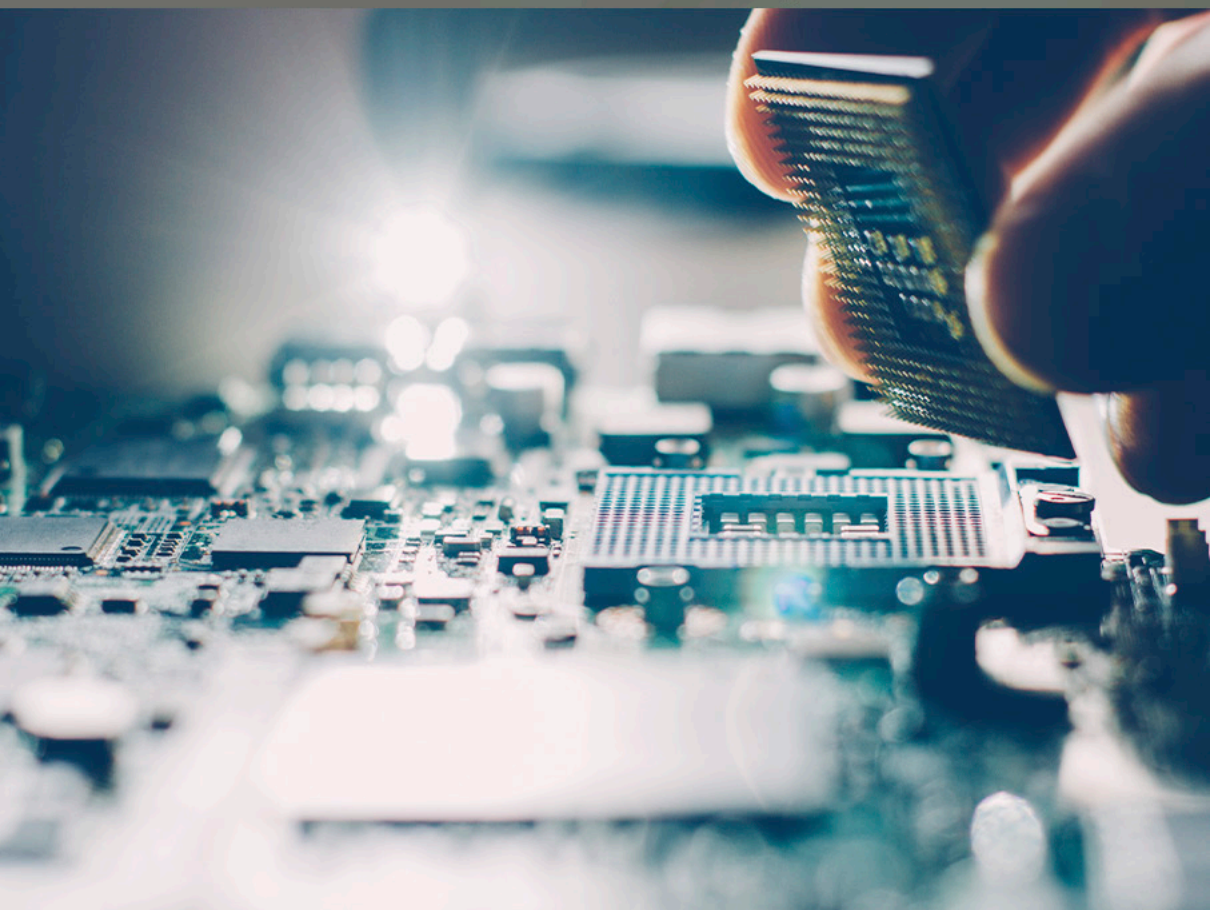


COLEÇÃO

DESAFIOS DAS ENGENHARIAS:

ENGENHARIA DE COMPUTAÇÃO 4

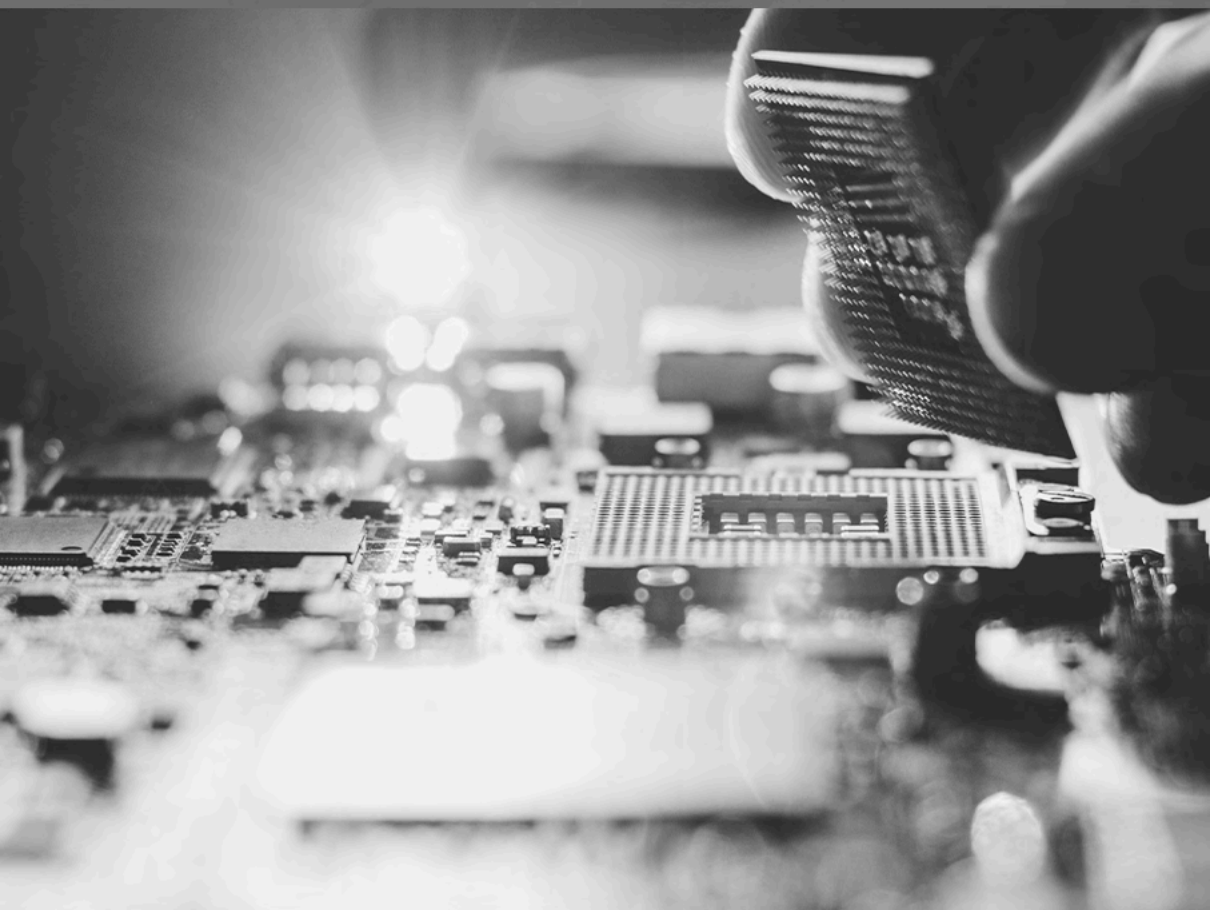


LILIAN COELHO DE FREITAS
(ORGANIZADORA)


Ano 2021

COLEÇÃO
DESAFIOS
DAS
ENGENHARIAS:

ENGENHARIA DE COMPUTAÇÃO 4



LILIAN COELHO DE FREITAS
(ORGANIZADORA)

Atena
Editora
Ano 2021

Editora chefe

Profª Drª Antonella Carvalho de Oliveira

Editora executiva

Natalia Oliveira

Assistente editorial

Flávia Roberta Barão

Bibliotecária

Janaina Ramos

Projeto gráfico

Camila Alves de Cremo

Daphynny Pamplona

Gabriel Motomu Teshima

Luiza Alves Batista

Natália Sandrini de Azevedo

Imagens da capa

iStock

Edição de arte

Luiza Alves Batista

2021 by Atena Editora

Copyright © Atena Editora

Copyright do texto © 2021 Os autores

Copyright da edição © 2021 Atena Editora

Direitos para esta edição cedidos à Atena Editora pelos autores.

Open access publication by Atena Editora



Todo o conteúdo deste livro está licenciado sob uma Licença de Atribuição Creative Commons. Atribuição-Não-Comercial-Não-Derivativos 4.0 Internacional (CC BY-NC-ND 4.0).

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores, inclusive não representam necessariamente a posição oficial da Atena Editora. Permitido o *download* da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

Todos os manuscritos foram previamente submetidos à avaliação cega pelos pares, membros do Conselho Editorial desta Editora, tendo sido aprovados para a publicação com base em critérios de neutralidade e imparcialidade acadêmica.

A Atena Editora é comprometida em garantir a integridade editorial em todas as etapas do processo de publicação, evitando plágio, dados ou resultados fraudulentos e impedindo que interesses financeiros comprometam os padrões éticos da publicação. Situações suspeitas de má conduta científica serão investigadas sob o mais alto padrão de rigor acadêmico e ético.

Conselho Editorial**Ciências Exatas e da Terra e Engenharias**

Prof. Dr. Adélio Alcino Sampaio Castro Machado – Universidade do Porto

Profª Drª Ana Grasielle Dionísio Corrêa – Universidade Presbiteriana Mackenzie

Prof. Dr. Carlos Eduardo Sanches de Andrade – Universidade Federal de Goiás

Profª Drª Carmen Lúcia Voigt – Universidade Norte do Paraná

Prof. Dr. Cleiseano Emanuel da Silva Paniagua – Instituto Federal de Educação, Ciência e Tecnologia de Goiás

Prof. Dr. Douglas Gonçalves da Silva – Universidade Estadual do Sudoeste da Bahia

Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná



Profª Drª Érica de Melo Azevedo – Instituto Federal do Rio de Janeiro
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará
Profª Dra. Jéssica Verger Nardeli – Universidade Estadual Paulista Júlio de Mesquita Filho
Prof. Dr. Juliano Carlo Rufino de Freitas – Universidade Federal de Campina Grande
Profª Drª Luciana do Nascimento Mendes – Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte
Prof. Dr. Marcelo Marques – Universidade Estadual de Maringá
Prof. Dr. Marco Aurélio Kistemann Junior – Universidade Federal de Juiz de Fora
Profª Drª Neiva Maria de Almeida – Universidade Federal da Paraíba
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Profª Drª Priscila Tessmer Scaglioni – Universidade Federal de Pelotas
Prof. Dr. Sidney Gonçalo de Lima – Universidade Federal do Piauí
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista



Coleção desafios das engenharias: engenharia de computação 4

Diagramação: Gabriel Motomu Teshima
Correção: Yaidy Paola Martinez
Indexação: Amanda Kelly da Costa Veiga
Revisão: Os autores
Organizadora: Lilian Coelho de Freitas

Dados Internacionais de Catalogação na Publicação (CIP)

C691 Coleção desafios das engenharias: engenharia de computação 4 / Organizadora Lilian Coelho de Freitas. – Ponta Grossa - PR: Atena, 2021.

Formato: PDF

Requisitos de sistema: Adobe Acrobat Reader

Modo de acesso: World Wide Web

Inclui bibliografia

ISBN 978-65-5983-752-6

DOI: <https://doi.org/10.22533/at.ed.526211012>

1. Engenharia de computação. I. Freitas, Lilian Coelho de (Organizadora). II. Título.

CDD 621.39

Elaborado por Bibliotecária Janaina Ramos – CRB-8/9166

Atena Editora

Ponta Grossa – Paraná – Brasil

Telefone: +55 (42) 3323-5493

www.atenaeditora.com.br

contato@atenaeditora.com.br



Atena
Editora
Ano 2021

DECLARAÇÃO DOS AUTORES

Os autores desta obra: 1. Atestam não possuir qualquer interesse comercial que constitua um conflito de interesses em relação ao artigo científico publicado; 2. Declaram que participaram ativamente da construção dos respectivos manuscritos, preferencialmente na: a) Concepção do estudo, e/ou aquisição de dados, e/ou análise e interpretação de dados; b) Elaboração do artigo ou revisão com vistas a tornar o material intelectualmente relevante; c) Aprovação final do manuscrito para submissão.; 3. Certificam que os artigos científicos publicados estão completamente isentos de dados e/ou resultados fraudulentos; 4. Confirmam a citação e a referência correta de todos os dados e de interpretações de dados de outras pesquisas; 5. Reconhecem terem informado todas as fontes de financiamento recebidas para a consecução da pesquisa; 6. Autorizam a edição da obra, que incluem os registros de ficha catalográfica, ISBN, DOI e demais indexadores, projeto visual e criação de capa, diagramação de miolo, assim como lançamento e divulgação da mesma conforme critérios da Atena Editora.



DECLARAÇÃO DA EDITORA

A Atena Editora declara, para os devidos fins de direito, que: 1. A presente publicação constitui apenas transferência temporária dos direitos autorais, direito sobre a publicação, inclusive não constitui responsabilidade solidária na criação dos manuscritos publicados, nos termos previstos na Lei sobre direitos autorais (Lei 9610/98), no art. 184 do Código Penal e no art. 927 do Código Civil; 2. Autoriza e incentiva os autores a assinarem contratos com repositórios institucionais, com fins exclusivos de divulgação da obra, desde que com o devido reconhecimento de autoria e edição e sem qualquer finalidade comercial; 3. Todos os e-book são *open access*, *desta forma* não os comercializa em seu site, sites parceiros, plataformas de *e-commerce*, ou qualquer outro meio virtual ou físico, portanto, está isenta de repasses de direitos autorais aos autores; 4. Todos os membros do conselho editorial são doutores e vinculados a instituições de ensino superior públicas, conforme recomendação da CAPES para obtenção do Qualis livro; 5. Não cede, comercializa ou autoriza a utilização dos nomes e e-mails dos autores, bem como nenhum outro dado dos mesmos, para qualquer finalidade que não o escopo da divulgação desta obra.



APRESENTAÇÃO

A Atena Editora tem a honra de presentear o público em geral com a série de *e-books* intitulada “*Coleção desafios das engenharias: Engenharia de computação 4*”. Em seu quarto volume, esta obra apresenta várias aplicações tecnológicas da Engenharia de Computação na automação industrial, na agricultura, no setor de energias renováveis, e no mercado financeiro.

Organizado em 07 capítulos, este volume objetiva facilitar a difusão do conhecimento científico produzido em várias instituições de ensino e pesquisa do país.

Dessa forma, esta obra contribuirá para aprimoramento do conhecimento de seus leitores e servirá de base referencial para futuras investigações.

Os organizadores da Atena Editora, agradecem especialmente os autores dos diversos capítulos apresentados, parabenizam a dedicação e esforço de cada um, os quais viabilizaram a construção deste trabalho.

Boa leitura.

Lilian Coelho de Freitas

SUMÁRIO

CAPÍTULO 1..... 1


RED NEURAL ARTIFICIAL PARA EL SEGUIMIENTO DE PLANTACIONES DE ARROZ A ALTAS TEMPERATURAS

Silvia Soledad Moreno Gutiérrez

Mónica García Munguía

Yesica Zamudio Briseño

Carlos Pérez Núñez

 <https://doi.org/10.22533/at.ed.5262110121>

CAPÍTULO 2..... 10

REDES NEURAIS USADAS NA PREVISÃO DE CURVAS CARACTERÍSTICAS DE UM SISTEMA FOTOVOLTAICO

Gabriel Mancini


Jose Airton Azevedo dos Santos

Hugo Andrés Ruiz Flórez

Gloria Patricia Lopez Sepúlveda

Cristiane Lionço Zeferino

Leandro Antonio Pasa


 <https://doi.org/10.22533/at.ed.5262110122>

CAPÍTULO 3..... 23

UMA APLICAÇÃO PARA CONTROLE DE TEMPERATURA BASEADO EM SENSORES DE BAIXO CUSTO

Ana Carolina Mariath Magalhães Corrêa e Castro

Mário Mestría


 <https://doi.org/10.22533/at.ed.5262110123>

CAPÍTULO 4..... 35

SISTEMA ELETRÔNICO DE RECOMENDAÇÃO AGNÓSTICO E ONLINE DE APLICAÇÃO EM FUNDOS DE INVESTIMENTOS

Antonio Newton Licciardi Junior

Paulo Henrique Barros de Moura


 <https://doi.org/10.22533/at.ed.5262110124>

CAPÍTULO 5..... 57

CONTROLE ADAPTATIVO PID USADO EM DOIS ELOS DE UM ROBÔ DE três GRAUS DE LIBERDADE

José Antonio Riul

Paulo Henrique de Miranda Montenegro


 <https://doi.org/10.22533/at.ed.5262110125>

CAPÍTULO 6..... 69

PROCESAMIENTO Y GRAFICACIÓN DE SEÑALES ELETROMIOGRÁFICAS CON

RASPBERRY-PI 2 PARA LA REHABILITACIÓN DE MUÑECA


Mario Alberto García Martínez
Daniel Ivann Arias Guevara
Ingrid Lizette Sánchez Carmona

 <https://doi.org/10.22533/at.ed.5262110126>

CAPÍTULO 7..... 80

COMPARAÇÃO DE APIS DE OCR PARA RECONHECIMENTO DE DÍGITOS EM IMAGENS DE MOSTRADOR DE SETE SEGMENTOS

Jonathan Ribeiro da Silva
Leandro Colombi Resendo
Jefferson Oliveira Andrade
Karin Satie Komati

 <https://doi.org/10.22533/at.ed.5262110127>

SOBRE A ORGANIZADORA 93

ÍNDICE REMISSIVO..... 94

COMPARAÇÃO DE APIS DE OCR PARA RECONHECIMENTO DE DÍGITOS EM IMAGENS DE MOSTRADOR DE SETE SEGMENTOS

Data de aceite: 01/12/2021

Data da submissão: 06/09/2021

Jonathan Ribeiro da Silva

IFES – Coordenação de Informática
Viana – Espírito Santo
<http://lattes.cnpq.br/7538036826754459>

Leandro Colombi Resendo

IFES – Programa de Pós-graduação em
Computação Aplicada (PPComp)
Serra – Espírito Santo
<http://lattes.cnpq.br/8108487234297364>

Jefferson Oliveira Andrade

IFES – Programa de Pós-graduação em
Computação Aplicada (PPComp)
Serra – Espírito Santo
<http://lattes.cnpq.br/7138275599443632>

Karin Satie Komati

IFES – Programa de Pós-graduação em
Computação Aplicada (PPComp)
Serra – Espírito Santo
<http://lattes.cnpq.br/9860697624155451>

RESUMO: As plataformas de computação em nuvem tornam acessível a tecnologia de reconhecimento de texto. Entretanto, a escolha da plataforma mais adequada para uma determinada aplicação nem sempre é evidente. Este trabalho avaliou cinco soluções diferentes para reconhecimento de texto: AWS Rekognition, Microsoft Azure, Cloudmersive, Google OCR e OCRSpace. Foi utilizada uma base de dados

de imagens de mostradores de sete segmentos em medidores de energia elétrica, a “YUVA EB Dataset”. Não houve pré-processamento para a melhoria da qualidade da imagem, para melhoria de iluminação ou para eliminação de ruídos. O Google Cloud apresentou melhores resultados acertando 100 resultados das 169 imagens de entrada, com acurácia de 86,5% considerando os 965 dígitos. Os resultados obtidos sugerem que as soluções oferecidas comercialmente não são adequadas para uso em produção sem uma etapa anterior de pré-processamento das imagens.

PALAVRAS-CHAVE: Rekognition. Azure cognitive services. Cloudmersive. Google Cloud. OCRSpace. Processamento de imagens. Reconhecimento de texto. OCR.

COMPARISON OF OCR APIS FOR DIGIT RECOGNITION IN SEVEN SEGMENT DISPLAY IMAGES

ABSTRACT: Cloud computing platforms make text recognition technology accessible. However, the most suitable solution for a given application is not always evident. This paper evaluated five different text recognition solutions: AWS Rekognition, Microsoft Azure, Cloudmersive, Google OCR, and OCRSpace. A database of images of seven-segment displays in electricity meters, the “YUVA EB Dataset”, was used. There was no pre-processing to improve image quality, to improve lighting or to eliminate noise. Google Cloud showed better results, hitting 100 results of the 169 input images, with an accuracy of 86.5 % considering the 965 digits. The results

obtained suggest that the use of the solutions offered commercially are not suitable for use in production without a previous stage of pre-processing of the images.

KEYWORDS: Rekognition. Azure cognitive services. Cloudmersive. Google Cloud. OCRSpace. Image processing. Text recognition. OCR.

1 | INTRODUÇÃO

O mostrador (ou *display*) de sete segmentos, como o seu próprio nome diz, é composto de sete elementos, os quais podem ser ligados ou desligados individualmente de forma que a combinação desses elementos produz representações alfanuméricas. O uso mais comum do mostrador de 7 segmentos é na representação de algarismos arábicos, embora seja possível representar algumas das letras do alfabeto romano. A Figura 1 apresenta uma das representações dos algarismos arábicos (de 0 a 9) usando o mostrador de 7 segmentos. Variações destas representações são possíveis, tal como a retirada de um dos segmentos para os dígitos 6, 7 e 9.



Figura 1. Algarismos arábicos representados em 7 segmentos

Atualmente, mostradores com dígitos de sete segmentos são encontrados em calculadoras, relógios digitais, bem como em equipamentos de vídeo e áudio. Além disso, também são usados em muitos equipamentos de medição, tais como medidores de consumo de água, vazão, termômetros e controladores de temperatura digitais na indústria e medidores de consumo de energia elétrica. Eles são uma convenção de escrita amplamente utilizada e facilmente reconhecível.

Vários equipamentos de medição, que utilizam o mostrador de 7 segmentos, que estão em atividade são antigos. Apesar de funcionarem perfeitamente, não possuem alguma forma de comunicação para transferir os dados medidos para um sistema de informação. Como em muitos casos, o custo para a troca destes equipamentos é alto, prefere-se usar soluções computacionais para o reconhecimento automático destes dígitos [Finnegan *et al.* 2019] [Bonačić *et al.*, 2020].

Destaca-se o trabalho de [Kanagarathinam and Sekar 2019]. Neste trabalho, a solução foi aplicada em medidores de consumo de energia elétrica, que dependem de interação humana para coletar os dados de medição e realizar a transferência desses dados para um sistema que possa fazer o cálculo da fatura. A troca desses aparelhos antigos por mais modernos se mostrou inviável, acarretando um custo muito alto, na parada de processos de

produção, e/ou fornecimento de serviço. Nesse trabalho, as imagens capturadas passam por operações de processamento de imagens, pelos algoritmos MSER (*Maximally Stable Extremal Regions*) e posteriormente pela operação de dilatação de imagem. Estas imagens processadas é que são avaliadas pelo OCR (do inglês *Optical Character Recognition*) do MATLAB, que reconhece caracteres em uma imagem, convertendo as imagens dos caracteres em texto. Os resultados obtidos são bem satisfatórios alcançando uma taxa de acurácia superior a 90% para maioria dos dígitos entre 0 e 9. Além disso, o trabalho disponibiliza a base de dados “YUVA EB Dataset”, que contém uma coleção de imagens de medidores digitais de energia. As imagens foram capturadas em condições de luz diurna e noturna.

Existem várias ferramentas OCR disponíveis no mercado capazes de realizar o reconhecimento de caracteres, criando opções para solucionar problemas de transferência de dados de medidores digitais. A questão de pesquisa é: será que as ferramentas de OCR de mercado teriam um resultado melhor que o trabalho [Kanagarathinam and Sekar 2019], usando a mesma base de dados?

A proposta deste trabalho é comparar ferramentas de OCR usando a base de dados “YUVA EB Dataset” sem a etapa de pré-processamento das imagens. A escolha das ferramentas foi via pesquisa Google das APIs¹ mais populares e que fossem de uso gratuito. Assim, foram selecionadas 5 APIs: (i) AWS Rekognition², (ii) Microsoft Azure³, (iii) Cloudfmersion⁴, (iv) Google OCR⁵ e (v) OCRSpace⁶.

Para cada imagem da base de dados, serão coletados os resultados de cada API selecionada. As análises e comparações serão feitas levando em consideração as características da base de dados, como condição de captura das imagens, posição dos dígitos, reconhecimento do ponto decimal e dígito após o ponto decimal. Além disso, será feita a comparação com o trabalho base [Kanagarathinam and Sekar 2019].

2 | OCR

O reconhecimento óptico de caracteres é, em um sentido amplo, um ramo da inteligência artificial e também um ramo da visão computacional. No entanto, é uma disciplina distinta por direito próprio, análoga ao reconhecimento de fala. Os processos que ocorrem no OCR podem ser divididos aproximadamente em três partes: pré-processamento, extração de recursos e classificação. Sendo o pré-processamento a etapa onde o tratamento da imagem é feito, aplicando filtros para retirar ruídos e equalizar a iluminação

1 API é um conjunto de rotinas e padrões de programação para acesso a um aplicativo de software ou plataforma baseado na Web. A sigla API refere-se ao termo em inglês “*Application Programming Interface*” que significa em tradução para o português “Interface de Programação de Aplicativos”.

2 <https://docs.aws.amazon.com/rekognition/latest/dg/text-detection.html>

3 <https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/concept-recognizing-text>

4 <https://cloudmersion.com/ocr-api>

5 <https://cloud.google.com/google/ocr>

6 <https://ocr.space/>

na imagem e também a definição da área de interesse da imagem. A etapa de extração de recursos ou características é onde são aplicadas técnicas para identificar os contornos dos possíveis caracteres presentes na imagem. A terceira etapa, de discriminação também conhecida como classificação, é onde são aplicados métodos para classificar os caracteres encontrados a partir das informações geradas pela extração de características na segunda etapa.

No campo de detecção de texto em imagens ainda existem muitos desafios para se obter uma boa precisão tanto para detecção de texto quanto para reconhecimento de texto. Em particular, o reconhecimento óptico de caracteres (OCR) frequentemente falha em fornecer resultados exatos. A detecção de texto precisa e eficiente em imagens de cenas naturais ainda é um problema em aberto neste campo devido às características imprevisíveis do texto da cena, se assemelhando com as complexidades do problema de resolver CAPTCHAs (do inglês *Completely Automated Public Turing tests to tell Computers and Humans Apart*, em português, Testes Automatizados de Turing para Diferenciar Humanos e Computadores) automaticamente, que envolve a extração de texto na imagem.

Muitos fatores influenciam a precisão do texto reconhecido por OCR. Esses fatores incluem resolução da digitalização, qualidade do *scanner*, tipo de documento impresso fotocopiado ou impressora a laser, qualidade do papel, complexidades linguísticas, fontes usadas no texto e dicionário usado. A iluminação não uniforme e as marcas d'água são exemplos de problemas que influenciam a precisão do OCR em comparação com um texto limpo em um fundo branco [Kanagarathinam and Sekar 2019].

A maioria das ferramentas OCR comerciais são “caixas pretas” proprietárias, não fornecendo contexto para entender seu funcionamento ou personalizar os parâmetros sob os quais operam. Desta forma, o que nos interessa é a acurácia e a aceitação dos resultados que essas ferramentas entregam.

Uma das vantagens de se utilizar ferramentas OCR comerciais é a não necessidade de estudo aprofundado sobre aprendizado de máquina, assim como o desenvolvimento de uma aplicação mais complexa para realizar o reconhecimento de caracteres baseado em padrões previamente definidos. As três principais ferramentas que estão no mercado hoje são dos grandes serviços de nuvem, AWS, Microsoft Azure e Google Cloud. A expertise em aprendizado de máquina que esses serviços possuem permite a disponibilização de diversas ferramentas de reconhecimento de padrões, reconhecimento de fala, e o objeto de estudo deste trabalho, o reconhecimento de caracteres em imagens.

A evolução das técnicas de aprendizado de máquina melhorou os resultados das ferramentas de OCR, mas ainda possuem a barreira do pré-processamento e delimitação da área a ser lida em alguns casos mais específicos. Ferramentas diferentes usam recursos e técnicas variadas, apresentando um resultado variado também.

2.1 Amazon rekognition

Amazon Web Services (AWS) é uma plataforma de serviços de computação em nuvem. Esses serviços são oferecidos sob demanda para os clientes da AWS. Com o plano de conta gratuita é possível explorar mais de 85 produtos que a AWS fornece. São ferramentas de análise, armazenamento, aprendizado de máquina, robótica, segurança entre outras.

A Nuvem AWS abrange 77 zonas de disponibilidade em 24 regiões geográficas em todo o mundo, com planos já divulgados para mais 15 zonas de disponibilidade e mais 5 regiões da AWS na Índia, Indonésia, Japão, Espanha e Suíça.

O Amazon Rekognition é uma plataforma de visão computacional baseada em nuvem que oferece serviços e ferramentas para análise de imagens e vídeos utilizando aprendizado de máquina. Através dos serviços do Rekognition é possível moderar conteúdo de páginas web, detectar texto, detectar e analisar faces, pesquisar e confirmar faces e, detectar equipamentos de proteção individual, dentre outros.

O Amazon Rekognition pode ler texto inclinado e distorcido para capturar informações como nomes de lojas, narrativas forçadas sobrepostas à mídia, placas de rua e texto na embalagem do produto. As APIs de detecção de texto de imagem e vídeo suportam a maioria das fontes, incluindo as altamente estilizadas. Depois de detectar o texto, o Amazon Rekognition cria uma representação das palavras e linhas de texto detectadas, mostra a relação entre elas e informa onde o texto está em uma imagem ou quadro de vídeo.

2.2 Azure cognitive services

O Azure é uma plataforma de serviços que fornece capacidades de processamento e armazenamento em nuvem para solucionar diversos problemas ou ampliar escopo de aplicações. Essa plataforma entrega diversas ferramentas e serviços de aprendizado de máquina, banco de dados, computação, rede, segurança e outras mais.

Dentre os vários serviços que o Azure fornece existe uma família de serviços de IA e APIs cognitivas que auxiliam em moderação, detecção de sentimentos, reconhecimento de voz, transcrita de fala em texto, pesquisa visual computacional, detecção facial, extração de texto e outros mais.

O serviço Pesquisa Visual Computacional é um serviço de IA que analisa o conteúdo de imagens e vídeos sem a necessidade prévia de conhecimento de aprendizado de máquina por parte do usuário. Nele está a função OCR para extrair texto de imagens e documentos. É possível extrair texto de imagens, como fotos de placas de veículos ou contêineres com números de série, bem como de documentos - faturas, contas, relatórios financeiros, artigos.

2.3 Cloudfmersive

A Cloudfmersive é uma empresa que desenvolve APIs e fornece como serviço para

seus clientes. A Cloudmersive traz aos seus clientes um portfólio completo de APIs em escaneamento de vírus, conversão e processamento de documentos, OCR utilizando *Deep Learning*, reconhecimento e processamento de imagens, processamento de linguagem natural, processamento de código de barras e quaisquer outras áreas-chave. A empresa entrega aos clientes esses serviços através de um modelo de planos e preços, oferecendo também suporte. A API de reconhecimento de texto utiliza *Deep Learning* para converter imagens, fotos de documentos e formulários, em texto. Isso é feito com suporte a mais de 90 idiomas.

2.4 Google cloud

A GoogleCloud é uma plataforma de computação oferecida pelo Google, compartilhando a mesma infraestrutura que a empresa utiliza para seus produtos oferecidos ao usuário comum, como o Google e o Youtube. A plataforma oferece diversos produtos em nuvem, dentre eles os que mais se destacam são os de processamento através de máquinas virtuais, armazenamento de objetos, serviço de banco de dados, reconhecimento de fala e transcrição e processamento visual (Vision AI). A API Vision do Google Cloud oferece modelos avançados de aprendizado de máquina pré-treinados por meio das APIs REST e RPC, fornecendo serviços de detecção de objetos, detecção de faces e reconhecimento de texto, dentre outros.

2.5 OCRSPACE

A OCRSpace fornece serviços de conversão de digitalizações ou imagens de documentos de texto em arquivos editáveis usando o reconhecimento ótico de caracteres. O serviço online é gratuito para usar, não tem necessidade de registro.

Entretanto, o serviço via API necessita de cadastro, e possui planos de uso, sendo um deles o plano gratuito, mas com limites de chamadas por mês.

A API OCR fornece uma maneira simples de analisar imagens e documentos PDF de várias páginas (PDF OCR) e obter os resultados de texto extraídos retornados em um formato JSON. A API OCR tem três camadas/níveis. O plano de OCR API gratuito tem um limite de taxa de 500 solicitações em um dia por endereço IP para evitar spam acidental.

3 | MATERIAIS E MÉTODOS

Nesta seção serão detalhadas as características da base de imagens “YUVA EB Dataset”, bem como o desenvolvimento do ambiente de testes usando as ferramentas de OCR.

3.1 Base de imagens “yuva eb dataset”

O conjunto de dados utilizado neste trabalho tem o nome “YUVA EB Dataset”⁷ o mesmo utilizado em [Kanagarathinam and Sekar 2019]. Esta base de dados consiste em

⁷ Disponível em https://drive.google.com/drive/folders/1J9TYUiLKdJKfSeotL-_ElyvXQ-3pE282

uma coleção de imagens de telas de medidores digitais de energia de sete segmentos, coletadas na região de Tamil Nadu na Índia. As imagens presentes nessa base de dados estão todas no formato JPG, também conhecido como JPEG, é um formato comum para fotografias digitais por armazenar imagem em quantidade menor de bytes.

As imagens estão divididas em subconjuntos de acordo com as condições de captura, diurna, noturna, inclinada e baixa resolução. A base de dados tem um total de 169 imagens, divididas da seguinte forma: base *Day Time* com 50 imagens capturadas à luz do dia, base *Night time* 49 noturnas, base *Tilted* com 50 imagens inclinadas, e base *Blurred* com 20 imagens desfocadas, alguns exemplares são mostrados na Figura 2, Figura 3, Figura 4 e Figura 5 respectivamente. As imagens noturnas, diurnas e inclinadas foram capturadas usando uma câmera digital no modo de alta resolução e as imagens de baixa resolução capturadas no modo de baixa resolução. Pelas imagens, é possível verificar que a cor do fundo varia entre várias cores: verde, laranja, lilás, azul, em todas as cores apresentam diferentes intensidades.



Figura 2. Exemplos de imagens capturadas à luz do dia.



Figura 3. Exemplos de imagens capturadas à noite.



Figura 4 Exemplos de imagens inclinadas



Figura 5. Exemplos de imagens de baixa resolução.

Todas as imagens dos medidores, além dos dígitos da medida, contêm outras informações/caracteres, tais como “kWh”, às vezes a palavra “CUM”, esses caracteres não são considerados para a avaliação de reconhecimento de caracteres. Todas as amostras dos medidores têm o ponto decimal antes do último dígito, e em muitas amostras este último dígito está em tamanho menor. Mesmo em tamanho menor, este dígito será considerado na avaliação do OCR. A quantidade de dígitos varia por imagem. Há uma predominância de telas que apresentam 5 dígitos sendo 74 amostras de imagens, e apenas 8 amostras de imagens com 8 dígitos. Não há imagens com menos de 4 dígitos, nem com mais de 8 dígitos.

3.2 Ambiente de experimentos

Para o processamento das imagens foi desenvolvido um projeto em .NET Core do tipo “Console Application”, o projeto se encontra no GitHub⁸. Na Figura 6 é possível ver o diagrama de classes da aplicação, mostrando as relações entre as classes. Basicamente a aplicação consiste em uma classe **Program** que chama cada serviço OCR.

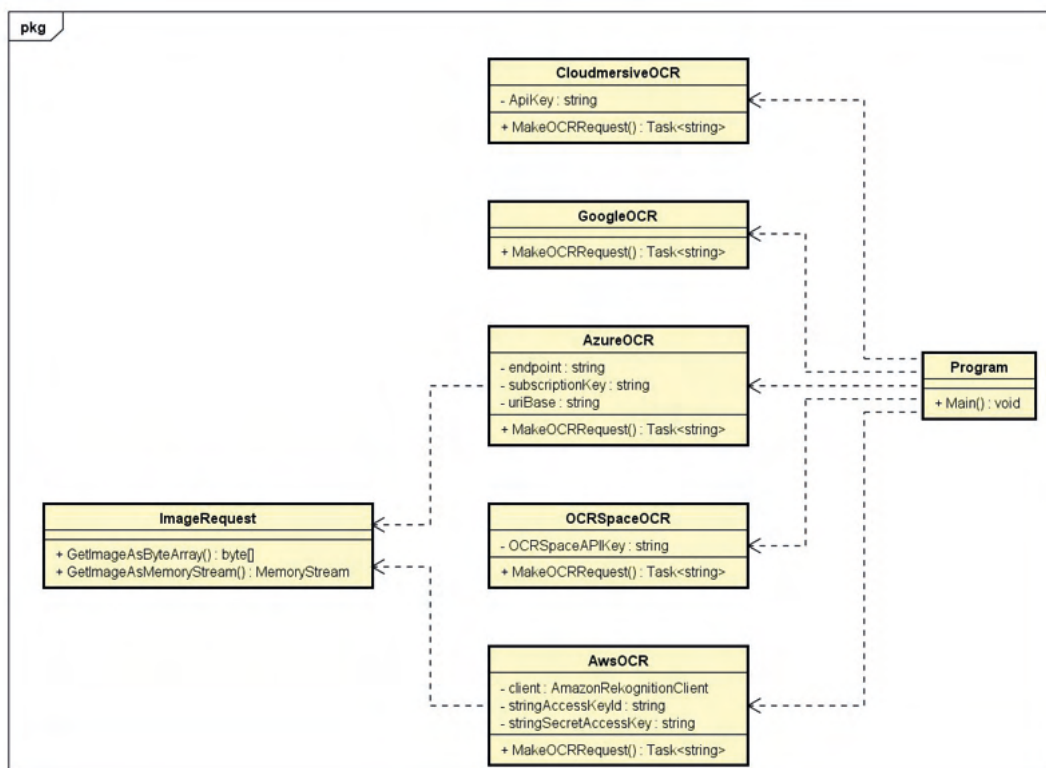


Figura 6. Diagrama de Classes da Aplicação de Testes.

⁸ <https://github.com/JonathanRibeiro92/OCRComparer.git>

Para cada arquivo de imagem processado é criado um caminho de arquivo de resultado por serviço consultado, isto é uma pasta específica para cada API seguindo o mesmo padrão de nomenclatura do caminho da base de dados, uma pasta para cada condição de captura de imagem. A aplicação realiza as chamadas às APIs que fornecem o serviço OCR e armazena os resultados em formato JSON, e depois armazena os arquivos de resultado utilizando os caminhos de pasta de resultados montados anteriormente.

Serviço OCR	Formatos aceitos	Suporte às tecnologias	Custo
AWS	Arquivos PDF, JPEG, PNG. Vídeos armazenados no próprio serviço	REST, Javascript, Python, PHP, .NET, Ruby, Java, Go, Node.js, C++	Nível Gratuito 12 meses 5.000 imagens por mês
Azure	PDF, TIFF, JPEG, PNG, BMP	REST, .NET, Python, Java, Node.js, Go	Gratuito até 5.000 transações por mês
Cloudmersive	PNG, JPEG	REST, .NET, Java, Node.js, Python, PHP, Objective-C, Ruby, Zapier	Gratuito até 800 transações por mês
Google Cloud	Documentos PDF e TIFF, Imagens codificadas em base64	REST, C#, Go, Java, Node.js, PHP, Python, Ruby	Gratuito para primeiras 1.000 imagens por mês
OCRSpace	PDF, JPG, GIF, PNG	Rest, C#, C++, Java, Javascript, Node.js, PHP, Python, Ruby, Swift, Objective-C	Gratuito para até 25.000 requisições no mês, com restrição de 500 por dia

Tabela 1. Relação dos Serviços OCR escolhidos

Para facilitar a comparação das cinco ferramentas selecionadas para este trabalho, um resumo das características é apresentado na Tabela 1. A primeira coluna da tabela tem o nome do serviço/ ferramenta, seguido pelos formatos aceitos, quais as tecnologias suportadas e o custo. Todas as ferramentas são gratuitas, mas cada uma apresenta uma limitação diferenciada, determinando restrições de uso.

4 | RESULTADOS E DISCUSSÃO

A análise dos experimentos será feita comparando os erros na identificação dos dígitos presentes na tela do medidor de energia entre cada serviço. Entende-se como erro qualquer dígito não identificado, ou identificado erroneamente. Por exemplo, para um *display* com o número “03525.9”, foram considerados corretos os resultados: “03525.9” ou “035259” ou “035259*” ou “035259kW” ou “035:259” ou “035 259” ou “03525 9kW”. Isto é, o ‘.’ (ponto decimal) pode ou não ser reconhecido, mas ainda se conta como correto; a inclusão de um ou mais caracteres (podendo ser um espaço em branco ou caracteres especiais) no meio ou após a sequência de dígitos, também foi considerado correto. Foi considerado incorreto “35259”, pois não reconheceu o ‘0’ (zero) à esquerda. Na base de imagens não está presente um gabarito para verificação dos dígitos de cada imagem, então para a comparação e verificação dos resultados foi feita de forma manual, comparando cada resultado com os dígitos presentes na imagem original. O desempenho geral das

ferramentas testadas é apresentado na Figura 7. A ordem de apresentação foi a mesma ordem dos experimentos realizados.

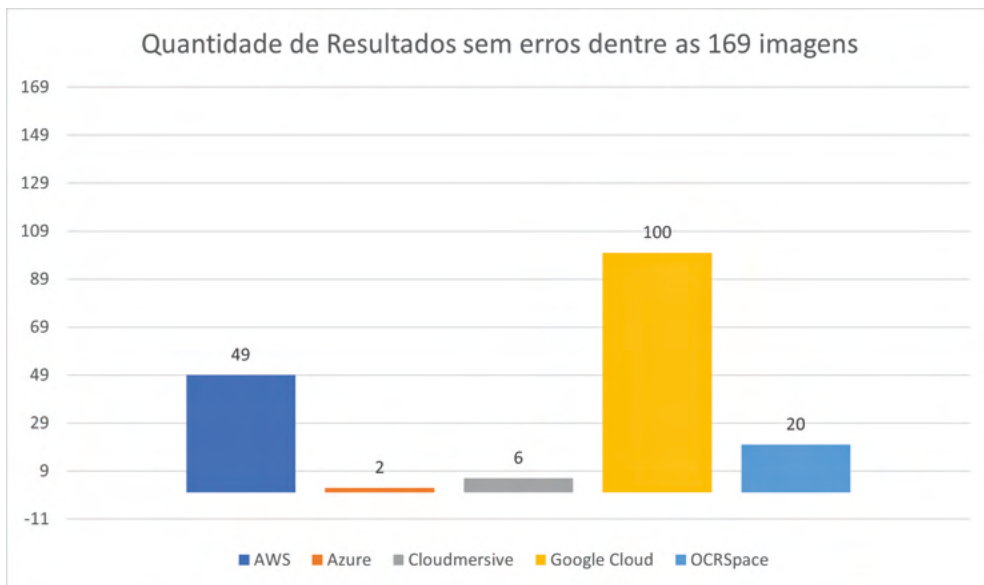


Figura 7. Gráfico de resultados sem erros por ferramenta.

Fazendo uma análise geral das ferramentas comparando os resultados sem erros, o melhor resultado foi do Google Cloud com 100 acertos dentre as 169 imagens da base de dados (59,17% de acerto), o segundo foi a AWS com 49 imagens que tiveram todos os dígitos identificados corretamente (28,99%), OCRSpace com 11,83% de acertos (20 das 169 imagens), o Cloudmersive com 6 reconhecimentos corretos (3,55% de acerto) e a AZURE com apenas 2 imagens (1,18% de acerto).

4.1 Comparação por dígito

Nesta seção a análise de erros será contada de forma diferenciada que foi na Figura 7. Nesta análise, serão contabilizados os acertos de cada reconhecimento, assim, se na seção anterior a população total era de 169 imagens, nesta seção, a população total é de 985 dígitos. Assim, para o número “035259”, se a resposta for “03525” ou “35259” ou “085259”, a contagem é de 1 (um) erro, em “03525” faltou o dígito ‘9’ no final, em “35259” faltou o ‘0’ no início, em “085259” houve a troca do ‘3’ pelo ‘8’, mas conta-se um acerto para os dígitos ‘0’, ‘2’ e ‘9’ e dois acertos para o dígito ‘5’. O resultado, tanto por dígito e por ferramenta de OCR, é apresentado na Tabela 2, valores maiores que 75% foram marcados em negrito.

De novo, Google Cloud teve um melhor desempenho apresentando uma quantidade

maior de acertos que as demais ferramentas, tanto no geral (acurácia de 86,5%, acertando 852 dígitos dos 985), quanto por dígito (o percentual de acerto foi maior que 75% em todos os dígitos). O Google Cloud erra mais o dígito '1' que os demais, pois muitas das vezes responde como o dígito '7', e acerta mais o dígito '5'. Em segundo lugar, segue o AWS (68,2% de taxa de acerto no geral) e em terceiro o OCRSpace (42,84%). O pior resultado foi da Azure que apresentou um número de acertos consideravelmente baixo (apenas 50 dos 985, taxa de acerto de 5,08%) seguido do Cloudmersive (234 dos 965, taxa de acerto de 23,76%).

Dígitos	Total Dígitos	Serviço									
		AWS		AZURE		Cloudmersive		Google Cloud		OCRSpace	
		Acertos	%acerto	Acertos	%acerto	Acertos	%acerto	Acertos	%acerto	Acertos	%acerto
0	243	134	55,14%	8	3,29%	52	21,40%	209	86,01%	81	33,33%
1	108	57	52,78%	8	7,41%	26	24,07%	83	76,85%	26	24,07%
2	101	85	84,16%	2	1,98%	30	29,70%	91	90,10%	49	48,51%
3	113	81	71,68%	5	4,42%	19	16,81%	103	91,15%	58	51,33%
4	85	70	82,35%	4	4,71%	18	21,18%	75	88,24%	39	45,88%
5	69	57	82,61%	10	14,49%	30	43,48%	63	91,30%	41	59,42%
6	53	29	54,72%	2	3,77%	18	33,96%	46	86,79%	30	56,60%
7	68	42	61,76%	1	1,47%	19	27,94%	54	79,41%	27	39,71%
8	62	47	75,81%	4	6,45%	15	24,19%	56	90,32%	23	37,10%
9	83	70	84,34%	6	7,23%	7	8,43%	72	86,75%	48	57,83%
Total	985	672	68,22%	50	5,08%	234	23,76%	852	86,50%	422	42,84%

Tabela 2. Performance das ferramentas em relação aos dígitos por serviço

4.2 Comparação por condição de captura

Esta análise é feita levando em consideração as condições em que a imagem foi capturada: Luz do Dia, Noturna, Inclinada ou Baixa Resolução. A Tabela 3 apresenta um comparativo do percentual de acertos de cada uma das ferramentas, valores acima dos 75% foram marcados em negrito.

Serviço	Luz do Dia (%)	Noturna (%)	Inclinada (%)	Baixa Resolução (%)
AWS	80,65	83,00	39,86	71,30
AZURE	5,38	5,67	4,12	5,22
Cloudmersive	29,39	27,33	18,90	13,04
Google Cloud	87,81	84,00	91,41	77,39
OCRSpace	43,01	51,33	42,96	20,00

Tabela 3. Percentual de acertos por tipo de condição na captura da imagem, considerando os 985 dígitos

Os dados confirmam que o Google Cloud é melhor em todas as condições, e apresentam os melhores resultados em imagens inclinadas. Na base Noturna, a diferença com relação ao AWS é bem pequena equivalente à uma única imagem. O AWS tem bons

resultados nas bases de Luz do Dia e Noturna, mas tem baixa taxa de acerto para a base Inclinada e na de Baixa Resolução. De uma forma diferente, o AWS reconhece mais corretamente as imagens noturnas do que as diurnas, supõe-se que essa ferramenta lide melhor com menos incidência de luz. Apesar da primeira impressão ser de que as imagens de Baixa Resolução teriam um resultado pior em todas as ferramentas, percebe-se que isso não é verdadeiro para o AWS.

5 | CONSIDERAÇÕES FINAIS

Foram abordadas as ferramentas da AWS Rekognition, Microsoft Azure, Cloudmersive, Google OCR e OCRSpace. O reconhecimento de caracteres em imagens é muito útil em diversos cenários, sendo uma delas a que foi abordado neste trabalho que é quando o parque instalado é tão grande que o custo de substituição por equipamentos mais modernos com interface de comunicação torna a troca de equipamentos praticamente inviável. A ferramenta que possui um desempenho melhor dentre as cinco avaliadas é a da Google Cloud, o Google OCR, com taxa de acerto geral de 86,5% considerando cada dígito ou 59,17% considerando a imagem como um todo. O pior resultado foi da ferramenta da AZURE com taxa de acerto de apenas 5,08% considerando cada dígito ou 1,18% considerando a imagem como um todo (acertou o *display* completo de apenas duas imagens da base de dados).

Quando os resultados dessas ferramentas são comparados com resultados de [Kanagarathinam and Sekar 2019] (93,17%) por dígito, em que há o desenvolvimento de um pré-processamento de imagens e de um sistema de aprendizado de máquina treinado na mesma base de dados, a diferença é muito discrepante, o mais próximo é o Google Cloud com 86,50% de acurácia. As ferramentas comerciais testadas aqui não tiveram como entrada imagens pré-processadas, nem foi realizado treinamento das ferramentas usando as imagens da base de dados.

Com isso, considera-se que utilizar as ferramentas de reconhecimento de imagens sem um pré-processamento pode não ser uma boa opção quando é necessário realizar análises mais assertivas como a captação de dados de medição. Uma abordagem em que o usuário possa fazer a delimitação de área, por exemplo, em que o usuário indicasse o local correto (por exemplo, movendo um retângulo) em um aplicativo de celular provavelmente melhoraria em muito os resultados dos OCRs analisados.

REFERÊNCIAS

ANDA, F., LILLIS, D., LE-KHAC, N., and SCANLON, M. (2018). **Evaluating automated facial age estimation techniques for digital forensics**. In Proceedings of 2018 IEEE Security and Privacy Workshops (SPW), pages 129–139. IEEE.

BONAČIĆ, I., HERMAN, T., KRZMAR, T., MANGIĆ, E., MOLNAR, G., and ČUPIĆ, M. (2015). **Optical character recognition of seven-segment display digits using neural networks**. In 32st International Convention on Information and Communication Technology, Electronics and Microelectronics, volume 3.

FINNEGAN, E., VILLARROEL, M., VELARDO, C., and TARASSENKO, L. (2019). **Automated method for detecting and reading seven-segment digits from images of blood glucose metres and blood pressure monitors**. Journal of Medical Engineering & Technology, 43(6):341–355.

KANAGARATHINAM, K. and SEKAR, K. (2019). **Text detection and recognition in raw image dataset of seven segment digital energy meter display**. Energy Reports, 5:842–852.

TORRES, W., VAN DEN BRAND, M. G., and SEREBRENIK, A. (2020). **Suitability of optical character recognition (ocr) for multi-domain model management**. In International Conference on Systems Modelling and Management, pages 149–162. Springer.

SOBRE A ORGANIZADORA

MA. LILIAN COELHO DE FREITAS- Professora do Instituto Federal de Educação, Ciência e Tecnologia do Pará (IFPA). Possui graduação em Engenharia da Computação pela Universidade Federal do Pará (2007) e mestrado em Computação Aplicada pelo Programa de Pós-Graduação em Engenharia Elétrica (PPGEE) da UFPA, obtido em 2009. Em 2008, realizou estágio de mestrado no Instituto de Engenharia de Sistemas e Computadores (INESC/Porto - Portugal). Atuou como pesquisadora membro do Laboratório de Eletromagnetismo Aplicado (LEA/UFPA) de 2004 a 2014 e do Laboratório de Sensores e Sistemas Embarcados (LASSE/UFPA) de 2008 a 2012. Atuou como Pesquisadora Visitante no *Georgia Institute of Technology* (Atlanta, Georgia, Estados Unidos), no período de Março/2012 a Fev/2013. Tem diversos trabalhos publicados, envolvendo publicações em livros, revistas e eventos científicos. Seus interesses de pesquisa são: telecomunicações (comunicações sem fio, rádio cognitivo) e *machine learning* e educação.

ÍNDICE REMISSIVO

A

Aplicativo 23, 25, 26, 28, 29, 30, 32, 33, 82, 91
Arduino 23, 25, 26, 27, 28, 32, 34, 69, 70, 71, 72, 77
Azure cognitive services 80, 81, 84

C

Cambio climático 1, 2
Cereal básico 1, 2, 3, 4
Cloudmersive 80, 81, 82, 84, 85, 88, 89, 90, 91
Comunicação sem fio 23, 24, 25
Controle Adaptativo 3, 57, 58
Controle de temperatura 3, 23, 24, 25, 30, 31, 32, 33

D

Daño a cultivo 1

E

Energia Solar Fotovoltaica 10

F

Fontes Alternativas de Energia 10
Fundos de investimento 35, 36, 37, 38, 39, 48, 54, 55

G

Google Cloud 45, 80, 81, 83, 85, 88, 89, 90, 91

I

Identificação de sistemas 57, 68
Inovação 23, 25, 32, 33

M

Modelo Computacional 10, 11, 13, 14, 16, 19, 20

O

OCR 82, 92
OCRSpace 80, 81, 82, 85, 88, 89, 90, 91

P

Perfil de investidor 35, 40, 45, 46

Placa de desenvolvimento 23, 26

Processamento de imagens 80, 82, 85, 91

R

Raspberry-Pi 4, 69, 70, 72, 74, 77, 78

Reconhecimento de texto 80, 83, 85

Redes Neurais Artificiais 2, 10, 11, 14, 15, 16, 17, 19, 20, 21

Rehabilitación de muñeca 4, 69, 70, 77

Rekognition 80, 81, 82, 84, 91

Robótica 57, 84


S

Sistema eletrônico agnóstico 35, 54

COLEÇÃO

DESAFIOS DAS ENGENHARIAS:





ENGENHARIA DE COMPUTAÇÃO 4

-  www.atenaeditora.com.br
-  contato@atenaeditora.com.br
-  [@atenaeditora](https://www.instagram.com/atenaeditora)
-  www.facebook.com/atenaeditora.com.br

COLEÇÃO

DESAFIOS DAS ENGENHARIAS:

ENGENHARIA DE COMPUTAÇÃO 4

-  www.atenaeditora.com.br
-  contato@atenaeditora.com.br
-  [@atenaeditora](https://www.instagram.com/atenaeditora)
-  www.facebook.com/atenaeditora.com.br