UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

JOÃO PEDRO MARQUES BASSO

METODOLOGIA DE PROCESSAMENTO DE IMAGENS PARA LEITURA E RECONHECIMENTO DIGITAL SELETIVO DE TEXTO

CORNÉLIO PROCÓPIO

2022

JOÃO PEDRO MARQUES BASSO

METODOLOGIA DE PROCESSAMENTO DE IMAGENS PARA LEITURA E RECONHECIMENTO DIGITAL SELETIVO DE TEXTO

IMAGE PROCESSING METHODOLOGY FOR SELECTIVE DIGITAL TEXT READING AND RECOGNITION

Trabalho de conclusão de curso de graduação apresentada como requisito para obtenção do título de Bacharel em Engenharia Elétrica da Universidade Tecnológica Federal do Paraná (UTFPR).
Orientador: Prof. Dr. Wagner Endo

CORNÉLIO PROCÓPIO

2022

Universidade Tecnológica Federal do Paraná
Campus Cornélio Procópio
Departamento Acadêmico de Elétrica
Curso de Engenharia Elétrica

FOLHA DE APROVAÇÃO

João Pedro Marques Basso

METODOLOGIA DE PROCESSAMENTO DE IMAGENS PARA LEITURA E RECONHECIMENTO DIGITAL SELETIVO DE TEXTO

Trabalho de conclusão de curso apresentado às 10:30hs do dia 23/06/2022 como requisito parcial para a obtenção do título de Engenheiro Eletricista no programa de Graduação em Engenharia Elétrica da Universidade Tecnológica Federal do Paraná. O candidato foi arguido pela Banca Avaliadora composta pelos professores abaixo assinados. Após deliberação, a Banca Avaliadora considerou o trabalho aprovado.

_____
Prof(a). Dr(a). Wagner Endo - Presidente (Orientador)

_____
Prof(a). Dr(a). Cristiano Marcos Agulhari - (Membro)

_____
Prof(a). Dr(a). Paulo Rogério Scalassara - (Membro)

A folha de aprovação assinada encontra-se na coordenação do curso.

Dedico este trabalho à minha família, por sempre me apoiar em minha jornada.

AGRADECIMENTOS

Agradeço primeiramente a Deus, à minha mãe Lucimeire Marques Basso e meu pai Silvio Sérgio Basso, pois eles foram essenciais no meu desenvolvimento como ser humano e profissional.

A minha amada Giovana Cardoso, por sempre estar ao meu lado, nos momentos de sucesso e nos momentos difíceis, sendo meu porto seguro para todas as decisões que tomei até aqui.

Imensamente ao meu orientador Prof. Dr. Wagner Endo, pela sabedoria, paciência e atenção em que me guiou nesta trajetória e a banca avaliadora Paulo Scalassara e Cristiano Agulhari.

Aos meus amigos de graduação João Biachi, Yuri Piccolo, Anderson, Guilherme, Victor, Tarumoto e todos outros que fizeram parte de minha jornada na graduação. Aos meus amigos Diego Santo, Paula Pedrozo, Lucas e Pedro Odorizzi, Leilaine, José, Lucas e todos os outros de minha cidade natal.

Enfim, a todos os que por algum motivo contribuíram para a realização deste trabalho e de meu desenvolvimento.

RESUMO

BASSO, João Pedro Marques. **METODOLOGIA DE PROCESSAMENTO DE IMAGENS PARA LEITURA E RECONHECIMENTO DIGITAL SELETIVO DE TEXTO.** 2022. 32 f. Trabalho de Conclusão de Curso (Graduação) – Engenharia Elétrica. Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2022.

O trabalho teve por objetivo implementar uma metodologia de seleção de textos a partir da diferença de componentes de cores afetadas em imagens digitais após o uso de marcadores de textos e a aplicação em reconhecimento digital do texto selecionado utilizando uma biblioteca de OCR no Software MatLab. As atividades foram realizadas em duas imagens digitais, uma criada através de um software de edição e a segunda obtida em um banco de dados, ambas marcadas com marcador de texto digital. Foram utilizadas as ferramentas de manipulação de imagens do MatLab, para identificar a componente de cor afetada nas imagens e isolar esta componente a fim de obter uma seleção do texto em ambas as imagens. Para a imagem ideal, obteve-se um resultado satisfatório, com o texto selecionado e reconhecido pela biblioteca OCR. Ao passo que ao aplicar a metodologia na imagem de um banco de dados os ruídos presentes na imagem dificultaram a seleção, impossibilitando a leitura pela biblioteca OCR. Em uma tentativa de contornar a falha, utilizou-se da técnica de detecção de bordas para uma definição melhor do texto selecionado. Os resultados obtidos com a imagem ideal foram novamente satisfatórios, todavia ainda sem reconhecimento de texto para a imagem do banco de dados. O estudo forneceu uma visão positiva sobre a seleção de textos a partir de marcadores, espera-se que em trabalhos futuros seja possível aplicar a técnica de forma mais precisa e automatizada a fim de obter-se um resultado preciso para imagens reais e em larga escala.

Palavras-chave: seleção; segmentação; marcadores; reconhecimento; digitalização.

ABSTRACT

BASSO, João Pedro Marques. **IMAGE PROCESSING METHODOLOGY FOR SELECTIVE DIGITAL TEXT READING AND RECOGNITION.** 2022. 32 p. Final paper (Graduate) – Electrical Engineering. Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2022.

The objective of this work was to implement a text selection methodology based on the difference of affected color components in digital images after the use of text markers and the application in digital recognition of the selected text using an OCR library in MatLab Software. The activities were carried out in two digital images, one created using an editing software and the second obtained from a database, both marked with a digital text marker. MatLab image manipulation tools were used to identify the affected color component in the images and isolate this component in order to obtain a selection of the text in both images. For the ideal image, a satisfactory result was obtained, with the text selected and recognized by the OCR library. On the other hand, when applying the methodology in the image of a database, the noise present in the image made the selection difficult, making it impossible to read it by the OCR library. In an attempt to circumvent the flaw, the edge detection technique was used to better define the selected text. The results obtained with the ideal image were again satisfactory, but still without text recognition for the database image. The study provided a positive view on the selection of texts from markers, it is expected that in future works it will be possible to apply the technique in a more precise and automated way in order to obtain an accurate result for real and large-scale images.

Keywords: selection; segmentation; bookmarks; recognition; scanning.

LISTA DE FIGURAS

LISTA DE SIGLAS

OCR *Optical Character Recognition*

RGB *Red, Green, Blue*

PNG *Portable Network Graphics*

SUMÁRIO

1 INTRODUÇÃO

O uso de imagens digitais tem se tornado cada vez mais frequente em diversas áreas. Com o passar dos anos, as informações vêm sendo produzidas em uma escala acelerada. É imprescindível que essas informações estejam registradas em documentos, para contextualizar o conhecimento e serem disponibilizadas para consulta.

De acordo com Rodrigo José (2013, p. 7) o grande volume de papel nas empresas requer uma gestão efetiva dos documentos gerados e recebidos. Tê-los de uma forma mal estruturada pode causar prejuízos para a organização, uma vez que a perda, extravio ou até mesmo o tempo excessivo para encontrá-los tem um custo.

Assim faz-se necessário organizar digitalmente os arquivos e como cita Barboza (2013), o processamento digital de imagens é uma área de pesquisa em constante expansão. A cada dia surgem novos dispositivos, com maior capacidade de processamento e memória, possibilitando novas aplicações e trazendo dificuldades e problemas originais a serem resolvidos. As imagens digitais hoje estão presentes em todas as áreas da atividade humana, facilitando cada vez mais a obtenção e armazenamento destes arquivos.

Além de digitalizar os documentos, outra prática bastante utilizada é o reconhecimento dos caracteres, a fim de facilitar a manipulação do conteúdo armazenado. Como cita Campestrini (2013, p. 65), apesar de ser uma tarefa corriqueira para o ser humano, o reconhecimento de caracteres, ação necessária para o armazenamento correto de textos, é uma tarefa complexa, o que torna qualquer imperfeição no texto digitalizado uma grande dificuldade do ponto de vista computacional.

É comum também realizarmos marcações no texto, visando resumir ou destacar partes essenciais. Como citado por Barboza (2013), uma caneta selecionadora, caneta marcadora de feltro, ou simplesmente um marcador, é uma caneta que tem a sua própria fonte de tinta, e geralmente, uma ponta feita de um material poroso, tal como o feltro ou nylon. Logo ao realizar marcações utilizando as mesmas, alteramos as cores do papel sendo possível realizar a identificação dessa alteração através de algoritmos e posteriormente o recorte e reconhecimento seletivo.

A maioria dos marcadores de texto disponível comercialmente afeta um ou mais componentes de cores do documento original. Análises realizadas no trabalho de Barboza (2013) demonstraram que o marcador diminui o valor da intensidade do componente de cor original, em relação às áreas não marcadas como mostrado na Figura 1.

Figura 1: Componentes afetadas por marcadores

Marcador	Cor	Componente Afetada
	Amarela	Azul
	Azul	Vermelho/Verde
	Verde	Vermelho/Azul
	Laranja	Verde/Azul
	Magenta	Vermelho/Verde/Azul

Fonte: Barboza (2013)

Sabendo-se das componentes afetadas pelos marcadores é possível utilizar esta componente para identificar certa região do documento, digitalizando-se apenas o que foi destacado.

Barboza (2013) cita que a filtragem de um ruido com conhecimento do que se deseja retirar e/ou manter na imagem resulta na diminuição dos dados que serão processados nas etapas posteriores, diminuindo o tempo necessário para o processamento e melhorando a eficiência destes processos.

Desta forma, este trabalho visa utilizar de maneira inteligente a diferença nas componentes causada pelos marcadores, inicialmente tratando digitalmente a imagem a ser analisada. É possível assim selecionar a componente causada pelo marcador e a correção das imperfeições que o mesmo causa na imagem com a finalidade de se obter um documento digitalizado e resumido.

Foi realizada então uma verificação da possibilidade de se trabalhar com o texto obtido visto que imagens textuais como jornais, panfletos e documentos antigos geralmente possuem letras cursivas e fundos coloridos ou heterogêneo, o que aumenta os erros no reconhecimento digital do texto.

1.1 Objetivo Geral

Desenvolver um algoritmo capaz de identificar a diferença de componente afetada por marcador de texto amarelo em textos gerados e obtidos por banco de dados digitalizados a fim obter um arquivo de texto seletivo através de algoritmos de OCR - *Optical Character Recognition,* em português Reconhecimento Ótico de Caracteres.

1.2 Objetivos Específicos

- Estudar a composição de imagens digitais;
- Consultar ferramentas para seleção de imagens via reconhecimento de cores;
- Discorrer sobre o tratamento de imagens utilizando ferramentas computacionais;
- Explorar algoritmos de reconhecimento ótico de caracteres (OCR);
- Implementar uma solução automatizada para OCR seletivo.

1.3 Estrutura do trabalho

O trabalho está dividido em sete capítulos, no Capítulo 2 está à conceituação teórica dos temas abordados neste trabalho e o Capítulo 3 expõe a ferramenta ser utilizada para a implementação do algoritmo.

O Capítulo 4 descreve uma proposta de implementação do algoritmo, enquanto no Capítulo 5 são apresentados os resultados preliminares e por fim no Capítulos 6 e 7 tem-se o resultado das atividades desenvolvidas e as considerações finais.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo está presente a fundamentação teórica do trabalho, com enfoque no reconhecimento de cores RGB (*Red, Green, Blue*) em imagens digitais, também serão abordadas técnicas e etapas de um algoritmo de OCR.

2.1 Processamento digital de imagens

A imagem digital surgiu em meados de 1957, a partir de *scanners* com resoluções abaixo de 176x176 pixels, por conta de limitações de *hardware*. Desde então, devido ao advento da tecnologia, encontram-se disponíveis no mercado métodos de obtenção de imagens digitais cada vez mais fidedignas à realidade, com resoluções expressivamente maiores advindas das melhorias em *hardware*.

Este avanço, possibilitou a criação de estratégias de armazenamento a longo prazo de acervos físicos que são codificados para formatos digitais, facilitando o armazenamento e transmissão das informações em meios eletrônicos.

Jayaraman et al (2009, p.43) descreve imagens, no âmbito computacional, como matrizes compostas por elementos chamados pixels, sendo esses a menor amostra de uma imagem que representa o brilho ou intensidade em um ponto.

De acordo com Petrou e Petrou (2010), uma imagem pode ser definida como uma função bidimensional dada por:

$$f(x, y) \tag{1}$$

Sendo:
x e y – coordenadas espaciais;
f – amplitude, chamada intensidade;

Uma imagem é caracterizada como digital quando f(x,y) for discretizada tanto em coordenadas espaciais como em brilho. O valor de brilho digitalizado é chamado de escala de cinza (Petrou e Petrou, 2010), que consiste em uma média das componentes RGB da imagem.

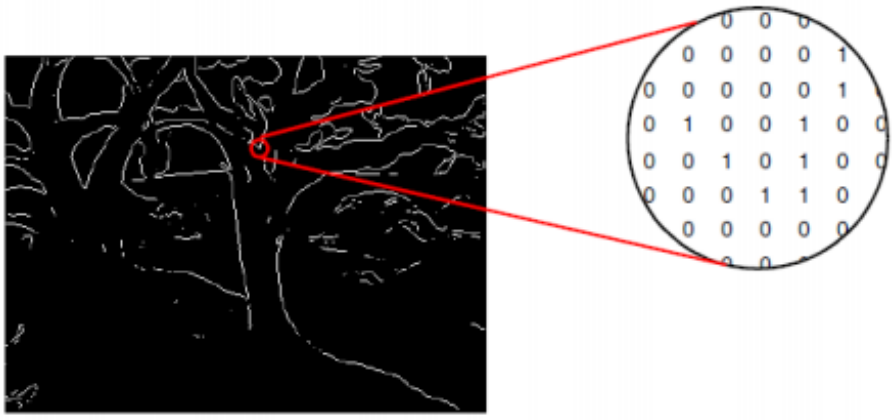
2.1.1 Representação de uma imagem binária

Na Figura 2 tem-se uma imagem binária onde cada pixel assume apenas um de dois valores discretos possíveis: 0 (off) e 1 (on), sendo o bit alto branco e o bit baixo preto, com essa configuração são representadas apenas duas cores ou uma imagem monocromática.

Estas imagens são geralmente geradas a partir de imagens em tons de cinza ou RGB e são em geral com a lógica mostrada do pseudo-código abaixo:

```
se "tom de cinza" > limite
"pixel" igual a 1
senão "pixel" igual a 0
```

Figura 2: Representação de uma imagem binária



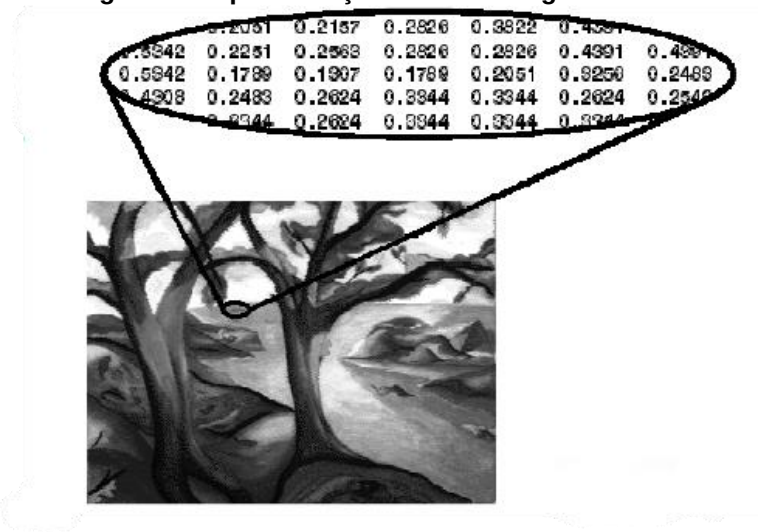
Fonte: Nobre (2013)

2.1.2 Imagens de intensidade (em escala de cinza)

A Figura 3 mostra uma imagem com diferentes níveis de intensidade, que consiste apenas de uma matriz cujos valores representam intensidades dentro de alguma faixa.

Uma correlação com as cores RGB acontece quando as três componentes são iguais, gerando a cor cinza. Nesse ponto é interessante relembrar o conceito de luminosidade, é dada pela soma das componentes RGB, resultando em uma intensidade de cinza, onde os valores estão entre 0 e 1.

Figura 3: Representação de uma imagem em escala de cinza



Fonte: Nobre (2013)

2.1.3 Representação de uma imagem RGB

Na Figura 4 é representada uma imagem RGB que é armazenada em uma matriz $m \times n \times 3$ onde cada camada $m \times n$ define uma das componentes R (Vermelho), G (Verde) e B (Azul) para cada pixel.

São imagens mais pesadas pois requerem 24 bits/pixel, sendo 8 bits para cada componente RGB, representando 256 tons de cada componente ou 16.777.216 tons de cores distintos obtidos através da combinação das componentes.

Figura 4: Representação de uma imagem RGB

Fonte: Nobre (2013)

2.2 Reconhecimento Ótico de Caracteres

Um sistema de reconhecimento ótico de caracteres (OCR – *Optical Character Recognition*) consiste em reconhecer padrões em imagens que contenham textos e que, após uma série de etapas, possibilita a cópia dos caracteres para um arquivo de texto.

Como cita Eikvil (1993, p. 11), o princípio básico para o reconhecimento automático de caracteres é ensinar à máquina padrões que podem ocorrer e como eles se assemelham. O ensino da máquina é feito através de amostragem de exemplos de textos com as devidas correspondências.

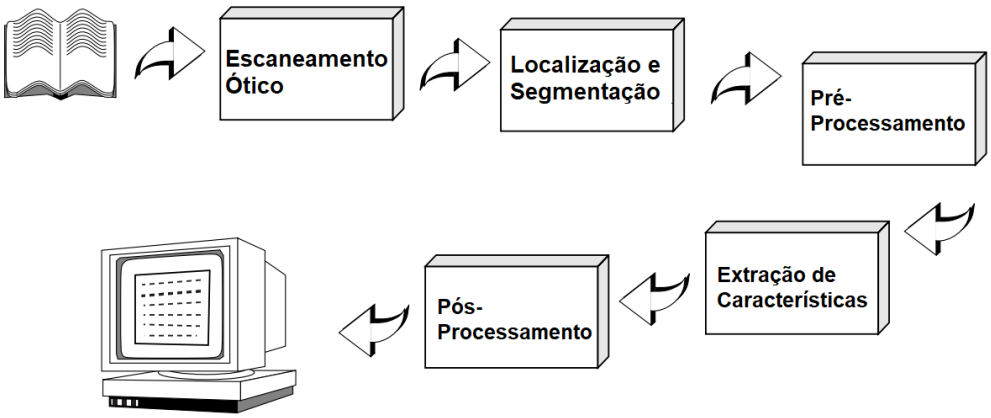
Com base nesses exemplos a máquina constrói um protótipo ou uma descrição de cada classe de caracteres. Então, durante o reconhecimento, os caracteres desconhecidos são comparados com os anteriormente aprendidos, e então se atribui a melhor correspondência (Eikvil, 1993).

2.2.1 Componentes de um sistema OCR

Segundo Eikvil (1993, p. 11), um sistema de OCR é composto pelos seguintes componentes: escaneamento ótico, localização e segmentação, pré-

processamento, extração de características e pós-processamento, como mostra a Figura 5. Cada etapa é melhor descrita nos itens que seguem.

Figura 5: Componentes de um sistema OCR



Fonte: Eikvil (1993)

2.2.2 Escaneamento Ótico

Normalmente documentos impressos estão em preto e branco. Quando estes estão em múltiplos níveis, ou seja, coloridos, o *scanner* executa um processo conhecido como limiar ou *thresholding* que realiza essa conversão para escala de cinza, tendo como objetivo diminuir o esforço computacional e a redução na utilização da memória (EIKVIL, 1993).

É válido salientar que, por conta dos objetivos deste trabalho, esta etapa sofrerá alterações já que é inviável realizar uma conversão para escala de cinza sem antes selecionar o campo marcado pelos marcadores, estes que geralmente são coloridos e afetam outras componentes RGB. Sendo assim, antes da primeira etapa do processo de OCR, será aplicada a segmentação seletiva do texto.

2.2.3 Localização e Segmentação

A segmentação de uma imagem consiste em dividir uma imagem nas áreas significativas, ou seja, nas partes relevantes da mesma. É uma tarefa simples de ser

descrita, porém é das mais difíceis de ser implementada (KHOSHAFIAN & BAKER, 1996).

A maior dificuldade de um algoritmo de segmentação é encontrar um meio de medir a diferença entre os pixels que pertencem ao texto e os pixels do fundo (CHEN, LUETTIN, SHEARER, 2000). O processo procura diferir textos de números e gráficos. Quando a segmentação é aplicada ao texto, ela isola caracteres ou palavras, o grande problema que ocorre no processo é que muitas vezes acontece uma confusão por conta de textos e gráficos que apresentem caracteres muito unidos.

Se o documento foi escaneado com o limiar muito baixo uma perda de caracteres e espaços pode ocorrer. Se o limitar for alto demais, o sistema de OCR também pode falhar durante a segmentação e ignorar espaços, além de unir caracteres aos fundos (MITHE, SUPRIYA, DIVEKAR, 2013).

2.2.4 Pré-processamento

Conforme Leondes (1998), a imagem resultante do processo de análise contém sempre certa quantidade de ruído e o processo de pré-processamento é necessário para reduzir o seu efeito. Ruído pode ser descrito como qualquer coisa que impede um sistema de reconhecimento de padrões cumprirem o seu papel, não importando como este ruído afeta os dados.

O pré-processamento consiste em diversos métodos de filtragem que podem ser aplicados sobre os dados. No reconhecimento de caracteres, a imagem escaneada pode ser filtrada para remover os ruídos que podem dificultar o processo de segmentação. Segundo Khoshafian & Baker (1996), a imagem original terá uma qualidade inferior à resultante desse processo realizado na imagem digitalizada.

Para este trabalho o ruído gerado pelas marcações será utilizado para identificar as áreas a serem reconhecidas, sendo filtrado após prover esta seleção.

2.2.5 Extração de características e pós processamento

Essas etapas não serão abordadas com profundidade neste trabalho, visto que a proposta se baseia nas etapas de seleção e pré-processamento. Todavia a extração de características consiste em analisar geometricamente as formas resultantes das primeiras etapas, dá-se aí a importância de através do pré-processamento ter-se em mãos uma imagem nítida e sem “defeitos”.

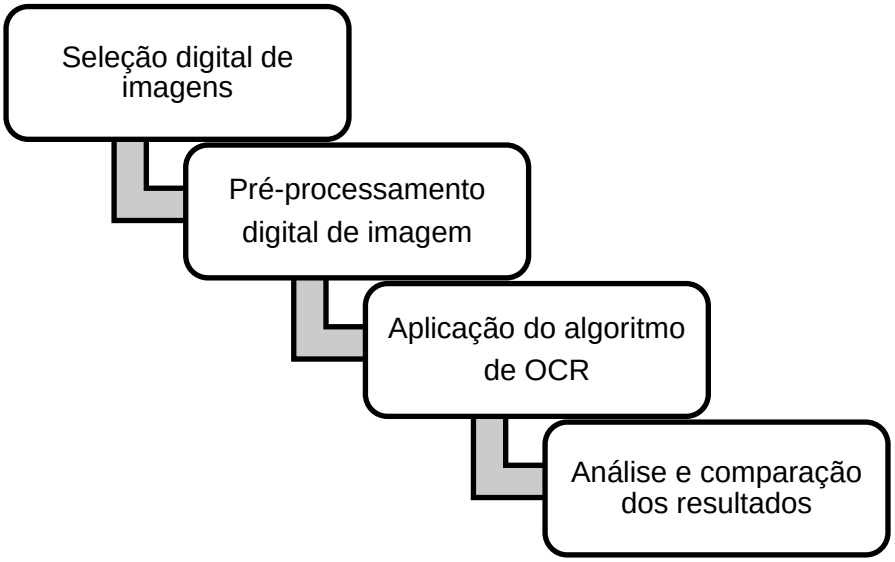
Por fim cabe ao pós-processamento a análise dos dados obtidos e conferência dos resultados, feita geralmente em comparação com uma referência humana para comprovar a eficácia do método aplicado.

3 DESENVOLVIMENTO DE UM ALGORITMO OCR SELETIVO

Dentre as diversas ferramentas cabíveis para realizar o tratamento digital de imagens, o *software* MatLab mostrou ser interessante por possuir uma vasta biblioteca, possibilitando através de seus *toolkits* realizar a implementação do algoritmo tanto para o processamento das imagens quanto para o reconhecimento digital de textos.

A presente proposta espera apresentar uma alternativa viável para o reconhecimento seletivo de caracteres. A Figura 6 apresenta de forma resumida os passos a serem utilizados na metodologia deste trabalho, mais bem descrito nos próximos itens.

Figura 6: Fluxograma da metodologia



Fonte: Autoria própria

3.1 Metodologia

Através da utilização do *software MatLab,* pretende-se desenvolver o algoritmo para seleção e pré-processamento de imagem a fim de obter resultados para imagens de difícil tratamento.

Em seguida faz-se necessário uma interpretação dos componentes de um sistema OCR a fim de classificar as melhores técnicas de processamento de imagens

que possibilitem o reconhecimento com precisão de todos os caracteres presentes nas imagens.

Por fim, realizou-se o desenvolvimento de um método de melhoria dos resultados para, aplicando-se a técnica de detecção de bordas para melhorar a detecção do texto pela biblioteca OCR.

3.1.1 Seleção digital de imagens

A seleção será feita com base na região afetada pelo marcador identificada através da alteração das componentes de cor mostrada no Capítulo 1. A lógica é apresentada no pseudocódigo a seguir:

```
se “pixel” igual componente afetada
“pixel” igual a branco
senão se “pixel” igual a componente de texto afetada
“pixel” igual a preto
senão “pixel” igual a preto
```

O principal problema desta seleção se dá em sua dificuldade em tratar imagens com baixa definição e muito ruído, visto que ao tratar apenas as componentes afetadas, o texto resultante perde suas características, dificultando a identificação por algoritmos de OCR.

Tal problema faz necessária a aplicação de filtros e manipulações que facilitem a identificação dos caracteres, tema que será tratado no próximo item.

3.1.2 Pré-processamento digital de imagem

O pré-processamento consiste em aplicar alterações que facilitem a identificação das imagens, como aplicado neste trabalho a detecção de componentes afetadas por marcado para isolamento do texto e a técnica de detecção de bordas, após a conversão da imagem para tons de cinza, a fim de melhor definir os contornos das formas do texto facilitando assim o seu reconhecimento pela biblioteca OCR.

3.1.3 Aplicação do algoritmo OCR e análise dos resultados

Existe uma vasta biblioteca de algoritmos OCR existente na bibliografia, em suma a maior parte trabalha com a extração de características e comparação via banco de dados para identificar os caracteres.

Neste trabalho a biblioteca utilizada será a biblioteca de OCR do MATLAB, onde é possível através de seu *toolbox* isolar as regiões onde foram identificados caracteres. Com isso é criada uma matriz com as palavras reconhecidas, podendo ser também exportada para um arquivo de texto.

A partir do texto reconhecido é possível verificar a eficácia das etapas anteriores. Uma análise preliminar será apresentada no capítulo que segue, utilizando primeiramente uma imagem com características ideais para este fim e outra mais próxima da realidade, onde será necessário um pré-processamento para o reconhecimento dos textos.

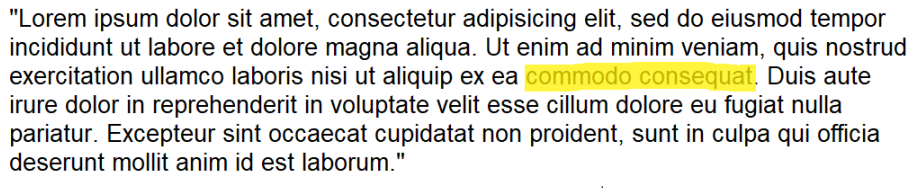
4 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os resultados, adquiridos através da biblioteca disponível no software MatLab.

4.1 Processamento de imagem textual de boa qualidade

Para um primeiro teste foi utilizado um texto criado em um programa de edição de imagens e, para simular o marcador, coloriu-se parte do texto com a cor amarela, característica dos marcadores encontrados no mercado. A imagem resultante foi salva em um formato de imagem comum (PNG - Portable Network Graphics) e pode ser vista na Figura 7.

Figura 7: Imagem textual de boa qualidade



Fonte: Autoria Própria

4.1.1 Seleção digital do texto marcado

Em seguida, utilizando-se da ferramenta "*imtool*" do MatLab, extraiu-se os valores de cores da área afetada pelo marcador, sendo obtido os valores de red = 254, green = 244 e blue = 65 pela ferramenta. Podendo, através de operações com imagens descritas, isolar no item 3.1.1, o texto nesta área como apresentado na Figura 8.

Figura 8: Imagem textual de boa qualidade selecionada

Fonte: Autoria Própria

4.1.2 Reconhecimento digital dos caracteres

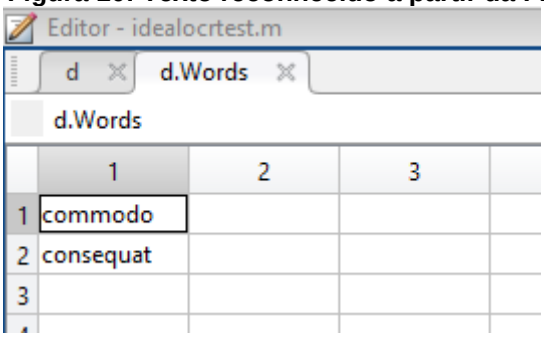
Após isolar o texto, utilizando-se da biblioteca de OCR do MatLab, foram identificados as janelas de possibilidades de texto, mostradas na Figura 9, sendo que a imagem toda é definida como uma possibilidade (linha amarela no contorno da imagem) e cada palavra grifada foi também identificada por uma janela amarela com um número de referência não utilizado no trabalho.

Figura 9: Imagem após a aplicação do algoritmo OCR

Fonte: Autoria Própria

Como mostra a Figura 10, a biblioteca nativa de reconhecimento foi capaz de identificar as duas palavras sem erros. Isso se deve à qualidade das imagens utilizadas no teste, produzidas idealmente para este fim.

Figura 10: Texto reconhecido a partir da Figura 9

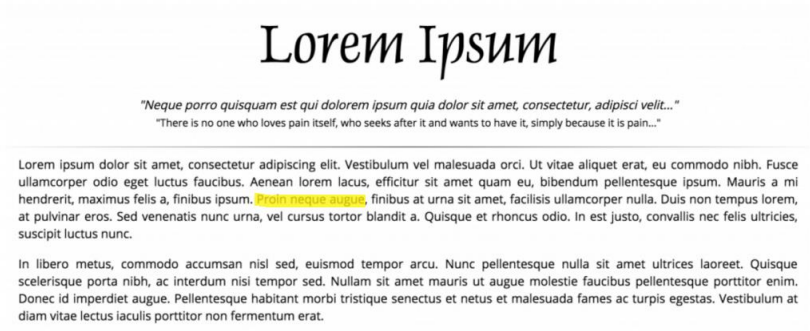


Fonte: Autoria Própria

4.2 Processamento de uma imagem textual genérica

No segundo teste foi utilizado um texto digital não ideal encontrado em um banco de imagens. Da mesma forma que no primeiro coloriu-se parte do texto com a cor amarela, característica dos marcadores encontrados no mercado. A imagem salva pode ser vista na Figura 11.

Figura 11: Imagem textual digital genérica



Fonte: Site "Autoria própria"

4.2.1 Seleção digital do texto marcado

Seguindo os mesmos passos, extraiu-se manualmente os valores de cores da área afetada pelo marcador, utilizando a ferramenta IMTOOL, com retorno dos valores RGB para isolamento red = 254, green = 244 e blue = 65. Assim é possível, isolar o texto nesta área, como visto na Figura 12.

Figura 12: Imagem digital genérica selecionada

Fonte: Autoria Própria

4.2.2 Reconhecimento digital dos caracteres

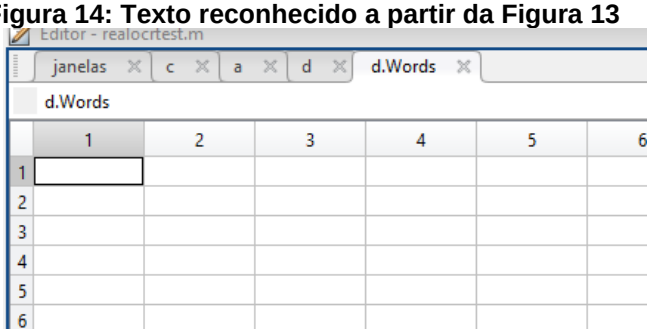
Por se tratar de uma imagem mais próxima da realidade e não uma imagem de boa qualidade do ponto de vista do reconhecimento de texto, a biblioteca OCR não foi capaz de identificar as palavras, como mostrado na Figura 13. Deste modo a imagem toda é definida como uma possibilidade e as palavras grifadas não puderam ser identificadas.

Figura 13: Imagem após a aplicação do algoritmo OCR

Fonte: Autoria Própria

Como a biblioteca nativa de reconhecimento não foi capaz de identificar as palavras o vetor estava vazio, como pode ser visto na Figura 14. Isso se deve à má qualidade da imagem, sendo necessária assim aplicação de filtros, segmentação e outras ferramentas a fim de melhorar a definição do texto recuperado após a aplicação do marcador.

Figura 14: Texto reconhecido a partir da Figura 13



Fonte: Autoria Própria

4.3 Processamento de uma imagem textual genérica a partir da técnica de detecção de bordas para uma imagem ideal

A partir dos testes realizados com um algoritmo simples de OCR, foi implementada então a técnica de detecção de bordas, citada no item 3.1.2, onde regiões candidatas a leitura do texto são identificadas para isolar estas regiões e tornar a leitura mais precisa.

Primeiramente, utilizando a mesma imagem do primeiro teste foi aplicado o isolamento das 3 matrizes da imagem, para que seja realizada a escolha de qual matriz é a melhor para a seleção do texto, foi escolhida então a matriz de isolamento da cor azul, pois ela traz mais destaque ao texto grifado.

A seguir aplicou-se a escolha da cor padrão do texto grifado, utilizando a ferramenta do MATLAB impixel, que isolou a cor padrão red = 254, green = 244 e blue = 65. Esta configuração deve ser realizada manualmente dentro do código e possui possibilidade de melhoria para que seja um passo automatizado no futuro, possibilitando a utilização em massa do código para leitura de diversos textos em simultâneo.

As 3 imagens resultantes e a seleção realizada estão mostradas na Figura 15:

Figura 15: Imagens geradas para as composições RGB e resultante após o isolamento da cor padrão.

Fonte: Autoria Própria

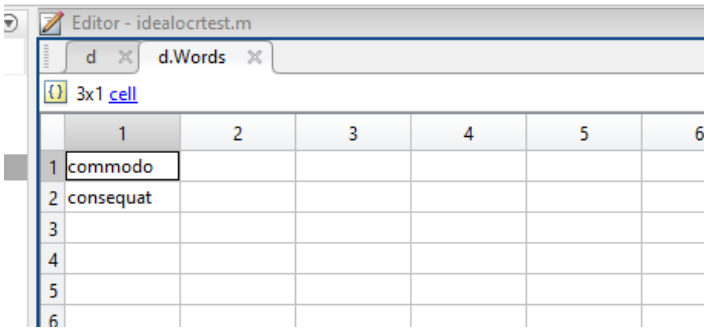
A seguir, aplicou-se o código para a leitura OCR a partir do isolamento de áreas candidatas, que através da leitura de bordas, realiza uma melhoria no isolamento do texto, retornando a imagem mostrada na Figura 16, onde é possível verificar as bordas criadas pela técnica de isolamento de bordas, estas auxiliam a identificação do texto pela biblioteca OCR.

Figura 16: Imagem gerada utilizando remoção de áreas candidatas

Fonte: Autoria Própria

Aplicando-se então esta imagem ao processo de OCR, o retorno se mostrou semelhante ao OCR simples, validando o processo de remoção de áreas candidatas, sendo o texto mostrado na Figura 17.

Figura 17: Texto reconhecido a partir da Figura 16

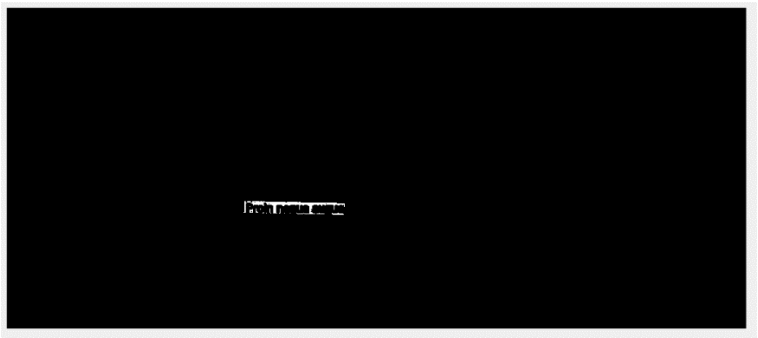


Fonte: Autoria Própria

4.4 Processamento de uma imagem textual genérica a partir da técnica de detecção de bordas para uma imagem real

No segundo teste com detecção de bordas foi utilizado um texto digital não ideal encontrado em um banco de imagens, o mesmo aplicado no primeiro teste real. Todavia, a imagem resultante no passo de seleção do texto causou muitos ruídos ao texto original, como é possível verificar na Figura 18. A ferramenta IMTOOL retornou para a imagem a seleção das cores red = 253, green = 230 e blue = 0;

Figura 18: Imagem gerada após a exclusão das cores do marca texto, utilizando IMTOOL



Fonte: Autoria Própria

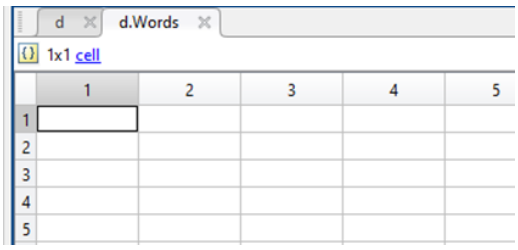
Assim, mesmo após a aplicação da técnica de detecção de bordas, o resultado com bordas, não possibilita uma leitura perfeita da fonte o que é possível verificar na Figura 19.

Figura 19: Imagem gerada após a aplicação da técnica de detecção de bordas

Fonte: Autoria Própria

O que ainda assim, aplicado a biblioteca OCR, não gera retorno de textos na matriz de retorno. Como mostra a Figura 20.

Figura 20: Texto reconhecido a partir da Figura 19



Fonte: Autoria Própria

5 CONSIDERAÇÕES FINAIS

O trabalho, para o caso de uma imagem ideal, ou seja, gerada e marcada digitalmente, obteve sucesso em identificar e isolar digitalmente a presença de marcadores, obtendo um arquivo de texto seletivo através de um algoritmo de OCR em conjunto com um algoritmo de reconhecimento de diferença de cores RGB.

Para o segundo caso, onde uma imagem marcada foi obtida de um banco de dados, o algoritmo obteve sucesso em isolar a área marcada, todavia, os dados do texto foram comprometidos com o método de isolamento utilizado, dificultando a leitura pela biblioteca OCR. Desta maneira, entende-se que melhorias podem ser implementadas futuramente de modo a possibilitar o uso do algoritmo em imagens reais, melhorando o limiar de decisão para a seleção, evitando o comprometimento do texto selecionado.

A automatização do sistema também pode ser implementada futuramente, de modo a retornar uma maior quantidade de resultados em um menor tempo pode ser uma melhoria para trabalhos futuros, visto que, este trabalho visou a implementação em imagens únicas.

REFERÊNCIAS

BARBOZA, Ricardo da Silva. **Filtragem de Ruídos em Imagens Digitais.** 2013. Disponível em: https://repositorio.ufpe.br/bitstream/123456789/29325/1/TESE%20Ricardo%20da%20Silva%20Barboza.pdf. Acesso em: 14 abr. 2019.

EIKVIL, Line. **OCR Optical Character Recognition.** 1993. Disponível em: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.25.3684&rep=rep1&type=pdf. Acesso em: 27 abr. 2019.

CAMPESTRI, Marcio Rodrigo. **Protótipo de um sistema de reconhecimento de caracteres baseado em redes neurais.** 2000. Disponível em: http://dsc.inf.furb.br/arquivos/tccs/monografias/20001marciorodrigocampestrinivf.pdf. Acesso em: 25 abr. 2019.

CHEN, D.; LUETTIN, J.; SHEARER, K.; 2000; **A Survey of Texto Detection and Recognition**. Disponível em: http://www.cs.cmu.edu/~datong/survey.pdf. Acessado em: 22 mai. 2019.

FALCÃO, Alexandre Xavier; MENOTTI, David. **Segmentação por Limiarização.** Disponível em: <http://www.ic.unicamp.br/~afalcao/mo443/slides-aula25a.pdf>. Acesso em: 17 abr. 2019.

JAYARAMAN, S; ESAKKIRAJAN, S; VEEWAKUMAR T. **Digital Image Processing. New Delhi**: Tata McGraw Hill Education, 2009. 723p.

KHOSHAFIAN, S.; BAKER, A. B. **Multimedia and Imaging Databases**. Morgan Kaufmann, 1996.

LEONDES, C. T. **Image Processing and Pattern Recognition**, 1998.

MITHE, R.; SUPRIYA, I; DIVEKAR, N. **Optical Character Recognition**: International Journal of Recent Technology and Engineering (IJRTE), 2013.

NOBRE, Marcello. **O USO DO SOFTWARE MATLAB PARA O ESTUDO DE ALGUNS TÓPICOS DE ÁLGEBRA LINEAR.** Disponível em: <http://www.ucb.br/sites/100/103/TCC/22005/MarcelloNobreCardoso.pdf>. Acesso em: 20 maio 2019.

OLIVEIRA NETTO, A. A. de. **Metodologia da pesquisa científica:** guia prático para a apresentação de trabalhos acadêmicos. 3. ed. rev. e atual. Florianópolis: Visual Books, 2008.

PETROU E PETROU, M.; PETROU E PETROU C. **Image processing: the fundamentals**. 2nd ed. Wiley; 2010.

APÊNDICES

Código para o ocr sem detecção de bordas:

```
clc; clear all;

a = imread('idealtest.png');
imshow(a)

c = imread('idealtestcut.png');
d = ocr(c);

janelas = insertObjectAnnotation(c, 'rectangle',...
                                 d.WordBoundingBoxes,...
                                 d.WordConfidences);
figure
imshow(janelas);
```

Código para o teste com detecção de bordas:

```
clc; clear all; close all;

a = imread('idealtest.png');
imshow(a)

red = a(:,:,1); green = a(:,:,2); blue = a(:,:,3);
imshow(red)
imshow(green)
imshow(blue)
% d = impixel(a);

out = red>254 & green>244 & blue<65 & blue>62;
figure(81)
imshow(out)
figure(82)
subplot(2,2,1), imshow(red), subplot(2,2,2), imshow(green), subplot(2,2,3),
imshow(blue), subplot(2,2,4), imshow(out)

test = imread('idealtestcut.png');
figure(999)
subplot(2,3,1), imshow(test);

b = rgb2gray(test);
subplot(2,3,2), imshow(b);

figure
e = edge(b,'canny');
imshow(e)
matriz texto = ocr(e);

janelas = insertObjectAnnotation(c, 'rectangle',...
                                 d.WordBoundingBoxes,...
                                 d.WordConfidences);
figure
imshow(janelas);
```