

**UNIVERSIDADE PRESBITERIANA MACKENZIE**

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**Rafael Félix Alves**

**UMA METODOLOGIA DE BINARIZAÇÃO PARA ÁREAS DE IMAGENS DE  
CHEQUE UTILIZANDO ALGORITMOS DE APRENDIZAGEM SUPERVISIONADA**

São Paulo

2015

**UNIVERSIDADE PRESBITERIANA MACKENZIE**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**Rafael Félix Alves**

**UMA METODOLOGIA DE BINARIZAÇÃO PARA ÁREAS DE IMAGENS DE  
CHEQUE UTILIZANDO ALGORITMOS DE APRENDIZAGEM SUPERVISIONADA**

Documento de dissertação apresentado ao Programa de Pós-Graduação em Engenharia Elétrica e Computação da Universidade Presbiteriana Mackenzie, como requisito para a obtenção do título de Mestre em Engenharia Elétrica e Computação.

**Orientador: Prof. Dr. Leandro Nunes de Castro**

**Coorientador: Prof. Dr. Leandro Augusto da Silva**

São Paulo

2015

A474m Alves, Rafael Felix

Uma metodologia de binarização para áreas de imagens de cheque utilizando algoritmos de aprendizagem supervisionada / Rafael Felix Alves - 2015.

63f.: il., 30 cm

Dissertação (Mestrado em Engenharia Elétrica) –  
Universidade Presbiteriana Mackenzie, São Paulo, 2015.

Orientação: Prof. Dr. Leandro Nunes de Castro

Bibliografia: f. 58-63

1. Binarização. 2. Aprendizagem de máquina. 3. Classificação. 4. Imagens de cheque. 5. Montante de cortesia. 6. CMC-7. 7. Redes neurais artificiais. 8. Árvore de decisão. Título.

CDD 006.42

Dedico este Mestrado à Família Félix, pelo incentivo e apoio ao longo da minha vida. Especialmente a Mêrces Félix.

## AGRADECIMENTOS

Meus sinceros agradecimentos,

Ao meu orientador Prof. Dr. Leandro Nunes de Castro, pelas diversas conversas, conselhos e por todo o tempo cedido para a minha formação.

Ao meu coorientador Prof. Dr. Leandro Augusto da Silva, pelo companheirismo e conselhos ao longo do processo.

Um especial agradecimento ao meu amigo e Prof. Dr. Renato Dourado Maia por ter me apresentado ao Laboratório de Computação Natural e ao meu orientador.

Ao Macedo, Lincoln e toda a equipe da empresa Sincronica Sistemas Integrados LTDA pela confiança depositada.

À minha querida família que me apoia e me ajuda incondicionalmente.

À Letícia Andrade pela importante ajuda na reta final desse projeto de Mestrado.

Ao Laboratório de Computação Natural da Universidade Presbiteriana Mackenzie e seus membros pela amizade e compartilhamento de ideias. Agradeço em especial a Dávila Patrícia pelo companheirismo e paciência nos projetos em conjunto.

À CAPES, FAPESP, CNPq, MackPesquisa e Sincronica Sistemas Integrados LTDA pelo suporte financeiro.

## RESUMO

O processo de binarização de imagens consiste na transformação de uma imagem colorida em uma nova imagem com apenas duas cores: uma que representa o fundo, outra o objeto de interesse. Este processo é uma importante etapa de diversas aplicações modernas, como a Compensação de Cheque, o Reconhecimento Ótico de Caracteres (do inglês *Optical Character Recognition*) e o Reconhecimento de Texto Manuscrito (do inglês *Handwritten Recognition*, HWR). Dado que melhorias no processo automático de binarização de imagens representam impactos diretos nas aplicações que dependem desta etapa o presente trabalho propõe uma metodologia para realizar a binarização automática de imagens. A proposta realiza a binarização de forma automática baseado no uso de algoritmos de aprendizagem supervisionada, tais como redes neurais artificiais e árvore de decisão. O processo como um todo consiste das seguintes etapas: construção do banco de imagens; extração da região de interesse; construção da matriz de padrões; rotulação dos padrões; amostragem da base; e treinamento do classificador. Resultados experimentais são apresentados utilizando uma base de imagens de cheques de bancos brasileiros (CMC-7 e montante de cortesia) e a base de imagens da competição DIBCO 2009. Em conclusão, a metodologia proposta apresentou-se competitiva aos métodos da literatura destacando-se em aplicações onde o processamento de imagens está restrito a uma categoria de imagens, como é o caso das imagens de cheques de bancos brasileiros. A presente metodologia apresenta resultados experimentais entre as três primeiras posições e melhores resultados em relação a medida *F-Measure* quando comparada com as demais.

**Palavras-chave:** *binarização; aprendizagem de máquina; classificação; imagens de cheque; montante de cortesia; CMC-7; redes neurais artificiais; árvore de decisão.*

## ABSTRACT

The process of image binarization consists of transforming a color image into a new one with only two colors: black and white. This process is an important step for many modern applications such as Check Clearance, Optical Character Recognition and Handwriting Recognition. Improvements in the automatic process of image binarization represent impacts on applications that rely on this step. The present work proposes a methodology for automatic image binarization. This methodology applies supervised learning algorithms to binarize images and consists of the following steps: images database construction; extraction of the region of interest; patterns matrix construction; pattern labelling; database sampling; and classifier training. Experimental results are presented using a database of Brazilian bank check images and the competition database DIBCO 2009. In conclusion, the proposal demonstrated to be superior to some of its competitors in terms of accuracy and F-Measure.

**Keywords:** *binarization; machine learning; classification; check images; courtesy amount; CMC-7; artificial neural network; decision tree;*

## Lista de Figuras

Figura 1: Exemplo de corte do histograma, a) imagem a ser binarizada; b) histograma da imagem. ....	9
Figura 2: Fluxo macro de compensação de cheques no Brasil Fonte: (FEBRABAN, 2014) ..	15
Figura 3: Representação da imagem nos eixos $x$ e $y$ . ....	15
Figura 4: Exemplos de cheques de bancos brasileiros. ....	17
Figura 5: <i>Layout</i> de divisão de cheques de bancos brasileiros. ....	18
Figura 6: Exemplos de binarização por limiar. (a) Imagem de fundo complexo em escala de cinza. (b) Imagem binarizada por um método global. (c) Imagem binarizada por um método local. ....	19
Figura 7: Fluxo do método proposto. ....	25
Figura 8: Metodologia pixel a pixel para binarização de imagens ....	26
Figura 9: Exemplo de imagem do montante de cortesia de cheques de bancos brasileiros. ....	27
Figura 10: Representação do pixel $p(x, y)$ dada uma janela de dimensão $5 \times 5$ com base na vizinhança de Moore. ....	28
Figura 11: Vizinhança extrapola dimensão da imagem $I$ . ....	29
Figura 12: Rotulação de cada pixel $p(x, y)$ utilizando imagem binarizada manualmente, a) imagem em escala de cinza e b) imagem binarizada manualmente; ....	30
Figura 13: Nível de homogeneidade baseado no desvio padrão da vizinhança, a) imagem em escala de cinza; b) e c) região com pixels homogêneos; e d) região com pixels heterogêneos. ....	32
Figura 14: Exemplo de tipos de <i>pixels</i> para cálculo das medidas de avaliação.....	35
Figura 15: Exemplos de oito imagens de montantes de cortesia de diferentes cheques de bancos brasileiros. ....	38



Figura 16: Exemplos de quatro imagens de CMC-7 de diferentes cheques de bancos brasileiros. .....	39
Figura 17: Aplicação da extração da <i>RoI</i> a partir de imagens de cheques de bancos brasileiros. .....	40
Figura 18: Resultado visual do processo de binarização. (a) imagem em escala de cinza. (b) Otsu (1979), (c) Niblack (1986), (d) Sauvola (2000), (e) MLP, (f) SVM, (g) DT.....	46
Figura 19: Resultado visual do processo de binarização. (a) imagem em escala de cinza. (b) Otsu (1979), (c) Niblack (1986), (d) Sauvola (2000), (e) MLP, (f) SVM, (g) DT.....	49

## Lista de Tabelas

Tabela 1: Principais trabalhos envolvendo a binarização automática de imagens de cheques.	21
Tabela 2: Resultados utilizando medidas de avaliação para o exemplo.	35
Tabela 3: Acurácia ( <i>Acc</i> ), Precisão ( <i>Pr</i> ) e Revogação ( <i>Re</i> ) dos classificadores utilizando amostragem das imagens de montante de cortesia.	43
Tabela 4: Comparação da binarização das imagens de montante de cortesia utilizando a metodologia proposta.	44
Tabela 5: Acurácia ( <i>Acc</i> ), Precisão ( <i>Pr</i> ) e Revogação ( <i>Re</i> ) dos classificadores utilizando amostragem da base para imagens de CMC-7.	47
Tabela 6: Comparação da binarização das imagens de CMC-7 utilizando a metodologia proposta e métodos da literatura.	48
Tabela 7: Resultados avaliativos para todos os métodos submetidos ao DIBCO obtidos em Gatos (2009) e metodologia proposta.	51

## Sumário

1. INTRODUÇÃO .....	8
1.1 Objetivos .....	12
1.1.1 Objetivo Geral .....	12
1.1.2 Objetivos Específicos .....	13
1.2 Organização do Documento.....	13
2. REFERENCIAL TEÓRICO .....	14
1.3 Compensação de Cheques de Bancos Brasileiros.....	14
1.4 Imagens Digitais .....	15
1.4.1 Imagens Digitais de Cheques de Bancos Brasileiros.....	16
1.5 Binarização de Imagens Digitais .....	18
1.5.1 Binarização de Imagens de Cheques de Bancos.....	20
3. UMA METODOLOGIA DE BINARIZAÇÃO UTILIZANDO ALGORITMOS DE APRENDIZAGEM SUPERVISIONADA .....	25
1.5.2 Construção do Banco de Imagens .....	26
1.5.3 Extração da Região de Interesse.....	26
1.5.4 Construção da Matriz de Padrões .....	27
1.5.5 Rotulação dos Padrões.....	29
1.5.6 Amostragem da Base .....	31
1.5.7 Treinamento do Classificador.....	32
1.5.8 Medidas de Avaliação .....	32
4. APLICAÇÃO DA METODOLOGIA PROPOSTA E RESULTADOS EXPERIMENTAIS .....	36
1.6 Aplicação da Metodologia Proposta a Imagens De Cheques De Bancos Brasileiros	36
1.6.1 Caracterização do Problema .....	36
1.6.2 Aplicação e Análise Preliminar da Metodologia Proposta.....	39
1.6.3 Resultados Experimentais e Discussões .....	42
1.7 Aplicação da Metodologia Proposta a Imagens Do Banco de Imagens do DIBCO 2009	49
5. CONCLUSÕES E TRABALHOS FUTUROS .....	52
REFERÊNCIAS BIBLIOGRÁFICAS .....	54

## 1. INTRODUÇÃO

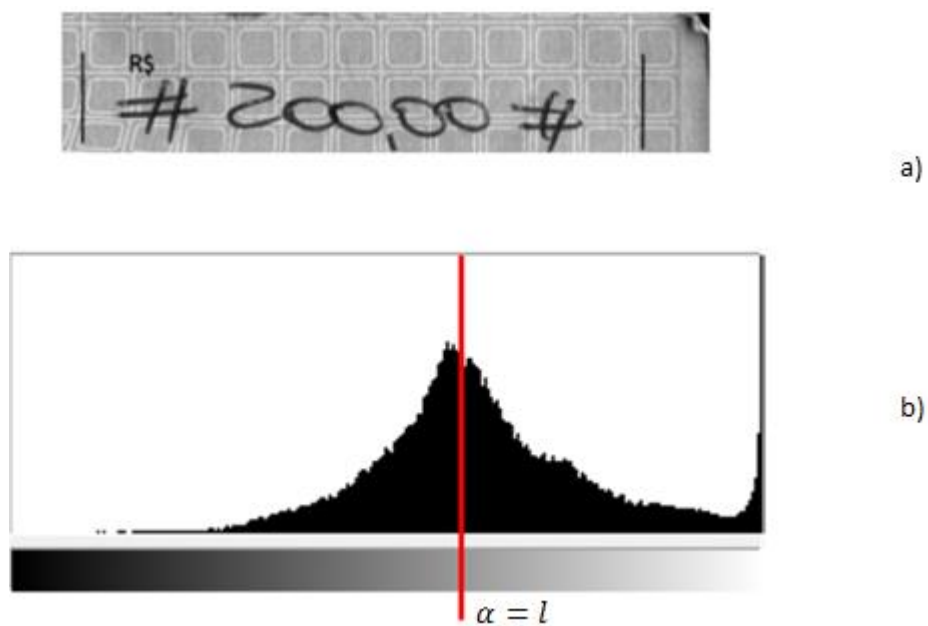
A *binarização* é uma técnica de processamento de imagens digitais correspondente a uma etapa fundamental para diversas aplicações de *visão computacional*, presente principalmente em aplicações como segmentação de imagens, detecção de objetos e reconhecimento de padrões (GONZALEZ e WOODS, 1992; RABELO, ZANCHETTIN, *et al.*, 2011). O processo de binarização consiste em converter as tonalidades de pixel de uma imagem, em escala de cinza ou padrão de cores, para uma imagem com tonalidades de pixel em duas cores, que representam o fundo e objeto da imagem. As cores escolhidas convencionalmente para denotar fundo e objeto de uma imagem são as cores branco e preto, respectivamente.

Para exemplificar a binarização, considere uma aplicação computacional que reconheça automaticamente os caracteres de uma imagem referente à página de livro. Para que o reconhecimento seja possível é necessária a separação dos caracteres das partes que formam o fundo da página, detalhes gráficos, marcas de caneta, ilustrações, dentre outros. O processo de binarização é responsável por realizar a filtragem do objeto desejado, neste exemplo os caracteres, que devem ser utilizados nas etapas subsequentes do processo computacional, como por exemplo, o reconhecimento do caractere. Portanto, uma vez que a binarização é o processo responsável pela filtragem da informação a ser processada subsequentemente, a falha nesse processo tem impacto direto no resultado final desejado.

Para uma imagem em escala de cinza, considere uma função bidimensional de intensidade da luz  $f(x,y)$ , na qual  $x$  e  $y$  são coordenadas especiais e o valor de  $f$  em qualquer ponto  $(x, y)$  é proporcional ao brilho da imagem naquele ponto. A imagem digital é uma imagem discretizada tanto em coordenadas espaciais, quanto em brilho, podendo ser representada por uma matriz cujos índices de linha e coluna correspondem a um ponto na imagem e o valor do elemento da matriz define o brilho naquele ponto (GONZALEZ e WOODS, 1992). Para a imagem digital em escala de cinza, cada pixel possui um valor escalar entre 0 e 255. Os métodos tradicionais de binarização são baseados em algoritmos que, a partir de um valor de corte ou limiar que deve estar no mesmo domínio da escala da imagem, é capaz de separar um objeto da imagem do seu fundo. Logo, dado um limiar  $\sigma$ , todos os pixels da imagem com intensidade de cor

igual ou superior a esse limiar pertencem ao fundo e inferior a ele pertencem ao objeto (de interesse) (GONZALEZ e WOODS, 1992).

Considere a imagem  $I$  que possui a intensidade dos pixels definidos pela função  $f(x,y)$  e a função  $B(x,y)$ , expressa na Equação (1), que representa a função de binarização de cada pixel da imagem  $I$ . A Figura 1(a) ilustra uma parte de uma imagem de cheque de um banco brasileiro, chamada de montante de cortesia, e (b) seu histograma. O histograma informa, para cada nível de cinza da imagem, que varia entre 0 e 255, a quantidade de pixels que existe com a mesma tonalidade. Para cada intensidade de pixel  $f(x,y)$  da Figura 1(a), o valor será transformado em 1 caso a intensidade seja maior ou igual ao limiar  $l$ , e 0 caso contrário, como mostra a Equação (1). Desta forma os valores abaixo do limiar  $l$  no histograma representam os *pixels* objeto, da mesma forma os valores acima representam os *pixels* fundo.



**Figura 1:** Exemplo de corte do histograma, a) imagem a ser binarizada; b) histograma da imagem.

$$b(x,y) = \begin{cases} 1, & f(x,y) \geq l \\ 0, & f(x,y) < l \end{cases} \quad (1)$$

Os métodos de binarização por limiarização são classificados em duas abordagens: global ou local (GONZALEZ e WOODS, 1992). A abordagem global consiste na definição de um único limiar para todos os *pixels* da imagem e, desta maneira, cada pixel da imagem é comparado ao valor de limiar. A abordagem local encontra valores de limiares adaptáveis para cada conjunto de pixels (RABELO, ZANCHETTIN, *et al.*, 2011; ALGINAHI, SID-AHMED e AHMADI, 2004).

Comumente, os algoritmos de binarização por limiarização estão em busca de um ou mais valores suficientemente capazes de realizar o corte do histograma dos tons de cinza de uma imagem. O objetivo destes valores é minimizar ruídos no fundo da imagem e maximizar a detecção do objeto. Apesar do conceito básico de binarização estar ligado ao processo de corte de histograma por limiar, existe na literatura algoritmos com abordagens que não são baseadas em limiar. Dentre estas abordagens destacam-se aquelas que são baseadas em medidas de agrupamento (KIM, KIM, *et al.*, 2001), metodologias que utilizam informações como entropia do histograma (HORNG, 2011), metodologias que utilizam atributos de imagem (HONGGANG, GUANG, *et al.*, 2005), abordagens com informações espaciais da imagem (DE JESUS e FACON, 2000) e as que utilizam características locais (RABELO, ZANCHETTIN, *et al.*, 2011).

Em função da facilidade de aquisição de imagens e, conseqüentemente, do aumento no volume deste tipo de dado, inúmeras aplicações com imagens e que necessitam do processo de binarização têm surgido na literatura como, por exemplo, reconhecimento de texto em imagens digitais, placas de automóveis, dentre outros (SEZGIN e SANKUR, 2004; GAO, LI, *et al.*, 2011; KARRAS, 2011).

Um exemplo de aplicação moderna que inclui a etapa de binarização em imagens digitais é o processo conhecido como *compensação automática de cheques* (FEBRABAN, 2014; AHMAD e MAHMOUD, 2012). A compensação de cheques é um processo crítico que possibilita que o cheque de um determinado banco possa ser cobrado através de outro banco, bastando que o beneficiário o deposite em sua conta. O uso de imagens nesse processo elimina possibilidades de clonagem, extravio, perdas e roubos dos cheques físicos (FEBRABAN, 2014). Dentro do fluxo do processo de compensação automática de cheques, a binarização é uma etapa fundamental para os processos subsequentes, como reconhecimento do valor monetário do cheque, reconhecimento de assinatura, CMC-7, dentre outras informações contidas nos

cheques brasileiros. A existência de diversos elementos compondo o fundo dos cheques de bancos brasileiros, diversos tipos de transcrição e demais detalhes das folhas de cheque demandam métodos de binarização automática e adaptáveis capazes de realizar a correta separação entre objeto e fundo. A binarização deste tipo de imagem coopera com processos de reconhecimento de caracteres e textos manuscritos (do inglês *optical character recognition* - OCR e *handwritten character recognition* - HCR) e, conseqüentemente, traz melhoria associada ao processo de compensação de cheques por imagem, instituído pela Federação Brasileira de Bancos – FEBRABAN (FEBRABAN, 2014).

Considere como um problema que envolve o uso do OCR a digitalização de um livro antigo que não possua uma versão em formatos convencionais como *PDF*, *DOC* ou *TXT*. Dentre as diversas maneiras de se criar uma versão digital deste livro está a digitação de todo o conteúdo do livro por um humano. Avalia-se que o processo de digitalização de um livro por meio de sua digitação por um humano pode ser aceitável para um livro, porém, para um conjunto de livros este processo já pode ser considerado lento e enfadonho. Logo, percebe-se a necessidade de uma metodologia que seja capaz de realizar o reconhecimento dos caracteres e conseqüentemente do texto de maneira automática e acurada. A utilização de vídeos e imagens como entradas dos processos de reconhecimento de caracteres apresentou-se como uma alternativa viável para a resolução deste problema. Uma vez que o processo de aquisição de imagem tornou-se comum por meio destas tecnologias, possibilitou-se realizar o processo computacional de digitalização de livros antigos, identificação de texto e outros utilizando algoritmos de visão computacional (MORI, NISHIDA e YAMADA, 1999).

O reconhecimento óptico de caracteres tem contribuído com diversas aplicações em várias áreas, tais como, auxílio para pessoas com cegueira, digitalização de jornais, artigos e livros, reconhecimento automático de documentos, processo de compensação de cheque, aplicações de reconhecimento automático de placas de carro, dentre outras (GOVINDAN e SHIVAPRASAD, 1990). Recentes pesquisas em reconhecimento de caracteres estão baseadas em três ramos, são eles: (1) os algoritmos utilizados para o reconhecimento dos padrões de caracteres, destaque para as aplicações utilizando *Deep Learning* (KANNAN e SUBRAMANIAN, 2015); (2) a variação do tipo dos caracteres, como, por exemplo, caracteres manuscritos e caracteres digitados (WANG, FU, *et al.*, 2015); e (3) caracteres de diferentes línguas, por exemplo os caracteres tâmil (KANNAN e SUBRAMANIAN, 2015) e caracteres

chineses (WANG, FU, *et al.*, 2015). O processo de binarização influencia diretamente o resultado do OCR e HWR.

Neste contexto, o presente trabalho propõe uma metodologia de binarização utilizando técnicas de reconhecimento de padrões e características locais da imagem. Dada uma técnica de reconhecimento de padrões, cada *pixel* da nova imagem binarizada é definido a partir da utilização de informações do *pixel* e sua vizinhança na imagem de origem. Com o objetivo de avaliar a metodologia proposta foram realizados testes utilizando duas áreas de imagens de folhas de cheques de bancos brasileiros, são elas: montante de cortesia; e CMC-7. As imagens de cheques foram obtidas por meio de um banco de imagens privado fornecido pela empresa Sincronica Serviços Integrados Limitada. O montante de cortesia consiste na região do cheque designada para a transcrição do valor monetário e o CMC-7 consiste na região de tarja magnética de identificação do cheque.

Foram realizados testes utilizando as técnicas de reconhecimento de padrões para realizar classificação, incluindo Redes Neurais Artificiais (ALGINAHI, SID-AHMED e AHMADI, 2004; DE CASTRO, 2006; DUDA, HART e STORK, 2000), *Support Vector Machines (SVM)* (DE CASTRO, 2006; DUDA, HART e STORK, 2000; CHANG e LIN, 2011), e Árvores de Decisão (DUDA, HART e STORK, 2000). A presente metodologia mostrou capacidade de realizar a separação do objeto em relação ao fundo e manter as principais características da imagem para outros processamentos fundamentais na compensação de cheque, como OCR e HWR. A metodologia proposta foi testada utilizando a base de imagens DIBCO (2009) obtida a partir da Conferência Internacional em Análise e Reconhecimento de Documentos (ICDAR) (GATOS, NTIROGIANNIS e PRATIKAKIS, 2009) e apresentou resultados competitivos com os métodos utilizados na competição.

## **1.1 OBJETIVOS**

### **1.1.1 Objetivo Geral**

Propor uma representação padronizada de um pixel que possibilite a utilização de técnicas de reconhecimento de padrões para realizar a binarização automática de imagens sem a necessidade de definição de um limiar.



### 1.1.2 Objetivos Específicos

- Propor uma representação padronizada para cada *pixel* utilizando informações de pixels vizinhos como entrada dos algoritmos de reconhecimento de padrões;
- Definir algoritmos a serem testados utilizando a metodologia proposta;
- Apresentar medidas de avaliação para o processo de binarização utilizando a metodologia proposta; e
- Avaliar a metodologia proposta utilizando regiões de folhas de cheques de bancos brasileiros.

## 1.2 ORGANIZAÇÃO DO DOCUMENTO

Esta dissertação está dividida em cinco capítulos, cujas descrições encontram-se a seguir.

O Capítulo 2 apresenta os trabalhos relacionados da área de binarização de imagens e contém informações sobre imagens digitais, imagens de cheques, binarização de imagens digitais e binarização de imagens de cheques.

O Capítulo 3 faz a descrição detalhada da metodologia proposta.

O Capítulo 4 apresenta a caracterização do problema, a parametrização e os resultados experimentais para a binarização de imagens utilizando a metodologia proposta.

O Capítulo 5 apresenta as conclusões e discussões sobre trabalhos futuros.

## 2. REFERENCIAL TEÓRICO

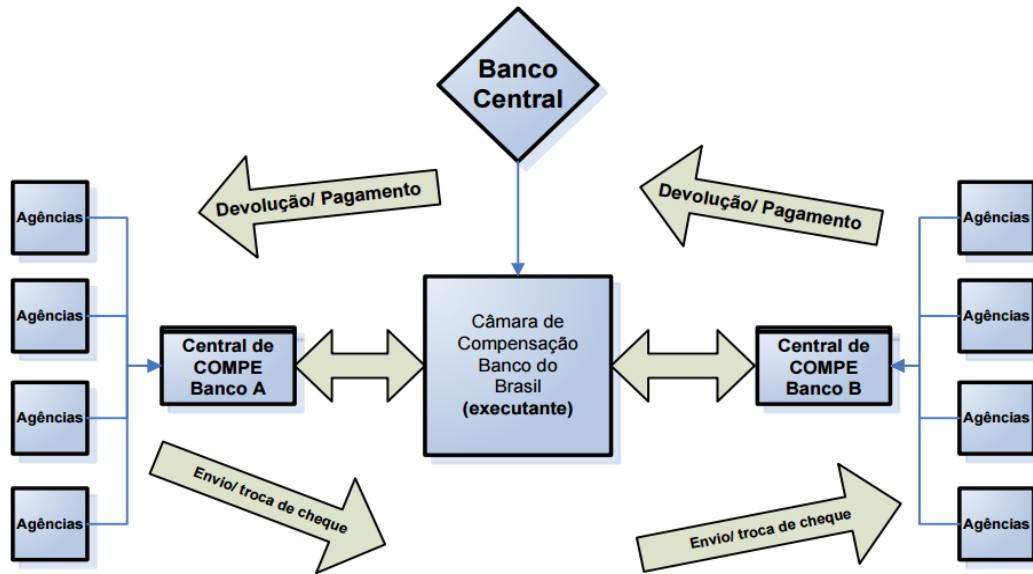
Este capítulo revisa os principais conceitos utilizados neste trabalho, incluindo imagens digitais e técnicas de binarização. Em relação às imagens digitais, serão enfatizadas as imagens de folhas de cheques de bancos brasileiros por se tratar de objetos de estudo da dissertação. Sobre as técnicas de binarização, a Seção 1.5 faz uma breve introdução conceitual e revisa os principais trabalhos da literatura que propõem e aplicam algoritmos de binarização em especial às imagens de cheques.

### 1.3 COMPENSAÇÃO DE CHEQUES DE BANCOS BRASILEIROS

A compensação de cheques consiste em um serviço de ordem pública regulamentado pelo Banco Central do Brasil que visa a realização do acerto de contas entre instituições (FILHO, NUNES, *et al.*, 2014). O acerto de contas refere-se aos cheques depositados em instituições financeiras diferentes das sacadas.

O processo de compensação representa um alto custo para instituições financeiras, uma vez que inclui gastos como transporte, tratamento, troca, microfilmagem, reprodução e guarda de documentos físicos (FEBRABAN, 2014). Na busca por alternativas que viabilizem a redução de custos, modernização e automação do processo da compensação de cheque, idealizou-se um processo automático de compensação com o uso de imagens digitais. Este processo também é conhecido como *Compe por Imagem de Cheques*. A compensação de cheques por imagem corrobora com ganhos de qualidade e redução de custo com transporte e tratamento dos papéis para armazenamento (FEBRABAN, 2014).

A Figura 2 apresenta o fluxo dos documentos no processo de compensação de cheques. Com o advento da Compe por Imagens o processo de transporte se tornou obsoleto, logo não há mais a troca física de cheques, como apresentado na Figura 2. Porém, a Figura 2 é capaz de ilustrar o processo genérico da compensação por imagens. Na parte inferior esquerda da Figura 2 as agências bancárias que receberam os cheques enviam estes à centralizadora do Banco do Brasil que, por sequência, envia ao banco pagador ou sacador. O processo inverso acontece na parte superior direita, quando o banco sacado envia a resposta para os cheques apresentados. Neste processo o Banco do Brasil assume o papel de mediador (FEBRABAN, 2014; FILHO, NUNES, *et al.*, 2014).

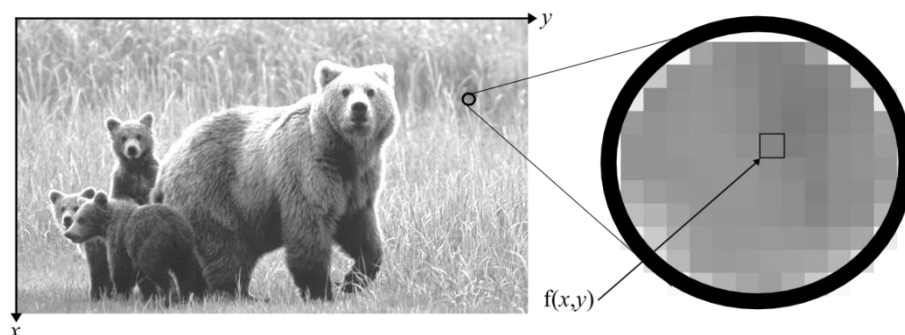


**Figura 2:** Fluxo macro de compensação de cheques no Brasil

**Fonte:** (FEBRABAN, 2014)

#### 1.4 IMAGENS DIGITAIS

O termo *imagem digital* consiste na representação e armazenamento digital de uma imagem (GONZALEZ e WOODS, 1992). A representação de uma imagem é realizada por meio de uma função de intensidade de luz,  $f(x,y)$ , onde  $x$  e  $y$  são coordenadas de um ponto (*pixel*) com intensidade  $f$ . O armazenamento de uma imagem digital consiste em uma representação matricial e discretizada da função  $f(x,y)$ , onde  $x$  e  $y$  são coordenadas dos índices em linhas e colunas e  $f$  é um valor definido comumente na escala entre 0 e 255 (GONZALEZ e WOODS, 1992). Cada ponto na matriz de representação possui uma intensidade e é conhecido como *pixel* (DUDA, HART e STORK, 2000; SOLOMON e BRECKON, 2013). A Figura 3 ilustra a representação de um pixel nos eixos  $x$  e  $y$ .



**Figura 3:** Representação da imagem nos eixos  $x$  e  $y$ .

As imagens digitais são comumente utilizadas em sistemas de reconhecimento de padrões, em processos de classificação e recuperação de conteúdo (DUDA, HART e STORK, 2000; MAHMOUD e AL-KHATIB, 2011; RABELO, ZANCHETTIN, *et al.*, 2011). As operações básicas para sistemas que utilizam imagens digitais são (GONZALEZ e WOODS, 1992):

- **Aquisição:** pode acontecer por meio de câmeras, scanners e raio X, dentre outros.
- **Armazenamento:** pode ser classificado em armazenamento por curto tempo, ou seja, armazenamento de imagens a serem utilizadas em um processo; armazenamento *online*, aquele de rápido acesso; e armazenamento em arquivo, para imagens acessadas com menor frequência.
- **Processamento:** procedimento algorítmico capaz de realizar filtragem, segmentação, reconhecimento de padrões, dentre outros.
- **Comunicação:** troca de informações que ocorre entre sistemas.
- **Exibição das imagens:** representação visual das imagens digitais, por exemplo, via monitores.

#### 1.4.1 Imagens Digitais de Cheques de Bancos Brasileiros

As imagens de cheques de bancos brasileiros são imagens de processamento complexo e apresentam desafios para métodos de processamento de imagens digitais. Os desafios são consequência, principalmente, da composição deste tipo de imagem, que possui diversidade de textura, cores e diferenças entre as composições das diferentes instituições financeiras.

Apesar das características inerentes da folha de cheque é necessário considerar a atuação humana no preenchimento de um cheque. Além das diferentes tipografias entre as pessoas, ainda existem fatores como inserção de elementos ruidosos e de segurança (RABELO, ZANCHETTIN, *et al.*, 2011) que têm o objetivo de dificultar a adulteração de uma folha de cheque.

Observa-se também a relevância crítica do processamento deste tipo de imagem visto a importância monetária inerente. A Figura 4 apresenta exemplos de algumas imagens de cheques utilizados no Brasil.



**Figura 4:** Exemplos de cheques de bancos brasileiros.

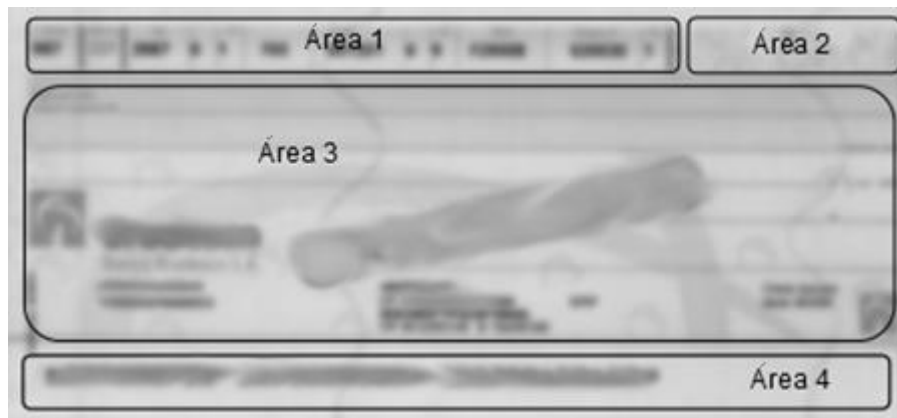
A Figura 4 também ilustra algumas das dificuldades de processamento das imagens de cheques de bancos brasileiros. É possível destacar duas perspectivas que influenciam e determinam diretamente a dificuldade do processamento deste tipo de imagem digital. A primeira perspectiva é de que a observação destas imagens isoladamente mostra a diferença nos padrões de fonte, cor, textura e elementos de segurança dentro de cada tipo de cheque. A segunda perspectiva pela qual estas imagens podem ser observadas é a diferença entre os cheques, o que aumenta a complexidade do processamento dessas imagens.

Por definições da Federação Brasileira de Bancos (Febraban<sup>1</sup>) os cheques brasileiros são divididos em quatro áreas (FEBRABAN, 2014):

- Área 1: código de banco, número de conta e outros números reservados;
- Área 2: montante de cortesia;
- Área 3: transcrição do montante legal; e
- Área 4: código magnético codificado em 7 barras, conhecido como CMC-7.

<sup>1</sup> [www.febraban.org.br](http://www.febraban.org.br)

No Brasil são mais de 192 modelos de cheques e todos obedecem a esta divisão em áreas. A Figura 5 apresenta o *layout* padrão dos cheques de bancos brasileiros.



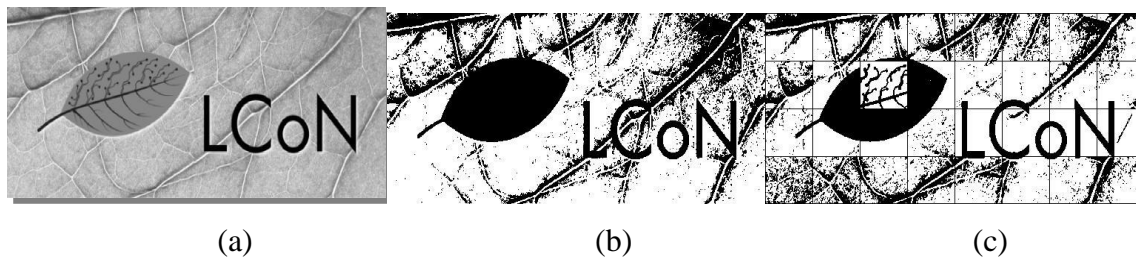
**Figura 5:** *Layout* de divisão de cheques de bancos brasileiros.

### 1.5 BINARIZAÇÃO DE IMAGENS DIGITAIS

A *binarização* é um método de processamento de imagens digitais e consiste na transformação de uma imagem em escala de cinza em uma imagem binária (ALGINAHI, SID-AHMED e AHMADI, 2004; GONZALEZ e WOODS, 1992). Este processo pode ocorrer simplesmente pela definição de um valor constante, chamado de *limiar*, usado na comparação da intensidade de cor de cada pixel da imagem original (SOLOMON e BRECKON, 2013). O objetivo desta técnica é realizar a filtragem dos elementos que são necessários para um processamento posterior da imagem, sendo assim a informação relevante deve ser destacada como “objeto”, ou objeto de interesse, e a informação a ser destacada deve ser expressa como “fundo” da imagem. Convencionalmente a classe fundo consiste no grupo de pixels de valor superior ao limiar, a estes é atribuída a cor branca; enquanto a classe objeto contém os pixels de valor inferior ao limiar, a estes é atribuída a cor preta (GONZALEZ e WOODS, 1992; LI e LEE, 1993; NEVES, MELLO, *et al.*, 2008). Este processo é usado para separar a imagem em fundo e objeto, ou objeto de interesse. A escolha do limiar pode ser feita, por exemplo, por meio do uso do histograma da imagem (GONZALEZ e WOODS, 1992; SEZGIN e SANKUR, 2004).

Os métodos de binarização por histograma podem ser baseados no espaço de atuação, que, por sua vez, podem ser divididos em métodos de binarização *global* e *local* (GONZALEZ e WOODS, 1992). Com o intuito de exemplificar a diferença entre os métodos de binarização local e global, considere o resultado binário na imagem em escala de cinza da Figura 6(a). A

binarização global consiste no processo de binarização da imagem utilizando um único processamento para toda a imagem, seja ele, por exemplo, o processo de corte por limiar (PALACIOS e GUPTA, 2008). A Figura 6(b) ilustra um exemplo de binarização por histograma global usando como limiar o resultado da média de níveis de cinza de toda a imagem. O método de binarização local, por outro lado, consiste na divisão da imagem em blocos (ou regiões) e o processamento de cada bloco separadamente (ALGINAHI, SID-AHMED e AHMADI, 2004). Na Figura 6(c), por exemplo, o processo de limiarização por histograma é aplicado localmente a cada um dos blocos destacados na imagem e é possível perceber que o método local identifica detalhes específicos dos blocos que são ignorados pelo método global. Neste caso o valor do limiar também pode ser calculado pela média de níveis de cinza, porém tendo como base cada bloco e não a imagem completa.



**Figura 6:** Exemplos de binarização por limiar. (a) Imagem de fundo complexo em escala de cinza. (b) Imagem binarizada por um método global. (c) Imagem binarizada por um método local.

Alguns algoritmos de binarização de imagens são historicamente importantes para as pesquisas desse assunto em processamento de imagens, dentre eles: Otsu (1979), Niblack (1986) e Sauvola (2000). Otsu (1979) é um algoritmo que utiliza a normalização do histograma de uma imagem e a distribuição de probabilidade para a escolha de um limiar ótimo. Niblack (1986) consiste no algoritmo de binarização local que varia o limiar para cada pixel baseado na média e no desvio padrão dos pixels de sua vizinhança. Sauvola e Pietikainen (2000) apresentam uma melhoria ao algoritmo de Niblack (1986), onde eles consideram que *pixels* pertencentes ao objeto e *pixels* pertencentes ao fundo devem ser tratados de maneira diferente.

Devido ao crescente volume de imagens e aplicações diretamente influenciadas pelo processo de binarização observa-se a necessidade de métodos capazes de realizar a binarização automática de imagens (CHEN, 2008; ALGINAHI, SID-AHMED e AHMADI, 2004). Diferentes tipos de aplicações de binarização são encontradas para variados tipos de imagens, dentre

estes, destacam-se aplicações que visam o processamento de imagens de cheques (NEVES, MELLO, *et al.*, 2009; CHEN, 2008).

Estes algoritmos de binarização automática podem ser categorizados em seis tipos (SEZGIN e SANKUR, 2004): (1) baseados na forma do histograma; (2) baseados em medidas de agrupamento; (3) baseados na entropia do histograma; (4) baseados nos atributos da imagem; (5) baseados na informação espacial; e (6) baseados em características locais. Observa-se que nem todos os algoritmos trabalham com a forma tradicional, ou seja, corte por limiar. A seção seguinte aborda algumas das técnicas utilizadas para a binarização automática de imagens de cheques de bancos.

### **1.5.1 Binarização de Imagens de Cheques de Bancos**

As imagens de cheques de bancos estão entre os principais desafios relatados pelas pesquisas de binarização automática de imagens e percebe-se a necessidade de métodos cada vez mais robustos e adaptativos que sejam capazes de realizar a separação entre o fundo e o objeto (AHMAD e MAHMOUD, 2012; SEZGIN e SANKUR, 2004).

A Tabela 1 apresenta alguns dos principais trabalhos realizados em binarização de imagens, especificamente usando imagens de cheques de bancos. Os trabalhos listados são categorizados com base em sua atuação espacial e algoritmos utilizados para realizar a binarização das imagens (SEZGIN e SANKUR, 2004).



**Tabela 1:** Principais trabalhos envolvendo a binarização automática de imagens de cheques.

Referência	Método		Categorias de Algoritmo
	Global	Local	
(DE JESUS e FACON, 2000)		X	Baseado na informação espacial
(KIM, KIM, <i>et al.</i> , 2001)		X	Baseado em medidas de agrupamento
(HONGGANG, GUANG, <i>et al.</i> , 2005)	X		Baseados nos atributos da imagem
(CAVALCANTI, SILVA, <i>et al.</i> , 2006)	X		Baseado no histograma
(CHEN, ZHANG e GUO, 2007)	X		Baseado no histograma
(CHANG, PAI e RUAN, 2008)		X	Baseado no histograma e em características locais
(PALACIOS e GUPTA, 2008)	X		Baseado no histograma
(NEVES, MELLO, <i>et al.</i> , 2009)	X		Baseado no histograma e na entropia do histograma
(LIU e CHEN, 2011)	X		Baseado no histograma e em características locais
(MAHMOUD e AL-KHATIB, 2011)		X	Baseado em características locais
(RABELO, ZANCHETTIN, <i>et al.</i> , 2011)		X	Baseado em características locais
(AHMAD e MAHMOUD, 2012)		X	Baseado em infos. espaciais e características locais

Os algoritmos utilizados nos métodos de binarização foram classificados em seis categorias (SEZGIN e SANKUR, 2004), da seguinte forma: (1) baseados em histograma (PALACIOS e GUPTA, 2008; CAVALCANTI, SILVA, *et al.*, 2006; CHEN, ZHANG e GUO, 2007; NEVES, MELLO, *et al.*, 2009; CHANG, PAI e RUAN, 2008; LIU e CHEN, 2011); (2) baseados em medidas de agrupamento (KIM, KIM, *et al.*, 2001; CHERIET, SAID e SUEN, 1998); (3) baseados na entropia do histograma (NEVES, MELLO, *et al.*, 2009; HORNG, 2011); (4) baseados nos atributos da imagem (HONGGANG, GUANG, *et al.*, 2005); (5) baseados na informação espacial (DE JESUS e FACON, 2000; AHMAD e MAHMOUD, 2012); e (6) baseados em características locais (RABELO, ZANCHETTIN, *et al.*, 2011; AHMAD e

MAHMOUD, 2012; LIU e CHEN, 2011; MAHMOUD e AL-KHATIB, 2011; CHANG, PAI e RUAN, 2008; YEN-LIN CHEN, 2008).

De Jesus e Falcon (2000) apresentam uma metodologia para realizar a segmentação em imagens de cheques de bancos brasileiros. A metodologia apresentada utiliza um processo de binarização local, realizado considerando técnicas de morfologia matemática que são responsáveis pela limpeza de ruídos das imagens processadas. Posteriormente, a imagem resultante é utilizada para realizar a identificação e segmentação do logotipo das imagens de cheques. Diversos experimentos são apresentados utilizando a metodologia proposta para a segmentação de imagens de cheques de bancos brasileiros.

Uma metodologia que propõe o reconhecimento de montante de cortesia e montante legal utilizando a segmentação de palavras é proposta por Kim *et al.*, (2001) para o processo de compensação de cheques de bancos. A metodologia utiliza uma técnica de binarização morfológica local e posteriormente a combinação de diferentes redes neurais artificiais. As redes neurais HMM (no inglês *Hidden Markov Model*) e MLP (no inglês *Multi-Layer Perceptron*) são utilizadas como um classificador híbrido para aumentar a acurácia do processo de reconhecimento de caracteres manuais após o processo de binarização utilizado. O processo de binarização é utilizado como uma importante etapa da tarefa de reconhecimento de caracteres manuais em imagens de cheque.

Uma metodologia de binarização global utilizando atributos das imagens de cheques de bancos chineses é proposta por Honggang *et al.* (2005). Para eles a análise na relação sinal-ruído entre imagens de cheques não preenchidas e as imagens preenchidas são a principal chave para o processo de binarização deste tipo de imagem. O objetivo era encontrar um limiar ótimo a partir da observação da diferença entre imagens de cheques preenchidas e sua imagem fonte.

Cavalcanti *et al.* (2006) descrevem uma metodologia que realiza binarização de documentos, denominados documentos de fundo complexo. Para realizar a binarização a metodologia proposta utiliza uma heurística baseada em dois processos de binarização e características globais do histograma de imagens de cheques. Os processos de binarização são utilizados para encontrar um limiar estimado, depois as imagens geradas por esses limiares são combinadas após um processamento morfológico baseado na dilatação das imagens. Uma imagem binarizada é resultante deste processo de combinação e corresponde a saída do algoritmo. Foram

realizados experimentos utilizando imagens de montante de cortesia e CMC-7 de imagens de cheques de bancos brasileiros.

A metodologia proposta por Chen, Zhang e Jun Guo (2007) consiste no reconhecimento de caracteres manuscritos chineses em imagens de cheque. A metodologia utiliza a distância de *Mahalanobis* e computação progressiva para realizar a binarização e reconhecimento de caracteres manuscritos. Um processo de binarização global das imagens é utilizado para a geração de vetores característicos utilizados na distância de *Mahalanobis*. A computação progressiva é utilizada como critério de parada do algoritmo. O objetivo do algoritmo é encontrar o caractere com menor distância de *Mahalanobis* dado o banco de dados de caracteres chineses.

Chang, Pai e Ruan (2008) apresentam uma metodologia de binarização baseada na detecção inteligente de blocos. O objetivo deste trabalho é encontrar um valor ótimo de limiar capaz de realizar a binarização de documentos em processo de degradação. De acordo com os autores, uma das principais características de um documento é ser dividido em blocos. Com base nesta premissa, é proposto um algoritmo global de binarização capaz de binarizar imagens de documentos por meio da derivada de uma superfície de limiares locais. Os experimentos desta metodologia são apresentados em imagens obtidas de artigos, jornais e revistas.

Palacios e Gupta (2008) apresentam um sistema capaz de realizar o processamento automático de campos manuscritos em imagens de cheques. O sistema é dividido em três processos: detecção de textos na imagem, reconhecimento do texto e pós-processamento do reconhecimento para assegurar assertividade do processo de reconhecimento. Para a primeira etapa, detecção de texto na imagem, o sistema utiliza uma metodologia de binarização dinamicamente ajustada baseada no ruído residual da imagem. Diversos limiares são calculados globalmente para a imagem e posteriormente é realizada a análise de conectividade dos componentes existentes na imagem binária. O limiar ótimo é encontrado quando o nível de componentes conexos é baixo, o que implica possibilidade de caracteres mais limpos e livres de ruído. Os experimentos realizados para o teste deste sistema foram realizados em imagens de cheques de bancos norte-americanos.

A proposta de uma metodologia de binarização para montantes de cortesia de imagens de cheques de bancos brasileiros é apresentada por Neves *et al.* (2009). A metodologia proposta pelos autores utiliza um processo de binarização global baseada na análise da entropia *Tsallis* do histograma das imagens, que consiste no cálculo de probabilidade de uma dada cor estar em

uma imagem. Experimentos são realizados utilizando imagens de cheques de bancos brasileiros.

Liu e Chen (2011) propõem uma metodologia de binarização de montante de cortesia de imagens coloridas de cheques de bancos chineses. A metodologia apresentada é composta por duas etapas denominadas remoção de fundo e remoção de selo. O processo de binarização apresentado realiza a combinação de dois métodos de binarização global, a saber Otsu (1979) e Niblack (1986). Na sequência as imagens geradas por meio dos algoritmos de binarização são utilizadas para realizar a análise das informações de cores das imagens de cheque. Os experimentos deste trabalho são realizados aplicando a metodologia proposta a 1000 imagens de cheques de bancos chineses.

Uma técnica de reconhecimento automático de caracteres árabes é proposta por Mahmoud e Al-Khatib (2011). A metodologia proposta apresenta um processo de binarização local que utiliza características locais da imagem, extraídas por meio dos filtros de *Gabor*. Após o processo de binarização são utilizados os classificadores SVM (em inglês *Support Vector Machines*), KNN (em inglês *K-Nearest Neighbor*) e HMM (em inglês *Hidden Markov Models*). Diversos experimentos são apresentados utilizando a metodologia aplicada ao banco de dados CENPARMI, que consiste em imagens de dígitos e cheques de bancos árabes.

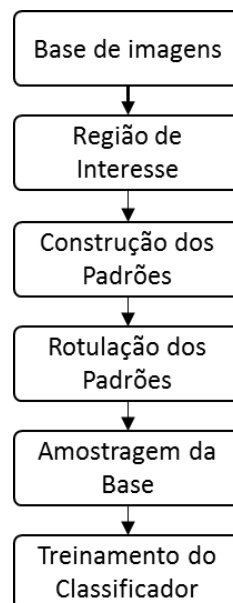
Rabelo *et al* (2011) abordam uma metodologia de binarização local para imagens de montante de cortesia e CMC-7 de cheques de bancos brasileiros utilizando uma rede neural MLP (em inglês *Multi-Layer Perceptron*). O processo de binarização local ocorre por meio da utilização de características locais da imagem para alimentar a rede neural artificial configuradas pelos autores. O trabalho contém experimentos sobre imagens de cheques de bancos brasileiros e sobre o banco de imagem DIBCO 2009 (GATOS, NTIROGIANNIS e PRATIKAKIS, 2009).

Os autores, Ahmad e Mahmoud (2012) propõem a análise e extração de áreas de imagens de cheques de bancos árabes. A metodologia apresentada utiliza um método de processamento adaptativo para realizar a binarização local destas imagens a partir da análise de componentes conexos. A etapa de binarização tem um importante papel na metodologia para a extração de áreas de imagens de cheques de bancos árabes. O trabalho apresenta experimentos utilizando a metodologia proposta aplicada ao banco de dados CENPARMI.

### 3. UMA METODOLOGIA DE BINARIZAÇÃO UTILIZANDO ALGORITMOS DE APRENDIZAGEM SUPERVISIONADA

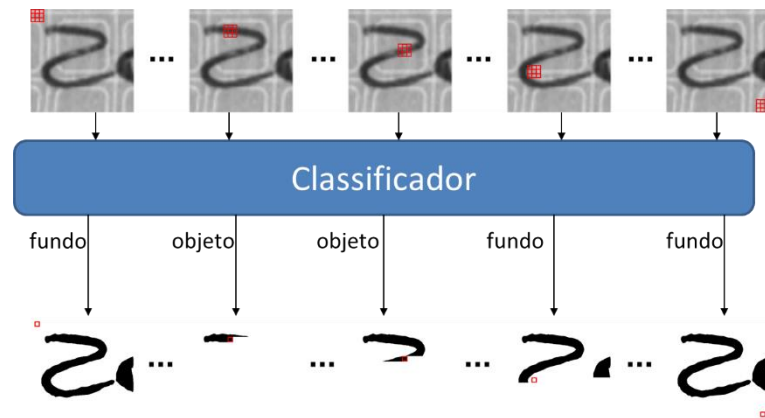
Este capítulo tem como objetivo apresentar uma metodologia genérica para o processo de binarização automática de imagens utilizando algoritmos supervisionados de aprendizagem de máquina. As seções seguintes apresentam as etapas metodológicas propostas, sendo elas: geração do banco de imagens; determinação da região de interesse; construção dos padrões; rotulação dos padrões; amostragem da base; e treinamento do classificador.

A metodologia permite o uso de diversos algoritmos de reconhecimento de padrões, que devem ter a capacidade de aprender a diferença entre *pixels* de fundo e objeto. O processo proposto para realizar a binarização segue o fluxo tradicional da aprendizagem de máquina supervisionada: treinamento; e teste. A Figura 7 resume os principais passos do método proposto. Cada etapa será detalhada na sequência.



**Figura 7:** Fluxo do método proposto.

No intuito de desenvolver uma metodologia de binarização de imagens serão utilizados algoritmos capazes de realizar uma etapa de aprendizagem por meio da utilização de características locais, operando *pixel a pixel*, e classificando cada *pixel* como pertencente ao fundo ou ao objeto. A Figura 8 ilustra o processo de binarização *pixel a pixel* utilizando informações da vizinhança de cada *pixel*.



**Figura 8:** Metodologia pixel a pixel para binarização de imagens

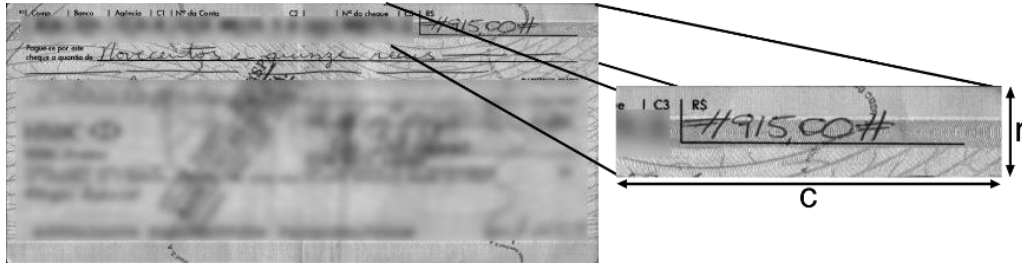
### 1.5.2 Construção do Banco de Imagens

O banco de imagens consiste no conjunto de imagens que serão utilizadas para o treinamento e teste da metodologia de binarização proposta. Este conjunto de imagens deve ser composto por imagens de mesma categoria, como imagens de documentos, imagens de veículos, imagens de pessoas, imagens de trânsito, imagens médicas, e imagens de cheque.

### 1.5.3 Extração da Região de Interesse

O primeiro passo do método proposto consiste na extração da região de interesse das imagens do banco de imagens, caso exista a necessidade. A extração da região de interesse é feita para uma dada classe de imagens de acordo com a necessidade da aplicação da binarização. Como exemplo de extração de regiões de interesse pode-se considerar a extração de placas de imagens de veículos para processos de reconhecimento de placa, identificação de rostos em imagens de pessoas e extração de áreas de cheque em uma aplicação de compensação de cheque (SOLOMON e BRECKON, 2013).

Considere como ilustração o processo de extração da região de interesse do montante de cortesia de uma imagem de cheque de banco brasileiro, como apresentado na Figura 9. Para este exemplo de região de interesse podem-se utilizar informações padronizadas do *layout* dos cheques de bancos brasileiros que determinam a região aproximada na imagem de cheque que contém o montante de cortesia (FEBRABAN, 2014).



**Figura 9:** Exemplo de imagem do montante de cortesia de cheques de bancos brasileiros.

Dado este tipo de aplicação considere a região de interesse,  $RoI \in Z^{c \times r}$ , a ser especificada para uma imagem de cheque. Seja  $d$  o número de pixels da  $RoI$  ( $d = c \cdot r$ ), onde  $c$  é a dimensão horizontal e  $r$  a dimensão vertical. Portanto, cada categoria de imagem pode ter uma  $RoI$  que deverá ser extraída de acordo com a necessidade da aplicação. Cada  $RoI$  consiste em uma nova imagem de dimensão diferente da imagem original, ou seja, consiste em um segmento de interesse para uma dada aplicação da imagem original. Se o interesse da aplicação é realizar a binarização da imagem original, então a extração da região de interesse não se faz necessária (FELIX, SILVA e DE CASTRO, 2015).

#### 1.5.4 Construção da Matriz de Padrões

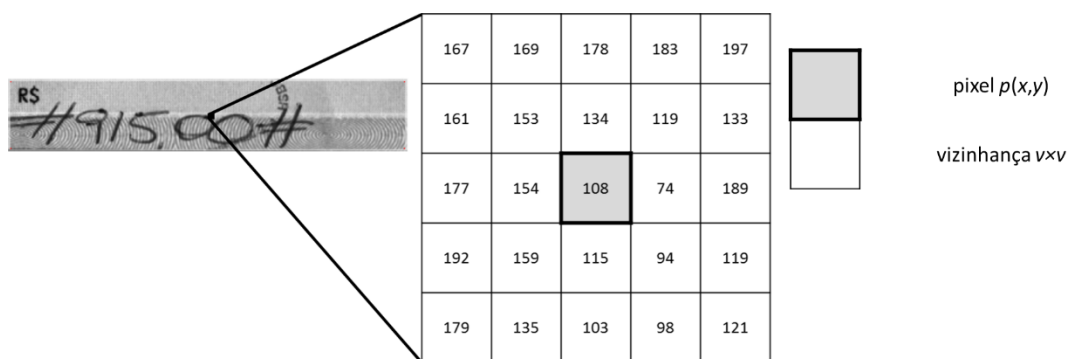
A metodologia propõe que a binarização seja realizada por meio de características locais da imagem. As características são definidas para cada pixel da imagem em análise, a qual chamaremos aqui de *padrão*. Sendo assim, a construção de cada padrão, nesta metodologia, é a dada pela utilização de uma *vizinhança* de pixels em torno do padrão, ou seja,  $v \times v$ , onde  $v$  é a uma quantidade de *pixels* (GONZALEZ e WOODS, 1992). A abordagem para definir uma vizinhança em torno do pixel  $p(x,y)$  consiste em uma sub imagem quadrada centrada no próprio *pixel*. O padrão e sua vizinhança são utilizados para compor um vetor característico que comporá uma matriz de padrões.

Por convenção, uma imagem, que pode ser representada por uma função de intensidade de luz  $f(x,y)$ , será representada aqui por uma matriz  $\mathbf{I} \in Z^{c \times r}$ , onde  $r$  é o número de linhas da imagem e  $c$  o número de colunas. Cada padrão será dado pela construção de um vetor,  $s_i \in Z^V$ , com  $V$  posições, onde  $V = v \times v$ . A cardinalidade de padrões da matriz de padrões construída a partir da imagem  $\mathbf{I}$  é dada pelo seu número de *pixels*, dado por  $D$ , onde  $D = r \cdot c$ . Sendo assim a matriz de padrões  $\mathbf{M}$  da imagem  $\mathbf{I}$  é expressa por  $\mathbf{M}_{\mathbf{I}} \in Z^{D \times V}$ .

$$s_i = \begin{bmatrix} p\left(x - \left(\frac{v}{2}\right), y - \left(\frac{v}{2}\right)\right), p\left(x - \left(\frac{v}{2}\right), y - \left(\frac{v}{2}\right) + 1\right), \\ \dots, p(x, y), \dots, \\ p\left(x + \left(\frac{v}{2}\right), y + \left(\frac{v}{2}\right) - 1\right), p\left(x + \left(\frac{v}{2}\right), y + \left(\frac{v}{2}\right)\right) \end{bmatrix} \quad (2)$$

A Equação (2) apresenta a construção de um padrão, onde o padrão  $s_i$  é dado pela matriz quadrada em volta do pixel de referência  $p(x,y)$ . A mesma representação pode ser vista na Figura 10, onde, por exemplo,  $p\left(x - \left(\frac{v}{2}\right), y - \left(\frac{v}{2}\right)\right)$  é representado pelo nível 167, e  $p\left(x + \left(\frac{v}{2}\right), y + \left(\frac{v}{2}\right)\right)$  é representado pelo nível 121 em uma vizinhança onde  $v = 5$ .

Para ilustrar a construção de um padrão desta matriz considere o caso em que cada pixel é representado pelo valor escalar de seu nível de cinza. Dado o *pixel*  $p(x,y)$  e uma *vizinhança*  $5 \times 5$ , onde a vizinhança abrange um total de 25 *pixels* incluindo o *pixel* referência  $p(x,y)$ . A Figura 10 representa a construção de um padrão a partir do pixel  $p(x,y)$  que possui intensidade com valor 108 e a construção de uma *vizinhança*  $5 \times 5$  em torno deste pixel de referência. O padrão resultante do *pixel*  $p(x,y)$  é construído como um vetor característico  $s_i$  onde cada elemento do vetor é preenchido por um valor de intensidade de cinza dos vizinhos de  $p(x,y)$ . O preenchimento do vetor é realizado percorrendo a vizinhança da esquerda para direita e de cima para baixo.



$$s_i = [167, 169, 178, 183, 197, 161, 153, 134, 119, 133, 177, 154, 108, 74, 189, 192, 159, 115, 94, 119, 179, 135, 103, 98, 121]$$

**Figura 10:** Representação do pixel  $p(x, y)$  dada uma janela de dimensão  $5 \times 5$  com base na vizinhança de Moore.

Como cada *pixel* e seus vizinhos comporá um padrão de treinamento do classificador, de forma genérica uma imagem  $\mathbf{I}$  ( $\mathbf{I} \in I$ , onde  $I$  é o conjunto de  $N$  imagens de dimensão  $D$  cada) será representada por uma matriz  $\mathbf{M}_I \in Z^{D \times V}$ , onde  $D$  é o número de *pixels* da imagem original,



e  $V = v \times v$  é a dimensão da vizinhança do pixel  $p(x,y)$ . Nota-se, portanto, que cada imagem gera um conjunto de padrões, representados nas linhas de  $\mathbf{M}_I$ , igual ao número de pixels da imagem original. O número total de padrões é igual ao produto do número de pixels de cada imagem (assumindo que ele será sempre o mesmo) pelo número de imagens a serem utilizadas.

O número de padrões resultantes de cada imagem é dado pela contagem de cada pixel representado excluindo uma quantidade igual a  $v$  da vizinhança do *pixel* de referência das primeiras e últimas linhas e a mesma quantidade das primeiras e últimas colunas. A exclusão destas linhas e colunas é necessária, pois elas extrapolam o tamanho da imagem para a vizinhança  $v \times v$  definida. A Figura 11 ilustra a extrapolação da imagem para uma vizinhança  $5 \times 5$  onde o número de linhas e colunas exclusas é igual 2;



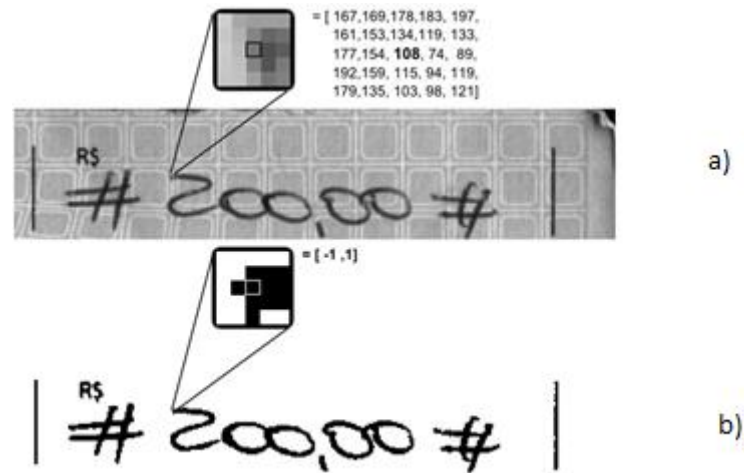
**Figura 11:** Vizinhança extrapola dimensão da imagem  $I$ .

### 1.5.5 Rotulação dos Padrões

A etapa de rotulação consiste na atribuição de um *rótulo de classe* (ou *padrão ouro*) a cada padrão gerado na etapa de construção da matriz de padrões. Logo, cada *pixel* da imagem deve ser atribuído a uma das classes: fundo ou objeto. Para cada imagem  $I \in I$ , a ser usada no processo de treinamento do classificador, é criada uma imagem  $\mathbf{B}$  binarizada e o rótulo para cada padrão obtido da imagem  $I$  consiste na classe do *pixel*  $b(x,y)$  na imagem  $\mathbf{B}$ . A duas etapas para

a binarização da imagem **B** são: (1) utilização do algoritmo Otsu (1979) para a binarização da imagem **I**; e (2) o aprimoramento manual da imagem (FELIX, SILVA e DE CASTRO, 2015).

O exemplo da Figura 12 ilustra o processo de representação e rotulação de cada padrão utilizando uma vizinhança  $5 \times 5$  e cujo rótulo é definido com base na classe do *pixel* equivalente na imagem binarizada.



**Figura 12:** Rotulação de cada *pixel*  $p(x, y)$  utilizando imagem binarizada manualmente, a) imagem em escala de cinza e b) imagem binarizada manualmente;

Esta abordagem possibilita a obtenção de padrões com informações relevantes sobre a vizinhança de cada *pixel* e mostra que as informações contidas na vizinhança revelam uma variação do gradiente na região do *pixel*. A Figura 12 (a) ilustra a suavização do gradiente dos pixels à medida que se afastam do objeto e também mostra que o processo de rotulação determinou que o *pixel* analisado pertence à classe objeto.

A segunda etapa do processo de geração da Figura 12 (b), aprimoramento manual da imagem, pode ser considerado custoso computacionalmente. No entanto, este processo consiste na geração de um padrão ouro que é utilizado como rótulo e resultado desejado para avaliação da metodologia. Além disso, é um processo necessário para que os algoritmos de reconhecimento de padrões se tornem capazes de identificar as diferenças entre fundo e objeto em uma dada classe de imagem a partir de uma perspectiva humana.

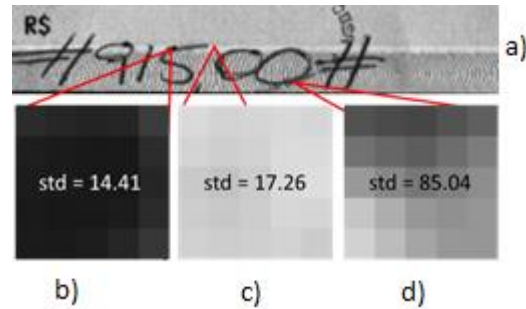
### 1.5.6 Amostragem da Base

Como cada pixel da imagem original a ser binarizada se transforma em um padrão de treinamento para o classificador, mesmo um conjunto de poucas imagens com baixa resolução facilmente geraria uma imensa matriz de padrões de treinamento. Além disso, uma parcela significativa desses padrões pode ser considerada trivial para o classificador, ou seja, padrões cuja vizinhança do pixel de referência é toda representada por um único valor (0 ou 255).

Para reduzir significativamente e de forma automática o número de objetos de treinamento do classificador e focar em casos que apresentam complexidade para o classificador, optou-se por uma metodologia de amostragem da base de dados (DUDA, HART e STORK, 2000). Nesta proposta são abordadas duas metodologias de amostragem da base de dados para otimizar o treinamento do classificador e diminuir o tempo de processamento, são elas: amostragem aleatória (*rand*); e amostragem baseada no desvio padrão (*std*) do nível de cinza dos padrões. A abordagem aleatória (*rand*) consiste na seleção aleatória de  $np$  padrões da base  $\mathbf{M}_I$  para utilização no treinamento e teste do classificador. A abordagem baseada no desvio padrão (*std*) consiste na seleção dos  $np$  primeiros padrões com maior desvio padrão na base  $\mathbf{M}_I$ .

A abordagem baseada no desvio padrão (*std*) é um índice que avalia a homogeneidade entre os níveis de cinza da vizinhança do padrão. Dado o desvio padrão de um padrão  $s_i$  é possível inferir a homogeneidade, uma vez que padrões com baixo desvio padrão consistem em vizinhanças mais homogêneas. Os padrões homogêneos consistem naqueles em que os níveis de cinza de seus pixels contêm valores mais próximos e, conseqüentemente, possuem desvio padrão baixo. Os padrões homogêneos apresentam baixa complexidade no processo de classificação. Já os padrões que apresentam desvio padrão elevados possuem vizinhanças mais heterogêneas e indicam altas distâncias entre os valores dos *pixels* no padrão, conseqüentemente, possuem desvio padrão alto. Os padrões heterogêneos apresentam maior complexidade para os classificadores do que os padrões homogêneos.

A Figura 13 ilustra a homogeneidade de três diferentes vizinhanças: a Figura 13, primeira (a) e a segunda (b) apresentam vizinhanças homogêneas com baixo desvio padrão e alta similaridade entre os *pixels*, respectivamente a (b) terceira representa uma vizinhança heterogênea com baixa similaridade entre os *pixels* e alto desvio padrão.



**Figura 13:** Nível de homogeneidade baseado no desvio padrão da vizinhança, a) imagem em escala de cinza; b) e c) região com pixels homogêneos; e d) região com pixels heterogêneos.

### 1.5.7 Treinamento do Classificador

Quando cada padrão possui um *rótulo de classe* ou um valor de saída associado que representa o resultado desejado (padrão ouro), o objetivo do treinamento é construir um *modelo* que possa ser usado para prever qual seria essa saída para novos padrões. Esse tipo de tarefa é chamado genericamente de *reconhecimento de padrões* (DUDA, HART e STORK, 2000) e é dividido em duas etapas principais: 1) treinamento: na primeira etapa o classificador é gerado, tal que ele se torne capaz de reconhecer um conjunto pré-determinado de classes. O classificador é gerado usando um conjunto de pares entrada-saída  $\{(s_i, b_i)\}_{i=1,\dots,D}$ , onde  $s_i$  e  $b_i \forall i$ , são os padrões de entrada e as respectivas saídas desejadas; e 2) teste: uma vez que o classificador foi treinado é preciso avaliar seu desempenho quando aplicado a padrões não usados no processo de treinamento. O desempenho do classificador quando aplicado a dados de teste oferece uma estimativa de sua capacidade de generalização (DUDA, HART e STORK, 2000).

A metodologia proposta não faz restrição quanto ao tipo de classificador a ser utilizado na binarização (FELIX, SILVA e DE CASTRO, 2015). Como será visto na análise experimental, neste trabalho serão usados classificadores baseados em redes neurais artificiais implementadas (DE CASTRO e VON ZUBEN, 1998), técnicas clássicas de aprendizagem de máquinas (DUDA, HART e STORK, 2000) e máquinas de vetores suporte (CHANG e LIN, 2011).

### 1.5.8 Medidas de Avaliação

A metodologia de validação cruzada em  $k$  pastas foi utilizada para estimar o desempenho dos algoritmos de reconhecimento de padrões (DUDA, HART e STORK, 2000). As medidas escolhidas para avaliar o desempenho dos algoritmos são baseadas na matriz de confusão, a partir da qual se obtém o número de verdadeiros positivos (TP), que indica os *pixels* do objeto que

foram acertados pelo classificador; o número de falsos positivos (FP), que indica os *pixels* de fundo classificados como objeto; o número de verdadeiros negativos (TN), que indica os *pixels* de fundo que foram acertados pelo classificador; e o número de falsos negativos (FN), que indica os *pixels* de objeto classificados como fundo, como ilustra a Figura 14. As medidas quantitativas de avaliação de desempenho dos algoritmos são: acurácia (*Acc*), precisão (*Pr*) e revocação (*Re*), formalmente definidas como:

- **Acurácia:**

$$Acc = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (3)$$

- **Precisão:**

$$Pr = \frac{TP}{(TP + FP)} \quad (4)$$

- **Revocação:**

$$Re = \frac{TP}{(TP + FN)} \quad (5)$$

- **F-Measure:**

$$FMeasure = \frac{2 \times Re \times Pr}{Re + Pr} \quad (6)$$

- **PSNR:** a medida *Peak Signal-to-Noise Ratio* mede a similaridade entre duas imagens: quanto maior for o valor de PSNR, maior será a similaridade entre as imagens. O valor máximo de intensidade de cinza de pixel na imagem é dado pela variável *C*, e o erro quadrático médio é dado por *MSE*:

$$PSNR = 10 \log \frac{C^2}{MSE} \quad (7)$$

, onde:

$$MSE = \frac{\sum_{x=1}^M \sum_{y=1}^N (I(x, y) - I'(x, y))^2}{MN} \quad (8)$$

- **NRM:** a *Negative Rate Metric* consiste na taxa de erro pixel a pixel entre a imagem desejada e a imagem obtida (YOUNG e FERRYMAN, 2005).

$$NRM = \frac{NR_{FN} + NR_{FP}}{2} \quad (9)$$

, onde:

$$NR_{FN} = \frac{FN}{FN + TP} \quad (10) \quad NR_{FP} = \frac{FP}{FP + TN} \quad (11)$$

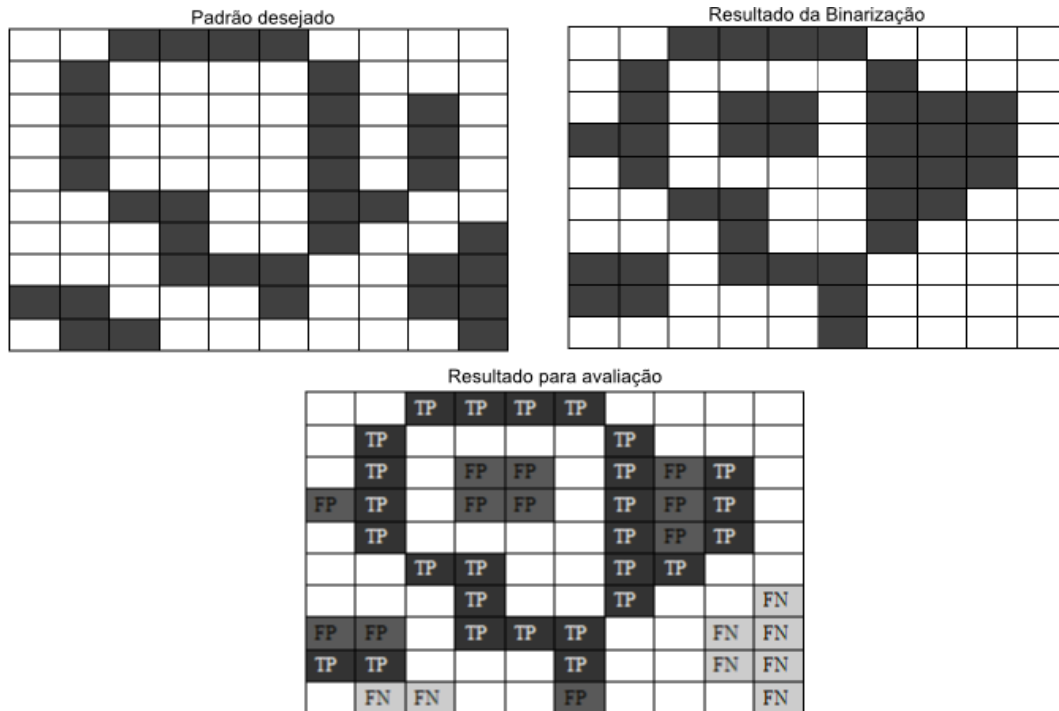
- **MPM:** a *Misclassification Penalty Metric* consiste na avaliação de similaridade entre as bordas do objeto da imagem desejada e as bordas do objeto da imagem obtida. Onde,  $d_{FN}^i$  e  $d_{FP}^j$  são as distâncias do  $i^{\text{ésimo}}$  pixel falso negativo e do  $j^{\text{ésimo}}$  pixel falso positivo do contorno do objeto segmentado no padrão desejado, e  $D$  é a soma de todos os pixels de contorno do objeto segmentado (YOUNG e FERRYMAN, 2005).

$$MPM = \frac{MP_{FN} + MP_{FP}}{2} \quad (12)$$

, onde:

$$MP_{FN} = \frac{\sum_{i=1}^{N_{FN}} d_{FN}^i}{D} \quad (13) \quad MP_{FP} = \frac{\sum_{j=1}^{N_{FP}} d_{FP}^j}{D} \quad (14)$$

A Figura 14 apresenta o padrão desejado que consiste na imagem desejada ou padrão ouro, o resultado da binarização que consiste na imagem obtida pelo processo de binarização e o resultado para a avaliação que consiste em uma imagem resultante da diferença entre a imagem desejada e a imagem obtida. Os valores de  $TP$ ,  $TN$ ,  $FP$  e  $FN$  são dados pela somatria dos *pixels* no resultado para a avaliação.



**Figura 14:** Exemplo de tipos de *pixels* para cálculo das medidas de avaliação.

Para o exemplo apresentado na Figura 14, a somatória dos indicadores de pixel são,  $TP = 27$ ,  $FP = 11$ ,  $FN = 8$  e  $TN = 0$ . Logo, os cálculos das medidas de avaliação são exibidos na Tabela 2.

**Tabela 2:** Resultados utilizando medidas de avaliação para o exemplo.

Medida	Cálculo	Resultado
Acurácia (ACC):	$Acc = \frac{27}{(27 + 11 + 8)}$	0.58
Precisão (Pr):	$Pr = \frac{27}{(27 + 11)}$	0.71
Revocação (Re):	$Re = \frac{27}{(27 + 8)}$	0.77
F-Measure:	$FMeasure = \frac{2 \times 0.77 \times 0.71}{0.77 + 0.71}$	0.73
PSNR:	$PSNR = 10 \log \frac{1^2}{0.22}$	6.57
NRM:	$NRM = \frac{0.61 + 0.11}{2}$	0.366
MPM:	$MPM(\times 1000) = \frac{0 + 0.10}{2}$	54.94

## 4. APLICAÇÃO DA METODOLOGIA PROPOSTA E RESULTADOS EXPERIMENTAIS

Este capítulo tem como objetivo apresentar o conjunto de resultados obtidos por meio da utilização do método proposto para o processo de binarização automática de imagens, utilizando algoritmos supervisionados de aprendizagem de máquina. As seções subsequentes apresentam três conjuntos de experimentos, sendo eles: a configuração e utilização do método proposto para a binarização de imagens de montante de cortesia; binarização do código CMC-7; e a binarização das imagens da competição DIBCO 2009.

### 1.6 APLICAÇÃO DA METODOLOGIA PROPOSTA A IMAGENS DE CHEQUES DE BANCOS BRASILEIROS

Cada área do *layout* dos cheques de bancos brasileiros tem uma importância para o processo de gerenciamento desse recurso para as instituições financeiras. Duas destas áreas assumem um papel importante na compensação de cheques por imagens digitais: o montante de cortesia, que representa uma região que destaca o valor monetário do cheque e tem sido alvo de diversas pesquisas (AHMAD e MAHMOUD, 2012; MAHMOUD e AL-KHATIB, 2011; NEVES, MELLO, *et al.*, 2008; RABELO, ZANCHETTIN, *et al.*, 2011); e o CMC-7, que consiste no número de identificação que contém todas as informações em relação ao cheque e ao cliente do banco (FEBRABAN, 2014).

Este trabalho utiliza estas duas importantes regiões de cheques de bancos brasileiros. A restrição do problema a estas duas áreas permite a diminuição da dimensionalidade da base de dados, consequentemente, diminuindo o custo computacional do processo.

#### 1.6.1 Caracterização do Problema

Esta seção apresenta a análise experimental da metodologia proposta aplicada à binarização automática de imagens de montante de cortesia (Figura 5, Área 2) utilizando algoritmos de reconhecimento de padrões. Dada a importância econômica presente nos cheques, o processo de binarização é crucial para o desenvolvimento de aplicações de compensação de cheque (AHMAD e MAHMOUD, 2012).

Na metodologia proposta a vizinhança de um pixel possui informações inerentes de localidade (GONZALEZ e WOODS, 1992). Intuitivamente é possível encontrar três tipos predominantes de pixels em uma vizinhança quando observado o problema de binarização: o primeiro tipo consiste em *pixels* predominantemente pertencentes à classe fundo; o segundo tipo consiste

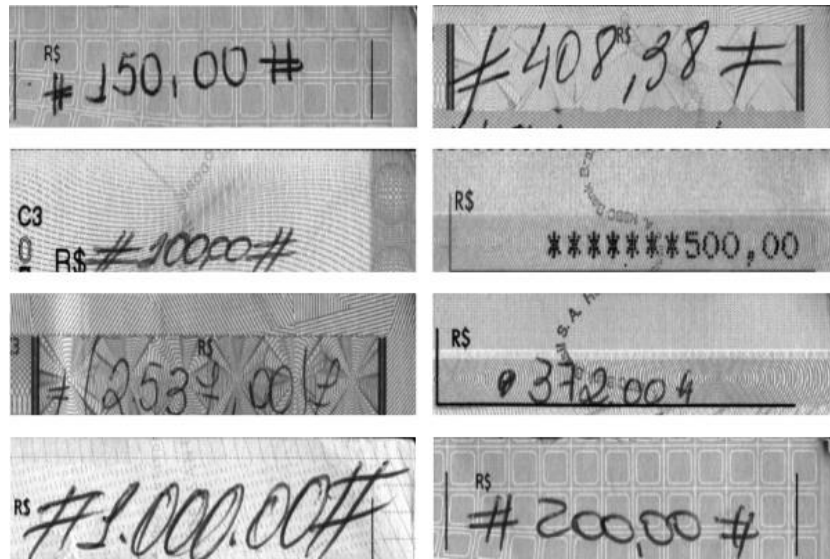


em *pixels* predominantemente pertencentes à classe objeto; e o terceiro tipo está baseado na transição de *pixels* de fundo para *pixels* de objeto e vice-versa. As duas primeiras predominâncias de *pixels* em uma vizinhança não apresentam grandes desafios para os métodos de binarização, pois quando analisados seus histogramas é possível determinar mais claramente um limiar. No entanto, os principais desafios para esse processo encontram-se no terceiro tipo de vizinhança devido à variação de valores entre os *pixels*, o que dificulta por exemplo a determinação de um limiar. A metodologia proposta trata exatamente a vizinhança com maior transição de *pixels* entre fundo e objeto, como será apresentado ao longo deste capítulo.

### 1.6.1.1 Montante de Cortesia

Alguns dos problemas que dificultam a binarização desta área do cheque são os elementos de segurança, como linhas aleatórias e fatores culturais observados na Figura 15, tais como os caracteres ruidosos ‘#’ e ‘\*’, comumente utilizados para dificultar a ação de terceiros na adulteração do valor do cheque. A Figura 15 também ilustra outros elementos desafiadores para o processo de binarização do montante de cortesia de bancos brasileiros, por exemplo, diferentes tipos de caligrafia, fontes digitadas, diferença no posicionamento espacial da informação, múltiplos tipos de fundo, diferentes composições de fundo, dentre outros.

A existência de tais elementos sugere a necessidade de aplicação de um método local de binarização capaz de identificar as diferenças entre fundo e objeto. A percepção humana entre fundo e objeto possivelmente está associada a elementos como a mudança abrupta no gradiente entre o fundo e o objeto, cores, contrastes e conhecimento adquirido previamente pelo observador (RABELO, ZANCHETTIN, *et al.*, 2011). Acredita-se que esses elementos ajudem no processo de separação entre fundo e objeto dos sistemas naturais (RABELO, ZANCHETTIN, *et al.*, 2011; DE CASTRO, 2006). Essa variação de gradiente é percebida pela visão humana sempre de forma local e serve de motivação para o método proposto.

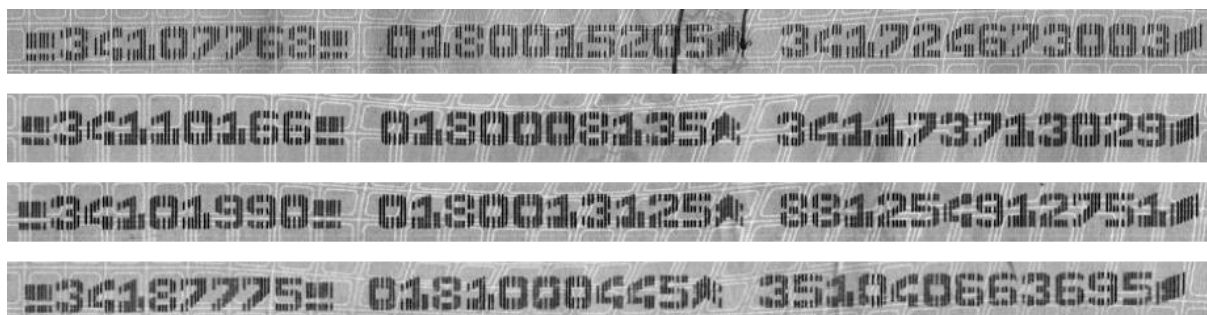


**Figura 15:** Exemplos de oito imagens de montantes de cortesia de diferentes cheques de bancos brasileiros.

#### 1.6.1.2 CMC-7

A área de cheques de bancos brasileiros conhecida como Caracteres Magnéticos Codificados em 7 Barras (CMC-7) consiste em um tipo de código de barras impresso utilizando toner magnético e pretende dificultar a ação de terceiros em fraudes bancárias utilizando cheques.

A Figura 16 apresenta exemplos de CMC-7 de cheques de bancos brasileiros. Dentre alguns dos desafios que podem dificultar o processo de binarização desta área estão a divisão em 7 barras, a variedade de fundos de cheques de bancos brasileiros e ruídos no processo de aquisição da imagem. O processo de binarização e reconhecimento dos caracteres da área CMC-7 é um passo importante para a verificação das informações das imagens de cheque no processo de compensação do mesmo.



**Figura 16:** Exemplos de quatro imagens de CMC-7 de diferentes cheques de bancos brasileiros.

## 1.6.2 Aplicação e Análise Preliminar da Metodologia Proposta

Esta seção tem como objetivo ilustrar o uso da metodologia proposta no contexto de binarização automática de imagens do montante de cortesia e CMC-7 de cheques de bancos brasileiros. Cada etapa do processo é detalhada nessa aplicação e são apresentados as medidas de desempenho e os resultados dos experimentos.

### 1.6.2.1 Construção do Banco de Imagens

Na avaliação feita nessa pesquisa utilizou-se para ambas as áreas do cheque um banco de imagens com 60 folhas de cheques privados em escala de cinza de seis diferentes bancos brasileiros. Cada uma destas imagens tem resolução de 200 *ppp* (pontos por polegada) e dimensão de  $1575 \times 600$  *pixels*. O banco de imagens foi cedido pela empresa Sincronica Serviços Integrados Limitada.

### 1.6.2.2 Extração da Região de Interesse

Os experimentos realizados na etapa de análise da metodologia utilizam máscaras que obedecem aos limites definidos pelo *layout* de divisão dos cheques de bancos brasileiros. Como resultado do processo de segmentação são obtidas imagens do montante de cortesia e CMC-7, como ilustra a Figura 17.

Para a extração da região de interesse referente ao montante de cortesia, a máscara possui dimensão  $500 \times 120$  *pixels* a partir da extremidade superior direita da imagem do cheque e, portanto, cada *RoI* possui  $6 \times 10^4$  *pixels*. No caso da extração da região de interesse referente ao CMC-7, a máscara aplicada para a segmentação da área pretendida é dada pela dimensão

920×50 pixels a partir da extremidade inferior esquerda da imagem do cheque e, portanto, cada *RoI* possui  $4,5 \times 10^4$ .



**Figura 17:** Aplicação da extração da *RoI* a partir de imagens de cheques de bancos brasileiros.

### 1.6.2.3 Construção da Matriz de Padrões

No processo de construção da matriz de padrões adotou-se uma *vizinhança*  $5 \times 5$ . Logo, cada padrão será composto pela concatenação dos valores de níveis de cinza de uma vizinhança, sendo  $v = 5$ . Logo, o número de atributos consiste em  $V = 25$ . Sendo assim, para cada imagem de montante de cortesia, onde  $c = 120 \text{ pixels}$  e  $r = 500 \text{ pixels}$ , a metodologia constrói uma matriz de padrões,  $\mathbf{M}_I \in Z^{(120 \cdot 500) \times 25}$ . Similarmente, cada imagem de CMC-7 onde  $c = 920 \text{ pixels}$  e  $r = 50 \text{ pixels}$  dá origem a uma matriz de padrões,  $\mathbf{M}_I \in Z^{(920 \cdot 50) \times 25}$ .

### 1.6.2.4 Rotulação dos Padrões

A etapa de rotulação consiste na atribuição de um rótulo de classe a cada padrão gerado na etapa de construção da matriz de padrões. Logo, cada pixel da imagem deve ser atribuído a uma das classes: fundo ou objeto. Para cada imagem de montante de cortesia obtida foi gerada uma imagem equivalente binarizada em duas etapas. A primeira etapa consiste na aplicação da metodologia de limiarização de Otsu (1979) nas regiões de interesse de montante de cortesia e CMC-7. A segunda etapa consiste no aprimoramento manual das imagens resultantes da limiarização de Otsu (1979) utilizando filtros morfológicos para remoção de ruídos do tipo *sal e*

*pimenta* e aprimoramento manual. Estas duas etapas geram imagens que são os padrões ouro para treinamento dos algoritmos de reconhecimento de padrão e para avaliar o processo de binarização do método proposto.

### 1.6.2.5 Amostragem da Base

Como cada pixel da imagem original a ser binarizada, juntamente com sua vizinhança, se transforma em um padrão de treinamento para o classificador, mesmo um conjunto de poucas imagens com baixa resolução facilmente geraria uma imensa matriz de padrões de treinamento. Além disso, uma parcela significativa desses padrões pode ser considerada trivial para o classificador, ou seja, padrões cuja vizinhança do pixel de referência é toda representada por um único valor (0 ou 255) ou valores muito próximos. Assim, foram utilizadas seis diferentes amostragens da base referente ao montante de cortesia para treinamento do classificador. O número de amostras é definido pela variável  $n_{sample}$ .

- i. Aleatória (*rand*):
  1. Seleção de  $n_{sample} = 10^3$  padrões aleatórios para treinar o classificador.
- ii. Baseado no desvio padrão (*std*):
  2. Seleção de  $n_{sample} = 10^3$  padrões baseados no desvio padrão dos níveis de cinza de cada um deles;
  3. Seleção de  $n_{sample} = l$ , onde  $l$  representa o número de padrões com  $std > 80\%$  da média dos desvios padrões da base;
  4. Seleção de  $n_{sample} = 0,1\%$  do total de padrões da base;
  5. Seleção de  $n_{sample} = 0,05\%$  do total de padrões da base;
  6. Seleção de  $n_{sample} = m_{std}$ , onde  $m$  representa o número de padrões com desvio padrão superior à média dos desvios padrões da base.

A mesma abordagem de amostrar a base de dados foi utilizada para os testes realizados com a região de montante de cortesia e CMC-7. Entretanto, os testes realizados utilizando  $n_{sample} = m_{std}$  não foram executados para o CMC-7 por limitações de recursos físicos e tecnológicos referentes ao equipamento (*Intel® Core™ i5-3230M CPU @ 2.60GHz 4,00 GB (RAM)*) utilizado para o desenvolvimento deste trabalho.

### 1.6.2.6 Escolha e Treinamento dos Classificadores

Foram selecionados três classificadores tradicionais da literatura para avaliar a metodologia de binarização proposta nas áreas selecionadas, são eles:

- **Redes Neurais do tipo Perceptron Multicamadas (MLP):** foi utilizada uma rede neural do tipo Perceptron de múltiplas camadas treinada com um algoritmo de segunda ordem, o *gradiente conjugado escalonado* (do inglês *Scaled Conjugate Gradient*, SCGM) (DE CASTRO e VON ZUBEN, 1998) com uma camada intermediária contendo 10 neurônios com função de ativação tangente hiperbólica,  $5 \times 10^3$  épocas de treinamento e dois neurônios bipolares na camada de saída. O algoritmo foi implementado no interpretador de linguagem matricial *Octave*<sup>2</sup>.
- **Support Vector Machine (SVM):** é considerado um dos métodos de melhor desempenho na solução de problemas de classificação binária (CHANG e LIN, 2011). Este trabalho utilizou a biblioteca de código aberto LIBSVM. Neste trabalho foi utilizada uma SVM com *kernel* linear (CHANG e LIN, 2011) implementada no *Octave*.
- **Decision Tree (DT):** a *Decision Tree* ou árvore de decisão utilizada neste trabalho foi configurada com o ganho de informação como critério de seleção dos nós, 20 nós de profundidade máxima e confiança de  $2,5 \times 10^{-1}$  (DUDA, HART e STORK, 2000) nativa no *software* Rapidminer 5<sup>3</sup>.

### 1.6.3 Resultados Experimentais e Discussões

Esta seção apresenta resultados da aplicação do método proposto utilizando 60 imagens de cheques de bancos brasileiros. Para cada uma das imagens foram realizados testes utilizando o montante de cortesia e CMC-7. Os experimentos foram executados em uma mesma estação de trabalho com a seguinte configuração: *Intel® Core™ i5-3230M CPU @ 2.60GHz, 4,00 GB (RAM)*.

Para ambos os grupos de imagens os resultados experimentais aqui apresentados consistem em dois conjuntos. O primeiro conjunto de testes representa a análise de desempenho de cada um dos três classificadores selecionados. O objetivo deste teste é avaliar o desempenho e

---

<sup>2</sup> <https://www.gnu.org/software/octave/>

<sup>3</sup> <https://rapidminer.com/>

a capacidade de generalização de cada um dos três classificadores selecionados. O segundo conjunto representa o desempenho de binarização da metodologia proposta para imagens de cheques de bancos brasileiros. Para tal conjunto de testes foi aplicado todo o fluxo metodológico proposto neste trabalho, a fim de comparar os resultados dos classificadores.

### 1.6.3.1 Resultados Experimentais para Imagens de Montante de Cortesia

A Tabela 2 mostra a média  $\pm$  o desvio padrão das medidas para cada um dos classificadores utilizando validação cruzada em 10-pastas (DUDA, HART e STORK, 2000). Os classificadores foram treinados e testados utilizando amostras da base de dados. Para cada classificador foram calculadas as medidas de desempenho *Acc*, *Pr* e *Re*.

**Tabela 3:** Acurácia (*Acc*), Precisão (*Pr*) e Revocação (*Re*) dos classificadores utilizando amostragem das imagens de montante de cortesia.

		<i>RAND</i>	<i>STD</i>				
		$n_{sample} = 10^3$	$n_{sample} = 10^3$	$n_{sample} = l$	$n_{sample} = 0,1\%$	$n_{sample} = 0,05\%$	$n_{sample} = m$
MLP	Acc	0.95 $\pm$ 0.004	0.97 $\pm$ 0.002	0.95 $\pm$ 0.001	0.93 $\pm$ 0.001	0.93 $\pm$ 0.001	0.93 $\pm$ 0.001
	Pr	0.94 $\pm$ 0.004	0.96 $\pm$ 0.003	0.93 $\pm$ 0.001	0.91 $\pm$ 0.001	0.91 $\pm$ 0.001	0.91 $\pm$ 0.001
	Re	0.94 $\pm$ 0.006	0.98 $\pm$ 0.003	0.97 $\pm$ 0.001	0.95 $\pm$ 0.001	0.95 $\pm$ 0.001	0.95 $\pm$ 0.001
SVM	Acc	0.94 $\pm$ 0.001	0.96 $\pm$ 0.002	0.93 $\pm$ 0.001	0.91 $\pm$ 0.001	0.91 $\pm$ 0.001	0.92 $\pm$ 0.001
	Pr	0.93 $\pm$ 0.002	0.95 $\pm$ 0.003	0.93 $\pm$ 0.001	0.91 $\pm$ 0.001	0.90 $\pm$ 0.001	0.92 $\pm$ 0.001
	Re	0.95 $\pm$ 0.002	0.97 $\pm$ 0.003	0.93 $\pm$ 0.001	0.92 $\pm$ 0.001	0.91 $\pm$ 0.001	0.93 $\pm$ 0.001
DT	Acc	0.95 $\pm$ 0.004	<b>0.98<math>\pm</math>0.001</b>	0.94 $\pm$ 0.001	0.91 $\pm$ 0.000	0.92 $\pm$ 0.000	0.93 $\pm$ 0.000
	Pr	0.95 $\pm$ 0.002	<b>0.98<math>\pm</math>0.003</b>	0.90 $\pm$ 0.001	0.86 $\pm$ 0.001	0.86 $\pm$ 0.001	0.88 $\pm$ 0.001
	Re	0.95 $\pm$ 0.001	0.99 $\pm$ 0.002	<b>1.00<math>\pm</math>0.002</b>	1.00 $\pm$ 0.001	1.00 $\pm$ 0.001	0.99 $\pm$ 0.001

A Tabela 3 apresenta os valores obtidos pela aplicação da metodologia e dos algoritmos de reconhecimento de padrões. Os valores médios obtidos entre os classificadores apontam que a metodologia atingiu valores médios de acurácia superiores a 92%, precisão em torno de 0.91

e revocação em torno de 0.91. Dentre os resultados apresentados observa-se que o melhor resultado de classificação foi obtido pela árvore de decisão utilizando a amostragem  $n_{sample} = 10^3$  baseada no desvio padrão, atingindo acurácia de 98%, 0.98 de precisão e 0.99 de revocação.

O segundo conjunto de testes consiste na aplicação do método proposto para realizar a binarização de imagens de montante de cortesia. Para tal, o treinamento do classificador foi realizado utilizando todos os padrões contidos nas amostragens, sendo que cada teste consiste no treinamento do classificador utilizando uma dada amostragem e a aplicação do classificador no conjunto de imagens a serem binarizadas. Os resultados apresentados na

Tabela 4 consistem nos resultados obtidos para 60 regiões de interesse, montante de cortesia, extraídas de imagens de cheques de bancos brasileiros.

A medida de desempenho do método proposto para a binarização de imagens de montante de cortesia é dada pelo desempenho de cada classificador treinado utilizando cada uma das seis abordagens de amostragem da base de dados. Cada classificador treinado foi utilizado para realizar a binarização das 60 imagens de montante de cortesia. A

Tabela 4 apresenta a medida de desempenho de cada um dos classificadores no processo de limiarização, dada uma base de teste. O desempenho da metodologia proposta é comparado ao desempenho do método de limiarização automática utilizando o filtro de Otsu (1979), Niblack (1986) e Sauvola e Pietikainen (2000).

**Tabela 4: Comparação da binarização das imagens de montante de cortesia utilizando a metodologia proposta.**

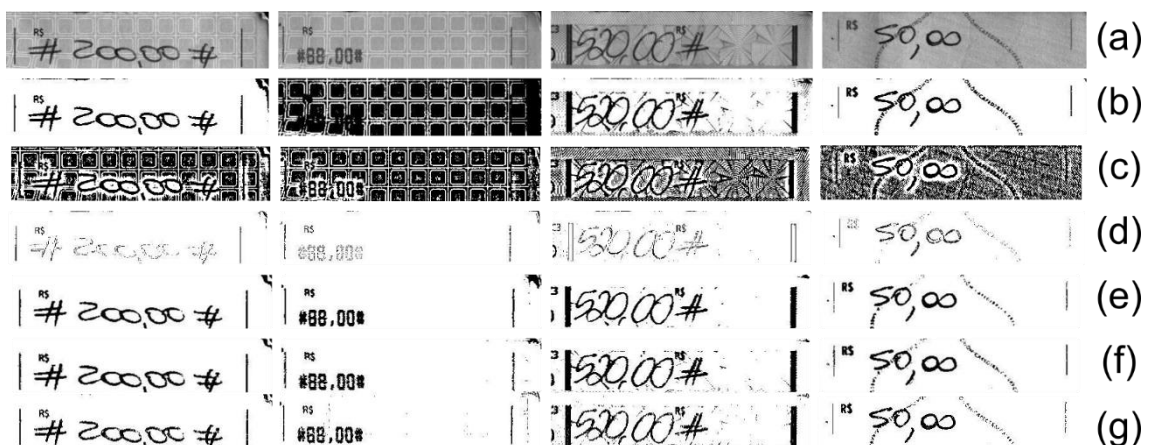
		<i>RAND</i>	<i>STD</i>				
		$n_{sample} = 10^3$	$n_{sample} = 10^3$	$n_{sample} = l$	$n_{sample} = 0,1\%$	$n_{sample} = 0,05\%$	$n_{sample} = m$
MLP	Acc	0.96	0.94	0.94	0.95	0.95	0.96
	Pr	0.71	0.67	0.67	0.67	0.71	0.76
	Re	0.95	0.97	0.95	0.93	0.93	0.95
SVM	Acc	0.96	0.95	0.93	0.93	0.93	0.96
	Pr	0.73	0.69	0.58	0.60	0.62	0.74
	Re	0.96	0.98	0.96	0.93	0.93	0.93
DT	Acc	<b>0.97</b>	0.94	0.96	0.97	0.97	0.97
	Pr	<b>0.99</b>	0.99	0.98	0.98	0.98	0.98



	Re	<b>0.97</b>	0.94	0.97	0.98	0.98	0.97
(OTSU, 1979)	Acc				0.89		
	Pr				0.56		
	Re				0.98		
(NIBLACK, 1986)	Acc				0.54		
	Pr				0.51		
	Re				0.99		
(SAUVOLA e PIETIKÄINEN, 2000)	Acc				<b>0.94</b>		
	Pr				<b>0.98</b>		
	Re				<b>0.95</b>		

A Tabela 4 mostra os resultados para a binarização utilizando a metodologia proposta. O melhor resultado foi atingido utilizando a rede neural DT com acurácia de 97%, precisão de 0.99 e revocação de 0.97. Observa-se também que a metodologia proposta obteve resultados acima de 90% de acurácia em todos os testes realizados.

Para reforçar os resultados apresentados nas tabelas são apresentadas algumas imagens binarizadas utilizando a metodologia proposta e outros métodos da literatura. As imagens utilizadas como exemplo de binarização utilizando o método proposto são apresentadas na Figura 18. Essas imagens consistem nos resultados visuais da binarização e ilustram a eficiência do método proposto. Para a limiarização dos exemplos contidos na Figura 18(a) foram apresentadas as imagens ilustrativas do processo de limiarização manual na Figura 18(b) e também são apresentados os resultados utilizando o filtro de OTSU (OTSU, 1979), na Figura 18(c) e na Figura 18(d) são apresentados os resultados utilizando a metodologia proposta com os classificadores *MLP* e *SVM* treinados com a abordagem  $STD_{n_{sample}} = 10^3$ .



**Figura 18:** Resultado visual do processo de binarização. (a) imagem em escala de cinza. (b) Otsu (1979), (c) Niblack (1986), (d) Sauvola (2000), (e) MLP, (f) SVM, (g) DT.

Os resultados apresentados nas Tabela 2 e Tabela 4 mostram que a metodologia proposta é robusta e sofisticada para a partir de uma amostragem da base de dados atingir elevados níveis de acurácia, que representa a capacidade do classificador em distinguir entre fundo e objeto. Desta perspectiva obtiveram-se resultados competitivos quando comparados a métodos consolidados na literatura. Quando observada a Figura 18(e), (f) e (g) nota-se que os resultados provenientes da metodologia proposta são melhores visualmente também.

A precisão avalia o nível de acerto do fundo: quanto maior for a precisão, maior será o nível de acerto do algoritmo em relação ao fundo. Neste aspecto o algoritmo Sauvola (2000) atingiu melhores resultados por conseguir realizar de maneira mais precisa a limpeza no fundo das imagens. No entanto, o seu valor de revocação foi baixo, implicando que o algoritmo Sauvola não conseguiu identificar os pixels do objeto corretamente, como mostra a Figura 18(d).

A revocação apresenta o nível de acerto do objeto: quanto maior for a precisão, maior a taxa de acerto do algoritmo em relação ao objeto. Niblack (1986) atingiu os maiores níveis de revocação. No entanto, seus níveis de precisão foram baixos, implicando que o algoritmo não conseguiu identificar corretamente os pixels de fundo, como ilustra a Figura 18(c).

### 1.6.3.2 Resultados Experimentais para Imagens de CMC-7

A Tabela 5 mostra a média  $\pm$  desvio padrão das medidas para cada um dos classificadores utilizando validação cruzada em 10-pastas. Os classificadores foram treinados e testados utilizando amostras da base de dados. Para cada classificador foram calculadas as medidas *Acc*, *Pr* e *Re*.

**Tabela 5: Acurácia (Acc), Precisão (Pr) e Revogação (Re) dos classificadores utilizando amostragem da base para imagens de CMC-7.**

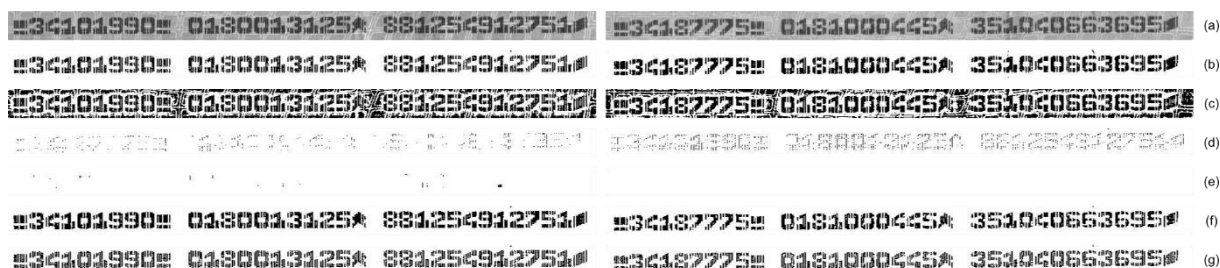
		<i>RAND</i>		<i>STD</i>		
		$n_{sample} = 10^3$	$n_{sample} = 10^3$	$n_{sample} = l$	$n_{sample} = 0,1\%$	$n_{sample} = 0,05\%$
MLP	Acc	0.94±0.003	0.96±0.003	0.96±0.002	0.96±0.001	0.96±0.001
	Pr	0.94±0.002	0.96±0.001	0.96±0.001	0.96±0.004	0.96±0.002
	Re	0.92±0.002	0.94±0.001	0.96±0.002	0.96±0.002	0.96±0.002
SVM	Acc	0.94±0.03	0.96±0.04	0.96±0.01	<b>0.97±0.02</b>	0.96±0.02
	Pr	0.95±0.02	0.95±0.04	<b>0.97±0.01</b>	<b>0.96±0.03</b>	0.96±0.03
	Re	0.93±0.08	<b>0.98±0.03</b>	0.97±0.01	<b>0.97±0.02</b>	0.97±0.02
DT	Acc	0.94±0.03	0.94±0.03	0.96±0.01	0.96±0.01	0.95±0.02
	Pr	0.95±0.02	0.94±0.03	0.97±0.009	0.95±0.02	0.95±0.03
	Re	0.95±0.05	0.94±0.05	0.97±0.01	0.95±0.03	0.94±0.05

A Tabela 5 apresenta os valores obtidos para a análise dos algoritmos de reconhecimento de padrões. Os valores médios obtidos entre os classificadores apontam que a utilização da metodologia no processo de classificação atingiu valores médios de acurácia superiores a 94%, precisão em torno de 0.94 e revogação em torno de 0.92. Dentre os resultados apresentados observa-se que o melhor resultado de classificação de padrões foi obtido utilizando o algoritmo *Support Vector Machine*. A SVM apresentou melhor resultando quando utilizada a amostragem  $n_{sample} = 0.1\%$ , atingindo acurácia de 97%, 0.96 de precisão e 0.97 de revogação.

Tabela 6: Comparação da binarização das imagens de CMC-7 utilizando a metodologia proposta e métodos da literatura.

		<i>RAND</i>		<i>STD</i>		
		$n_{sample} = 10^3$	$n_{sample} = 10^3$	$n_{sample} = 1$	$n_{sample} = 0,1\%$	$n_{sample} = 0,05\%$
MLP	Acc	0.94	0.58	0.43	0.77	0.57
	Pr	0.89	0.35	0.27	0.52	0.34
	Re	0.90	0.86	0.81	0.84	0.83
SVM	Acc	0.96	0.71	0.91	0.62	0.71
	Pr	0.97	0.93	0.94	0.92	0.93
	Re	0.97	0.67	0.95	0.54	0.66
DT	Acc	0.94	0.94	<b>0.96</b>	0.96	0.95
	Pr	0.95	0.94	<b>0.97</b>	0.95	0.95
	Re	0.95	0.94	<b>0.97</b>	0.95	0.94
(OTSU, 1979)	Acc			<b>0.97</b>		
	Pr			<b>0.98</b>		
	Re			<b>0.97</b>		
(NIBLACK, 1986)	Acc			0.54		
	Pr			0.51		
	Re			0.99		
(SAUVOLA e PIETIKÄINEN, 2000)	Acc			0.77		
	Pr			0.72		
	Re			0.98		

A Tabela 6 apresenta a acurácia (Acc), a precisão (Pr) e a revocação (Re) para os resultados do processo de binarização das imagens de CMC-7. Observa-se que a metodologia proposta manteve a acurácia superior à 92% entre a Tabela 4 e a Tabela 6. Entretanto, o algoritmo de Otsu (1979) apresentou os melhores resultados de binarização para as imagens de CMC-7. A metodologia proposta apresentou-se como segunda melhor opção entre os resultados apresentados para realizar a binarização de imagens de CMC-7, com acurácia 21% e 45% superior em relação aos algoritmos de Sauvola (2000) e Niblack (1986), respectivamente.



**Figura 19:** Resultado visual do processo de binarização. (a) imagem em escala de cinza. (b) Otsu (1979), (c) Niblack (1986), (d) Sauvola (2000), (e) MLP, (f) SVM, (g) DT.

A Figura 19 apresenta alguns dos resultados visuais de binarização utilizando a metodologia proposta e algoritmos da literatura. Por meio dessa figura observa-se que o resultado Otsu (1979) foi levemente superior aos demais. Os baixos valores de precisão na Figura 19(d) e (e) apresentam imagens que possuem objetos não bem definidos, em sua maioria objetos suavizados e que não permitem o reconhecimento dos caracteres. Por sua vez, as imagens na Figura 19(c) apresenta imagens com taxa de revocação baixa com fundo contém muitos *pixels* ruidosos.

## 1.7 APLICAÇÃO DA METODOLOGIA PROPOSTA A IMAGENS DO BANCO DE IMAGENS DO DIBCO 2009

A Competição Internacional de Binarização de Imagens de Documentos tem como objetivo avaliar metodologias para binarização de imagens de documentos. Para a avaliação destas metodologias é proposta a utilização de um banco de dados padronizado criado por meio de imagens de impressão de escritos a máquina e escritas a mão (GATOS, NTIROGIANNIS e PRATIKAKIS, 2009). As dez imagens utilizadas para a competição são cinco imagens de impressão de escritos por máquina e cinco imagens de escritos à mão. As imagens de testes foram criadas em duas etapas: 1) escâner de imagens de impressão de máquina ou escrita à mão em fundo branco; e 2) sobreposição destas imagens em imagens de fundos históricos (GATOS, NTIROGIANNIS e PRATIKAKIS, 2009). Dado o processo de criação das imagens do banco de imagens do DIBCO (2009) os padrões desejados, ou rótulos, consistem nas próprias imagens escaneadas de escritos em fundo branco.

Para os experimentos utilizando esta base de dados não é necessária a extração da região de interesse, uma vez que o resultado proposto é a própria imagem. O processo de construção da matriz de padrões é dado pela utilização de uma vizinhança de tamanho 5. Dado o processo

de construção do banco de dados onde os padrões desejados foram disponibilizados pelas imagens escaneadas contendo os escritos por impressão de máquina ou a mão o processo de rotulação dos padrões é dado pela utilização destas imagens.

A amostragem da base de dados foi feita com  $n_{sample} = 10^3$  padrões baseados no desvio padrão dos níveis de cinza de cada uma das 10 imagens para o treinamento do classificador. O classificador utilizado para este experimento foi a Árvore de Decisão (DT) com 20 nós de profundidade máxima e confiança de  $2,5 \times 10^{-1}$ . A Árvore de Decisão foi escolhida por ter apresentado os melhores resultados nos experimentos anteriores.

A metodologia proposta foi comparada aos métodos apresentados na competição por Gatos et. al. (2009), tal qual as medidas apresentadas, são elas: *F-Measure*, *PSNR*, *NRM* e *MPM*. Os resultados experimentais são apresentados na Tabela 7.

A metodologia proposta apresentou melhores resultados considerando a medida *F-Measure* que foi utilizada para criar o ranking dos resultados para Gatos (2009). Para as demais medidas a metodologia proposta atingiu valores competitivos em relação ao demais métodos da competição DIBCO (2009).

Diferente dos métodos apresentados por Gatos (2009) na competição a metodologia proposta possui uma etapa de treinamento. Esta etapa consiste na aquisição de uma porção da imagem a ser binarizada para que um dado classificador seja capaz de rotular cada um dos *pixels* da imagem. De fato, esta etapa permite que a metodologia alcance melhores taxas avaliativas uma vez que ela utiliza informações especialistas da própria imagem a priori para a classificação dos demais *pixels*.

**Tabela 7:** Resultados avaliativos para todos os métodos submetidos ao DIBCO obtidos em Gatos (2009) e metodologia proposta

<b>Metodologia</b>	<b>F-Measure (%)</b>	<b>PSNR</b>	<b>NRM (<math>\times 10^{-2}</math>)</b>	<b>MPM (<math>\times 10^{-3}</math>)</b>
<b>Proposta</b>	<b>97.60</b>	<b>14.35</b>	<b>12.10</b>	<b>0.43</b>
26	91.24	18.66	4.31	0.55
14	90.06	18.23	4.75	0.89
24	89.34	17.79	5.32	1.90
10	87.03	17.21	7.03	0.57
9 (a)	87.89	17.12	7.73	0.97
8	87.71	16.86	5.99	2.19
33 (c)	86.35	16.66	6.03	1.45
9 (b)	87.16	17.08	8.5	0.74
4	86.53	16.47	5.41	1.76
34 (a)	87.49	16.83	7.76	1.57
33 (b)	85.66	17.01	11.37	0.52
6	86.93	16.61	7.29	2.58
11	85.72	16.44	8.94	1.12
34 (b)	85.99	16.37	8.28	1.46
35	85.11	15.75	5.38	2.22
33 (a)	84.59	16.66	11.48	0.61
1	85.06	16.36	6.49	3.78
34 (c)	84.78	16.02	8.73	1.50
25	83.99	15.58	4.18	4.60
3	85.30	15.68	7.59	4.18
7 (c)	85.17	16.04	9.93	1.93
17	83.98	15.81	4.51	5.48
34 (d)	84.03	15.86	8.78	1.40
29	84.69	16.33	7.96	3.83
18	83.74	15.22	4.62	3.86
23	82.50	15.11	4.47	3.62
12	83.53	15.59	4.91	5.34
22	83.54	15.53	7.62	3.54
7 (a)	84.57	15.67	7.81	5.84
28	84.25	16.42	9.13	7.46
30	83.62	15.57	7.67	5.53
2	83.10	14.74	5.18	7.11
19	79.71	16.62	9.93	4.55
7 (b)	80.74	14.86	5.98	9.60
5	80.90	14.64	8.17	4.22
15	74.12	15.05	18.07	2.57
16	82.27	14.96	8.04	41.30
21	75.86	13.34	15.45	2.51
20	80.43	14.37	8.21	7.70
27	82.74	14.78	10.12	56.22
13	35.28	12.44	36.60	2.68
31	61.48	9.22	14.69	86.03
32	58.77	9.27	18.77	118.02

## 5. CONCLUSÕES E TRABALHOS FUTUROS

O presente trabalho propôs uma metodologia de binarização de áreas de imagens de cheque utilizando algoritmos de aprendizagem supervisionada. As áreas, montante de cortesia e CMC-7, foram utilizadas para realizar a avaliação da metodologia de binarização em contraste as técnicas consolidadas na literatura. As duas áreas selecionadas para o teste da metodologia são consideradas importantes no processo de compensação de cheques de bancos brasileiros. O processo de binarização é a primeira etapa de processamento de imagens que é utilizado para o reconhecimento de informações das imagens de cheque.

Foram realizados testes para avaliar o desempenho de três tipos de classificadores utilizando o fluxo metodológico proposto, são eles: rede neural artificial do tipo MLP; *Máquinas de Vetores Suporte* (SVM); e Árvore de Decisão (DT). Este primeiro conjunto de testes teve por objetivo a análise do desempenho dos classificadores utilizando o fluxo metodológico proposto. O classificador que apresentou melhor desempenho geral foi o algoritmo baseado em Árvore de Decisão (DT), entretanto, todos os classificadores apresentaram desempenho com acurácia superior a 92%. Esta é uma importante etapa para validação da metodologia e mostra sua capacidade de adaptação a diversos tipos de classificadores baseados em aprendizagem supervisionada.

O segundo conjunto de testes foi realizado a fim de analisar o desempenho da metodologia para a tarefa de binarização de imagens. Esta etapa consistiu em realizar a binarização de imagens a partir do treinamento da metodologia utilizando porções amostrais destas. Os resultados mostraram que a presente metodologia é robusta para resolver casos que necessitam de uma binarização especialista, como as imagens de cheques de bancos brasileiros. Considerando-se um escopo previamente conhecido, como o das imagens de cheque, a metodologia foi capaz de aprender com o processo de binarização manual tornando-se eficiente e eficaz a partir de poucas amostras de um banco de imagens. A metodologia apresentou resultados compatíveis e competitivos quando comparados aos algoritmos existentes na literatura.

Um terceiro conjunto de testes comparou a metodologia a métodos usados na competição DIBCO (2009). A metodologia proposta ficou entre as três primeiras posições dentre as medidas avaliativas da competição e apresentou o melhor resultado em relação à medida F-M. Como fruto deste trabalho obteve-se uma publicação associada (FELIX, SILVA e DE CASTRO, 2015).



Como trabalhos futuros para esta pesquisa, propõem-se:

- Aplicação da metodologia a outros bancos de imagens, como documentos históricos, placas de carro, envelopes de correio, mapas, jornais velhos, livros digitalizados, dentre outros;
- Análise de sensibilidade da metodologia ao tamanho da vizinhança (parâmetro  $V$ ) na construção dos padrões e na busca de um valor ótimo de vizinhança;
- Verificação da correlação entre o nível de homogeneidade de uma vizinhança e a acurácia do processo de classificação; e
- Utilização de algoritmos de aprendizagem não supervisionada para eliminar a necessidade de interação humana no processo.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AHMAD, I.; MAHMOUD, S. A. **Arabic Bank Check Analysis and Zone Extraction**. 9th International Conference, ICIAR 2012. Aveiro, Portugal: [s.n.]. 2012. p. 141-148.
- ALGINAHI, Y.; SID-AHMED, M. A.; AHMADI, M. **Local thresholding of composite documents using multi-layer perceptron neural network**. Circuits and Systems, 2004. MWSCAS '04. The 2004 47th Midwest Symposium on. [S.l.]: [s.n.]. 2004.
- BHANDARI, A. K.; KUMAR, A.; SINGH, G. K. Modified artificial bee colony based computationally efficient multilevel thresholding for satellite image segmentation using Kapur's, Otsu and Tsallis functions. **Expert Systems with Applications**, 15 February 2014. 1573-1601.
- CASH, G. L.; HATAMIAN, M. Optical character recognition by the method of moments. **Computer Vision, Graphics, and Image Processing**, 39, September 1987. 291-310.
- CAVALCANTI, G. D. C. et al. **"A Heuristic Binarization Algorithm for Documents with Complex Background"**. Image Processing, 2006 IEEE International Conference on. [S.l.]: [s.n.]. 2006. p. 389-392.
- CHANG, C.-C.; LIN, C.-J. LIBSVM: A Library for Support Vector Machines. **ACM Trans. Intell. Syst. Technol.**, p. 1-27, 2011.
- CHANG, Y.-F.; PAI, Y.-T.; RUAN, S.-J. **An efficient thresholding algorithm for degraded document images based on intelligent block detection**. Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on. [S.l.]: [s.n.]. 2008. p. 667-672.
- CHEN, G.; ZHANG, H.-G.; GUO, J. **Efficient Computation of Mahalanobis Distance in Financial Hand-Written Chinese Character Recognition**. Machine Learning and Cybernetics, 2007 International Conference on. [S.l.]: [s.n.]. 2007. p. 2198-2201.
- CHEN, Y.-L. **A robust technique for character string extraction from complex document images**. Information Technology, 2008. ITSIm 2008. International Symposium on. Kuala Lumpur: IEEE. 2008. p. 1-9.
- CHERIET, M.; SAID, J. N.; SUEN, C. Y. . **"A recursive thresholding technique for image segmentation"**. Image Processing, IEEE Transactions on. [S.l.]: [s.n.]. 1998. p. 918,921.

DE CASTRO, L. N. Fundamentals of natural computing: an overview. **Physics of Life Reviews**, v. 4, n. 1, p. 1-36, 2006.

DE CASTRO, L. N. **Fundamentals of Natural Computing: Basic Concepts, Algorithms, and Applications**. [S.l.]: CRC Press, 2006.

DE CASTRO, L. N.; VON ZUBEN, F. J. **Optimised Training Techniques For Feedforward Neural Networks**. [S.l.]. 1998.

DE JESUS, A.; FACON, J. **Segmentation of Brazilian bank check logos without a priori knowledge**. Information Technology: Coding and Computing, 2000. Proceedings. International Conference on. [S.l.]: [s.n.]. 2000. p. 259,263.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**. 2nd. ed. New York: Wiley-Interscience, 2000.

FEBRABAN. FEBRABAN, 24 set. 2014. Disponível em: <[www.febraban.org.br](http://www.febraban.org.br)>. Acesso em: 24 set. 2014.

FELIX, R.; SILVA, L. A.; DE CASTRO, L. N. Thresholding the Courtesy Amount of Brazilian Bank Checks Using a Local Methodology. In: BAJO, J., et al. **Highlights of Practical Applications of Agents, Multi-Agent Systems, and Sustainability - The PAAMS Collection**. Salamanca: Springer International Publishing, v. 524, 2015. p. 213-221. Disponível em: <[http://link.springer.com/chapter/10.1007/978-3-319-19033-4\\_18#](http://link.springer.com/chapter/10.1007/978-3-319-19033-4_18#)>.

FILHO, J. M. D. O. et al. Os Impactos da Reengenharia no processo de compensação: Estudo em uma Instituição Financeira. **ABCustos Associação Brasileira de Custos**, v. IX, n. 2, p. 80-110, maio-agosto 2014.

GAO, G. et al. **Fast detecting and tracking algorithm of Infrared target under complex background**. Electronics and Optoelectronics (ICEOE), 2011 International Conference on. Dalian, Liaoning: IEEE. 2011. p. V2-430 - V2-434.

GATOS, B.; NTIROGIANNIS, K.; PRATIKAKIS, I. **ICDAR 2009 Document Image Binarization Contest (DIBCO 2009)**. Document Analysis and Recognition, 2009. ICDAR '09. 10th International Conference on. Barcelona: IEEE. 2009. p. 1375 - 1382.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. Boston, MA, USA: ADDISON-WESLEY PUBLISHING COMPANY, INC, 1992.

GOVINDAN, V. K.; SHIVAPRASAD, A. P. Character Recognition - A Review. **Pattern Recognition**, v. 23, n. 7, p. 671-683, 1990.

HONGGANG, Z. et al. **Bank Check Image Binarization Based on Signal matching**. Information, Communications and Signal Processing, 2005 Fifth International Conference on. [S.l.]: [s.n.]. 2005. p. 1430-1433.

HORNG, M.-H. Multilevel thresholding selection based on the artificial bee colony algorithm for image segmentation. **Expert Systems with Applications**, 38, October 2011. 13785-13791.

ILACHINSKI, A. **ellular Automata: A Discrete Universe**. [S.l.]: World Scientific, 2001.

KANNAN, R. J.; SUBRAMANIAN, S. **An Adaptive Approach of Tamil Character Recognition Using Deep Learning with Big Data-A Survey**. Emerging ICT for Bridging the Future - Proceedings of the 49th Annual Convention of the Computer Society of India (CSI) Volume 1. [S.l.]: [s.n.]. 2015. p. 557-567.

KARRAS, D. A. **Face detection in complex backgrounds based on the integration of neural classifiers through similarity analysis of higher order representation patterns**. Imaging Systems and Techniques (IST), 2011 IEEE International Conference on. Penang: IEEE. 2011. p. 335 - 340.

KIM, K. K. et al. **Legal amount recognition based on the segmentation hypotheses for bank check processing**. Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on. [S.l.]: [s.n.]. 2001. p. 964-996.

KIM, P.-K. **Automatic text location in complex color images using local color quantization**. TENCON 99. Proceedings of the IEEE Region 10 Conference (Volume:1 ). Cheju Island: IEEE. 1999. p. 629 - 632 vol.1.

KURBAN, T. et al. Comparison of evolutionary and swarm based computational techniques for multilevel color image thresholding. **Applied Soft Computing**, 23, October 2014. 128-143.

LI, C. H.; LEE, C. K. Minimum cross entropy thresholding. **Pattern Recognition**, v. 26, n. 4, p. 617-625, 1993.

LI, L.; GONG, R.; CHEN, W. Gray level image thresholding based on fisher linear projection of two-dimensional histogram. **Pattern Recognition**, v. 30, n. 5, p. 743-749, May 1997.

LIU, D.; CHEN, Y. **Binarizing the Courtesy Amount Field on Color Chinese Bank Check Images**. Document Analysis and Recognition (ICDAR), 2011 International Conference on. [S.l.]: [s.n.]. 2011. p. 774-778.

MAHMOUD, S.; AL-KHATIB, W. Recognition of Arabic (Indian) bank check digits using log-gabor filters. **Applied Intelligence** , 01 dez. 2011. 445-456.

MELLO, C.A.B. ; BEZERRA, B.L.D. ; ZANCHETTIN, C. ; MACARIO, V. **An Efficient Thresholding Algorithm for Brazilian Bank Checks**. Ninth International Conference on Document Analysis and Recognition, 2007. ICDAR 2007.. Parana: [s.n.]. 2007.

MORI, S.; NISHIDA, H.; YAMADA, H. **Optical Character Recognition**. 1. ed. New York: John Wiley & Sons, Inc, v. 1, 1999.

NEVES, R. F. P. et al. **A new algorithm to threshold the courtesy amount of Brazilian bank checks**. IEEE International Conference on Systems, Man and Cybernetics, 2008. SMC 2008. Singapore: IEEE. 2008. p. 1226 - 1230.

NEVES, R. F. P. et al. Thresholding the Courtesy Amount of Brazilian Bank Checks Based on Tsallis Entropy. **Latin America Transactions, IEEE (Revista IEEE America Latina)**, v. 7, n. 6, p. 726 - 731, 2009.

NIBLACK, W. **An Introduction to Digital Image Processing**. Englewood Cliffs: Prentice Hall, 1986.

O'DELL, W. G. et al. Traversing and labeling interconnected vascular tree structures from 3D medical images. **Proc. SPIE 9034, Medical Imaging 2014: Image Processing**, p. 90343C-90343C-15, 2014.

OTSU, N. **A Threshold Selection Method from Gray-Level Histograms**. IEEE Transactions on Systems, Man and Cybernetics. [S.l.]: [s.n.]. 1979. p. pp. 62-66.

PALACIOS, R.; GUPTA, A. A System for Processing Handwritten Bank Checks Automatically. **Image Vision Computing**, October 2008. 1297-1313.

PICTURE Thresholding Using an Iterative Selection Method. **Systems, Man and Cybernetics, IEEE Transactions on**, v. 8, n. 8, p. 630-632, 1978.

RABEE, A.; BARHUMI, I. **License plate detection and recognition in complex scenes using mathematical morphology and support vector machines**. Systems, Signals and Image Processing (IWSSIP), 2014 International Conference on. Dubrovnik: IEEE. 2014. p. 59-62.

RABELO, J. C. et al. **A Multi-Layer Perceptron approach to threshold documents with complex background**. In Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on. Chicago: IEEE. 2011. p. 2523-2530.

SAHOO, P. K.; S. SOLTANI, A. K. C. W.; CHEN, Y. C. A survey of thresholding techniques. **Comput. Vision Graph. Image Process.** **41**, p. 233-260, 1988.

SAHOO, P.; WILKINS, C.; YEAGER, J. Threshold selection using Renyi's entropy. **Pattern Recognition**, v. 30, n. 1, p. 71-84, 1997.

SAUVOLA, J.; PIETIKÄINEN, M. Adaptive document image binarization. **Pattern Recognition**, v. 33, n. 2, p. 225-236, 2000.

SEZGIN, M.; SANKUR, B. Survey over image thresholding techniques and quantitative performance evaluation. **Journal of Electronic Imaging**, v. 131, n. 1, p. 146-168, 01 Jan 2004.

SOLOMON, C.; BRECKON, T. **Fundamentos de Processamento Digital de Imagens: Uma abordagem prática com exemplos em matlab**. Rio de Janeiro: LTC, 2013.

STOCKMAN, G.; SHAPIRO, L. G. **Computer Vision**. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.

SZABO, A. **Agrupamento esperto**. São Paulo: Esperta, 2014.

TRUSSELL, H. J. Comments on "Picture Thresholding Using an Iterative Selection Method". **Systems, Man and Cybernetics, IEEE Transactions on**, v. 9, n. 5, p. 311-311, 1979.

WANG, Y. et al. **Importance sampling based discriminative learning for large scale offline handwritten Chinese character recognition**. Pattern Recognition. [S.l.]: [s.n.]. 2015. p. 1125-1234.

WOLFRAM, S. **A New Kind of Science**. [S.l.]: Wolfram Media, 2002.

WU, L.; MA, S.; LU, H. **An effective entropic thresholding for ultrasonic images**. Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on. Brisbane, Qld., Australia: [s.n.]. 1998. p. 1552-1554.

YAGER, R. R. ON THE MEASURE OF FUZZINESS AND NEGATION Part I: Membership in the Unit Interval. **International Journal of General Systems**, v. 5, n. 4, p. 221-229, 1979.

YEN-LIN CHEN. **A robust technique for character string extraction from complex document images**. Information Technology, 2008. ITSIM 2008. International Symposium on. [S.l.]: [s.n.]. 2008. p. 1-9.

YOUNG, D. P.; FERRYMAN, J. M. **PETS Metrics: On-Line Performance Evaluation Service**. Proceedings 2nd Joint IEEE International Workshop on VS-PETS. Beijing: IEEE. 2005. p. 317-324.