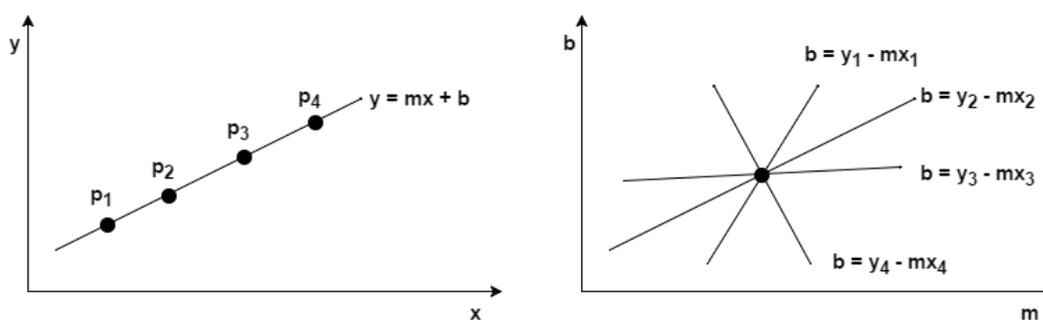
Assim como o plano xy é denominado de plano da imagem, o plano mb é chamado de

espaço de parâmetros.

Sendo assim, todas retas que passam pelo ponto p_1 são representadas no espaço de parâmetros pela equação $b = y_1 - mx_1$. De forma análoga, a equação $b = y_2 - mx_2$ representa todas as retas que passam pelo ponto p_2 no plano da imagem.

É interessante observar que o ponto (m, b) , localizado no espaço de parâmetros, é comum a essas duas retas associadas aos pontos p_1 e p_2 . Mais interessante ainda é notar que todos os pontos que são colineares no plano da imagem se interceptam em um mesmo ponto no espaço de parâmetros.

A Figura 2.13 ilustra essa afirmação. Na Figura 2.13(a), observa-se quatro pontos colineares, ou seja, pertencentes à mesma reta. Esses pontos, no espaço de parâmetros, dão origem a quatro retas que se encontram no mesmo ponto.



(a) Pontos colineares no plano da imagem. (b) Retas concorrentes no mesmo ponto no espaço de parâmetros.

Figura 2.13: Os quatro pontos colineares em (a) são mapeados em quatro retas que se cruzam no mesmo ponto no espaço de parâmetros em (b).

Esse entendimento é a base para a transformada de Hough que, para melhor representação, utiliza a equação de uma reta em sua forma polar. Ou seja, em vez de representar uma reta por meio da equação 2.11, representa pela seguinte expressão matemática

$$\rho = x.\cos\theta + y.\sen\theta, \quad (2.13)$$

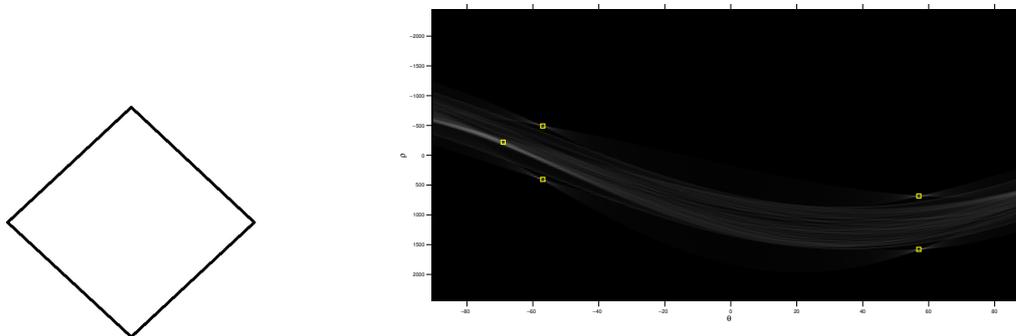
em que ρ é a distância da origem, ponto $(0,0)$, à reta e θ é o ângulo formado entre a reta perpendicular à reta em análise e o eixo x .

Portanto, agora, passa-se do espaço (m, b) para o espaço (ρ, θ) , denominado espaço de Hough. Mais do que isso, pontos colineares no espaço (x, y) correspondem a curvas senoidais que se interceptam no espaço de Hough.

Para ilustração, a Figura 2.14 mostra uma imagem com um losango e sua correspondente representação no espaço Hough, em que são destacados alguns pontos. Esses pontos são destacados por consistirem do cruzamento de várias curvas senoidais, ou seja, dizem respeito a vários pontos no plano imagem que potencialmente são colineares.

De fato, na imagem, são observados quatro conjuntos de pontos colineares, que determi-

nam os lados do losango. Esses quatro conjuntos correspondem a quatro pontos no espaço Hough, que estão destacados na figura. Há, ainda, um quinto ponto, que consiste de um falso positivo.



(a) Losango em um plano cartesi- (b) Representação de (a) no espaço de Hough, em que se destacam pontos
ano. referentes aos segmentos de reta que forma o losango daquela figura.

Figura 2.14: Ilustração de uma imagem no espaço de Hough.

Feita essa contextualização, a transformada Hough pode ser implementada por meio dos seguintes passos:

1. O espaço Hough (ρ, θ) é discretizado em intervalos finitos, criando para cada uma das células resultantes um acumulador $Ac(\rho, \theta)$;
2. Todas as células do acumulador $Ac(\rho, \theta)$ são iniciadas com o valor zero;
3. Para cada ponto (x, y) no plano da imagem, calcula-se os valores de ρ e θ que satisfazem a equação (2.13);
4. Encontrados os valores, incrementa-se de uma unidade o acumulador $Ac(\rho, \theta)$;
5. Uma vez determinados os parâmetros de todos os pontos do plano da imagem, as células com os maiores valores do acumulador $Ac(\rho, \theta)$ indicam potenciais retas na imagem.

Uma das dificuldades do método é determinar, dentre os maiores valores do acumulador, aqueles que são relevantes, ou seja, deve-se definir um limiar acima do qual uma célula será considerada como tendo os parâmetros de um segmento de reta na imagem. Além disso, a definição dos parâmetros de discretização também influencia nos resultados.

Por outro lado, algumas vantagens merecem destaque. A transformada Hough é capaz de detectar segmentos que apresentam regiões obstruídas na imagem. Além disso, ela é pouco sensível a ruído, uma vez que os pontos da imagem afetados pelo ruído dificilmente são mapeados na mesma célula do acumulador.

2.4.3 Abordagens para a Localização

A tarefa de localização de textos (palavras ou caracteres) em imagens não é um tema recente na literatura. Por conta disso, existem diversas abordagens para se atacar o problema, como, por exemplo, similaridade de padrões (*template matching*) [45], reconhecimento de objetos (*object recognition*) [46], propostas de objetos (*object proposals*) [47], análise de componentes conexas [48], entre outras.

Basicamente, elas podem ser divididas em dois grupos: métodos baseados em busca por padrões conhecidos *a priori* e métodos que encontram os caracteres partindo da premissa de que os *pixels* que o formam possuem características semelhantes, sendo possível a utilização de componentes conexas [48].

Sabendo que os melhores resultados têm sido obtidos com a utilização da abordagem de componentes conexas [14], optou-se por sua utilização. Contudo, antes de prosseguir com seu detalhamento, considera-se importante justificar a não utilização de outras abordagens.

A abordagem que utiliza similaridade de padrões (*template matching*), por exemplo, é custosa computacionalmente, visto que será necessária uma janela deslizante para percorrer toda a imagem. Mais do que isso, caso se queira robustez à rotação e variação de escala, objetivos deste trabalho, é necessário o processamento para diferentes escalas e, dependendo da situação, para diferentes ângulos de rotação, o que torna a abordagem ainda mais custosa em termos computacionais.

Técnicas de reconhecimento de objetos, como, por exemplo, SIFT e SURF, já não possuem esse problema quanto à variação de orientação e escala. Contudo, para a localização de um caractere como sendo um objeto, elas não se mostraram eficientes, visto que um caractere não possui pontos característicos o suficiente para ser localizado em uma imagem do banco criado.

Uma técnica que tem em seus princípios a utilização de componentes conexas é a técnica que procura por Regiões Extremas Maximamente Estáveis (MSER - *Maximally Stable Extremal Regions*) apresentada em [49].

Em visão computacional, esses tipos de regiões são utilizadas como uma forma de se detectar *blobs* em imagens. Um *blob* pode ser entendido como uma região da imagem em que algumas propriedades são constantes ou aproximadamente constantes, ou seja, todos os *pixels* em um *blob* podem, de alguma maneira, serem considerados similares entre eles.

Por conta dessa característica, é de grande valia para a busca por caracteres em imagens, visto que caracteres, por formação, são regiões que possuem características semelhantes.

2.5 Reconhecimento Ótico de Caracteres - *Tesseract*

A ideia principal do OCR é o reconhecimento de um caractere em uma determinada imagem digitalizada [2]. Para isso, as características desse caractere são comparadas com características de padrões de um determinado alfabeto, utilizando, dentre outras, ferramentas disponíveis de processamento digital de imagens e visão computacional.

Mais especificamente, o OCR é um processo de conversão de um documento que contenha números, letras ou símbolos em uma sequência de dados passíveis de serem interpretadas por um computador, como, por exemplo, um arquivo-texto.

Em linhas gerais, um sistema OCR deve desempenhar os seguintes passos: digitalização do documento, segmentação, pré-processamento, extração das características, classificação, pós-processamento, formatação da saída.

A finalidade da digitalização do documento é que se tenha um registro digitalizado dele próprio, que pode ser feito, por exemplo, por um *scanner* ou até mesmo por uma máquina fotográfica digital.

A segmentação consiste em determinar os constituintes da imagem, ou melhor, distinguir as regiões que contém os dados de interesses (texto) das regiões não interessantes para a situação, como, por exemplo, desenhos.

A imagem digitalizada, independentemente do meio de digitalização, pode conter certa quantidade de ruído. A qualidade de uma imagem capturada por meio de uma câmera digital, por exemplo, depende, dentre outros fatores, da resolução da máquina, da iluminação no momento da captura e da distância entre o documento e o equipamento. Como consequência desses fatores, é possível que haja baixas taxas de reconhecimento, que podem ser mitigadas com a utilização de um pré-processamento para suavizar e normalizar os caracteres digitais.

Na etapa de extração das características, captura-se as características essenciais dos símbolos detectados. Existem diversos algoritmos, que devem ser avaliados quanto à sua robustez (comportamento diante de ruído, distorção, variação de estilo, translação, rotação) e quanto ao seu uso prático (velocidade de reconhecimento, complexidade de implementação).

Por sua vez, a classificação é o processo de identificação de cada um dos caracteres de modo a atribuir-lhes uma classe. Para que haja essa atribuição, é necessário que os padrões sejam conhecidos do sistema. Portanto, deve haver uma fase prévia de aprendizado, na qual as características de cada um dos padrões é passada, ensinada ao sistema.

Após o reconhecimento de cada um dos caracteres, deve-se fazer a associação entre eles com a finalidade de se ter uma sequência de caracteres que tenha algum significado, ou seja, eles devem ser agrupados de modo a formar palavras, por exemplo.

Além disso, é importante ressaltar que se deve considerar a ocorrência de erros na etapa de classificação. Sendo assim, no pós-processamento pode ser realizado uma atividade de detecção e correção dos erros.

Por fim, torna-se necessário que se tenha uma interface de comunicação entre o sistema e o mundo externo a ele com o intuito de expor o resultado. Há sistemas que formatam sua saída em documentos de texto, tabelas, banco de dados, entre outros. Há também sistemas que expõem suas saídas a outros sistemas, fazendo parte de um sistema automatizado.

Feita essa introdução básica sobre o que se espera de uma ferramenta de OCR, faz-se importante citar que existem várias disponíveis no mercado e de vários tipos, grátis e comerciais; *on line* e *offline*; com código aberto e com código fechado etc. Contudo, poucas das ferramentas são grátis e de código aberto.

Aqui, será dado destaque ao *Tesseract* [50], um *software* de OCR livre e de código aberto, visto que é bastante utilizado pela literatura e mostra bom desempenho na tarefa que se propõe a fazer.

O *Tesseract* foi desenvolvido pela HP (Hewlett-Packard) entre os anos de 1984 e 1994, sendo que em 2005 o seu código se tornou aberto. A partir de 2006, passou a ser desenvolvido pelo Google [50].

É interessante ressaltar que esse software oferece suporte para várias línguas e funciona apenas por linha de comando. Caso se deseje uma aplicação com interface gráfica, deve-se procurar por um produto de terceiros que o utilize.

A Figura 2.15 mostra um fluxograma que traz, de forma representativa, como se dá o funcionamento da ferramenta em questão [23]. Nota-se que a imagem de entrada é primeiramente alvo de uma limiarização adaptativa, em que obtém uma imagem binária. Logo em seguida, é realizada uma análise de componentes conexas, em que são extraídos os contornos dos caracteres. Posteriormente, utiliza-se um algoritmo com a finalidade de encontrar linhas e palavras na imagem [51].

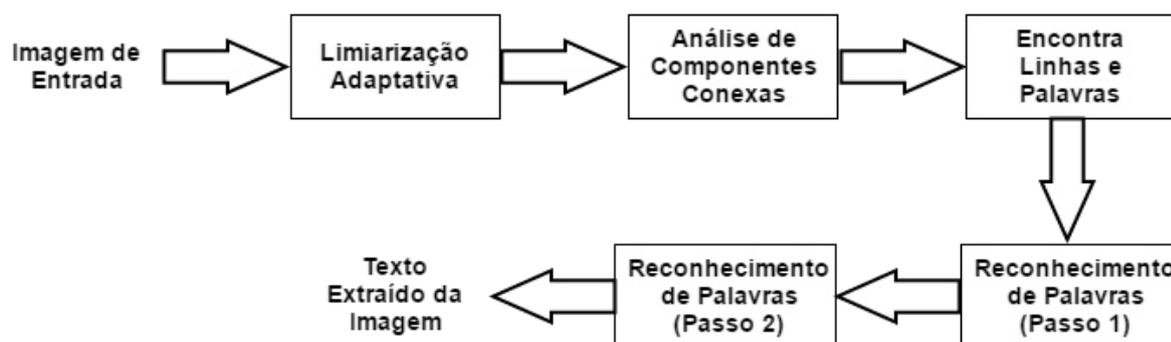


Figura 2.15: Fluxograma de funcionamento do *Tesseract*, adaptado de [2]

Sendo assim, é iniciada o reconhecimento do texto em dois passos. No primeiro, faz-se uma tentativa de reconhecimento de cada palavra do texto. Cada caso de sucesso é passado a um classificador adaptativo como um dado de treinamento. Como esse classificador recebe dados de treinamento, ele tenta reconhecer o texto de uma maneira mais eficiente no último passo. A referência [22] traz maiores detalhes sobre cada uma das etapas. Voltando a uma das principais características do *Tesseract*, a de funcionamento por linha de comando, a

Figura 2.16 mostra uma tela de comando em que se encontra como se utilizar o software.

Primeiramente, a linha de comando a ser digitada para que sejam reconhecidos os caracteres de uma imagem qualquer "imagenome" e sejam dispostos em um arquivo texto qualquer "outputbase" é:

```
"tesseract imagenome outputbase [-l lang] [-psm pagesegmode] [configfile...]",
```

em que, opcionalmente, a expressão *lang* deve ser substituída pela língua de interesse; *pagesegmode*, substituída pelo modo de operação da aplicação e *configfile*, pelas configurações desejadas.

Aqui, vale a pena destacar os modos de operação 8 ("Treat the image as a single word") e 10 ("Treat the image as a single character"), que são os modos utilizados para a obtenção dos resultados. Eles significam, respectivamente, que a aplicação irá tratar a imagem como uma única palavra e como um único caractere.

```
Usage:tesseract imagenome outputbase [-l lang] [-psm pagesegmode] [configfile...
]
pagesegmode values are:
0 = Orientation and script detection (OSD) only.
1 = Automatic page segmentation with OSD.
2 = Automatic page segmentation, but no OSD, or OCR
3 = Fully automatic page segmentation, but no OSD. (Default)
4 = Assume a single column of text of variable sizes.
5 = Assume a single uniform block of vertically aligned text.
6 = Assume a single uniform block of text.
7 = Treat the image as a single text line.
8 = Treat the image as a single word.
9 = Treat the image as a single word in a circle.
10 = Treat the image as a single character.
-l lang and/or -psm pagesegmode must occur before anyconfigfile.

Single options:
-u --version: version info
--list-langs: list available languages for tesseract engine
```

Figura 2.16: Tela de comando com as opções do *Tesseract*.

2.6 Trabalhos Correlatos

Como já abordado, o reconhecimento de caracteres em imagens digitais é uma área ampla com inúmeras aplicações, tais como, reconhecimento automático de textos digitalizados, reconhecimento de caracteres escritos à mão, reconhecimento de placas de trânsito, reconhecimento de placas de carro, entre outras.

O tema alvo específico deste trabalho, reconhecimento de caracteres que formam números seriais de cédulas, também possui sua importância na literatura. Alguns dos trabalhos mais relevantes são citados a seguir.

Em [9], utiliza-se SVM (*Support Vector Machine*) com múltiplos *kernels* para um sistema de reconhecimento de número serial. Observa-se que a localização do número serial é conhecida *a priori* e não há variação de escala nem de orientação. Destaca-se também que a aplicação é testada para o Dólar, o Euro e o Renminbi.

Por sua vez, em [12], os números seriais são extraídos de imagens escaneadas, sendo que, para a localização deles, considera-se a possível ocorrência de inclinação das imagens. Além disso, aqui, propõe-se um método para a localização do número serial e a extração dos caracteres. Contudo, não se propõe alternativa para o reconhecimento desses caracteres. Além disso, apenas a cédula chinesa é alvo do trabalho.

Já em [11], são testados diferentes tipos de extrator de características, diferentes técnicas de classificação e diferentes estratégias de combinação de classificadores com o intuito melhorar a precisão no reconhecimento de caracteres de números seriais de cédulas chinesas. Assim como em [9], a localização do número serial é conhecida bem como não há variações de escala e rotação. Além disso, é dado enfoque apenas à cédula de Renminbi.

Capítulo 3

Metodologia Proposta

Uma vez que a maioria dos estudos publicados relacionados a leitura de caracteres de cédulas são relativos a moedas de outros países que não o Brasil, as abordagens utilizadas não são diretamente aplicáveis às cédulas de Real. Isso ocorre por conta das diferenças existentes entre as cédulas, tais como, diferentes fontes de caracteres, diferentes posições para os números seriais, diferentes planos de fundos, entre outras.

Soma-se a essas diferenças o fato de que, neste trabalho, deve-se lidar também com a variação de escala e orientação, questões não abordadas pelos estudos existentes.

Sabendo disso e tendo em mente que uma das motivações do trabalho é a criação de um sistema de monitoramento de cédulas de Real baseado na detecção e reconhecimento do número serial no futuro e como parte inicial da metodologia desenvolvida, fez-se necessária a criação de um banco de imagens de cédulas de Real da Segunda Família. Essa foi a única forma encontrada para testar e avaliar as propostas do trabalho.

Antes de prosseguir com o texto, é essencial pontuar que, para se ter um ponto de partida em relação a desempenho e justificar as soluções e metodologias propostas, foram realizadas algumas medidas iniciais. A forma como isso foi realizado é descrito neste Capítulo nas Seções 3.2 e 3.3.

Dando sequência, uma vez criado o banco de imagens a ser utilizado, tornou-se viável a implementação do que é proposto, que consiste, em linhas gerais, de duas fases: A primeira pode ser entendida como uma etapa de detecção dos números seriais, que deve levar em consideração tanto a variação de escala como a variação de orientação. A segunda consiste da análise do reconhecimento das cadeias de caracteres encontradas na etapa anterior.

Em suma, a metodologia, aqui proposta, pode ser dividida em três etapas: *Elaboração do Banco de imagens*, *Detecção dos números seriais* e *Análise do reconhecimento da cadeia de caracteres*. Conforme pode ser visto pelo fluxograma da Figura 3.1, as duas últimas etapas são realizadas na ordem apresentada e em estrutura de cascata, sequenciais.

Seguindo a mesma linha do fluxograma, este capítulo está organizado da seguinte ma-

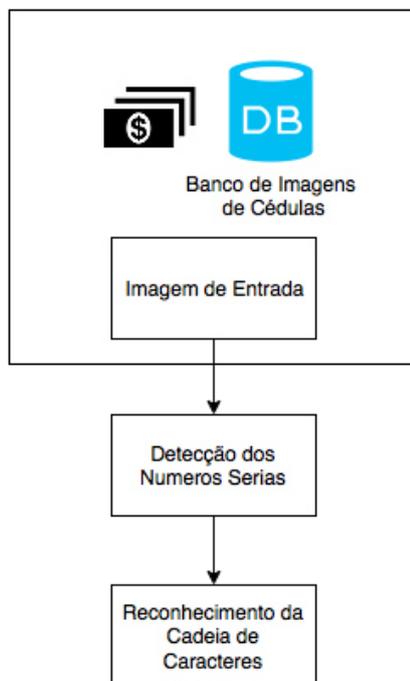


Figura 3.1: Fluxograma da metodologia proposta do trabalho.

neira. A seção 3.1 detalha o banco de imagens criado, ao passo que a seção 3.2 detalha os cenários de medidas iniciais. Já a seção 3.3 traz as abordagens utilizadas para a localização das cadeias de números seriais. Por fim, a forma como é feita o reconhecimento propriamente dito dos caracteres é tratada.

3.1 Elaboração do Banco de Imagens

Para a aplicação e avaliação das técnicas propostas, foi elaborado um banco de imagens de cédulas da Segunda Família do Real, composto por cédulas de 2 (dois) e de 5 (cinco) Reais e organizado em quatro classes. Cada classe foi desenvolvida para representar uma peculiaridade a ser encontrada durante a execução de um sistema de monitoramento de cédulas em situações reais. A ideia básica é a construção de uma base de dados que seja capaz de representar as principais situações, mesmo que em um conjunto reduzido, que se aproximam da realidade de uma aplicação prática desse tipo de sistema.

A seguir, está a denominação e o detalhamento de cada uma das classes:

1. Quantidade (Classe 1): varia-se a quantidade de cédulas por imagem, ou melhor, há aquelas com uma cédula, duas cédulas, três cédulas, quatro cédulas e cinco cédulas. Em todas as imagens, as cédulas estão horizontalmente dispostas e alinhadas entre si;
2. Rotacionadas (Classe 2): varia-se a orientação das cédulas presentes nas imagens em relação ao eixo x do plano imagem. Essa orientação é medida em graus e há conjuntos de imagens com orientações de: 0° , 30° , 60° e 90° ;

3. Variação de Escala (Classe 3): varia-se a distância das cédulas em relação à câmera, distância pequena, média e grande, que correspondem, respectivamente, à 7, 11 e 18 centímetros;
4. Ausência de Números Seriais (Classe 4): imagens com cédulas, mas com números seriais totalmente oclusos.

Observa-se que, para as três primeiras classes, os números seriais das cédulas aparecem nas imagens e, quando possuem mais de uma cédula, elas estão alinhadas entre si. Já a quarta classe (ausência de números seriais) se caracteriza por ter imagens em que nenhum número serial está visível. O objetivo desta classe é justamente avaliar a influência da detecção de possíveis falsos positivos, tanto na etapa de detecção, quanto na etapa de reconhecimento dos números de série das cédulas. Além disso, para todas as classes, com exceção da terceira, a distância à câmera é a distância média.

Tratando-se de quantidades, há para as classes 1 à 4, respectivamente, 161, 130, 98 e 34 imagens, totalizando mais de 980 números seriais visíveis, o que resulta em mais de dez mil caracteres a serem detectados e reconhecidos.

A Figura 3.2 apresenta uma amostra de cada uma das classes que podem ser encontradas no banco de imagens. Nenhuma informação *a priori* da câmera, incluindo qualquer processo de calibração e/ou ajustes, foram realizados para a captura das imagens que compõem essa base.

Por fim, a Tabela 3.1 lista as configurações da câmera utilizada para a captura das imagens do banco, que é a câmera de um celular Samsung Galaxy S3 GT-i9305T. A utilização de uma câmera de celular se justifica pela ideia da criação de um Sistema de Monitoramento de Cédulas, conforme já abordado no Capítulo 2, em que os usuários poderão utilizar câmeras que não sejam de alta resolução ou não tenham uma boa qualidade de conjunto ótico.

Tabela 3.1: Configurações da câmera utilizada para construção do banco de imagens.

| Parâmetro | Configuração |
|---------------------|------------------|
| Modo do foco | Automático |
| Resolução | 3264 x 2448 |
| Controle de branco | Azulado |
| Sensibilidade à luz | ISO 200 |
| Medição | Centro ponderado |
| Qualidade da imagem | Super fino |

3.2 Cenários Iniciais

Uma vez que não há pesquisa anterior com o mesmo objetivo específico, reconhecimento de caracteres de números seriais de cédulas de Real da Segunda Família, fez-se necessário



(a) Imagem da classe *Quantidade* com 3 cédulas. (b) Imagem da classe *Rotacionadas* com inclinação de 60° .



(c) Imagem da classe *Variação de Escala* com distância grande (18 cm). (d) Imagem da classe *Ausência de Números Seriais*.

Figura 3.2: Amostra do banco de imagens criado.

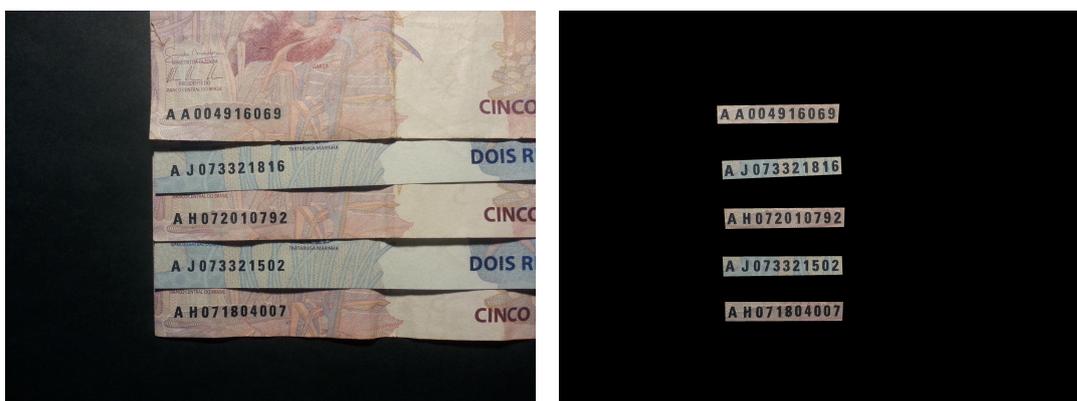
um ponto de partida em termos de desempenho. Para isso, criou-se dois cenários, doravante, denominados Cenário de Referência e Cenário de Controle.

No Cenário de Referência, todas as imagens do banco, sem nenhum processamento, assim como mostradas pela Figura 3.2, tornaram-se entradas do *Tesseract*. A ideia é mostrar que se faz necessário um pré-processamento da imagem, visto que são reconhecidos inúmeros caracteres que não são de interesse, tais como, os presentes em: "BANCO CENTRAL DO BRASIL", "REAIS", "PRESIDENTE DO BANCO CENTRAL DO BRASIL", dentre outros. Mais do que isso, tornou-se uma primeira oportunidade para avaliar um pouco do funcionamento do *Tesseract*.

Já para o Cenário de Controle, a ideia é ter uma primeira mensuração mais próxima de um dos objetivos do trabalho, o reconhecimento dos números seriais. Sendo assim, foram passadas como entrada ao *Tesseract* todas imagens do banco, mas com uma máscara preta aplicada nas regiões em que não existem números seriais. Com esse segundo cenário, mesmo sendo um cenário ideal em termos de localização dos seriais, já é possível ter uma ideia melhor a respeito do conhecimento dos caracteres, sobretudo quando são consideradas as diferentes escalas e orientações das cédulas.

A Figura 3.3 mostra um exemplo de imagem de cada um dos cenários iniciais. Nota-se que elas se diferem apenas pela aplicação da máscara na imagem utilizada para o segundo cenário. Ou melhor, os elementos de interesse, ou seja, os caracteres que formam os números seriais são os mesmos.

O Capítulo 4 explicita resultados relacionados a esses cenários, oportunidade em que se constata a melhora trazida pelas propostas que são explicadas na sequência deste capítulo.



(a) Exemplo de imagem utilizada para o Cenário de Referência.

(b) Exemplo de imagem utilizada para o Cenário de Controle.

Figura 3.3: Exemplos de imagens utilizadas no cenários iniciais

3.3 Detecção dos Números Seriais

Para a tarefa de localização de caracteres em uma imagem, geralmente utilizam-se basicamente duas abordagens: Quando se tem o conhecimento *a priori* do padrão desejado, no caso, o número serial, abordagens que utilizam *template matching* [45], permitindo que outras técnicas para reconhecimento de objetos sejam passíveis de serem empregadas [46]. Por outro lado, existem os métodos que encontram os caracteres partindo da premissa de que os pixels que o formam possuem características semelhantes, sendo possível a utilização de componentes conexos [48].

Sabendo que os melhores resultados têm sido obtidos com a utilização da abordagem de componentes conexos [14], optou-se neste trabalho por sua utilização. Ressalta-se, ainda, que além da localização dos caracteres, deve-se considerar também a variação da orientação das cédulas na imagem. Sendo assim, a tarefa de localização pode ser dividida em três etapas: determinação do ângulo de inclinação das cédulas nas imagens, determinação de prováveis regiões que contenham números seriais e suas posteriores validações, descritas nas subseções a seguir.

Antes, todavia, faz-se necessário pontuar que a abordagem de *template matching* para casos em que há variação de escala e variação de orientação se torna custosa computacionalmente e com baixa exatidão, uma vez que o *template* deve ser procurado na imagem em várias escalas e várias orientações.

Sabendo que há técnicas que são robustas a essas dificuldades, como, por exemplo, o uso da Transformada SIFT/SURF[38], tentou-se sua utilização para a procura tanto de caracteres do número serial quanto da cadeia de caracteres como um todo. No entanto, essa abordagem não se mostrou promissora, uma vez que os pontos característicos dos *templates* na situação (caracteres ou cadeia de caracteres) não se mostraram discriminadores ao ponto de serem localizados em uma imagem sem erro. É evidente que essa informação ganharia mais respaldo com a análise de algumas métricas, mas, em testes iniciais, essa opção se mostrou tão infrutífera que se decidiu não prosseguir na direção desta abordagem.

3.3.1 Determinação da Inclinação

Feita essa breve introdução, a metodologia utilizada para a determinação do ângulo de inclinação das cédulas da imagem para a posterior reversão da rotação será detalhada. Esta se divide em dois passos: determinação de retas contidas na imagem e escolha da reta que representa a inclinação desejada.

Assumindo que os grafismos presentes em uma cédula possuem vários elementos de alta frequência, mas que possuem dimensões reduzidas em comparação à extensão das bordas, e com a finalidade de se destacar somente os pontos que determinarão as retas associadas a estas bordas, utilizou-se a seguinte sequência de ações: Filtro passa-baixas (*blurring*),

conversão da imagem para tons de cinza, detector de bordas de Canny[43], operador de dilatação com elemento estruturante quadrado e Transformada Hough [44]. A Figura 3.4 ilustra as saídas obtidas com os passos descritos.

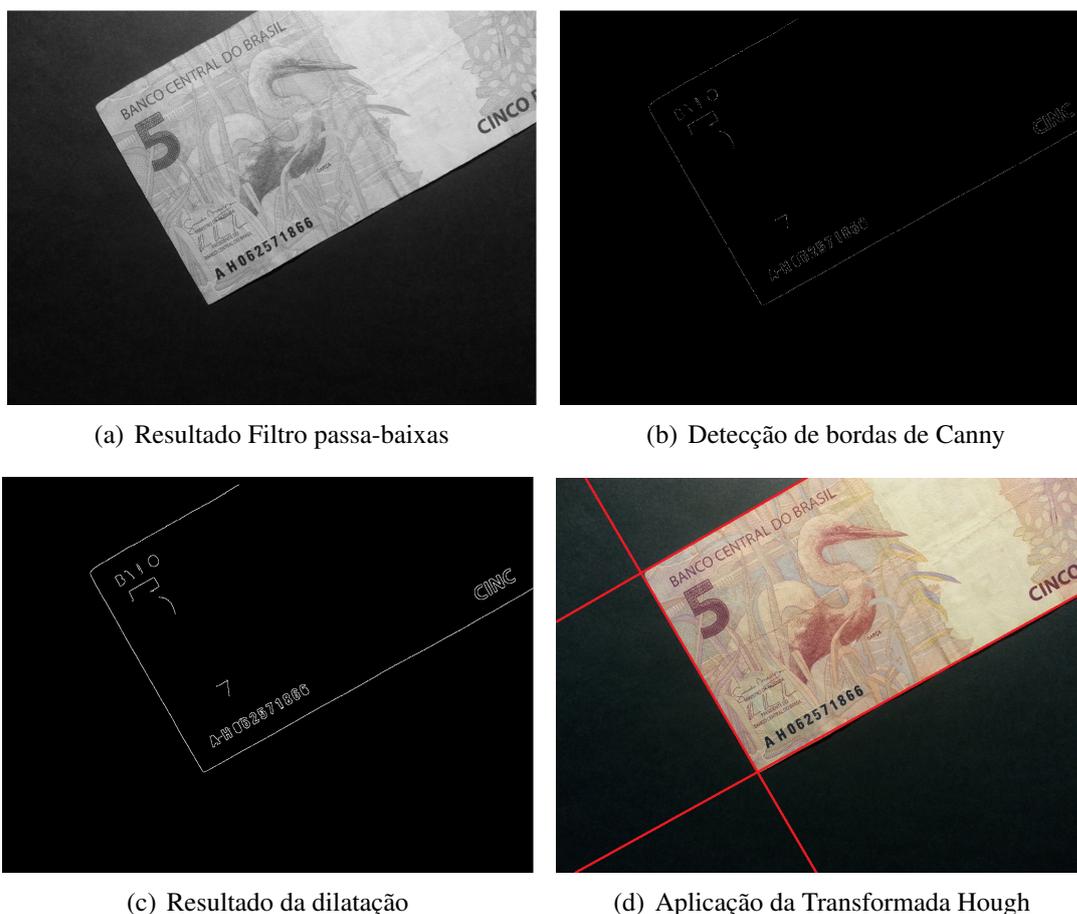


Figura 3.4: Ilustração dos passos para determinar as retas.

Definidas as retas, como pode-se ver na Figura 3.5(a), deve-se escolher aquela que guiará a reversão da rotação. Os passos realizados anteriormente são eficientes a ponto de localizarem retas que, em sua imensa maioria, estão associadas às bordas das cédulas presentes nas imagens. Sabendo disso, a ideia é encontrar aquela que está associada à maior borda.

Para isso, de um lado, a imagem, após a aplicação de um filtro passa baixas, é binarizada pelo método de Otsu (3.5(b)). De outro, faz-se uma máscara, em que cada reta obtida no passo anterior é referência para a formação de uma região com pixels diferente de zero (3.5(c)), da seguinte forma: tendo uma reta como base, os limites dessa região serão essa reta transladada de n pixels para cima e para baixo, tendo como referência de direção o seu vetor normal.

Faz-se então uma operação lógica *AND pixel a pixel* da imagem binarizada com a máscara gerada, resultando em uma imagem que destacará os pixels localizados nas regiões próximas às bordas das cédulas (3.5(d)).

É sensato supor que a região que está associada à maior borda terá a maior quantidade de

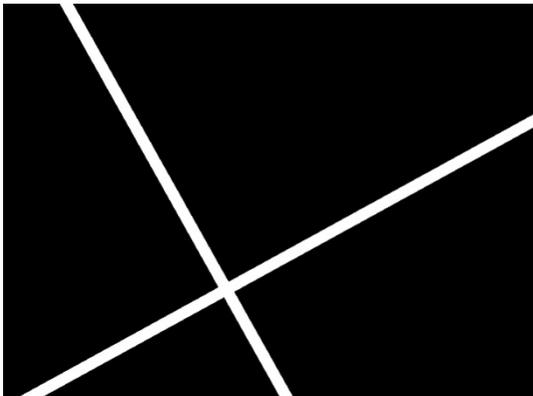
pixels diferentes de zero. Logo, a reta escolhida como base para a determinação do ângulo de inclinação será aquela que resultar na maior soma de *pixels* diferentes de zero, e o ângulo de inclinação será o ângulo que esta faz com a horizontal. A Figura 3.5 ilustra imagens dos passos descritos para a determinação do ângulo.



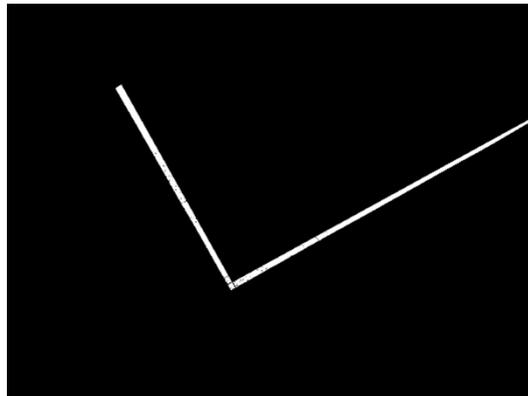
(a) Imagem com as retas detectadas na etapa anterior



(b) Imagem binarizada



(c) Máscara binária utilizada



(d) Imagem com as regiões definidas

Figura 3.5: Ilustração dos passos para determinação do ângulo.

3.3.2 Prováveis Regiões com Número Serial

Uma vez desfeita a rotação das cédulas na imagem, parte-se para a localização propriamente dita dos números seriais. O cerne da tarefa de determinação das prováveis regiões da imagem que contém um número serial passa pelo algoritmo de procura por Regiões Extremamente Máximamente Estáveis (MSER - *Maximally Stable Extremal Regions*) apresentada em [49].

Todavia, para maior eficiência e diferentemente de [52], que utilizou detector de bordas de Canny em paralelo, optou-se por um processamento da imagem que consiste em filtragem passa-baixa e binarização de Otsu. Assim, primeiramente, de um lado tem-se uma imagem com as regiões identificadas, do outro, a imagem após o processamento supracitado. Faz-se então uma operação *AND* lógico entre essas duas imagens, resultando em uma imagem que

passará pelas operações descritas no fluxograma da Figura 3.6 e ilustradas pela Figura 3.7.

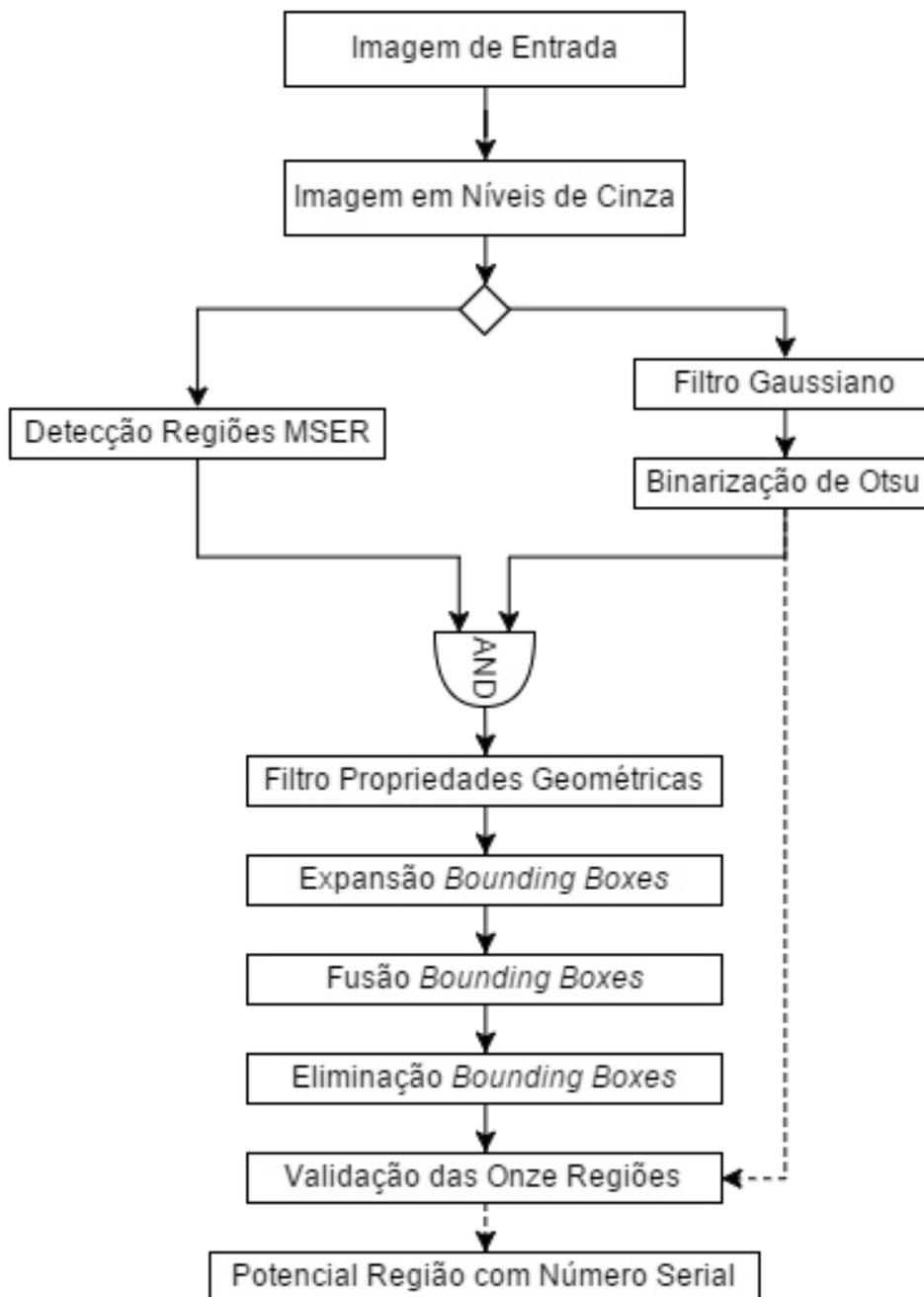
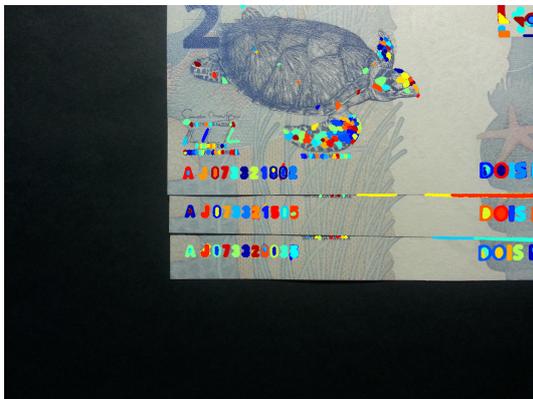


Figura 3.6: Fluxograma para a localização prováveis regiões com número serial.

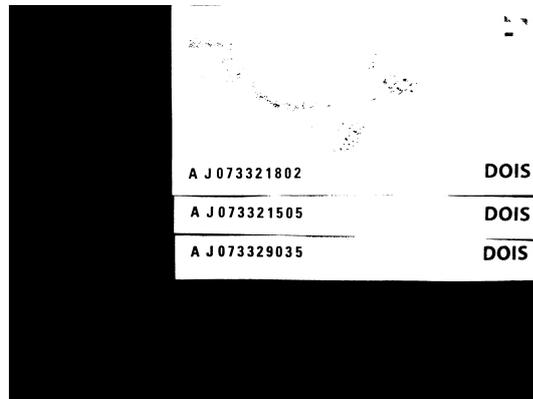
Com a finalidade de se discriminar as regiões que possuem caracteres das que não possuem, faz-se uma filtragem baseada nas seguintes propriedades geométricas, que são destacadas em [53], a saber: excentricidade, razão de aspecto, dimensão e solidez.

Logo após, com o intuito de se encontrar esses grupos de caracteres, as janelas (*bounding boxes*) das regiões não filtradas são estendidas e aquelas que se sobrepõem são fundidas. Observa-se que a extensão é maior na horizontal do que na vertical, visto que os caracteres que formam um número serial estão alinhados horizontalmente.

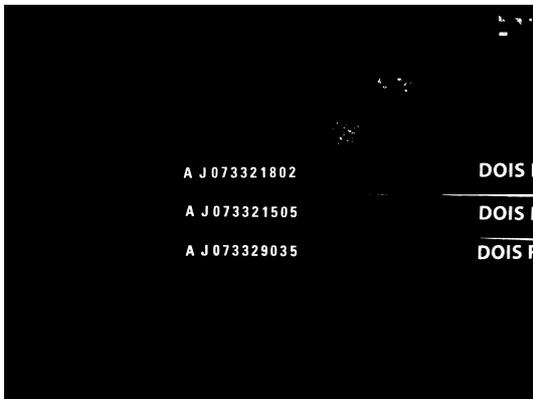
Por fim, as janelas que englobam menos de onze regiões extremas maximamente estáveis são excluídas. Importante frisar que esse número é decorrente da quantidade de caracteres



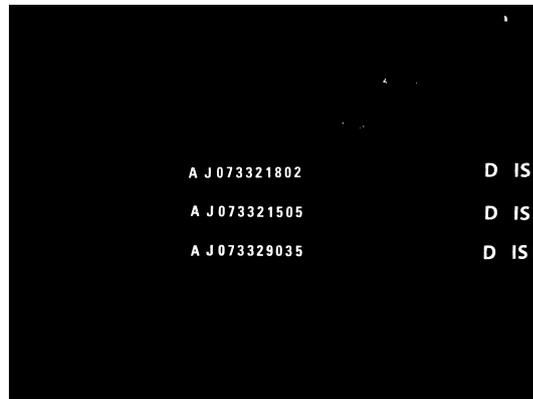
(a) Imagem com as Regiões Extremas Maximamente Estáveis.



(b) Imagem *blur* e binarização de Otsu.



(c) Imagem após operação AND.



(d) Imagem após filtragem pelas propriedades geométricas.



(e) Imagem após a fusão dos *bounding boxes*.



(f) Imagem com as prováveis regiões com números seriais detectadas.

Figura 3.7: Ilustração de imagens obtidas com o fluxograma ilustrado pela Figura 3.6.

que um número serial da cédula da Segunda Família do Real possui. Esse passo é feito com a finalidade de se evitar outros caracteres remanescentes em regiões com letras pequenas das cédulas.

Mesmo possuindo poucas regiões englobadas após todas as filtragens, essas janelas, quando levadas para a imagem alvo da validação no passo seguinte (validação das onze regiões), podem acabar englobando caracteres o suficiente para serem validados como um número serial, como pode acontecer, por exemplo, com uma região que contenha os grafismos de animais presentes em todas as cédulas da segunda família de Real.

3.3.3 Validação das Regiões com Caracteres do Número Serial

Nessa última etapa, faz-se uma validação das regiões que não foram eliminadas nas ações da etapa anterior utilizando uma contagem de picos e vales. Faz-se imperativo ressaltar que a metodologia adotada aqui pode ser estendida para quantas regiões quantas se queira, ou ainda, pode-se utilizar para um menor número de regiões. A utilização de onze se justifica pela quantidade de caracteres presentes em um número serial de uma cédula da segunda família de real.

O primeiro passo da validação consiste em, com base nas coordenadas das janelas (*bounding boxes*) obtidas na etapa anterior, recortar a imagem resultante das operações de filtragem passa baixa e binarização de Otsu.

Calcula-se então o quadrado da projeção vertical, definida como $P(x)$ desse recorte de acordo com a Equação 3.1.

$$P(x) = \left(\sum_{n=1}^k I(x, n) \right)^2, \quad (3.1)$$

em que $I(x, y)$ é o valor lógico de um pixel no recorte da imagem e k a quantidade dos pixels pertencentes ao recorte. Em seguida reescalonam-se de forma linear os valores resultantes para a faixa de zero até a unidade, obtendo $P_1(x)$ de acordo com a equação 3.2.

$$P_1(x) = \frac{P(x)}{\max P} \quad (3.2)$$

Ao final, limiariza-se esses valores de acordo com a Equação 3.3.

$$P_2(x) = \begin{cases} 0, & \text{if } P_1(x) < \delta \\ 1, & \text{caso contrário} \end{cases} \quad (3.3)$$

em que δ é um valor definido para o processo de limiarização, ajustado durante o processo de configuração do sistema. Sendo assim, espera-se que $P_2(x)$ possua onze trechos com valores unitários em sequência, cada trecho representando um caractere formador da cadeia de caracteres de um número serial. Logo, se uma região não possui esses onze trechos, não

será entendida como uma janela válida.

Mesmo havendo a etapa de eliminação pelo número de regiões, essa etapa de validação se faz necessária, pois descarta algumas partes da imagem que possuem, por exemplo, onze regiões extremas maximamente estáveis, mas que não formam um número serial. Como exemplo, cita-se as regiões em que se localizam a pata das tartarugas das cédulas de dois reais (vide Figura 3.2(b)).

A Figura 3.8 mostra, para um número serial em específico (3.8(a)), os gráficos das projeções verticais obtidos com as manipulações matemáticas descritas pelas equações acima: projeção vertical (3.8(b)), projeção vertical ao quadrado reescalada (3.8(c)) e gráfico de $P_2(x)$ (3.8(d)). Apesar de não ser escopo deste trabalho, nota-se que as projeções verticais possuem um padrão para cada caractere, o que pode ser utilizada para o seu reconhecimento. De forma mais específica, as projeções para os caracteres '3' possuem bastante semelhança.

3.4 Reconhecimento dos Caracteres

Como já posto anteriormente, para a tarefa de reconhecimento dos caracteres em si, foi utilizado o *framework* do *Tesseract*. Como este *framework* possui mais de uma forma de tratar a imagem para o reconhecimento de caracteres, detalhes já abordados na seção 2.5, optou-se pela utilização de apenas dois: tratamento da imagem como um única palavra e como uma única letra.

Para o primeiro modo, a imagem passada ao *software* consiste de uma região com toda a cadeia de caracteres que forma um número serial. Para o segundo, é passado caractere a caractere individualmente, sendo que, para isso, cada uma das regiões encontradas na etapa de validação é entendida como uma região que engloba um caractere.

Ressalta-se, contudo, que a imagem cuja região é passada ao *Tesseract* não é a colorida, mas aquela obtida após conversão para níveis de cinza, filtragem passa baixa e binarização de Otsu. Dessa forma, a tarefa de reconhecimento pelo *software* é facilitada, já não se observa muito ruído, os caracteres estão mais bem definidos e não há elementos de textura significativos.

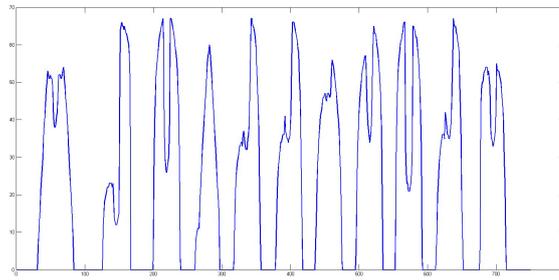
A título de ilustração, a Figura 3.9 mostra exemplos de imagens de entrada do *Tesseract*, em que é possível notar a diferença entre a imagem colorida (Figura 4.1(a)) e a imagem pré-processada (Figura 3.9(b)). Já a Figura 3.9(c) mostra um dos caracteres que compõem o número serial

Para ser ainda mais ilustrativo, as saídas obtidas (constantes em um arquivo texto) com essas entradas, foram, respectivamente, as seguintes sequências de caracteres: 'AJ0733Z1505', 'AJ073321505' e 'A'. Para as duas primeiras entradas, o modo utilizado é o que trata a imagem como um única palavra, para a terceira, como um único caractere.

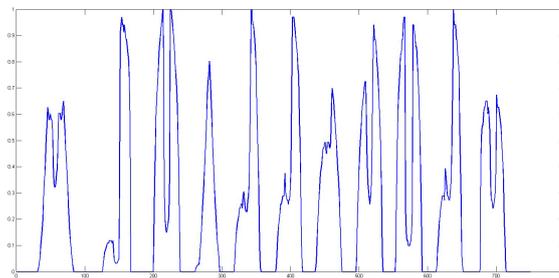
Com esse exemplo, já é possível notar uma diferença no resultado quando se faz o pré-

A J 073329035

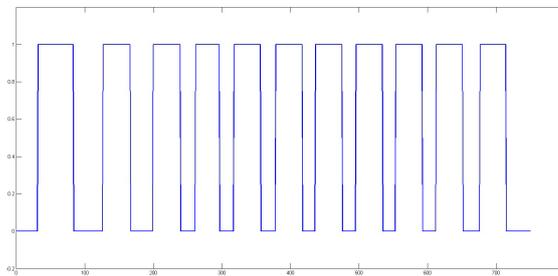
(a) Região com caracteres de um número serial.



(b) Gráfico da projeção vertical da região com o número serial.



(c) Gráfico da projeção vertical ao quadrado, reescalada, ou seja, $P_1(x)$.



(d) Gráfico de $P_2(x)$.

Figura 3.8: Ilustração dos passos para a validação das 11 regiões.

A J 073321505

A J 073321505

A

(a) Região com número serial sem qualquer processamento.

(b) Região com número serial após processamento.

(c) Região com um caractere do número serial após processamento.

Figura 3.9: Exemplos de imagens de entrada do *Tesseract*.

processamento da imagem antes de passá-la ao OCR: no primeiro caso, em vez do caractere '2', é reconhecido o caractere 'Z', erro que não ocorre com a imagem processada.

Pode-se questionar que o erro foi de apenas um caractere, mas é o bastante para alterar o número serial como um todo. É evidente que existem algumas soluções para esse contra-tempo em particular, como um pós-processamento em que se verifica se naquela posição do número serial é possível haver uma letra. Ou ainda, ser mais criterioso quando a saída for um caractere que possui outro com características de formação parecidas, o que ocorre na situação: caracteres '2' e 'Z'.

De qualquer forma, o erro ocorrido aqui não é o único que existe, o pré-processamento utilizado se justifica ainda mais para um maior número de imagens, o que será sustentado nos resultados, mostrados no Capítulo 4.

Capítulo 4

Resultados Experimentais

Primeiramente, para a obtenção dos resultados aqui relatados, é importante citar que parte das rotinas foi implementada em C++, utilizando-se a biblioteca OpenCV [54] em conjunto com o [55]. A outra parte foi desenvolvida por meio da ferramenta [56].

A partir da abordagem proposta por este trabalho para a localização e o reconhecimento dos números seriais presentes nas imagens do banco de cédulas criado, a avaliação dos resultados considerará o desempenho aferido em três situações: determinação do ângulo de inclinação, localização do número serial e melhorias no processo de reconhecimento obtidas por meio das propostas deste trabalho, que pode ser entendida como uma avaliação global do sistema.

Para evidenciar as melhorias no processo de reconhecimento, os resultados para o Cenário de Referência e para o Cenário de Controle, introduzidos no Capítulo 3, também são apresentados.

Pontua-se também que os resultados são dispostos por classe e global, situação em que são consideradas concomitantemente todas as classes na métrica de validação utilizada.

Sendo assim, este capítulo traz, na Seção 4.1, as métricas utilizadas para a avaliação dos resultados. Em seguida, são mostrados os resultados propriamente ditos, divididos de acordo com o que foi detalhado acima. Na Seção 4.2, mostra-se os resultados para a tarefa de determinação do ângulo; em 4.3, são mostrados os resultados obtidos para a localização dos números seriais. Por fim, na Seção 4.4, evidencia-se os resultados do processo como um todo, em que são considerados diferentes cenários.

4.1 Metodologia de Avaliação

Para a mensuração dos resultados, foi utilizada a *F-measure* [57] (também conhecida como norma *F1*), definida pela Eq. 4.1,

$$F\text{-measure} = 2 \cdot \frac{\textit{precision} \times \textit{recall}}{\textit{precision} + \textit{recall}}, \quad (4.1)$$

em que *precision* e *recall* são definidas por,

$$\begin{aligned} \textit{precision} &= \frac{tp}{(tp+fp)} \\ \textit{recall} &= \frac{tp}{(tp+fn)}, \end{aligned} \quad (4.2)$$

em que *tp*, *fp* e *fn* são as taxas de verdadeiros positivos (*true positive*), falsos positivos (*false positive*) e falsos negativos (*false negative*), respectivamente. Pontua-se também a existência dos verdadeiros negativos *tn*.

Aqui, é prudente ressaltar que os significados de *tp*, *fp* e *fn* são definidos para cada uma das situações em que há mensuração de resultado, conforme será visto a seguir.

4.2 Determinação da Inclinação

Neste cenário, na avaliação dos resultados obtidos considera-se que um *tp* ocorre quando o ângulo de inclinação é estimado de forma correta, havendo uma margem de erro de dois graus para cima e para baixo; um *fp*, quando essa margem é extrapolada. Além disso, as imagens da Classe 4 não são consideradas, pois, como é composta de imagens com cédulas em posições aleatórias, não há uma orientação definida para cada imagem.

A Tabela 4.1 mostra os resultados obtidos com a proposta apresentada para a estimação do ângulo de inclinação das cédulas do banco de imagens.

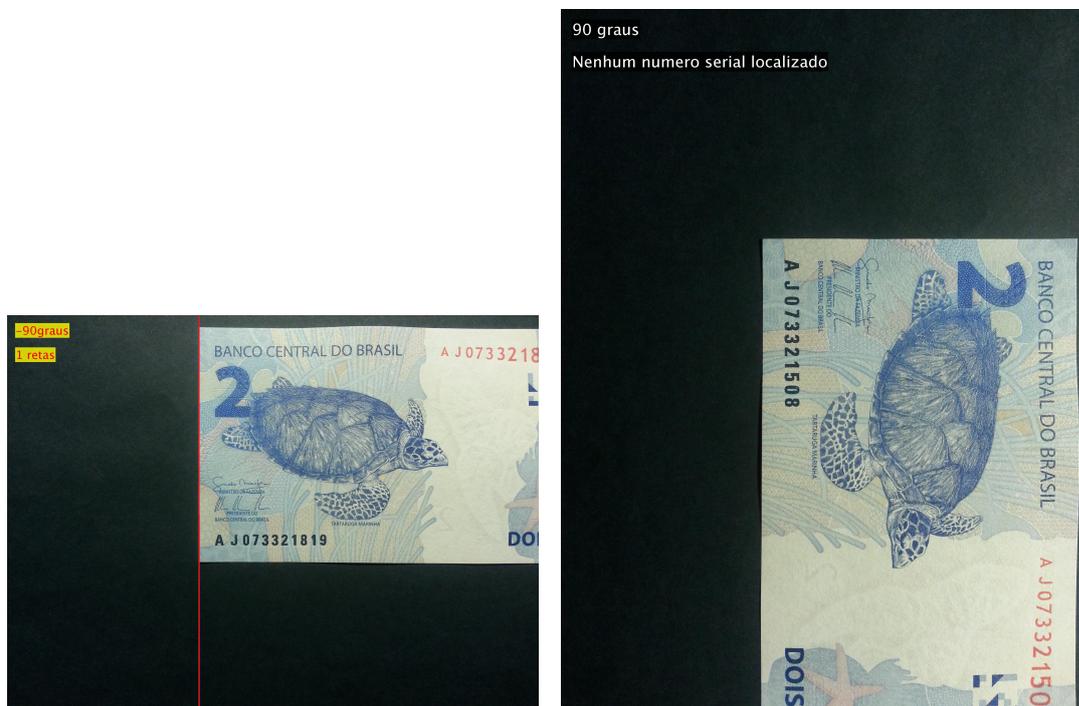
Tabela 4.1: Resultados da etapa de determinação da inclinação das cédulas

| | Classe 1 | Classe 2 | Classe 3 | Classe 4 | Global |
|------------------|-----------------|-----------------|-----------------|-----------------|---------------|
| F-measure | 0,98 | 0,96 | 0,96 | - | 0,9709 |

Analisando para cada classe, nota-se que a *F-measure* da Classe 1 é maior, o que é justificável por conter mais cédulas em uma imagem, o que aumenta as chances de retas relacionadas às bordas serem identificadas.

Além disso, destaca-se uma *F-measure* de 0,971 para o conjunto todo. Os erros na estimação do ângulo ocorrem em duas situações, ambas oriundas do primeiro passo, o de de-

terminação de retas na imagem: quando não se encontra nenhuma reta ou quando se encontra retas atreladas somente às bordas menores. A Figura 4.1 ilustra esta situação. Logo, se no primeiro passo, acha-se pelo menos uma reta atrelada à maior borda, os resultados mostram que a determinação do ângulo é correta.



(a) Imagem em que foi achada apenas uma reta relacionada à menor borda.

(b) Imagem com erro na correção do ângulo de inclinação encontrado com base na imagem à esquerda.

Figura 4.1: Exemplo de erro obtido na estimativa do ângulo de inclinação da cédula na imagem.

4.3 Localização dos Números Seriais

Na mensuração da proposta para localização dos números seriais nas imagens, a métrica proposta utilizada, define-se que um tp ocorre quando uma janela engloba um número serial; um fp , quando engloba uma região que não possui número serial; um fn , quando há um número serial, mas nenhuma janela o engloba. Um tn ocorre nas situações em que não há nenhum número serial na imagem (imagens da Classe 4), e a abordagem proposta não retorna nenhuma janela válida.

Como as propostas deste trabalho estão organizadas em cascata, os erros da etapa de estimação do ângulo de inclinação refletem nesta etapa. Contudo, para que a avaliação da proposta de localização não sofra essa influência, os resultados não levarão em conta os erros advindos da etapa anterior.

Assim, a Tabela 4.2 mostra os resultados obtidos com a proposta apresentada para a

localização dos números seriais nas imagens do banco.

Tabela 4.2: Resultados da etapa de localização dos números seriais

| | Classe 1 | Classe 2 | Classe 3 | Classe 4 | Global |
|------------------|----------|----------|----------|----------|--------|
| F-measure | 0,99 | 0,99 | 1,00 | - | 0,9930 |

Ressalta-se que não há ocorrência de falsos positivos. As ocorrências de falsos negativos ocorrem em imagens com mais de uma cédula, em que há sombras geradas pela cédula superior próximas ao número serial. A Figura 4.2 ilustra um desses casos. Isso afeta a validação das onze regiões, visto que, por conta dessas sombras, uma região que deveria ser de vale se torna região de pico, resultando em menos regiões.



Figura 4.2: Ilustração de imagem em que não se consegue localizar um número serial, por conta da sombra de uma cédula na outra.

4.4 Avaliação Global do Sistema

Nesta seção, avalia-se um cenário global para a metodologia proposta, ou seja, o processo como um todo, medindo a capacidade da abordagem proposta em reconhecer de forma correta um número serial da Segunda Família do Real presente em uma imagem. Para isso, são utilizadas duas formas de avaliação, doravante denominadas mensuração por cadeia e mensuração por caractere.

Na primeira, um número serial é tido como uma única entidade. Logo, ocorre um *tp* para toda cadeia de caracteres reconhecida de forma correta; um *fp* ocorre para toda cadeia reconhecida de forma errada, mesmo que só por um caractere, ou quando forem encontradas mais números seriais do que a imagem de fato possui; tem-se um *tn* quando são localizados a mesma quantidade de números seriais que a imagem tem; um *fn* é computado na situação em que se localiza menos cadeias do que a imagem possui.

Na abordagem de caractere, um número serial já não é mais encarado como uma única entidade, entende-se que são onze entidades, formada por caracteres do número serial. Portanto, tem-se um *tp* para cada caractere idêntico e de mesma posição na cadeia; computa-se um *fp* para cada caractere diferente do da mesma posição ou para a quantidade de caracteres que excederem a onze; um *tn* ocorre quando se acerta a quantidade de caracteres, ou seja, onze na situação em que há um número serial, zero, quando não há nenhum número serial; por fim, um *fn* é computado para cada caractere a menos do que onze. Esse tipo de mensuração encontra sentido em uma aplicação real quando pode-se ter interesse em trechos de um número serial, não apenas o serial completo.

Além disso, são avaliados três cenários, um de referência e dois de melhoria, sendo estes dois provenientes da metodologia proposta.

Como ponto de partida para a avaliação de desempenho, foi realizado uma mensuração inicial, denominada de Cenário de Referência, que ocorre quando se envia ao *Tesseract* uma imagem qualquer do banco de imagens sem processamento, a saída conterá inúmeros caracteres, já que uma cédula possui, além do número serial, algumas palavras impressas.

A avaliação deste cenário aqui será apenas qualitativa e apenas a título de exemplo, visto que as saídas obtidas não se encaixam nas métricas utilizadas aqui para a avaliação de desempenho e estão muito aquém do esperado. Para isso, a Figura 4.3 mostra um exemplo de imagem utilizada para o cenário e os caracteres reconhecidos nela.

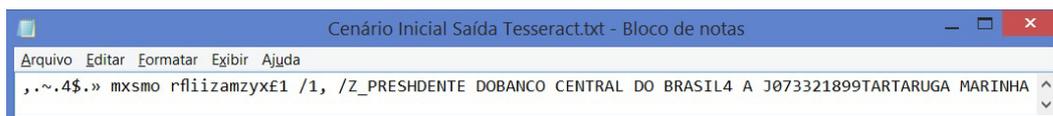
Nota-se que além dos caracteres referentes ao número serial, foram reconhecidos outros presentes na própria cédula, como, por exemplo, 'PRESIDENTE DO BANCO CENTRAL DO BRASIL'. Além disso, são reconhecidos outros inúmeros caracteres que não existem na imagem, provavelmente, o *Tesseract* interpreta erroneamente alguns detalhes da cédula como sendo um caractere.

Partindo dessas considerações e para se ter uma referência mais próxima do que se pretende mensurar, tem-se as imagens do Cenário de Controle, em que se aplica uma máscara com valores zero em todas as imagens do banco, de modo que apenas os números seriais das imagens fiquem visíveis. Ressalta-se que esse cenário já conta com uma informação *a priori* bem relevante para o resultado, a localização dos números seriais. Ou seja, tem-se um nível alto como referência. Uma vez que esse cenário já foi introduzido no Capítulo 3, uma imagem que o representa pode ser vista na Figura 3.3(b).

Por sua vez, as imagens dos cenários 1 e 2 consistem das imagens processadas com as abordagens descritas na seção 3.4. A diferença é que, para a etapa de reconhecimento



(a) Imagem de entrada para o Cenário de Referência.



(b) Arquivo texto de saída para o Cenário de Referência.

Figura 4.3: Exemplos de entrada e saída para o Cenário de Referência.

dos caracteres do cenário 1, utiliza-se o segundo modo, passa-se um caractere por vez ao *Tesseract*. Para o cenário 2, utiliza-se o primeiro modo, passa-se toda a cadeia de caracteres para o reconhecimento.

As Tabelas 4.3 e 4.4 mostram os resultados obtidos para os três cenários, considerando as abordagens de cadeia e de caractere, respectivamente.

Tabela 4.3: Resultados para mensuração por cadeia

| | Classe 1 | Classe 2 | Classe 3 | Classe 4 | Global |
|----------------------------|-----------------|-----------------|-----------------|-----------------|---------------|
| Cenário de Controle | 0,61 | 0,20 | 0,72 | - | 0,56 |
| Cenário 1 | 0,16 | 0,14 | 0,09 | - | 0,14 |
| Cenário 2 | 0,93 | 0,93 | 0,94 | - | 0,93 |

Tabela 4.4: Resultados para mensuração por caractere

| | Classe 1 | Classe 2 | Classe 3 | Classe 4 | Global |
|----------------------------|-----------------|-----------------|-----------------|-----------------|---------------|
| Cenário de Controle | 0,79 | 0,36 | 0,87 | - | 0,75 |
| Cenário 1 | 0,85 | 0,83 | 0,79 | - | 0,84 |
| Cenário 2 | 0,98 | 0,95 | 0,95 | - | 0,97 |

Ressalta-se, inicialmente, que a classe 4 não conta com medidas de *F-measure*, pois, de acordo com a modelagem para mensuração dos resultados, só foram obtidos verdadeiros negativos tn , o que mostra a total eficiência do sistema em não reconhecer um número serial em uma imagem que de fato não possui um (ausência de fp).

Entende-se que a mensuração por caracteres é mais justa, visto que não atribui um erro total do número serial caso o sistema seja capaz de acertar parte dos seus caracteres. A prova

disso é que, para todos os cenários, os resultados foram mais expressivos considerando tal abordagem.

Isso fica ainda mais evidente quando se considera o Cenário 1 que, por ter pelo menos um erro de caractere por número serial em inúmeros casos, tem-se F -*measures* de 0, 14 e 0, 84 para as mensurações por cadeia e caractere, respectivamente.

A Figura 4.4 ilustra um caso em que no reconhecimento se acerta todos os caracteres com exceção de um.



Figura 4.4: A cadeia reconhecida posta à direita do número serial difere por apenas um caractere dele, há um \ no lugar de um 8.

Considerando os cenários e tendo em mente a atipicidade citada para o Cenário 1, nota-se que o sistema traz uma melhora expressiva em relação ao cenário de referência, sobretudo quando se considera o Cenário 2. Tem-se, respectivamente, para o cenário 1 e 3, F -*measures* de 0, 56 e 0, 93 (mensuração por cadeia) e F -*measures* de 0, 75 e 0, 97 (mensuração por caractere).

Portanto, conclui-se que o Cenário 2, em que é enviado ao *framework* do *Tesseract* a cadeia completa de caracteres do número serial, é o mais eficiente para o objetivo proposto, apresentando os melhores resultados.

Capítulo 5

Conclusões e Trabalhos Futuros

Neste trabalho foi proposto um método para o reconhecimento de números seriais da Segunda Família do Real a partir de imagens. Dentre as várias questões enfrentadas para a implementação das propostas, destacam-se a falta de material na literatura voltado para as cédulas de Real; as características do fundo das cédulas de Real, com diversos artefatos, o que dificulta a localização dos números seriais e sua posterior leitura; a existência de variação de orientação e variação de escala das cédulas nas imagens.

Para atingir este objetivo, foi desenvolvido um conjunto de procedimentos que permitiram a utilização de *framework* para sistemas de reconhecimento ótico de caracteres (OCRs), como o *Tesseract*. Contudo, é importante frisar que a proposta não se limita somente a esse *framework*, esta pode ser utilizada em conjunto com qualquer outro que utilize uma imagem como entrada.

Em linhas gerais, o conjunto de procedimentos pode ser resumido em três partes. A primeira consiste da confecção de um banco de imagens de cédulas de Real da Segunda Família; na sequência, tem-se a localização de prováveis regiões com números seriais em uma imagem; por último, há a tarefa de reconhecimento dos caracteres.

Para se ter ideia do desempenho da proposta, as mensurações foram, basicamente, divididas em três situações. Na primeira situação, mensurou-se a capacidade da proposta em desfazer o rotação das cédulas em uma imagem; na segunda, foi avaliada a eficiência em se localizar regiões com números seriais em uma imagem.

Na terceira, realizou-se uma avaliação a nível global, em que a intenção é ter a ideia do funcionamento da proposta como um todo. Para isso, foram estipulados dois tipos de mensuração, por caractere e por cadeia, além de serem avaliados três cenários diferentes.

Os índices alcançados na análise de desempenho, considerando tanto as mensurações específicas como as globais, mostrou que melhorias consideráveis foram alcançadas, principalmente no tratamento de imagens de entradas rotacionadas e variando a escala. Ressalta-se também a existência de situações em que há mais de uma cédula por imagem.

Para sustentar essa afirmação, atentando-se para as mensurações globais, na mensuração por cadeia, partindo do Cenário de Controle, em que se obteve uma *F-measure* de 0,56, chegou-se a 0,93 para o Cenário 2. Analisando a mensuração por caractere, respectivamente para os mesmos cenários comparados anteriormente, partindo de 0,75, chegou-se a 0,93.

Dentre os trabalhos futuros, estão previstos testes aplicando a metodologia proposta utilizando outros *frameworks* disponíveis para o reconhecimento automático de caracteres e avaliação dos índices obtidos. Assim, é possível a análise da proposta deste trabalho como um todo, inclusive a escolha do Tesseract.

Há a sugestão também de desenvolvimento e concepção de um protótipo, em parceria com instituições de investigação e financeiras, de um sistema a ser empregado no processo de rastreamento de cédulas em território nacional.

Aqui, faz-se importante repisar a importância de um sistema desse tipo tanto na esfera criminal como na econômica, já que possui potencial para a determinação do fluxo de cédulas em território nacional, para detecção de contrafação, para auxílio ao combate ao pagamento de propina, entre outros.

Referências Bibliográficas

- [1] BRASIL, B. C. do. *Cartilha de Treinamento - Segunda Família do Real*. <https://www.bcb.gov.br/novasnotas/assets/downloads/material-apoio/2e5/Cartilha.pdf>. (Acessado em 23/11/2016).
- [2] BOROVIKOV, E. A survey of modern optical character recognition techniques. *arXiv preprint arXiv:1412.4183*, 2014.
- [3] <http://portal.ead.senasp.gov.br/>. (Acessado em 22/11/2016).
- [4] SUTHERLAND, E. H.; GEIS, G. *White collar crime*. [S.l.]: Dryden Press New York, 1949.
- [5] SILVA, P. M. d.; POLASTRO, M. d. C. An overview of nudetective forensic tool and its usage to combat child pornography in brazil. In: *XLIII Jornadas Argentinas de Informática e Investigación Operativa (43JAIIO)-VI Workshop de Seguridad Informática (Buenos Aires, 2014)*. [S.l.: s.n.], 2014.
- [6] FEDERAL, M. P. <http://lavajato.mpf.mp.br/atuacao-na-la-instancia/resultados/a-lava-jato-em-numeros-1>. (Acessado em 22/11/2016).
- [7] FEDERAL, M. P. junho 2016. <cdn-wp.oantagonista.net/2016/06/..../ntegra-da-denuncia-da-Operacao-Saqueador.odt>. (Acessado em 23/11/2016).
- [8] FALSIFICAÇÃO de Cédulas. <http://www.bcb.gov.br/pt-br/MECIRESTFALSAS>. (Accessed on 15/07/2016).
- [9] GAI, S. et al. New banknote number recognition algorithm based on support vector machine. In: IEEE. *2013 2nd IAPR Asian Conference on Pattern Recognition*. [S.l.], 2013. p. 176–180.
- [10] GUO, Y.-p.; DING, W.-s. An us dollar number recognition system for location and division based on a projection method [j]. *Aeronautical Computing Technique*, v. 5, p. 014, 2007.

- [11] FENG, B.-Y. et al. Automatic recognition of serial numbers in bank notes. *Pattern Recognition*, Elsevier, v. 47, n. 8, p. 2621–2634, 2014.
- [12] FENG, B.-Y. et al. Extraction of serial numbers on bank notes. In: IEEE. *2013 12th International Conference on Document Analysis and Recognition*. [S.l.], 2013. p. 698–702.
- [13] FENG, B.-Y. et al. Part-based high accuracy recognition of serial numbers in bank notes. In: SPRINGER. *IAPR Workshop on Artificial Neural Networks in Pattern Recognition*. [S.l.], 2014. p. 204–215.
- [14] ZHOU, X. et al. Icdar 2015 text reading in the wild competition. *arXiv preprint arXiv:1506.03184*, 2015.
- [15] KARATZAS, D. et al. Icdar 2015 competition on robust reading. In: IEEE. *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*. [S.l.], 2015. p. 1156–1160.
- [16] YAO, C. et al. Incidental scene text understanding: Recent progresses on icdar 2015 robust reading competition challenge 4. *arXiv preprint arXiv:1511.09207*, 2015.
- [17] OPTICAL Character Recognition (OCR) - How it works. <https://www.nicomsoft.com/optical-character-recognition-ocr-how-it-works/>. (Accessed on 15/07/2016).
- [18] TAN, Z. R.; TIAN, S.; TAN, C. L. Using pyramid of histogram of oriented gradients on natural scene text recognition. In: IEEE. *2014 IEEE International Conference on Image Processing (ICIP)*. [S.l.], 2014. p. 2629–2633.
- [19] GONZÁLEZ, A. et al. A character recognition method in natural scene images. In: IEEE. *Pattern Recognition (ICPR), 2012 21st International Conference on*. [S.l.], 2012. p. 621–624.
- [20] GAO, C.; HUANG, X. Character recognition in low quality document images using local and global features. In: IEEE. *Image and Signal Processing (CISP), 2014 7th International Congress on*. [S.l.], 2014. p. 676–680.
- [21] YE, Q.; DOERMANN, D. Text detection and recognition in imagery: A survey. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 37, n. 7, p. 1480–1500, 2015.
- [22] SMITH, R. An overview of the tesseract ocr engine. In: *Proceedings of the Ninth International Conference on Document Analysis and Recognition - Volume 02*. Washington, DC, USA: IEEE Computer Society, 2007. (ICDAR '07), p. 629–633. ISBN 0-7695-2822-8. Disponível em: <<http://dl.acm.org/citation.cfm?id=1304596.1304846>>.

- [23] PATEL, C.; PATEL, A.; PATEL, D. Optical character recognition by open source ocr tool tesseract: A case study. *International Journal of Computer Applications*, Foundation of Computer Science, v. 55, n. 10, 2012.
- [24] NAKAMOTO, S. *Bitcoin: A peer-to-peer electronic cash system*. 2008.
- [25] LEI nº 4.595, de 31 de dezembro de 1964. Dispõe sobre a Política e as Instituições Monetárias, Bancárias e Creditícias, Cria o Conselho Monetário Nacional e dá outras providências. *Diário Oficial da União*, 1964.
- [26] PROGRAM, U. C. E. *The Seven Denominations*. <https://uscurrency.gov/seven-denominations>. (Accessed on 24/11/2016).
- [27] BANK, E. C. *Euro Banknotes - Security Features*. <http://www.new-euro-banknotes.eu/Euro-Banknotes/Security-features/LOOK/THE-NEW-%E2%82%AC20>. (Accessed on 24/11/2016).
- [28] LEI nº 8.880, de 27 de maio de 1964. Dispõe sobre o Programa de Estabilização Econômica e o Sistema Monetário Nacional, institui a Unidade Real de Valor (URV) e dá outras providências. *Diário Oficial da União*, 1964.
- [29] COMUNICADO nº 020392. Divulga o lançamento das cédulas de 50 e 100 reais da Segunda Família de Cédulas do Real, de circulação comum. *Banco Central do Brasil*, 2010.
- [30] <https://www.bcb.gov.br/htms/mecir/cedulas/appdinheirobr/index.asp>. (Acessado em 27/09/2016).
- [31] HEIJ, H. de; GELDER, A. van. Numbers of banknotes. what is their use? *Keesing Journal of Documents and Identity*, DNB, v. 20, p. 14–18, 2006.
- [32] <https://en.eurobilltracker.com>. (Acessado em 15/09/2016).
- [33] <http://www.wheresgeorge.com/>. (Acessado em 15/09/2016).
- [34] <http://www.cdn-money.ca/>. (Acessado em 15/09/2016).
- [35] <http://www.whereswilly.com/>. (Acessado em 15/09/2016).
- [36] LI, H.; SHEN, C. Reading car license plates using deep convolutional neural networks and lstms. *arXiv preprint arXiv:1601.05610*, 2016.
- [37] PAN, Y.-F.; LIU, C.-L. Performance evaluation for text localization algorithms: An empirical study. In: IEEE. *Pattern Recognition (CCPR), 2010 Chinese Conference on*. [S.l.], 2010. p. 1–5.

- [38] LOWE, D. G. Object recognition from local scale-invariant features. In: IEEE. *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*. [S.l.], 1999. v. 2, p. 1150–1157.
- [39] BAY, H.; TUYTELAARS, T.; GOOL, L. V. Surf: Speeded up robust features. In: SPRINGER. *European conference on computer vision*. [S.l.], 2006. p. 404–417.
- [40] IFEACHOR, E. C.; JERVIS, B. W. *Digital signal processing: a practical approach*. [S.l.]: Pearson Education, 2002.
- [41] SOLOMON, C.; BRECKON, T. *Fundamentals of Digital Image Processing: A practical approach with examples in Matlab*. [S.l.]: John Wiley & Sons, 2011.
- [42] OTSU, N. A threshold selection method from gray-level histograms. *Automatica*, v. 11, n. 285-296, p. 23–27, 1975.
- [43] CANNY, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, IEEE Computer Society, Washington, DC, USA, v. 8, n. 6, p. 679–698, jun. 1986. ISSN 0162-8828. Disponível em: <<http://dx.doi.org/10.1109/TPAMI.1986.4767851>>.
- [44] DUDA, R. O.; HART, P. E. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, ACM, v. 15, n. 1, p. 11–15, 1972.
- [45] CHANTARA, W.; HO, Y.-S. Object detection based on fast template matching through adaptive partition search. In: IEEE. *Computer Science and Software Engineering (JCSSE), 2015 12th International Joint Conference on*. [S.l.], 2015. p. 1–6.
- [46] BAY, H. et al. Speeded-up robust features (surf). *Computer vision and image understanding*, Elsevier, v. 110, n. 3, p. 346–359, 2008.
- [47] GOMEZ, L.; KARATZAS, D. Object proposals for text extraction in the wild. In: IEEE. *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*. [S.l.], 2015. p. 206–210.
- [48] NEUMANN, L.; MATAS, J. Real-time scene text localization and recognition. In: IEEE. *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. [S.l.], 2012. p. 3538–3545.
- [49] MATAS, J. et al. Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, Elsevier, v. 22, n. 10, p. 761–767, 2004.
- [50] <https://github.com/tesseract-ocr/tesseract>. (Acessado em 11/11/2016).
- [51] SMITH, R. A simple and efficient skew detection algorithm via text row accumulation. In: IEEE. *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*. [S.l.], 1995. v. 2, p. 1145–1148.

- [52] CHEN, H. et al. Robust text detection in natural images with edge-enhanced maximally stable extremal regions. In: IEEE. *2011 18th IEEE International Conference on Image Processing*. [S.l.], 2011. p. 2609–2612.
- [53] GONZALEZ, A. et al. Text location in complex images. In: IEEE. *Pattern Recognition (ICPR), 2012 21st International Conference on*. [S.l.], 2012. p. 617–620.
- [54] <http://opencv.org/about.html>. (Acessado em 24/11/2016).
- [55] <https://www.qt.io/ide/>. (Acessado em 24/11/2016).
- [56] <https://www.mathworks.com/products/matlab/>. (Acessado em 24/11/2016).
- [57] RIJSBERGEN, C. J. van. *Information Retrieval*. [S.l.: s.n.], 1979.