



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
CENTRO DE ENGENHARIA ELÉTRICA E  
INFORMÁTICA  
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**LEANDRO DE SOUTO LIMA**

**Reconhecimento de Datas Manuscritas em Prescrições Médicas:  
Comparativo entre OCRs com Licença de Software Livre e uma  
Solução Comercial**

**CAMPINA GRANDE - PB**

**2024**

**LEANDRO DE SOUTO LIMA**

**Reconhecimento de Datas Manuscritas em Prescrições Médicas:  
Comparativo entre OCRs com Licença de Software Livre e uma  
Solução Comercial**

**Trabalho de Conclusão Curso  
apresentado ao Curso Bacharelado em  
Ciência da Computação do Centro de  
Engenharia Elétrica e Informática da  
Universidade Federal de Campina  
Grande, como requisito parcial para  
obtenção do título de Bacharela em  
Ciência da Computação.**

**Orientador:  
Herman Martins Gomes**

**CAMPINA GRANDE - PB**

**2024**

**LEANDRO DE SOUTO LIMA**

**Reconhecimento de Datas Manuscritas em Prescrições Médicas:  
Comparativo entre OCRs com Licença de Software Livre e uma  
Solução Comercial**

**Trabalho de Conclusão Curso apresentado  
ao Curso Bacharelado em Ciência da  
Computação do Centro de Engenharia  
Elétrica e Informática da Universidade  
Federal de Campina Grande, como  
requisito parcial para obtenção do título  
de Bacharel em Ciência da Computação.**

**BANCA EXAMINADORA:**

**Herman Martins Gomes**

**Orientador – UASC/CEEI/UFCG**

**Eanes Torres Pereira**

**Examinador – UASC/CEEI/UFCG**

**Francisco Vilar Brasileiro**

**Professor da Disciplina TCC – UASC/CEEI/UFCG**

**Trabalho aprovado em: 15 de Maio de 2024.**

**CAMPINA GRANDE - PB**

**2024**

# **Handwritten Date Recognition in Medical Prescriptions: A Comparison between Open Source OCRs and a Commercial Solution**

## **ABSTRACT**

Data recognition in medical prescriptions is a critical task to ensure patient safety and the efficiency of medical processes. This study proposes a comparative analysis between optical character recognition (OCR) models and handwritten text recognition models to identify and extract data from medical prescriptions. Accurate date recognition is essential to avoid medication errors and ensure proper administration of treatments. We evaluate the models' accuracy, speed, and robustness across different scenarios and handwriting types, considering common variations found in medical prescriptions. The results show that OCR models present advantages in terms of speed and accuracy in printed texts, while handwritten text recognition models stand out in the interpretation of varied handwriting. These findings provide valuable information for developing more effective date recognition systems in clinical settings.

## ABSTRACT

Data recognition in medical prescriptions is a critical task to ensure patient safety and the efficiency of medical processes. This study proposes a comparative analysis between optical character recognition (OCR) models and handwritten text recognition models to identify and extract data from medical prescriptions. Accurate date recognition is essential to avoid medication errors and ensure proper administration of treatments. We evaluate the models' accuracy, speed, and robustness across different scenarios and handwriting types, considering common variations found in medical prescriptions. The results show that OCR models present advantages in terms of speed and accuracy in printed texts, while handwritten text recognition models stand out in the interpretation of varied handwriting. These findings provide valuable information for developing more effective date recognition systems in clinical settings.

## RESUMO

O reconhecimento de datas em prescrições médicas é uma tarefa crítica para garantir a segurança do paciente e a eficiência dos processos médicos. Este estudo propõe uma análise comparativa entre modelos de reconhecimento óptico de caracteres (OCR) e modelos de reconhecimento de texto manuscritos para identificar e extrair datas de prescrições médicas. O reconhecimento preciso de datas é essencial para evitar erros de medicação e garantir a administração adequada de tratamentos. Avaliamos a precisão, a velocidade e a robustez dos modelos em diferentes cenários e tipos de caligrafia, considerando as variações comuns encontradas em prescrições médicas. Os resultados mostram que os modelos de OCR apresentam vantagens em termos de velocidade e precisão em textos impressos, enquanto os modelos de reconhecimento de texto manuscritos destacam-se na interpretação de caligrafias variadas. Essas descobertas fornecem insights valiosos para o desenvolvimento de sistemas de reconhecimento de datas mais eficazes em ambientes clínicos.

## Palavras chave

Reconhecimento de texto manuscrito, prescrições médicas, aprendizado de máquina, Tesseract OCR, EasyOCR, Google Cloud Vision API, OCR.

## 1. INTRODUÇÃO

As prescrições médicas são documentos essenciais na prática clínica, fornecendo orientações detalhadas sobre tratamentos e medicamentos prescritos para os pacientes. No entanto, a interpretação manual desses documentos pode ser propensa a erros, especialmente no que diz respeito à identificação e extração precisa de datas. Erros de interpretação de datas podem resultar em administração inadequada de medicamentos, comprometendo a segurança do paciente e aumentando os riscos associados ao tratamento médico. Portanto, o desenvolvimento de sistemas automatizados para o reconhecimento preciso de datas em prescrições médicas torna-se fundamental para melhorar a qualidade dos cuidados de saúde.

Atualmente, existem dois principais enfoques para o reconhecimento de datas em prescrições médicas: o reconhecimento óptico de caracteres (OCR) e o reconhecimento de texto manuscrito. O OCR é amplamente utilizado para digitalizar e extrair informações de textos impressos, enquanto o

reconhecimento de texto manuscrito lida especificamente com a interpretação de escrita à mão. Ambos os métodos têm suas vantagens e limitações, e a escolha entre eles depende das características específicas das prescrições médicas e das demandas do ambiente clínico. O escopo desta pesquisa envolve testar a adequabilidade ou não de OCRs com licença de software livre treinados com texto impresso na tarefa de reconhecer datas manuscritas. Além disso, comparamos as soluções anteriores com uma solução comercial que, conforme sua documentação, é capaz de reconhecer tanto texto impresso como manuscrito.

Neste estudo, focaremos especificamente na avaliação de três ferramentas amplamente utilizadas para o reconhecimento óptico de caracteres: Tesseract OCR<sup>1</sup> (uma opção gratuita), EasyOCR<sup>2</sup> (também gratuita) e o módulo de OCR da Google Cloud Vision API<sup>3</sup> (uma solução comercial que vai muito além do OCR). Investigaremos como essas ferramentas se comparam em termos de precisão, velocidade e robustez na identificação e extração de datas manuscritas de prescrições médicas. Compreender a eficácia relativa dessas ferramentas é importante para orientar a seleção e implementação de sistemas de reconhecimento de datas mais confiáveis e precisos, contribuindo assim para a melhoria dos processos de interpretação de prescrições médicas e, conseqüentemente, para a qualidade dos cuidados de saúde fornecidos aos pacientes.

## 2. OBJETIVO

O objetivo primário deste estudo é avaliar a precisão dos modelos de reconhecimento óptico de caracteres (OCR) na identificação e extração de datas manuscritas de prescrições médicas. Nosso foco será determinar quais modelos e técnicas demonstram maior precisão no reconhecimento das datas, visando ajudar na escolha de tecnologias adequadas para o reconhecimento do conteúdo de documentos manuscritos.

## 3. METODOLOGIA

Para avaliar a precisão dos modelos de reconhecimento de datas em prescrições médicas, utilizaremos a métrica de Levenshtein<sup>4</sup>. A métrica de Levenshtein calcula a distância entre duas strings, representando o número mínimo de edições necessárias para transformar uma string na outra. Neste contexto, consideramos a distância de Levenshtein entre as datas reconhecidas pelos modelos e as datas reais presentes nas prescrições médicas.

Recortes de datas em prescrições médicas foram selecionados para análise. Esses recortes representam uma amostra diversificada de datas encontradas em diferentes contextos e estilos de escrita, abrangendo uma variedade de formatos de data comumente encontrados em prescrições médicas. Essa abordagem visa garantir que os modelos sejam testados em uma ampla gama de cenários realistas, permitindo uma avaliação abrangente de sua capacidade de reconhecimento de datas.

<sup>1</sup> Tesseract OCR: <https://github.com/tesseract-ocr/tesseract>

<sup>2</sup> EasyOCR: <https://github.com/JaidedAI/EasyOCR>

<sup>3</sup> Google Cloud Vision API: <https://cloud.google.com/vision>

<sup>4</sup> Levenshtein: [https://pt.wikipedia.org/wiki/Dist%C3%A2ncia\\_Levenshtein](https://pt.wikipedia.org/wiki/Dist%C3%A2ncia_Levenshtein)

Definição da métrica de precisão: Utilizaremos a métrica de precisão baseada no algoritmo de Levenshtein para determinar o quão próximo o resultado do reconhecimento de datas está em relação à data real. Por exemplo, consideramos uma data reconhecida como correta se a distância de Levenshtein entre ela e a data real for menor ou igual a um determinado limiar, como 1, 0.8 ou 0.5.

Avaliação dos modelos: Aplicamos a métrica de precisão com diferentes limiares de distância de Levenshtein aos resultados dos modelos de reconhecimento de texto manuscrito e de OCR. Isso nos permitirá comparar a capacidade de cada modelo em identificar e extrair datas de prescrições médicas com diferentes níveis de tolerância a erros.

Ao seguir essa metodologia centrada na métrica de Levenshtein, seremos capazes de quantificar a precisão dos modelos de reconhecimento de datas em prescrições médicas.

#### 4. PRODUÇÃO E ANÁLISE DE DADOS

Para a produção e análise de dados neste estudo, selecionamos três ferramentas amplamente utilizadas para o reconhecimento óptico de caracteres (OCR): Tesseract OCR, EasyOCR e Google Cloud Vision API. A escolha dessas ferramentas foi baseada em critérios como popularidade, disponibilidade de recursos e desempenho comprovado em tarefas semelhantes. O Tesseract OCR é uma opção gratuita e de código aberto que tem sido amplamente adotada pela comunidade de pesquisa devido à sua eficácia e flexibilidade. EasyOCR é outra escolha gratuita que oferece uma interface fácil de usar e suporte para vários idiomas, tornando-o uma opção atraente para experimentos de reconhecimento de texto manuscrito. Por fim, a Google Cloud Vision API é uma solução paga que oferece recursos avançados de reconhecimento de texto e imagem, incluindo suporte para texto manuscrito e uma infraestrutura escalável. Nas próximas subseções, executaremos os modelos e compararemos os resultados entre modelos.

Além disso, cabe ressaltar que, para este estudo, foram garimpadas algumas dezenas de imagens específicas contendo datas para análise, totalizando um conjunto específico para esta finalidade. Essas imagens foram obtidas na internet por meio de pesquisa de imagens e censuradas para não incluir nenhum dado do usuário referido, pois o intuito é apenas recortar as datas para análise. Como exemplificado nas Figuras 1, 2 e 3.

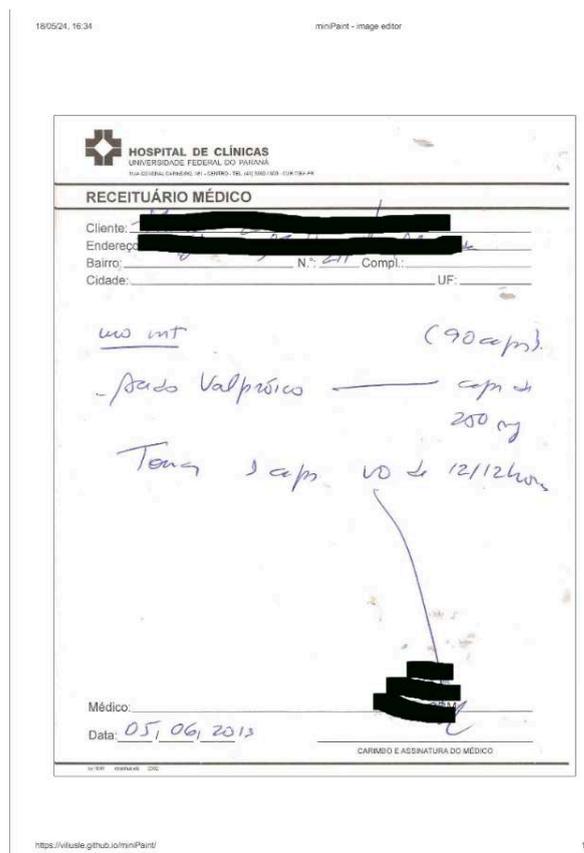


Fig. 1 Exemplo de imagem de receita médico

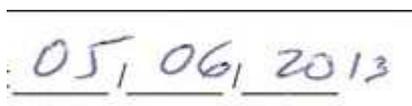


Fig. 2 Exemplo de imagem após o corte



Fig. 3 exemplo de amostragens de receita médico utilizadas após o corte

#### 4.1 TESSERACT OCR

O Tesseract OCR é uma biblioteca de reconhecimento óptico de caracteres (OCR, do inglês Optical Character Recognition) de código aberto, amplamente utilizada para converter imagens contendo texto em texto editável. Nesta seção, descrevemos nossa abordagem para utilizar o Tesseract OCR e discutimos os desafios encontrados durante esse processo.

##### Análise Geral

Inicialmente, aplicamos o Tesseract OCR para analisar as datas

presentes nas imagens de amostragem. No entanto, nos deparamos com uma série de desafios que impactaram a precisão e eficácia do OCR na interpretação desses documentos.

#### Desafios Encontrados

1. Variedade de Caligrafias e Estilos de Escrita: A diversidade na forma de escrever datas, tornou-se um desafio significativo para o Tesseract OCR. A falta de padronização na escrita manual dificultou a correta identificação e interpretação das informações.

#### Abordagem Alternativa

Devido à insatisfatória acurácia na interpretação de datas, optamos por explorar uma abordagem alternativa, concentrando-nos na análise específica de números. No entanto, mesmo com essa mudança de foco, os resultados finais ainda não alcançaram o nível de precisão desejado.

#### Exemplos de Desafios Encontrados

1. Formatos de Datas: O reconhecimento de números em diferentes formatos de datas (por exemplo, dia/mês/ano) continuou apresentando desafios significativos, contribuindo para imprecisões na interpretação das informações.
2. Complexidade na Identificação de Números de Datas: A complexidade em identificar números de datas reside na diversidade de formatos utilizados, como datas expressas no formato dia/mês/ano, mês/dia/ano, ou ainda em formatos abreviados ou com variações regionais. Isso demanda uma abordagem flexível e adaptável do sistema de OCR para reconhecer e interpretar corretamente essas diferentes representações numéricas de datas.

Comparando os resultados das duas configurações, notou-se que não houve muito ganho, inclusive na amostragem testada. A primeira abordagem se saiu melhor, em média, na interpretação geral.

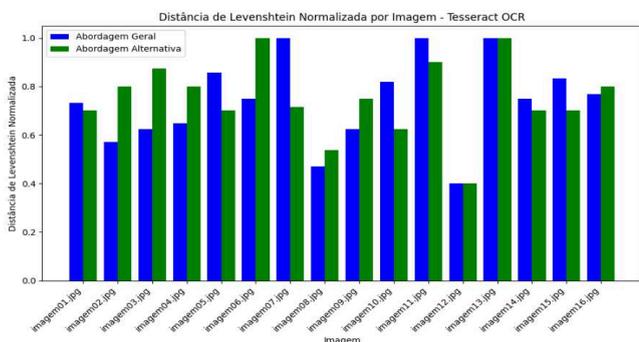


Fig. 4 Distância de Levenshtein Normalizada por Imagem - Tesseract OCR

Média de acertos da Abordagem Geral: 25.94%  
Média de acertos da Abordagem Alternativa: 24.98%  
A abordagem com melhor média de acertos é: Tesseract OCR Geral

Fig. 5 Porcentagem de acerto para amostragem.

## 4.2 EASYOCR

O EasyOCR é uma biblioteca de reconhecimento óptico de caracteres (OCR) que oferece uma abordagem alternativa para interpretar texto em imagens. Ao enfrentarmos dificuldades com a precisão e eficácia de outras bibliotecas OCR na interpretação de datas retiradas prescrições médicas, decidimos explorar o EasyOCR como uma possível solução.

#### Análise Geral

Inicialmente, utilizamos o modelo EasyOCR ao aplicar o algoritmo de Levenshtein para identificar precisamente as datas nas prescrições médicas. Ao concentrar nossa análise exclusivamente nas datas, nosso objetivo é aprimorar a capacidade do EasyOCR em identificar e interpretar corretamente essas informações cruciais. Ao focarmos especificamente nas datas, esperamos otimizar o desempenho do EasyOCR. Esse ajuste permitirá uma avaliação mais precisa da eficácia do EasyOCR na identificação de datas em prescrições médicas.

#### Desafios Encontrados

1. Números: Ainda existe uma confusão entre números e letras, o que resulta em dificuldades na identificação precisa dos dados.
2. Datas: Observamos uma melhora em relação à identificação de datas, pois agora conseguimos identificar caracteres especiais, porém ainda não alcançamos um resultado do final satisfatório.

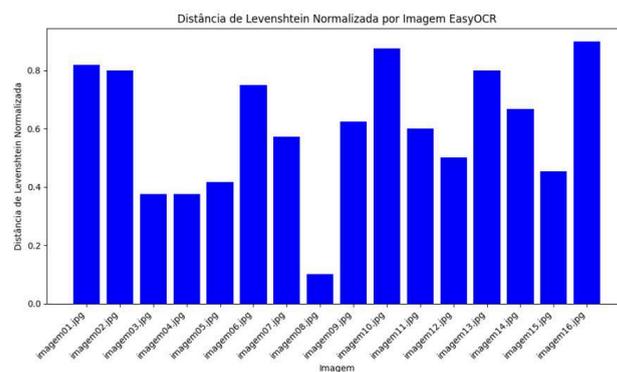


Fig. 6 Distância de Levenshtein Normalizada por Imagem - EasyOCR

**Média das Porcentagens de Acertos: 39.83%**

Fig. 7 Porcentagem de acerto EasyOCR

A abordagem foi mais promissora do que o Tesseract OCR, embora ainda não tenhamos alcançado uma porcentagem ideal de acurácia.

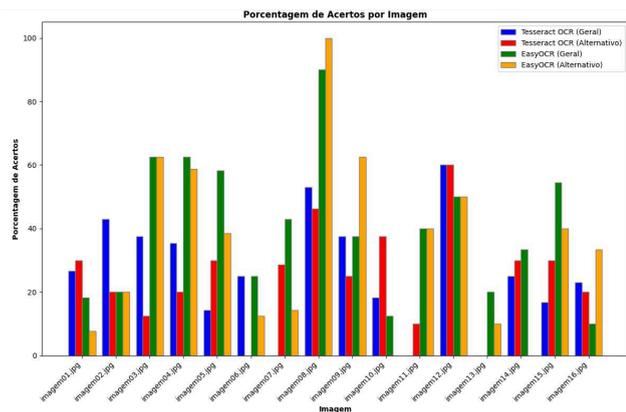


Fig. 8 Gráfico com a porcentagem de todos modelos.

Tesseract OCR - Média Abordagem Geral: 25.94%

Tesseract OCR - Média Abordagem Alternativa: 24.98%

EasyOCR : 39.83%

Fig. 9 Porcentagem de acerto para todos modelos.

### 4.3 GOOGLE CLOUD VISION API

O Google Cloud Vision API é uma ferramenta poderosa para análise de imagens por meio de algoritmos de aprendizado de máquina. Neste trabalho, o Cloud Vision API é avaliado no reconhecimento das datas de prescrições médicas. Discutiremos os desafios encontrados durante esse processo e as estratégias adotadas para lidar com eles.

Análise Geral:

Inicialmente, aplicamos o Google Cloud Vision API para analisar as datas presentes nas prescrições médicas. No entanto, nos deparamos com uma série de desafios que impactaram a precisão e eficácia do API na interpretação desses documentos, conforme discutido a seguir:

1. Variedade de Caligrafias e Estilos de Escrita: a diversidade na forma de escrever datas, tornou-se um desafio significativo para o Google Cloud Vision API. A falta de padronização na escrita manual dificultou a correta identificação e interpretação das informações.
2. Ajustes na Configuração do Google Cloud Vision API: após tentativas de ajustes e adaptações na configuração do API para melhor lidar com o idioma e os padrões de datas, ainda enfrentamos dificuldades significativas na interpretação.

Apesar dos desafios enfrentados ao utilizar a Google Cloud Vision API, notamos que a ferramenta se destacou pela sua precisão na identificação de datas em comparação com as outras duas soluções investigadas, apresentando uma acurácia na maioria dos casos acima de 60%.

A média geral de acurácia obtida com a Google Cloud Vision API foi de aproximadamente 70.98%, evidenciando sua eficácia em lidar com esse tipo específico de tarefa.

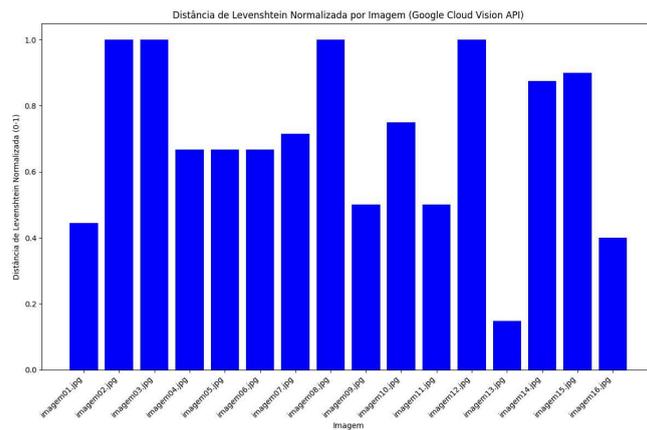


Figura 10: Distância de Levenshtein Normalizada por Imagem - Google Cloud Vision Api

Média Geral (Google Cloud Vision API): 70.98048941798942

Figura 11: média geral da análise da amostragem

Com algumas configurações adicionais, tais como a distinção entre vírgulas e traços em datas, a diferenciação entre letras como "l" e "i", e a identificação correta de barras como separadores de datas, acreditamos que seja possível aumentar ainda mais a acurácia da ferramenta, consolidando-a como uma escolha sólida para aplicações que exigem precisão na interpretação de datas em imagens, como a análise de prescrições médicas.

Com base nas médias de acurácia obtidas para cada ferramenta e abordagem:

A ferramenta Tesseract OCR apresentou uma média de acurácia de 25.94% na abordagem geral e 24.98% na abordagem alternativa.

O EasyOCR, por sua vez, registrou uma média de acurácia de 34.38% na abordagem geral.

Já a Google Cloud Vision API demonstrou uma significativa superioridade com uma média de acurácia de 70.98%.

Esses resultados ressaltam a eficácia da Google Cloud Vision API em comparação com as outras ferramentas, mostrando-se muito mais precisa na identificação de datas. Enquanto o EasyOCR mostrou um desempenho superior ao Tesseract OCR, em ambas as abordagens, sugerindo ser uma opção mais viável quando comparada ao Tesseract. Estes resultados evidenciam que modelos treinados com caracteres impressos não conseguem generalizar bem em tarefas envolvendo texto manuscrito. Como exemplificado nas Figuras 12.

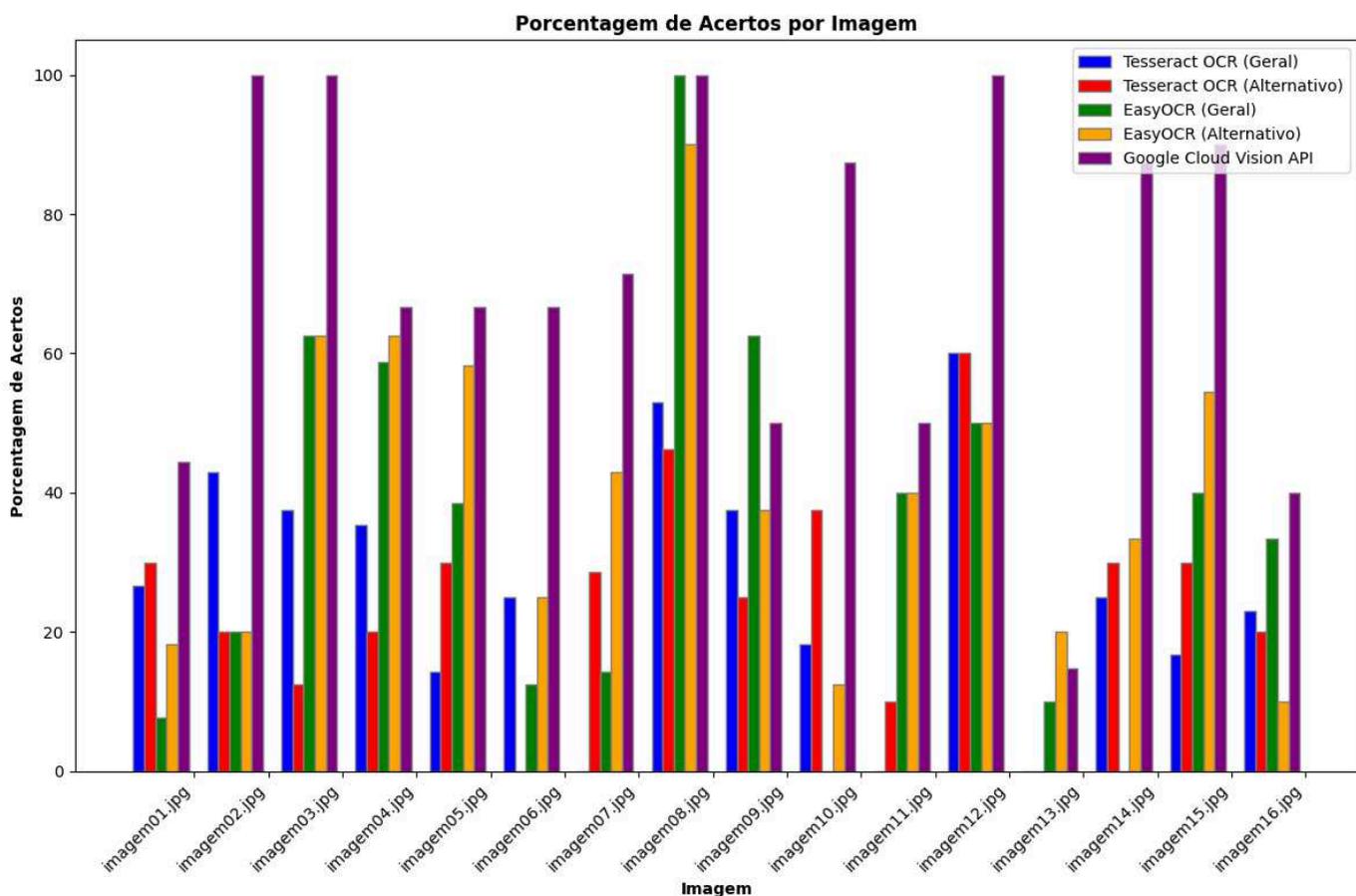


Figura 12: Gráfico com porcentagem de todos os modelos

Tesseract OCR - Média Abordagem Geral: 25.94%

Tesseract OCR - Média Abordagem Alternativa: 24.98%

EasyOCR : 39.83%

Google Cloud Vision API - Média : 70.98%

Figura 13: Média de acerto de todos os modelos

## 5. CONCLUSÕES

Neste estudo, exploramos a eficácia de três ferramentas de reconhecimento de escrita - o reconhecimento óptico de caracteres: Tesseract OCR (uma opção gratuita), EasyOCR (também gratuita) e a Google Cloud Vision API.

O Tesseract OCR, apesar de ser amplamente utilizado e de código aberto, enfrentou dificuldades na interpretação precisa das datas. Entre as três ferramentas analisadas, foi a que teve o pior desempenho, com uma porcentagem de acerto de Tesseract OCR - Média Abordagem Geral: 25.94%, Tesseract OCR - Média Abordagem Alternativa: 24.98%.

O EasyOCR, embora tenha apresentado alguma melhoria na identificação de datas após ajustes, também enfrentou desafios

semelhantes aos do Tesseract OCR. A confusão entre números e letras e a falta de opções explícitas para configurar o reconhecimento exclusivo de números limitaram sua eficácia, resultando em um desempenho de EasyOCR: 39.83%.

Por outro lado, a Google Cloud Vision API se destacou com uma média de acurácia de 70.98%, evidenciando sua superioridade na interpretação de datas retiradas de prescrições médicas. Apesar desses resultados promissores, reconhecemos a necessidade de um estudo mais aprofundado para explorar todo o potencial da API.

Além disso, observamos diferenças significativas entre ferramentas gratuitas e pagas, destacando a vantagem das soluções pagas em termos de desempenho e precisão. Enquanto na ferramenta da Google, pequenos ajustes já proporcionam resultados satisfatórios, nas ferramentas gratuitas como o Tesseract OCR e o EasyOCR, seria necessária uma tratativa muito mais intensa para alcançar resultados consideráveis.

Podemos concluir que não temos um bom uso de modelos de reconhecimento óptico de caracteres para reconhecer manuscritos, sendo o ideal o uso de ferramentas já treinadas para isso. Embora mais trabalhosas e lentas, elas oferecem uma segurança ao resultado final.

## 6. REFERÊNCIAS

1. Tesseract OCR. (s.d.). Recuperado de <https://tesseract-ocr.github.io/>. Acesso em: 01 de Maio de 2024.

2. Tesseract OCR Team. (2018). Tesseract OCR: Um sistema de reconhecimento óptico de caracteres de código aberto (Versão 4.0). [Software]. Disponível em: <https://github.com/tesseract-ocr/tesseract>. Acesso em: 02 de Maio de 2024.
3. Jaided AI. (s.d.). EasyOCR. Recuperado de <https://www.jaided.ai/easyocr/> Acesso em: 04 de Maio de 2024. Acesso em: 04 de Maio de 2024.
4. Google Cloud. Documentação oficial do Google Cloud Vision API em português. Disponível em: <https://cloud.google.com/vision/docs> Acesso em: 03 de Maio de 2024.
5. Google Cloud. Artigo "Como utilizar o Google Cloud Vision API em aplicações Python". Disponível em: <https://www.alura.com.br/artigos/como-utilizar-o-google-cloud-vision-api-em-aplicacoes-python> Acesso em: 04 de Maio de 2024.
6. Gonçalves, T., & Silva, A. (2023). Reconhecimento de texto em documentos médicos manuscritos: desafios e perspectivas. Revista Brasileira de Tecnologia Médica, 12(3), 45-62. [Disponível em: <https://teses.usp.br/teses/disponiveis/3/3141/tde-05032018-151842/publico/RodolfoValienteRomeroCorr18.pdf>]. Acesso em: 05 de Maio de 2024.
7. Wikipedia. Distância de Levenshtein. Disponível em: [https://pt.wikipedia.org/wiki/Dist%C3%A2ncia\\_Levenshtein](https://pt.wikipedia.org/wiki/Dist%C3%A2ncia_Levenshtein). Acesso em: 06 de Maio de 2024.