

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**CRISTOFER VIANTE HASS  
LEONALDO FERREIRA NUNES JUNIOR**

**USO DE VISÃO COMPUTACIONAL PARA OTIMIZAÇÃO DO TEMPO**

**PONTA GROSSA  
2023**

**CRISTOFER VIANTE HASS  
LEONALDO FERREIRA NUNES JUNIOR**

## **USO DE VISÃO COMPUTACIONAL PARA OTIMIZAÇÃO DO TEMPO**

### **The use of computer vision for time optimization**

Trabalho de conclusão de curso apresentada como requisito para obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador(a): Rogério Ranthum.

**PONTA GROSSA**

**2023**



[4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**CRISTOFER VIANTE HASS  
LEONALDO FERREIRA NUNES JUNIOR**

**USO DE VISÃO COMPUTACIONAL PARA OTIMIZAÇÃO DO TEMPO**

Trabalho de Conclusão de Curso apresentado como requisito para obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas da Universidade Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 29/junho/2023

---

Rogério Ranthum  
Mestrado  
Universidade Tecnológica Federal do Paraná

---

Geraldo Ranthum  
Mestrado  
Universidade Tecnológica Federal do Paraná

---

Augusto Foronda  
Doutorado  
Universidade Tecnológica Federal do Paraná

**PONTA GROSSA  
2023**

## **AGRADECIMENTOS**

Agradecemos primeiramente ao nosso orientador Rogério Ranthum que se mostrou sempre disponível em nos auxiliar e também sempre compreensível sobre as dificuldades da vida profissional.

Cristofer: “Gostaria de agradecer aos meus pais, em especial minha mãe que acompanhou minha trajetória acadêmica mais de perto, e a todos os amigos que fiz nessa faculdade, que tornaram essa graduação uma lembrança que sempre irei levar”.

Leonardo: “Gostaria de agradecer a Deus por ter me dado força em meus maiores desafios, sou grato minha mãe pela educação e todo o sacrifício feito em prol do meu progresso, aos professores por seus ensinamentos e paciência, e a todos meus amigos que fiz durante a faculdade pelo companheirismo.”

Por fim, queremos agradecer a todos os outros profissionais envolvidos na UTFPR sendo todos verdadeiramente comprometidos com a qualidade do curso e possibilitando dessa forma a nossa trajetória.

## RESUMO

O presente trabalho tem como tema o uso de visão computacional para otimização do tempo em operações de cadastro possuindo o objetivo de explorar o uso da visão computacional, mais especificamente a leitura óptica de caracteres, como uma solução para reduzir o tempo necessário em operações de cadastro. A proposta visa melhorar a experiência de pessoas que frequentemente enfrentam longas filas durante essas operações. Para alcançar o objetivo proposto, foram utilizadas uma interface web e uma API em Python, juntamente com as ferramentas OpenCV e *pytesseract*. Foram realizadas comparações entre as operações manuais e as operações automáticas, com o tempo gasto em cada uma sendo cronometrado. O intuito foi verificar se a utilização da leitura óptica de caracteres resulta em uma redução significativa do tempo necessário para a realização das operações. Com base nos experimentos realizados, constatou-se que a utilização da visão computacional para leitura óptica de caracteres pode resultar em um ganho de tempo nas operações de cadastro. No entanto, observou-se que a aplicação desse método em ambientes caóticos pode ser desafiadora. Dessa forma, apesar dos benefícios identificados, é necessário considerar cuidadosamente o contexto em que a solução será implementada, levando em conta fatores como a qualidade das imagens capturadas e a estabilidade do ambiente operacional.

Palavras Chave: Visão Computacional; Leitura Óptica de Caracteres; Tempo.

## ABSTRACT

The present study focuses on the use of computer vision for time optimization in registration operations, aiming to explore the use of computer vision, specifically optical character recognition (OCR), as a solution to reduce the time required for registration operations. The proposal aims to enhance the experience of individuals who frequently face long queues during these operations. To achieve the stated objective, a web interface and a Python API, along with OpenCV and *pytesseract* tools, were utilized. A comparison between manual and automatic operations was conducted, with the time spent on each method being measured. The goal was to determine whether the use of OCR leads to a significant reduction in the time required to perform these operations. Based on the conducted experiments, it was observed that the utilization of computer vision for optical character recognition can indeed save time in registration operations. However, it was noted that the application of this method in chaotic environments can be challenging. Therefore, despite the identified benefits, careful consideration of the implementation context is necessary, taking into account factors such as image quality and the stability of the operational environment.

Keywords: computer vision; Optical character reading; Time.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Documento fictício.....	27
Figura 2 - Tela de Login.....	29
Figura 3 - Tabela de Pessoas.....	29
Figura 4 - Tela para inserir pessoa.....	30
Figura 5 - Tela para selecionar documento.....	30
Figura 6 - Tela para ajustar documento.....	31
Figura 7 - Tela para conferir dados obtidos.....	31
Figura 8 - Operações Iniciais.....	32
Figura 9 - Pré Processamento.....	33
Figura 10 - Obtenção de Pontos.....	33
Figura 11 - Pós Processamento.....	34
Figura 12 - Leitura de Caracteres.....	34
Figura 13 - Retorno ao Cliente.....	34
Figura 14 - Documento limiarizado.....	36
Figura 15 - Documento cortado.....	37
Figura 16 - Documento com suavização de bordas.....	37
Figura 17 - Documento com remoção de ruído.....	38
Figura 18 - Documento com marcações.....	38

## LISTA DE TABELAS

<b>Tabela 1 - Tempos obtidos .....</b>	<b>39</b>
--	-----------

## LISTA DE ABREVIATURAS E SIGLAS

OCR      *Optical Character Reading*

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	10
<b>1.1 Objetivos</b>	12
1.1.1 Objetivo Geral	12
1.1.2 Objetivos Específicos	12
<b>1.2 Justificativa</b>	12
<b>1.3 Organização do trabalho</b>	12
<b>2 REFERENCIAL TEÓRICO</b>	14
<b>2.1 Tráfego de documento eletrônico</b>	15
<b>2.2 Pixel</b>	16
<b>2.3 Visão computacional</b>	17
<b>2.4 Trabalhos relacionados</b>	23
<b>3 DESENVOLVIMENTO</b>	27
<b>4 RESULTADOS OBTIDOS</b>	36
<b>5 CONSIDERAÇÕES FINAIS</b>	40
<b>REFERÊNCIAS</b>	43

## 1 INTRODUÇÃO

Na era atual, caracterizada por um ritmo acelerado, operações de cadastro frequentemente consomem uma quantidade significativa de tempo, resultando em longas filas e experiências frustrantes para os indivíduos envolvidos. Esse problema é particularmente evidente em ambientes de escritórios de Recursos Humanos, onde a admissão e o cadastro de novos colaboradores ocorrem regularmente e ainda são muitas vezes conduzidos manualmente.

Diante desse cenário, surgiu a ideia do presente estudo, que tem como objetivo explorar o uso da visão computacional, focando especificamente na leitura óptica de caracteres (OCR), como uma potencial solução para otimizar esses processos de cadastro. A motivação para esta pesquisa emerge da necessidade identificada de melhorar a experiência dos usuários que, frequentemente, se encontram diante de longas esperas durante tais operações.

Um estudo do IPEA (Instituto de Pesquisa Econômica Aplicada) realizado em 2011, aponta que a segunda maior preocupação dos usuários do SUS (Sistema Único de Saúde) é a demora no atendimento nas unidades de saúde, apontada por 35,9% dos entrevistados (FILHO E SANT'ANA, 2016).

A apreensão do brasileiro com a demora no atendimento do SUS é um dos principais vetores que impulsionam a busca pela saúde suplementar, os conhecidos planos de saúde. De acordo com a mesma pesquisa do IPEA, os usuários do plano de saúde apontam a busca por maior rapidez na realização de consultas e exames como principal razão de sua adesão à saúde suplementar. Essa foi a resposta de 40% dos entrevistados (FILHO E SANT'ANA, 2016, p.73)

Percebe-se portanto, uma inquietude por parte da pessoa que se propõe a pagar por um plano de saúde, quanto ao tempo de espera em filas no sistema de saúde, compreende-se aqui uma preocupação quanto a necessidade de acelerar processos. Existem ainda casos em que a demora na fila para o atendimento é levada até a justiça como pedido de indenização por danos morais ao cliente (CONJUR, 2017).

O dano moral é garantido por lei, Art. 14 do Código de Defesa do Consumidor - Lei 8078/90, onde está descrito: "O fornecedor de serviços responde, independentemente da existência de culpa, pela reparação dos danos causados aos

consumidores por defeitos relativos à prestação dos serviços, bem como por informações insuficientes ou inadequadas sobre sua fruição e riscos”. (BRASIL, 1990).

O tempo de resposta de um serviço é visto como uma medida para avaliar a responsividade e por conseguinte a qualidade de um serviço. (SORATTO, 2004). No contexto do estudo, a adoção da visão computacional, especificamente a leitura óptica de caracteres (OCR), é proposta como uma maneira de diminuir o tempo de resposta, tornando o processo de cadastro mais rápido e eficiente.

Não se pretende deste modo, aumentar a carga de trabalho do funcionário responsável pelo atendimento, mas sim, “eliminar do processo as atividades sem sentido, e remover os obstáculos ao bom desempenho” (SORATTO DA SILVA, 2004, p.31).

Para Harrington (1993) o tempo de execução de atividades é um fator importante na agilização de processos, e para que esse tempo seja reduzido, o autor apresenta a automatização, que faz com que “tarefas repetitivas e tediosas sejam feitas por ferramentas, equipamentos e computadores, deixando o empregado livre para tarefas mais criativas”.

Interessante ressaltar que o cadastro inicial, sob a perspectiva de um processo completo de atendimento no SUS, por exemplo, corresponde a uma pequena parte, porém ao automatizar-se o mesmo, diminui-se o tempo completo de espera, corroborando assim com Harrington (1993).

Para automatizar esse tempo, se faz interessante a utilização da visão computacional, que é um campo da área do conhecimento em crescimento exponencial, que tem por objetivo “determinar o que está acontecendo na frente de uma câmera e usar esse entendimento para controlar um computador ou sistema robótico” (PULLI et al, 2012, p.61).

A visão computacional agregada às técnicas de processamento de imagem, possibilita a análise e a manipulação da mesma. Dessa forma sugere-se o uso da leitura óptica de caracteres para analisar cinco documentos fictícios, processando assim a imagem destes documentos abstraindo seus dados para um cadastramento, e então, inserir de forma automática tais dados em um formulário de cadastro.

Pelo supracitado, a pergunta problema da presente pesquisa se caracteriza em como melhorar as operações de cadastro, utilizando o reconhecimento ótico de caracteres, para aquisição de documentos de identificação

## **1.1 Objetivos**

### **1.1.1 Objetivo Geral**

Utilizar o Reconhecimento Ótico de Caracteres (OCR), para aquisição de documentos de identificação.

### **1.1.2 Objetivos Específicos**

- Estudar bibliotecas de processamento de imagens;
- Processar e tratar uma imagem do documento, aplicando funções da biblioteca OpenCV;
- Estudar biblioteca de Reconhecimento Ótico de Caracteres.
- Reconhecer os caracteres dentro da imagem fazendo uso da biblioteca *pyTesseract*;
- Criar documentos de identificação fictícios para testar a aplicação;
- Testar a solução criada comparando com o cadastro por meio da digitação;

## **1.2 Justificativa**

Tempos de espera excessivos causam, tanto para quem aguarda, quanto para quem realiza o atendimento um acúmulo de estresse, prejudicando também a empresa que realiza o atendimento, uma tecnologia que automatize o tempo de atendimento consequentemente gera uma maior satisfação dos clientes e menor estresse para os funcionários. (SORATTO DA SILVA, 2004).

O projeto se justifica por implementar uma melhoraria na qualidade de vida das pessoas, tendo em vista que todos deparam-se com filas, e tempos de espera causados por processos, como um simples cadastro, ao automatizar esta parte do processo, espera-se contribuir para que o processo como um todo seja menos demorado.

## **1.3 Organização do trabalho**

Este trabalho está organizado em 5 capítulos. O capítulo 2 dispõe sobre a legislação que regulamenta o tráfego de documentos eletrônicos e todos os referenciais necessários para o entendimento do trabalho realizado, bem como alguns

casos aos quais este trabalho se relaciona. O capítulo 3 detalha o desenvolvimento do trabalho como as tecnologias utilizadas e descrição de métodos de código

O Capítulo 4 expõe os resultados obtidos com o sistema de cadastro automatizado através de OCR. Por fim o Capítulo 5 trata-se das conclusões e contribuições deste trabalho, além de sugestões para trabalhos futuros.

## 2 REFERENCIAL TEÓRICO

A evolução da tecnologia digital vem moldando diversos aspectos da vida moderna e impactando inúmeras áreas do conhecimento. Neste contexto, o presente trabalho explora três temas centrais: tráfego de documento eletrônico, *pixel* e visão computacional.

O tráfego de documento eletrônico diz respeito à tramitação digital de processos judiciais e à transmissão de peças processuais, conforme estabelecido pela Lei nº 11.419. Inclui-se aqui o armazenamento e transmissão de documentos e arquivos digitais, além da autenticação de identidades por meio de assinaturas digitais.

A seguir, aborda-se o conceito de *pixel*, a menor unidade de uma imagem com luminescência e cor específicas. Entender os *pixels* é vital para compreender como as imagens são formadas e exibidas em dispositivos digitais. Cada *pixel* é um elemento de uma matriz bidimensional que, em conjunto, forma a imagem que se vê. O valor numérico de cada *pixel* contém informações essenciais sobre a imagem.

Por fim, mas não menos relevante, o trabalho explora a visão computacional. Este campo de estudo busca replicar a visão humana através do uso de algoritmos de computador para interpretar e entender imagens e vídeos. A visão computacional engloba várias subdisciplinas, como técnicas de segmentação, limiarização, redução de ruído, transformação e remoção de bordas. Uma das aplicações mais notáveis da visão computacional é o Reconhecimento Óptico de Caracteres (OCR), que permite que computadores leiam e reconheçam texto em imagens.

Assim, este trabalho oferece uma compreensão desses três tópicos, todos interligados e fundamentais para o funcionamento de muitas tecnologias digitais modernas.

Agora será apresentado, mais detalhadamente nas disposições da Lei nº 11.419, que regula o tráfego de documento eletrônico. Essa lei estabelece o uso de meio eletrônico na tramitação de processos judiciais, comunicação de atos e transmissão de peças processuais, aplicando-se a diversos tipos de processos e juizados.

## 2.1 Tráfego de documento eletrônico

Quanto ao tráfego de documento eletrônico, e operações de informatização de qualquer processo judicial, a Lei nº 11.419 deixa claro que: “Art. 1º O uso de meio eletrônico na tramitação de processos judiciais, comunicação de atos e transmissão de peças processuais será admitido nos termos desta Lei”. (BRASIL, 2006).

§ 1º Aplica-se o disposto nesta Lei, indistintamente, aos processos civil, penal e trabalhista, bem como aos juizados especiais, em qualquer grau de jurisdição.

§ 2º Para o disposto nesta Lei, considera-se:

I - meio eletrônico qualquer forma de armazenamento ou tráfego de documentos e arquivos digitais;

II - transmissão eletrônica toda forma de comunicação a distância com a utilização de redes de comunicação, preferencialmente a rede mundial de computadores;

III - assinatura eletrônica as seguintes formas de identificação inequívoca do signatário:

a) assinatura digital baseada em certificado digital emitido por Autoridade Certificadora credenciada, na forma de lei específica;

b) mediante cadastro de usuário no Poder Judiciário, conforme disciplinado pelos órgãos respectivos.(BRASIL, 2006).

O trecho citado da Lei nº 11.419 apresenta uma importante disposição em relação ao tráfego de documento eletrônico e às operações de informatização de processos judiciais.

O artigo 1º deixa claro que o uso de meios eletrônicos na tramitação de processos judiciais, comunicação de atos e transmissão de peças processuais é admitido nos termos estabelecidos pela própria lei. É relevante observar que essa lei é aplicável de forma indistinta aos processos civil, penal e trabalhista, assim como aos juizados especiais, em qualquer grau de jurisdição. Isso demonstra a abrangência e a universalidade da utilização de meios eletrônicos no âmbito jurídico.

No que diz respeito às definições, o parágrafo 2º do texto citado estabelece alguns conceitos fundamentais. Ele define o meio eletrônico como qualquer forma de armazenamento ou tráfego de documentos e arquivos digitais. Além disso, a transmissão eletrônica é descrita como qualquer forma de comunicação a distância por meio de redes de comunicação, sendo a rede mundial de computadores preferencial.

Outro ponto relevante é a definição de assinatura eletrônica, que é considerada como uma forma de identificação inequívoca do signatário. Ela pode ser realizada através de uma assinatura digital baseada em certificado digital emitido por Autoridade Certificadora credenciada, conforme lei específica. Também é aceita a identificação

do signatário mediante cadastro de usuário no Poder Judiciário, conforme regulamentação dos órgãos competentes.

Essas definições estabelecidas pela Lei nº 11.419 são essenciais para garantir a autenticidade, a segurança e a validade dos documentos e processos judiciais realizados por meios eletrônicos. Ao reconhecer a importância dos meios digitais na tramitação e na comunicação processual, a legislação proporciona uma base sólida para a modernização e a eficiência do sistema judiciário.

No próximo tópico, será abordado o conceito de *pixel*. Os *pixels* são as unidades fundamentais que compõem uma imagem digital, sendo células retangulares com cores específicas e níveis de luminosidade. Cada *pixel* representa um elemento de uma matriz bidimensional que, quando combinados, formam a imagem visível em dispositivos digitais, como telas de computadores, *smartphones* e televisões. A compreensão do papel dos *pixels* é essencial para uma melhor compreensão da criação e exibição de imagens, bem como para explorar o potencial das tecnologias visuais modernas.

## 2.2 Pixel

É a menor unidade de uma imagem com luminescência e cor específicas. Sua proporção é determinada pelo número de linhas que compõem a varredura (padrão de ponto que compõe a imagem) e também pela resolução ao longo de cada linha, de forma geral os milhares *pixels* que compõem uma imagem individual, são projetados em uma tela como pontos iluminados, dessa forma, a distância parecem uma única imagem individual. Para que ocorra dessa forma um feixe de elétrons cria uma grade de *pixels* rastreando cada linha da esquerda para direita, um *pixel* por vez, da linha superior até a inferior, renderizando assim a imagem em um plano 2D. Um *pixel* também pode ser o menor elemento de um dispositivo sensível a luz como por exemplo câmeras. (ENCYCLOPEDIA BRITANNICA, 2017).

O termo *pixel* é resultado da junção das palavras inglesas "*picture element*" ou elemento de imagem. Trata-se de um elemento indexado em uma posição (x, y) ou coluna-linha (c, r) a partir da origem da imagem, um *pixel* representa o menor elemento constituinte de uma imagem digital e contém um valor numérico o qual corresponde a uma unidade básica de informação sobre a imagem. Portanto uma imagem é simplesmente um sinal 2D digitalizado em uma malha ou matriz de *pixels*. O conteúdo

de informação contido em *pixel* pode variar consideravelmente, dependendo do tipo da imagem em processamento (SOLOMON; BRECKON, 2013, p. 44).

A seguir, será abordado o campo da visão computacional. Esse campo de estudo busca reproduzir a capacidade de visão humana por meio da aplicação de algoritmos e técnicas computacionais para interpretar e compreender imagens e vídeos. O objetivo da visão computacional é extrair informações relevantes das imagens, como reconhecimento de objetos, detecção de padrões e análise de conteúdo visual. Essa área de pesquisa desempenha um papel fundamental em diversas aplicações, incluindo sistemas de vigilância, robótica, realidade aumentada e automação industrial. Será explorado como as técnicas avançadas da visão computacional têm impactado diversos setores e possibilitado o desenvolvimento de soluções baseadas em análise visual.

### **2.3 Visão computacional**

A visão computacional procura emular a visão humana, portanto também possui como entrada uma imagem, porém, a saída é uma interpretação da imagem como um todo, ou parcialmente. O espectro que vai do processamento de imagens até a visão computacional pode ser dividido em três níveis: baixo-nível, nível-médio e alto-nível. (MARENGONI, 2009).

- baixo-nível envolvem operações primitivas, tais como a redução de ruído ou melhoria no contraste de uma imagem.
- Os processos de nível-médio são operações do tipo segmentação (particionamento da imagem em regiões) ou classificação (reconhecimento dos objetos na imagem).
- Os processos de alto-nível estão relacionados com as tarefas de cognição associadas com a visão humana.

Com base nessas informações pode-se definir a visão computacional como um campo em rápido crescimento que se dedica a análise, modificação e entendimento de alto nível das imagens. Possui o objetivo de determinar o que está acontecendo na frente de uma câmera e usar esse entendimento para controlar um sistema robótico ou então um computador, ou para fornecer para as pessoas novas imagens que sejam mais informativas ou esteticamente melhores que a original. As áreas de aplicação para visão computacional incluem videovigilância, biometria, automotiva, fotografia,

produção de filmes, pesquisa na web, medicina, jogos de realidade aumentada, novas interfaces de usuário e entre outras. (PULLI et all, 2012).

Por fim SHAPIRO (2001) complementa que o objetivo da visão computacional é tomar decisões úteis a partir de imagens coletadas de uma câmera.

O campo de processamento de imagens é abordado, onde os princípios e técnicas fundamentais dessa área da visão computacional são explorados. Essa área engloba desde operações básicas de redução de ruído e melhoria de contraste até a segmentação e o reconhecimento de objetos. O processamento de imagens desempenha um papel crucial na interpretação visual e na obtenção de informações úteis.

As operações de processamento podem ser agrupadas em divergentes categorias, que depende da estrutura, nível ou finalidade. Existem operações que se destinam a melhorar a qualidade da imagem para o consumo humano, enquanto outras técnicas tem a finalidade de extrair informações para um processamento automático posterior, como por exemplo uma câmera que captura placas de veículos para comparar com uma biblioteca de placas posteriormente. Algumas operações geram novas imagens de saída, enquanto outras produzem uma descrição da imagem analisada (SHAPIRO, 2001).

As imagens se encontram no centro da visão, e existem muitas maneiras, indo do simples até o sofisticado, para processá-las e analisá-las, conforme destacado por Davies (2012, p.17).

O tema da segmentação no processamento de imagens é uma técnica fundamental para a divisão de uma imagem em objetos ou regiões distintas, levando-se em consideração características como proximidade e diferenças de cores. Serão discutidos os métodos de segmentação utilizados, bem como sua aplicação em diferentes cenários, abrangendo desde a identificação de veículos em imagens de garagens até a análise de áreas urbanas movimentadas em imagens aéreas.

O método de segmentação busca dividir determinada imagem em diferentes objetos ou regiões, isso de acordo com a proximidade e as diferenças de cores presentes na mesma. A resolução original da imagem e o objeto que se deseja identificar são os fatores que determinam o nível de detalhamento desse processo, por exemplo, ao efetuar a busca da região preenchida por um veículo em uma imagem fotografada dentro de uma garagem simples, ela iria ocupar uma grande porcentagem da imagem, por outro lado, se a busca da região preenchida por veículos for feita em

uma imagem aérea de uma avenida movimentada, por mais que seja a mesma tarefa, a resolução da imagem e possivelmente o processo de segmentação utilizado serão diferentes. (MARENGONI, 2009, p. 138).

Através da limiarização, é possível realizar a conversão de uma imagem em uma representação binária, onde a imagem é dividida em regiões com base em um valor de limiar. Serão exploradas as técnicas e a aplicação da limiarização na segmentação eficiente de imagens.

A limiarização é uma técnica pertencente ao grupo de segmentação de imagens. Por meio da limiarização, é possível realizar a binarização de uma imagem. Essa técnica envolve a escolha de um valor de limiar, no qual os pixels da imagem são divididos em duas regiões distintas. *Pixels* com valores acima do limiar são alocados em uma região, enquanto os *pixels* com valores abaixo do limiar são alocados em outra região. O resultado é uma imagem binária, em que os *pixels* são representados por apenas dois valores possíveis. (SOLOMON; BRECKON, 2013, p. 253).

A seleção do valor de Limiar é, essencialmente, baseada levando em consideração o histograma da imagem. Nas situações em que a aplicação de limiar é capaz de separar os objetos de um fundo, o histograma apresenta dois modos, onde um corresponde aos *pixels* do objeto e o outro aos *pixels* do fundo. O valor de limiar deve ser escolhido de forma em que esses dois modos sejam facilmente distinguíveis. (SOLOMON; BRECKON, 2013, p. 254).

Outra técnica frequentemente utilizada no processamento de imagens é a redução de ruído, a qual desempenha um papel essencial na eliminação de variações aleatórias de brilho e cor que podem comprometer a qualidade e a legibilidade das imagens. Nesse contexto, serão explorados métodos específicos empregados para a identificação e correção de *pixels* com valores discrepantes em relação à vizinhança. Essas abordagens têm como objetivo aprimorar a precisão da extração de informações, como texto, presente nas imagens.

O ruído em uma imagem é a variação aleatória de brilho e da cor que dificulta a leitura do texto. Alguns tipos de ruído não são removidos pelo OCR na fase de binarização, prejudicando a precisão da extração do texto na imagem. (TESSDOC, 2020).

O objetivo da redução é identificar em uma imagem, *pixels* que possuem um valor muito diferente da vizinhança em que se encontra, podemos definir a vizinhança

como o conjunto de *pixels* que rodeiam o *pixel* de valor anormal, após a identificação do *pixel* destoante uma média é extraída dessa vizinhança para assim substituir o valor do *pixel* por um valor mais adequado, dessa forma removendo picos de extrema intensidade que são improváveis na imagem. (DAVIES, 2012).

Além da redução de ruído existe a transformação, consistem em operações que permitem modificar as propriedades geométricas e espaciais de uma imagem digital. Elas englobam a translação, rotação e alteração de escala, que são capazes de alterar as coordenadas e dimensões dos objetos presentes na imagem.

De acordo com Solomon (2013, p. 152) a translação de um objeto da imagem de um lugar para outro, a rotação deste objeto e a alteração da escala das dimensões do mesmo, são operações que mudam as coordenadas do vetor de forma de um objeto contido na imagem, mas sua forma em essência não é alterada. Ele complementa que a forma de um objeto é definido basicamente pelo seu contorno, o qual deve ser invariante submetido a translação, rotação e mudança de escala das coordenadas que definem o contorno. Sendo assim definimos “forma” como toda informação geométrica que permanece em uma imagem após filtradas os efeitos de localização, mudança de escala e rotação.

E por fim, será explorada a detecção de bordas no processamento de imagens. Esse importante aspecto envolve a identificação, de forma automatizada, das transições bruscas de intensidade ou textura presentes em uma imagem. Tal técnica possibilita a segmentação da imagem e a análise das suas características visuais, desempenhando um papel fundamental na extração de informações relevantes.

Uma borda pode ser considerada uma descontinuidade na imagem, portanto para a detecção de bordas são utilizados filtros que detectem descontinuidades na imagem, chamados de filtros diferenciais, que são filtros feitos para responder, ou seja, retornar valores significativos em pontos onde há descontinuidade na imagem e não respondem em regiões onde predomina uma suavização, portanto, detectam bordas. (SOLOMON; BRECKON, 2013, p. 106).

A detecção de bordas é basicamente segmentar a imagem em regiões usando como base as suas descontinuidades, em outras palavras, permite o usuário analisar características de uma imagem onde há uma mudança brusca nos tons de cinza ou na textura, indicando assim o fim de uma região na imagem, e o início de outra. (SOLOMON; BRECKON, 2013, p. 259).

A proposta de aplicar Reconhecimento Óptico de Caracteres (OCR), para otimizar processos de cadastro pode ser aprimorada e complementada com várias técnicas de processamento de imagens. Antes da aplicação do OCR, pode-se implementar a limiarização na imagem, uma técnica que a converte em uma representação binária, facilitando a identificação dos caracteres. Esta abordagem ajuda a maximizar o contraste entre o texto e o fundo, facilitando a leitura por parte do sistema de OCR. Outra técnica útil é a redução de ruído.

Esta pode ser aplicada para eliminar variações aleatórias de brilho e cor que possam interferir na eficiência do OCR. Ao suavizar a imagem e remover ruídos, é possível aumentar a precisão do OCR, resultando em transcrições mais precisas e, conseqüentemente, em um processo de cadastro mais eficiente. As operações de transformação, por sua vez, podem ser usadas para corrigir problemas de orientação e escala na imagem.

Por exemplo, se a imagem do documento estiver inclinada, a rotação pode ser aplicada para alinhar o documento corretamente antes da extração do texto. Além disso, as alterações de escala podem ser úteis para padronizar o tamanho das imagens antes do processamento, garantindo que o OCR funcione da maneira mais eficaz possível. Por fim, a detecção de bordas pode ser útil para identificar e segmentar áreas específicas de um documento, como campos de formulário que precisam ser lidos pelo OCR. Isso é especialmente útil em processos de cadastro, onde os documentos têm um layout padrão.

Ao integrar estas técnicas de processamento de imagem ao OCR, é possível melhorar significativamente a eficiência e a qualidade do processo de cadastro, economizando tempo e reduzindo a possibilidade de erros. Isso resulta em uma operação mais eficiente e competitiva, beneficiando a organização como um todo.

OCR é um campo de estudo dentro de visão computacional focado na leitura e reconhecimento de caracteres, destinado a identificar um texto impresso ou digitalizado, dessa forma evitando a necessidade de redigitar um material já impresso em contexto de inserção de dados. Um *software* OCR procura identificar caracteres comparando-os com aquelas armazenadas na biblioteca do mesmo. O software busca identificar as palavras usando a proximidade de caracteres com a finalidade de reconstruir o *layout* da página original. A alta precisão da leitura é obtida usando varreduras nítidas e claras do material original de alta qualidade, no entanto a

qualidade diminui à medida que a qualidade do original diminui. (ENCYCLOPEDIA BRITANNICA, 2010).

De acordo com O'Gorman e Kasturi (1997), a *Optical Character Recognition* (OCR) é uma das tarefas de processamento textual realizadas na análise de imagens de documentos. Ela consiste em reconhecer o texto presente na imagem, por meio do reconhecimento óptico de caracteres, além de realizar outras tarefas, como determinar a inclinação da imagem, encontrar colunas, parágrafos, linhas de texto e palavras. A OCR é uma etapa essencial para extrair as informações desejadas de um documento de forma semelhante à leitura humana.

Segundo Rice, Nagy e Nartker (2012), os sistemas de reconhecimento de texto em máquinas surgiram no final da década de 1950 e estão amplamente presentes em computadores desde o início dos anos 1990. Esses sistemas, conhecidos como OCR, desempenham um papel importante ao converter o texto contido em documentos físicos para a forma eletrônica.

Essa conversão permite que sistemas de recuperação de informações localizem facilmente o conteúdo desejado e que softwares de processamento de texto editem o texto convertido. No entanto, é importante destacar que os sistemas de OCR não são perfeitos e podem cometer erros, resultando em uma versão eletrônica que não corresponde totalmente à versão em papel.

Em situações em que o sistema de OCR não consegue distinguir corretamente letras ou caracteres, pode ser mais eficiente digitar manualmente o conteúdo do documento do que corrigir os erros na saída do OCR. Portanto, a utilidade de um sistema de OCR está diretamente relacionada à sua precisão na conversão dos caracteres.

E. Chaudhuri et al. (2016) complementa que o Reconhecimento Óptico de Caracteres (OCR) é um processo de classificação de padrões ópticos contidos em uma imagem digital correspondente a caracteres alfanuméricos ou outros. O reconhecimento de caracteres é alcançado por meio de importantes etapas de segmentação, extração de características e classificação.

A tecnologia OCR permite converter diferentes tipos de documentos, como documentos de papel digitalizados, arquivos PDF ou imagens capturadas por uma câmera digital em dados editáveis e pesquisáveis. Embora muitos sistemas comerciais para a realização de OCR existam para uma grande variedade de aplicações, as máquinas disponíveis ainda não conseguem competir com as

capacidades humanas de leitura com os níveis de precisão desejados. O OCR pertence à família das técnicas de reconhecimento de máquinas que realizam identificação automática.

A identificação automática é o processo onde o sistema de reconhecimento identifica objetos automaticamente, coleta dados sobre eles e insere dados diretamente nos sistemas de computador, ou seja, sem a participação humana. O dado externo é capturado por meio da análise de imagens, sons ou vídeos. Para capturar dados, um transdutor é empregado que converte a imagem ou som real em um arquivo digital.

O arquivo é então armazenado e, posteriormente, pode ser analisado pelo computador. Existem várias técnicas para identificação automática que atendem às necessidades de diferentes áreas de aplicação.

## **2.4 Trabalhos relacionados**

Esta seção apresenta uma análise resumida dos trabalhos relacionados que exploram a aplicação da tecnologia de visão computacional em diferentes áreas. Os trabalhos abordados são: Câmeras de Segurança, Inspeções Automatizadas na Indústria e Diagnósticos de Doenças em Folhas de Milho.

No caso das Câmeras de Segurança, exemplificado pela câmera IP da Intelbras VIP 7208 LPR G2, a tecnologia de visão computacional e OCR é utilizada para a leitura de placas de veículos. Essas câmeras oferecem uma taxa de assertividade elevada na detecção de placas, agilizando o processo e armazenando dados de forma eficiente.

As Inspeções Automatizadas na Indústria têm se mostrado uma solução interessante para avaliar a conformidade de peças. Ao eliminar problemas associados à inspeção humana, como subjetividade, fadiga e lentidão, esses sistemas oferecem repetibilidade e precisão nas medições, resultando em melhorias na qualidade e eficiência do processo.

No campo dos Diagnósticos de Doenças em Folhas de Milho, o sistema DiagnoFoMil utiliza a visão computacional e redes neurais para analisar imagens de folhas de milho, identificar doenças e fornecer medidas para o combate às enfermidades. A segmentação da imagem é utilizada para extrair informações relevantes e descartar partes não essenciais.

Esses trabalhos relacionados demonstram a aplicabilidade e benefícios da visão computacional em diversas áreas, proporcionando avanços tecnológicos e contribuindo para o desenvolvimento de soluções eficientes e precisas.

Existem Câmeras com tecnologia de visão computacional embutidas as quais utilizam a tecnologia de OCR para leitura de placas de veículos, podendo citar como exemplo a câmera IP da intelbras VIP 7208 LPR G2. Possuindo um limite de velocidade para a leitura de 40Km/h ela consegue detectar placas de veículos através do vídeo possuindo uma assertividade de 90% para placas do padrão antigo e 70% para placas do padrão mercosul. (INTELBRAS, 2019).

Esta aplicação se relaciona com o trabalho pois utiliza a mesma tecnologia para processamento de caracteres e auxilia pessoas do setor de segurança agilizando e economizando o tempo pois faz a leitura e armazena os dados numa velocidade superior à do ser humano.

Outra aplicação relevante relacionada à visão computacional é o uso de inspeções automatizadas em indústrias. Através dessa tecnologia, é possível utilizar algoritmos e sistemas de visão para realizar verificações e análises precisas em processos industriais. Essas inspeções automatizadas podem identificar defeitos, realizar controle de qualidade, monitorar o desempenho de equipamentos e garantir a conformidade com os padrões estabelecidos. Isso resulta em maior eficiência, redução de erros e melhorias significativas nos processos de produção industrial.

Vem se tornando uma solução interessante e atrativa para analisar conformidade de peças devido a redução de preços das câmeras digitais. Processos de manufatura precisam avaliar: tolerâncias dimensionais, tolerâncias geométricas, ajustagem, incertezas e defeitos de fabricação. Sistemas desse tipo podem oferecer às empresas uma repetitividade e exatidão em medições que não exigem contato físico, devido também ao fato de eliminarem questões como subjetividade, fadiga, lentidão e o custo intrínseco na inspeção humana. (FELICIANO et al, 2005).

Em um sistema de medição automatizados por imagem, o principal problema se consiste em combinar a tecnologia com a aplicação em específico de uma maneira eficiente e econômica. Não existe um sistema que resolve todos os problemas pois cada problema é específico e exige uma solução específica. Nesse tipo de sistema existem duas abordagens comuns, a primeira se refere a operações de reconhecimento e a segunda a inspeções automatizadas, no primeiro tipo devemos extrair alguma característica do objeto e utilizando algum tipo de inteligência computacional observar distinção entre os objetos, portanto, não se torna essencial obter valores exatos. Por outro lado, em sistemas de inspeção a exatidão é fundamental, pois se trata de processos que visam manter a qualidade do objeto sem o

contato humano que utiliza algoritmos computacionais. (FELICIANO et al, 2005).

Nesse contexto podemos observar algumas aplicações como: identificação de erros em peças para controle de qualidade; orientar autômatos; rastreabilidade de materiais ou produtos na linha de produção; medição de peças; etc...(FELICIANO et al, 2005).

E por fim, um campo de aplicação interessante que também utiliza técnicas de visão computacional é o diagnóstico de doenças em folhas de milho. Através da análise de imagens das folhas, algoritmos podem identificar sintomas de doenças, como manchas, deformações ou coloração anormal. Essa abordagem permite um monitoramento mais eficiente e preciso das plantações, auxiliando os agricultores na detecção precoce de problemas fitossanitários. Com o diagnóstico rápido, é possível implementar medidas de controle e prevenção adequadas, contribuindo para a saúde e produtividade das lavouras de milho.

DiagnoFoMil é um sistema que consiste em envio de imagens para diagnóstico, onde a imagem é selecionada e então recortada, com o objetivo de realizar uma pré-segmentação e dessa forma excluindo partes irrelevantes da imagem, em seguida da pré-segmentação o sistema realiza de fato uma segmentação separando as manchas do restante da folha, então o sistema produz um diagnóstico informando a doença e fornece também algumas medidas para o combate à enfermidade. A ferramenta por intermédio da visão computacional e uma rede neural, extrai características encontradas na imagem para o diagnóstico da doença.(STAHNKE, 2013).

Comparando o DiagnoFoMil com o trabalho aqui apresentado ambos realizam a segmentação da imagem para extrair os pontos mais importantes e descartando então partes irrelevantes para a aplicação.

Os fundamentos do tráfego de documento eletrônico foram discutidos, destacando o amplo campo de aplicação da legislação brasileira que governa os processos judiciais digitais. A essência do elemento de imagem, conhecido como *pixel*, foi explorada, revelando sua importância na formação de imagens digitais. Foi dada uma atenção especial à visão computacional e suas técnicas de processamento de imagem, enfatizando métodos de segmentação, limiarização, redução de ruído, transformação e remoção de bordas.

Também foi analisada a tecnologia OCR, reconhecida pelo seu papel crucial na identificação de texto impresso ou digitalizado, fundamental para a eficiência e precisão na inserção de dados. Em cada um desses tópicos, a junção da tecnologia e

legislação se mostrou uma força poderosa para aumentar a eficiência, a precisão e a acessibilidade de várias tarefas relacionadas à manipulação e interpretação de dados.

Como seguimento ao embasamento teórico, o próximo passo na pesquisa é a aplicação prática desses conceitos no desenvolvimento de uma interface web utilizando o *framework* React. Essa interface será projetada para incorporar funcionalidades de cadastro manual e automatizado de pessoas, armazenamento de dados em uma tabela com informações pertinentes e comparação entre os tempos despendidos em cada tipo de cadastro. O objetivo é demonstrar se o tempo investido em cadastros automatizados realmente resulta em uma economia de tempo em comparação com os cadastros manuais.

### 3 DESENVOLVIMENTO

No desenvolvimento da solução foi utilizada a linguagem de programação Python, associada a biblioteca de funções OpenCV, para realizar o tratamento de uma imagem com o intuito de se obter um maior sucesso na identificação correta de caracteres, aplicando técnicas de identificação de contornos, inversão, desfoque, limiarização, binarização, redimensionamento, remoção de ruído, rotação, inclinação e remoção de bordas.

Como objeto de teste foram produzidos cinco documentos de identificação com dados fictícios, onde todos esses documentos por padrão irão possuir um número de registro com doze dígitos, a data da impressão do documento, nome da pessoa, nome da mãe, nome do pai, local de nascimento e data de nascimento como na imagem a seguir (Figura 1).



Após obter o resultado parcial da imagem tratada, a biblioteca de funções *pyTesseract* terá a finalidade de reconhecer caracteres presentes na imagem, e retornar as sequências de caracteres de cada um dos dados presentes no documento de identificação.

Comparar o tempo despendido na extração dos dados e a corretude de cada um dos 5 documentos de identificação, primeiramente digitando os dados e confrontando com o resultado obtido após enviar como parâmetro para a solução uma fotografia do mesmo documento.

Foram explorados dois principais componentes: uma API para processamento de imagens e OCR, e uma aplicação *frontend*. Ambas as soluções foram desenvolvidas e testadas usando um dispositivo desktop.

A API, escrita em Python, utiliza as bibliotecas *pytesseract* e OpenCV para processar imagens de documentos previamente digitalizados e extrair dados dos mesmos. O *pytesseract*, com licença Apache, é uma ferramenta de OCR que lê texto de imagens, enquanto o OpenCV, uma biblioteca de processamento de imagens licenciada pelo MIT, ajuda a preparar as imagens para a extração de texto.

Por outro lado, a interface de usuário foi desenvolvida usando o *framework* ReactJS, criando uma aplicação frontend que se conecta à API para facilitar a interação do usuário. A interface é estruturada para permitir tanto o cadastro manual quanto o automatizado de pessoas, armazenando e exibindo dados de cadastro de maneira eficiente. Além disso, oferece a funcionalidade de ajuste de imagem e verificação de dados extraídos pelo OCR.

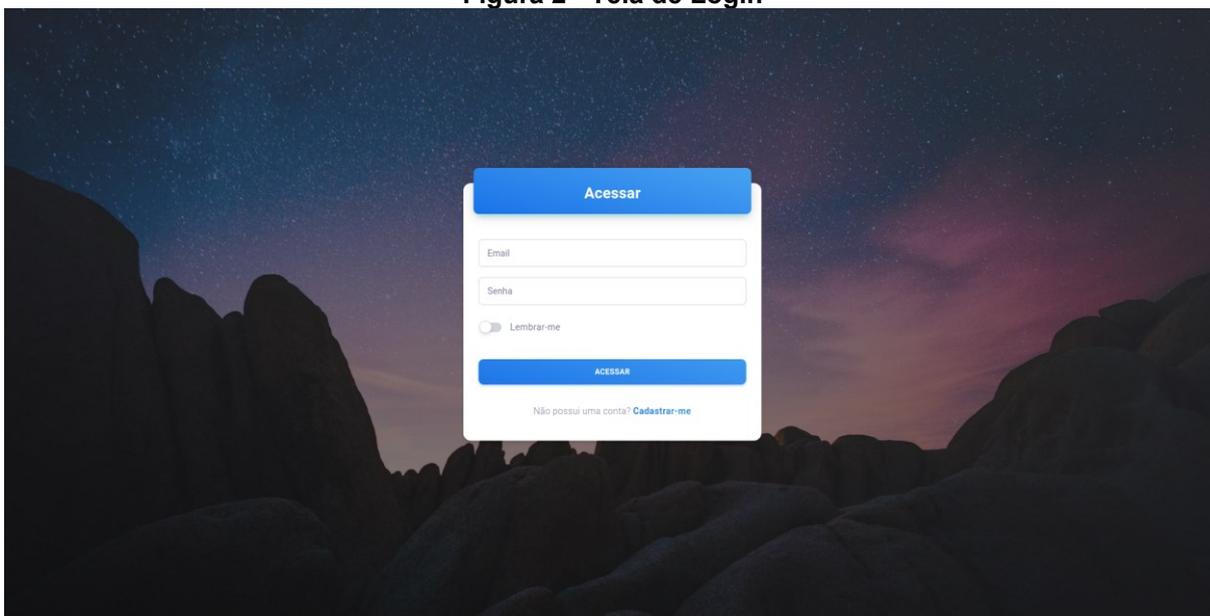
Vale ressaltar que esta aplicação foi projetada especificamente para o uso em computadores *desktop*, alinhando-se à necessidade de um ambiente de trabalho robusto e estável. O estudo visa entender se o tempo gasto em cadastros automatizados é, de fato, menor que em cadastros manuais, tendo como base a estrutura desenvolvida.

Para o Desenvolvimento do trabalho foi criado uma interface WEB usando o *framework* react, desenvolvida com os seguintes objetivos:

1. Realizar o cadastro manual e automatizado de pessoas.
2. armazenar os cadastros em uma tabela exibindo além dos dados da pessoa cadastrada, a forma de cadastro (manual ou OCR) e o tempo despendido na realização.
3. Comparar se o tempo gasto em cadastros automatizados é realmente menor que em cadastros manuais.

Interface onde o usuário pode realizar o acesso ao sistema (Figura 2).

**Figura 2 - Tela de Login**



**Fonte: Aatoria Própria (2023)**

Tabela onde fica armazenado os cadastros realizados bem como a forma de cadastro e o tempo gasto (Figura 3).

**Figura 3 - Tabela de Pessoas**

A interface de uma tabela de pessoas no sistema. À esquerda, há um menu lateral escuro com o logotipo "TCC" e opções para "Pessoas", "OCR" e "Sign In". O cabeçalho da página indica "Pessoas" e "Pessoas". Abaixo, há uma barra azul com o título "Lista de Pessoas". A tabela possui os seguintes cabeçalhos: TIPO, TEMPO, NOME, REGISTRO GERAL, CPF, NATURALIDADE, UF, PAI, MÃE, DATA DE NASCIMENTO e DATA DE ESPERAÇÃO. O corpo da tabela está atualmente vazio. Um botão "+" está visível no canto inferior direito da interface.

**Fonte: Aatoria Própria (2023)**

Tela onde é realizada o cadastro de uma pessoa (Figura 4).

**Figura 4 - Tela para inserir pessoa**

**Inserir pessoa**

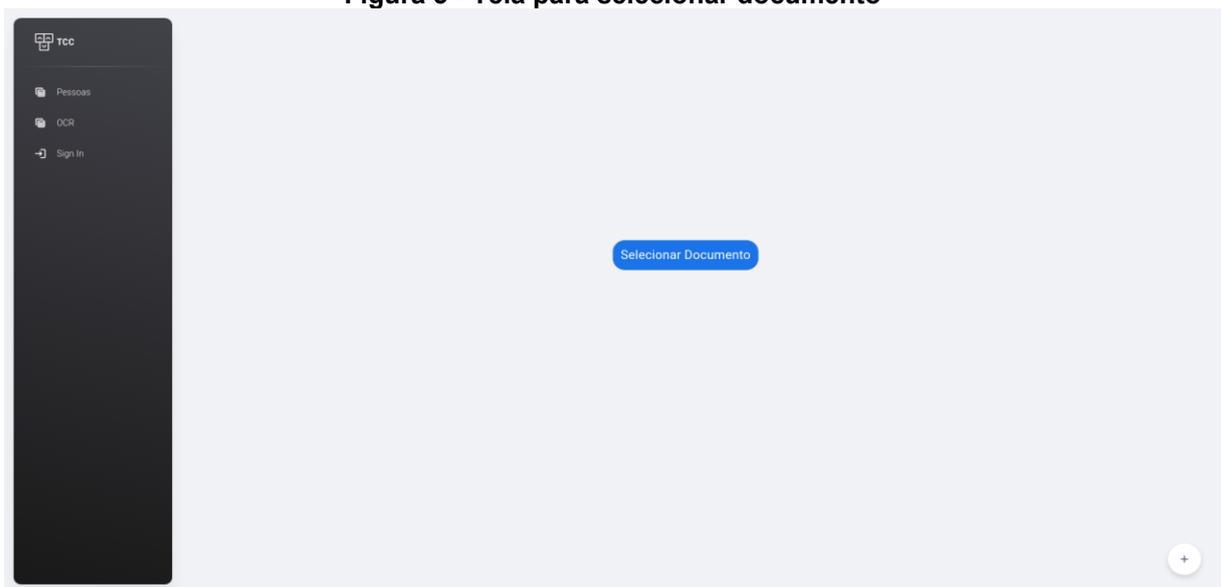
Nome Completo
RG
CPF
Naturalidade
UF
Nome do Pai
Nome da Mãe
Data de nascimento
Data de expedição

[CANCEL](#)   [SALVAR](#)

**Fonte: Autoria Própria (2023)**

Clicando na opção OCR do *dashboard* o usuário é levado a tela onde ele pode selecionar um documento já digitalizado (Figura 5).

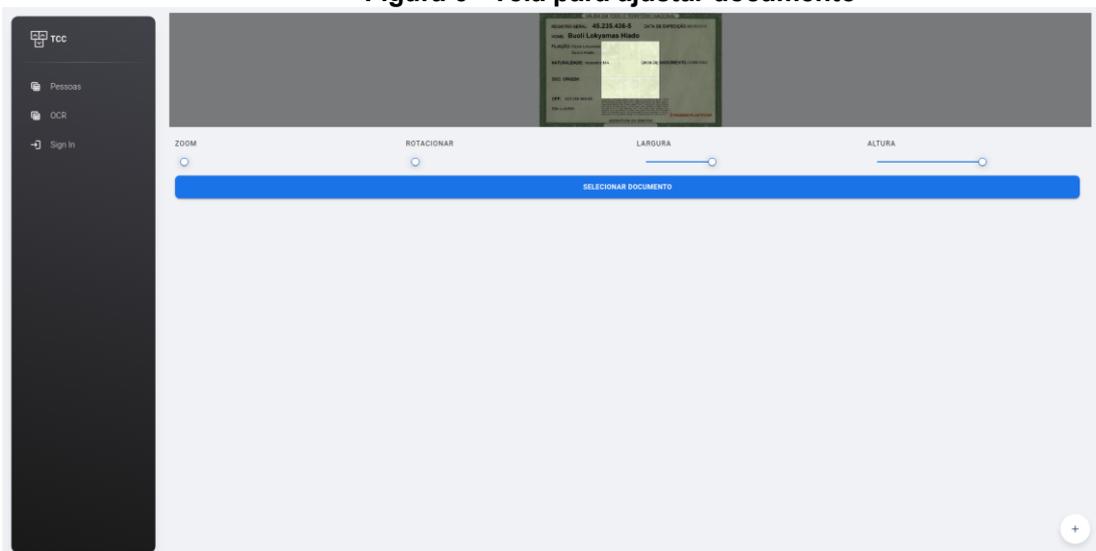
**Figura 5 - Tela para selecionar documento**



**Fonte: Autoria Própria (2023)**

Após selecionado o documento o usuário pode ajustar a imagem, rotacionando, dando zoom, e cortando apenas a parte necessária da imagem. (Figura 6).

**Figura 6 - Tela para ajustar documento**



Fonte: Aatoria Própria (2023)

Quando realizado o envio do documento para o OCR, após uma breve tela de carregamento, o usuário pode conferir se os dados obtidos pela visão computacional batem com o documento real, podendo realizar nesse passo qualquer ajuste necessário. (Figura 7).

**Figura 7 - Tela para conferir dados obtidos**

REGISTRO GERAL:	45.235.436-5	DATA DE EXPEDIÇÃO:	09/08/2010
NOME:	Buoli Lokyamas Hiado		
FILIAÇÃO:	Dayte Lokyamas Saeko Hiado		
NATURALIDADE:	Imperatriz/MA	DATA DE NASCIMENTO:	03/08/1992
DOC. ORIGEM:			
CPF:	921.066.893-65	Lorem ipsum condimentum egestas ornare donec convallis adipiscing sociisqu, vehicula	

Nome Completo:	Buoli Lokyamas Hiado		
RG:	452354365		
CPF:	921.066.893-65		
Naturalidade:	Imperatriz	UF:	MA
Nome do Pai:	Dayte Lokyamas		
Nome da Mãe:	Saeko Hiado		
Data de nascimento:	03/08/1992		
Data de expedição:	09/08/2010		

Fonte: Aatoria Própria (2023)

A ferramenta utilizada para o processamento de imagens foi o OpenCV extraído do pypi, plataforma de bibliotecas para python e possui licença MIT. E para a extração

de caracteres foi utilizado o *pytesseract* também extraído do pypi o qual possui licença apache, nas próximas subseções serão descritos alguns métodos de códigos.

Todas as imagens enviadas são salvas numa pasta já definida no projeto chamada “*uploads*”, em seguida todos os arquivos na pasta são apagados, sequencialmente é verificado no corpo da requisição enviada se realmente existe uma imagem, retornando um erro caso não exista, também verifica se a imagem enviada está no formato jpeg, após isso é capturado a data e hora atual utilizando esse dado para nomear a imagem na pasta uploads, e por fim ela é salva em formato png.

**Figura 8 - Operações Iniciais**

```
files = glob.glob('./uploads/*')
for f in files:
    os.remove(f)
r = request

if "image" not in r.files:
    response = {"message": "imagem não fornecida"}
    response_pickled = jsonpickle.encode(response)
    return Response(
        response=response_pickled,
        status=400,
        mimetype="application/json"
    )

file = request.files["image"]

if file.content_type != "image/jpeg":
    response = {"message": "this extension is not allowed"}
    response_pickled = jsonpickle.encode(response)
    return Response(
        response=response_pickled,
        status=400,
        mimetype="application/json"
    )

date = dt.datetime.now()
filename = date.strftime("%m%d%Y%H%M%S") + ".png"
file.save("./uploads/{}".format(filename))
```

**Fonte: Aatoria Própria (2023)**

Após a imagem ser armazenada na pasta uploads começamos a utilização do opencv onde o mesmo realizar a leitura da imagem em uma escala de cinza, redimensiona a imagem aumentando ela, para que a leitura de caracteres fique mais fácil, e por fim realiza a binarização da imagem onde o limiar de binarização foi definido como 127.

**Figura 9 - Pré Processamento**

```
img = cv.imread("./uploads/{}".format(filename), cv.IMREAD_GRAYSCALE)

img = cv.resize(
    img,
    None,
    fx=2,
    fy=2,
    interpolation=cv.INTER_CUBIC
)

ret, img = cv.threshold(
    img,
    127,
    255,
    cv.THRESH_BINARY
)
```

**Fonte: Autoria Própria (2023)**

Nesse trecho é capturado os 4 pontos do retângulo que contém as informações a serem extraídas, descartando o restante.

**Figura 10 - Obtenção de Pontos**

```
pontos, _ = cv.findContours(
    img,
    cv.RETR_TREE,
    cv.CHAIN_APPROX_SIMPLE
)

crop_img = img[pontos[1] : pontos[1] + pontos[2], pontos[0] :
pontos[0] + pontos[3]]
```

**Fonte: Autoria Própria (2023)**

Aqui está sendo realizada a suavização das bordas e a remoção de ruídos, visando entregar para o *pytesseract* uma imagem que com a menor quantidade possível de ruído, facilitando a extração de caracteres.

**Figura 11 - Pós Processamento**

```
crop_img = cv.GaussianBlur(crop_img, (5, 5), 1)

crop_img = cv.fastNlMeansDenoising(crop_img, 1.0, 7, 21)
```

**Fonte: Autoria Própria (2022)**

Para a leitura de caracteres, é realizado o corte da área onde contém a informação, como no exemplo do campo do cpf. Só então o *pytesseract* faz a leitura dos caracteres, evitando que seja realizada a leitura desnecessária de caracteres, esse passo é feito para todos os dados necessários totalizando 9.

**Figura 12 - Leitura de Caracteres**

```
imgCpf = crop_img[
    math.trunc(height * 0.94) : height,
    math.trunc(width * 0.085) : math.trunc(width * 0.27),
]

doc.setCpf(
    tess.image_to_string(
        imgCpf,
        config='--psm 10 --oem 3 -c tesseract_char_whitelist=0123456789'
    ).replace("\n\f", "")
)
```

**Fonte: Autoria Própria (2023)**

Por fim é construído o corpo da resposta para retornar ao cliente.

**Figura 13 - Retorno ao Cliente**

```
response = jsonpickle.encode(doc, unpicklable=False)
return Response(
    response=response,
    status=200,
    mimetype="application/json")
```

**Fonte: Autoria Própria (2023)**

O terceiro capítulo detalhou o desenvolvimento de uma interface web, utilizando o *framework* React. Esta interface, criada com o propósito de facilitar o cadastro de indivíduos, permite a realização deste processo tanto de maneira manual

como automatizada, por meio de uma tecnologia de Reconhecimento Óptico de Caracteres (OCR). O sistema conta com diversas seções, incluindo uma tela de login, uma tabela para exibição das pessoas cadastradas e áreas dedicadas para inserção de novas pessoas, ajustes de imagens de documentos e verificação de dados. Adicionalmente, foram empregadas as bibliotecas OpenCV e *pytesseract* para o processamento de imagens e extração de caracteres, respectivamente.

Avançando para o próximo capítulo, serão apresentados os resultados obtidos a partir da implementação e utilização do sistema desenvolvido. Esta seção, intitulada "Resultados Obtidos", irá explorar uma análise detalhada do tempo despendido em diferentes métodos de cadastro e uma avaliação do desempenho do *pytesseract*, a ferramenta de OCR empregada. A partir destes resultados, será possível obter uma compreensão mais clara do desempenho do sistema sob diversas situações, além de identificar possíveis áreas onde melhorias ou ajustes podem ser necessários.

## 4 RESULTADOS OBTIDOS

Avançando no trabalho acadêmico, o capítulo 4, intitulado "Resultados Obtidos", oferece uma visão dos efeitos práticos e do desempenho do sistema desenvolvido. Nesta seção, serão discutidos dois principais aspectos: o tempo despendido para a realização de cadastros, seja de forma manual ou automatizada com o uso de OCR, e a eficácia do *pytesseract*, a ferramenta de OCR utilizada. Os resultados desse capítulo fornecerão dados sobre os resultados das técnicas de processamento de imagens, tempo economizado ao utilizar o método de cadastro automatizado, bem como as limitações encontradas durante o uso da ferramenta *pytesseract* em diferentes condições de imagem.

Voltando-se para a API, quando enviado o documento para o mesmo, a primeira operação realizada é a de limiarização da imagem (Figura 14).

**Figura 14 - Documento limiarizado**

VÁLIDA EM TODO O TERRITÓRIO NACIONAL

REGISTRO GERAL: **45.235.436-5** DATA DE EXPEDIÇÃO: 09/08/2010

NOME: **Buoli Lokyamas Hiado**

FILIAÇÃO: Dayte Lokyamas  
Saeko Hiado

NATURALIDADE: Imperatriz/MA DATA DE NASCIMENTO: 03/08/1992

DOC. ORIGEM:

CPF: 921.066.893-65

São Luís/MA

Lorem ipsum: condimentum egestas ornare donec convallis adipiscing sociosque, vehicula suscipit faucibus at interdum lacinia volutpat; posuere sem atque hac nisi vulputate vivamus, nam vel ad tempor gravida eu nam erat mollis augue pulvinar cursus, donec lorem vestibulum ornare blandi sollicitudin arcu faucibus quisque morbi, ut massa quam congue consequat lacus semper nisi suscipit aenean ut sapien; tempor donec phasellus non habitant aliquam erat euismodi non a, class scelerisque dictum nibh volutpat; neque scelerisque lacus malesuada purus, varius facilisis vel habitant lorem euismodi eros porta vel ac himenaeos vivamus ut, ac bibendum ritus ut ultricesper torquent odio justo consectetur orci cubilia ut feugiat amet felis leo proin facilis a semper mollis semper fames habitant aptent; posuere e

ASSINATURA DO DIRETOR

É PROIBIDO PLASTIFICAR

Fonte: Autoria Própria (2023)

Com a imagem limiarizada é realizado o corte da imagem, visando remover dados irrelevantes para a leitura (Figura 15)

Figura 15 - Documento cortado

REGISTRO GERAL: **45.235.436-5** DATA DE EXPEDIÇÃO: 09/08/2010

NOME: **Buoli Lokyamas Hiado**

FILIAÇÃO: Dayte Lokyamas  
Saeko Hiado

NATURALIDADE: Imperatriz/MA

DATA DE NASCIMENTO: 03/08/1992

DOC. ORIGEM:

CPF: 921.066.893-65

Lorem ipsum condimentum egestas ornare donec convalis adipiscing sociosqu, vehicula suscipit faucibus et interdum lobortis vulputate, necesse non curae, hac nisi iudicata, utinam

**Fonte: Aatoria Própria (2023)**

A imagem cortada agora passa pelo processo de suavização de bordas tendo como objetivo suavizar as extremidades dos caracteres, as diferenças são pequenas e difíceis de perceber, se houver atenção prestada, as bordas são levementes desfocadas. (Figura 16).

Figura 16 - Documento com suavização de bordas

REGISTRO GERAL: **45.235.436-5** DATA DE EXPEDIÇÃO: 09/08/2010

NOME: **Buoli Lokyamas Hiado**

FILIAÇÃO: Dayte Lokyamas  
Saeko Hiado

NATURALIDADE: Imperatriz/MA

DATA DE NASCIMENTO: 03/08/1992

DOC. ORIGEM:

CPF: 921.066.893-65

Lorem ipsum condimentum egestas ornare donec convalis adipiscing sociosqu, vehicula suscipit faucibus et interdum lobortis vulputate, necesse non curae, hac nisi iudicata, utinam

**Fonte: Aatoria Própria (2023)**

O último processo pelo qual a imagem é submetida é o de remoção de ruído, existem alguns tipos de remoção de ruído, o utilizado foi o *fastNIMeansDenoising* gerando o resultado a seguir. (Figura 17).

Figura 17 - Documento com remoção de ruído

REGISTRO GERAL: **45.235.436-5** DATA DE EXPEDIÇÃO: 09/08/2010

NOME: **Buoli Lokyamas Hiado**

FILIAÇÃO: Dayte Lokyamas  
Saeko Hiado

NATURALIDADE: Imperatriz/MA

DATA DE NASCIMENTO: 03/08/1992

DOC. ORIGEM:

CPF: 921.066.893-65

Lorem ipsum condimentum egestas ornare donec convallis adipiscing sociosqu, vehicula

**Fonte: Autoria Própria (2023)**

Por fim, a imagem é submetida ao *pytesseract* para a extração do texto da imagem. A imagem é dividida em blocos para evitar leituras desnecessárias e garantir que o algoritmo não forneça informações incorretas de outros campos. A imagem a seguir foi criada apenas para exemplificar os blocos, sendo assim, apenas o conteúdo contido dentro de cada caixa é enviado para a leitura. (Figura 18).

Figura 18 - Documento com marcações

REGISTRO GERAL: **45.235.436-5** DATA DE EXPEDIÇÃO: 09/08/2010

NOME: **Buoli Lokyamas Hiado**

FILIAÇÃO: Dayte Lokyamas  
Saeko Hiado

NATURALIDADE: Imperatriz/MA

DATA DE NASCIMENTO: 03/08/1992

DOC. ORIGEM:

CPF: 921.066.893-65

Lorem ipsum condimentum egestas ornare donec convallis adipiscing sociosqu, vehicula

**Fonte: Autoria Própria (2023)**

A ferramenta utilizada *pytesseract* apesar de cumprir com o objetivo ao qual foi designado possui uma série de limitações quanto a imagem que lhe é submetida, ao lhe entregar uma imagem que não possua uma definição boa, a ferramenta realiza leituras erradas, trocando letras na maioria das vezes. Foi notado que uma simples barra no topo da imagem já dificulta a leitura da ferramenta sendo que a barra nem ao menos encostava nas letras.

Em relação ao tempo, foram realizados testes na utilização do sistema, fazendo o cadastro de forma manual e de forma automatizada com o OCR, capturando o tempo despendido em cada modalidade a fim de realizar uma comparação. O teste visou tanto pessoas com grande familiaridade com computadores quanto pessoas com pouco conhecimento, gerando uma diferença de em média 107 segundos quando utilizado o cadastro automatizado, pessoas com mais familiaridade na utilização de computadores possuem uma diferença de em média 52 segundos, e pessoas com mais dificuldades podem chegar a até 273 segundos de diferença.

**Tabela 1 - Tempos obtidos**

<b>Teste</b>	<b>Familiaridade com computadores</b>	<b>Tempo Manual</b>	<b>Tempo com OCR</b>	<b>Diferença</b>
1	Alta	90	30	60
2	Alta	87	29	58
3	Alta	88	31	57
4	Alta	85	33	52
5	Alta	75	28	47
6	Alta	77	28	49
7	Alta	78	30	48
8	Alta	86	27	59
9	Alta	74	33	41
10	Alta	79	30	49
11	Baixa	170	25	145
12	Baixa	180	30	150
13	Baixa	218	28	190
14	Baixa	298	25	273
15	Baixa	215	32	183
16	Baixa	213	27	186
17	Baixa	189	30	159
18	Baixa	137	31	106
19	Baixa	144	28	116
20	Baixa	145	33	112

**Fonte: Autoria Própria (2023)**

Os resultados apresentados no capítulo 4 demonstraram efetivamente os benefícios e limitações do sistema. Foi observado que o cadastro automatizado com o OCR, apesar de economizar tempo em comparação ao cadastro manual, possui suas próprias complexidades. A ferramenta *pytesseract*, embora útil, tem suas

restrições, principalmente ao lidar com imagens de baixa qualidade. Essas descobertas reiteram que, apesar do progresso na otimização do tempo de cadastro com o uso da visão computacional, ainda existem obstáculos que podem prejudicar a eficiência em cenários menos ideais.

Prosseguindo para o capítulo 5, serão abordadas as "Considerações Finais" deste trabalho. Aqui, o foco está na conclusão da pesquisa, recapitulando os resultados e os métodos utilizados. Além disso, serão discutidos os desafios encontrados, em especial os problemas relacionados ao uso de documentos de baixa definição e a eficácia do OCR em tempo real. Este capítulo também delineará possíveis direções para trabalhos futuros, sugerindo formas de melhorar o processo e superar as limitações identificadas. Isso pode incluir a experimentação com ferramentas de OCR mais avançadas ou aprimoramento da interface do usuário.

## 5 CONSIDERAÇÕES FINAIS

No presente trabalho acadêmico, foram abordadas técnicas de visão computacional em conjunto com um leitor óptico de caracteres, com o objetivo de comparar o tempo necessário para realizar tarefas de cadastro utilizando métodos manuais em contraste com métodos que empregam a leitura automática de caracteres. O propósito era investigar se haveria um aumento na agilidade do processo. Foram empregadas técnicas clássicas de visão computacional e OCR, que proporcionaram uma precisão aceitável dos documentos fictícios analisados.

Os resultados obtidos demonstraram um ganho de tempo significativo ao utilizar os métodos automatizados, o que leva à conclusão de que é possível empregar a visão computacional para otimizar o tempo necessário nessas tarefas. Entretanto, foram observados alguns desafios ao lidar com documentos em condições de definição inferior. Em cenários caóticos, a implementação de um leitor automático de caracteres poderia se tornar problemática, uma vez que a ferramenta apresentou dificuldades ao realizar leituras em tempo real. Se fosse necessário incluir a etapa de digitalização dos documentos no processo, o ganho de tempo alcançado na operação de cadastro poderia ser anulado pela nova tarefa de digitalização.

No que diz respeito a trabalhos futuros, considerando que essa tecnologia já foi explorada por outras empresas, existem diversas possibilidades de aprimoramento do processo. Seria viável realizar testes com ferramentas de OCR mais avançadas, visando obter uma leitura satisfatória e em tempo real dos documentos capturados por uma câmera. Além disso, melhorias relacionadas à interface do sistema são temas que podem ser discutidos e aprimorados no contexto deste estudo.

Neste trabalho foram detalhadas técnicas de visão computacional juntamente com um leitor óptico de caracteres com a proposta de comparar o tempo gasto em tarefas de cadastro quando utilizado métodos manuais e quando utilizado métodos com a leitura automática de caracteres, a fim de constatar se existe um ganho de agilidade no processo. Técnicas clássicas de visão computacional e OCR foram utilizados obtendo uma leitura satisfatória dos documentos fictícios.

Foi observado um ganho de tempo quando utilizado os métodos automatizados, concluindo que é possível utilizar-se de visão computacional para a otimização do tempo. Porém foram notados alguns problemas quando oferecidos documentos em pior estado de definição. Sendo assim em cenários caóticos ficaria

difícil a implementação de um leitor automático de caracteres, tendo em vista que a ferramenta possui uma deficiência ao ler documentos em tempo real, caso fosse necessário adicionar ao processo uma digitalização dos documentos, o tempo ganho na operação de cadastro poderia ser perdido na nova tarefa de digitalização.

Tratando-se de uma tecnologia já explorada por outras empresas, existem vários testes possíveis para a melhora do processo, testando ferramentas mais modernas de OCR, e conseguindo uma leitura satisfatória em tempo real dos documentos apresentados a uma câmera. Melhoras quanto a interface são passíveis de discussão

## REFERÊNCIAS

BRASIL. **Lei nº 11.419, de 19 de dezembro de 2006.** Dispõe sobre a informatização do processo judicial; altera a Lei nº 5.869, de 11 de janeiro de 1973 - Código de Processo Civil; e dá outras providências. Disponível em:

<[http://www.planalto.gov.br/ccivil\\_03/ato2004-2006/2006/lei/l11419.htm](http://www.planalto.gov.br/ccivil_03/ato2004-2006/2006/lei/l11419.htm)>

BRASIL. **Lei n. 8.078, de 11 de setembro de 1990.** Dispõe sobre a proteção do consumidor e dá outras providências. Brasília: Casa Civil, 1990. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/l8078compilado.htm#:~:text=do%20evento%20danoso.-,Art.,sobre%20sua%20frui%C3%A7%C3%A3o%20e%20riscos](https://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm#:~:text=do%20evento%20danoso.-,Art.,sobre%20sua%20frui%C3%A7%C3%A3o%20e%20riscos). Acesso em: 01 julho. 2023.

CONJUR. **Demora excessiva para atender emergência hospitalar causa dano moral**, diz TJ-RS. Disponível em: <<https://www.conjur.com.br/2017-dez-10/demora-excessiva-atender-emergencia-causa-dano-moral>>. Acesso em: 01 Ago. 2020.

CHAUDHURI, A. et al. **Optical Character Recognition Systems for Different Languages with Soft Computing.** Springer, 2016.

DAVIES, E. R. **Computer and Machine Vision: Theory, Algorithms, Practicalities.** 4. ed. Kidlington: Elsevier, 5 mar. 2012.

FELICIANO, F. SOUZA, I. LETA, F. Visão computacional aplicada à metrologia dimensional automatizada: considerações sobre sua exatidão. **Engevista.** Niterói. v. 7, n. 2, p. 38-50, dezembro 2005.

FILHO, Roberto Freitas. SANT'ANA, Ramiro Nóbrega. Direito Fundamental à Saúde no SUS e a Demora no Atendimento em Cirurgias Eletivas. **RDU**, Porto Alegre, Volume 2, n. 67, 2016, 70-102, jan-fev 2016.

O'GORMAN, L.; KASTURI, R. **Document Image Analysis.** Computer Society Executive Briefings (IEEE), 1997.

RICE, S. V.; NAGY, G.; NARTKER, T. A. **Optical Character Recognition.** Springer Science & Business Media, 2012.

STAHNKE, E.; SANTOS, F. Ferramenta para diagnóstico de Doenças Fúngicas em Folhas de Milho. **REAVI-Revista Eletrônica do Alto Vale do Itajaí** 2.2: pp.125-136, 2013.

TESSDOC. **Improving the quality of the output.** Disponível em: <<https://tesseract-ocr.github.io/tessdoc/ImproveQuality>>. Acesso em: 01 Ago. 2020.

INTELBRAS. **Ficha técnica – VIP 7208 LPR G2.** Disponível em: <[https://backend.intelbras.com/sites/default/files/2019-10/Datasheet\\_VIP\\_7208\\_LPR\\_G2\\_03-19.pdf](https://backend.intelbras.com/sites/default/files/2019-10/Datasheet_VIP_7208_LPR_G2_03-19.pdf)>. Acesso em: 23 abr. 2020

OCR. **Encyclopædia Britannica** Disponível em: <<https://www.britannica.com/technology/OCR>>. Acesso em: 26 jun. 2019.

**PIXEL. Encyclopedia Britannica.** Disponível em:

<<https://www.britannica.com/technology/pixel/>>. Acesso em: 03 jul. 2019

PULLI, K. BAKSHEEV, A. KORNYAKOV, K. ERUHIMOV, V. Real-time computer vision with OpenCV. **Communications of the ACM**. USA. v. 10, p 61-69, abr, 2019.

SHAPIRO, L; STOCKMAN, G. **Computer Vision**. Seattle: Prentice-Hall, Mar. 1 jan. 2001.

SOLOMON, Chris; BRECKON, Toby. **Fundamentos de Processamento Digital de Imagens: uma abordagem prática com exemplos em matlab**. Rio de Janeiro: LTC, 2013.

MARENGONI, M. STRINGHINI, D. **Tutorial: Introdução à Visão Computacional usando OpenCV**. Universidade Presbiteriana Mackenzie, 2009.

HARRINGTON, H. J. **Aperfeiçoando processos empresariais: estratégia revolucionária para o aperfeiçoamento da qualidade, da produtividade e da competitividade**. São Paulo: Makron Books, 1993.

MENDONÇA, R. 'Você não presta para nada': a rotina de estresse, xingamentos e pressão dos atendentes de telemarketing. **BBC News**, São Paulo, jul, 2018.

Disponível em:

<<https://www.bbc.com/portuguese/geral-44325034>>. Acesso em: 21 jun. 2019.

SORATTO DA SILVA, Alexandre Nixon Raulino. **Gerenciamento da responsividade de serviços: uma proposta para agilizar processos e moderar os efeitos de espera**. Dissertação de Mestrado. Universidade Tecnológica Federal de Santa Catarina. Florianópolis. 2004.