

MEC-SETEC
INSTITUTO FEDERAL MINAS GERAIS - *Campus* Formiga
Curso de Ciência da Computação

**ESTUDO COMPARATIVO SOBRE
FERRAMENTAS DE RECONHECIMENTO
ÓPTICO DE CARACTERES**

Renata Caroline Cunha

Orientador: Prof. Dr. Manoel Pereira
Junior

Formiga - MG

2018

RENATA CAROLINE CUNHA

**ESTUDO COMPARATIVO SOBRE
FERRAMENTAS DE RECONHECIMENTO
ÓPTICO DE CARACTERES**

Monografia do trabalho de conclusão de curso apresentado ao Instituto Federal Minas Gerais - Campus Formiga, como requisito parcial para a obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Manoel Pereira Junior

Formiga - MG

2018

004

Cunha, Renata Caroline.

Estudo comparativo sobre ferramentas de reconhecimento óptico de
Carateres / Renata Caroline Cunha.. -- Formiga : IFMG, 2018.
124p. : il.

Orientador: Prof. Dr. Manoel Pereira Júnior

Trabalho de Conclusão de Curso – Instituto Federal de Educação,
Ciência e Tecnologia de Minas Gerais – *Campus* Formiga.

1. Reconhecimento Óptico de Caracteres. 2. Ferramentas OCR.
3. Comparação. 4. Tesseract. 5. OCR Web Service. I. Título.

CDD 004

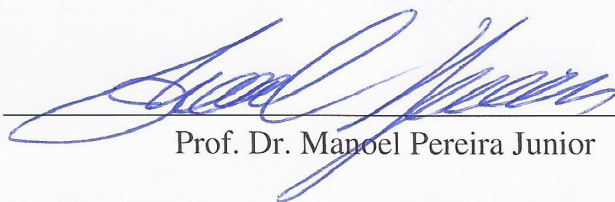
RENATA CAROLINE CUNHA

**ESTUDO COMPARATIVO SOBRE FERRAMENTAS DE
RECONHECIMENTO ÓPTICO DE CARACTERES**

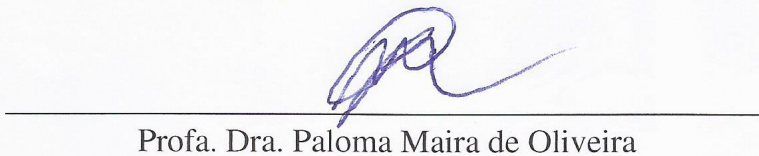
Trabalho de Conclusão de Curso apresentado ao
Instituto Federal de Minas Gerais-Campus Formiga,
como Requisito parcial para obtenção do título de
Bacharel em Ciência da Computação.

Aprovado em 21 de Novembro de 2018.

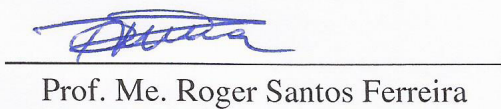
BANCA EXAMINADORA



Prof. Dr. Manoel Pereira Junior



Profa. Dra. Paloma Maira de Oliveira



Prof. Me. Roger Santos Ferreira

Agradecimentos

Agradeço primeiramente a Deus por me ter dotado de saúde, sabedoria, força e disposição para superar todas as dificuldades e alcançar mais uma etapa em minha vida; e a Nossa Senhora Aparecida, por todas as graças e intercessões concedidas a mim.

Aos meus pais, Moacir e Lediane, que nunca mediram esforços e sempre fizeram de tudo para que fosse possível alcançar os meus sonhos. Obrigada pelo amor incondicional, por todos os sacrifícios e pelo exemplo de vida que vocês são para mim.

Agradeço a toda minha família que sempre torceu muito por mim, aos meus avós, irmãos, padrinhos, tios e primos, em especial ao Leo, que desde criança sempre esteve ao meu lado em todos os momentos, me dando forças e apoio em todas minhas decisões.

Agradeço todos os meus amigos que sempre me apoiaram nessa jornada, em especial a Marcela, que esteve comigo em todos os momentos desde bem antes da faculdade, sempre me apoiando, me ouvindo, aconselhando e me ajudando, principalmente nos momentos mais difíceis, obrigada por tudo. Agradeço também de forma especial o Guilherme, por ter compartilhado comigo todos os momentos nesses últimos anos, por sempre estar disposto a me ajudar em tudo e por sempre acreditar no meu potencial.

Agradeço ao André por todo companheirismo, amor, carinho, incentivo e força que tem me dado ao longo de todo este ano, você também foi fundamental nesta conquista.

A todos professores do IFMG que contribuíram não só para o meu conhecimento acadêmico, mas também para meu caráter e formação profissional. Agradeço em especial ao meu orientador Manoel, por toda paciência, compreensão, pelas dicas, conselhos e por estar sempre pronto para ajudar durante todo o desenvolvimento do projeto. Agradeço também o Roger, por todo conhecimento transmitido a mim, sempre com muita paciência e disposição para ajudar. Agradeço ainda a Prof.^a Paloma que aceitou com disposição compor a banca e contribuir com este trabalho, juntamente com o Prof. Roger.

Por fim, agradeço de forma geral a todos que sempre torceram pelo meu sucesso e contribuíram para essa conquista, muito obrigada.

"Consagre ao Senhor tudo o que você faz, e os seus planos serão bem-sucedidos."

Provérbios 16:3

Resumo

Atualmente, há um grande volume de informação disponível digitalmente e a maior parte dessa informação é composta por imagens de documentos digitalizados. Devido a essa grande quantidade, se faz necessário que existam métodos para possibilitar a pesquisa nesses documentos, funções essas que são providas por ferramentas de Reconhecimento Óptico de Caracteres (OCR). Existem diversas ofertas de OCR disponíveis no mercado atualmente, porém, nenhuma dessas ferramentas domina como a melhor opção do mercado. A proposta deste trabalho é realizar um estudo comparativo da qualidade em função do tempo de tais ferramentas, para que seja um apoio à tomada de decisão de qual solução de OCR aplicar a um determinado negócio. Para isso, foi criado um *dataset* com 108 imagens e realizados testes individuais com todas as ferramentas consideradas com cada uma dessas imagens, a fim de verificar a acurácia e eficiência das ferramentas para diferentes tipos de imagens. Por fim, os resultados obtidos são apresentados e discutidos, demonstrando que as ferramentas OCR Web Service e Tesseract apresentaram os melhores resultados.

Palavras-chave: Reconhecimento Óptico de Caracteres, Comparação, Ferramentas OCR.

Abstract

Currently, there is a large volume of information available digitally and most of this information is composed of scanned document images. Due to this large amount, it is necessary to have methods to enable the search in these documents, which are provided by the Optical Character Recognition (OCR) tools. There are several OCR offerings available in the market today, however, none of these tools dominates as the best option on the market. The purpose of this work is to conduct a comparative study of the quality of these tools in order to support the decision-making of which OCR solution is most suitable for a particular business. For that, a dataset with 108 images was created and individual tests were conducted with all the tools considered with each of these images, in order to verify the accuracy and efficiency of the tools for different types of images. Finally, the results are presented and discussed, showing that the tools OCR Web Service and Tesseract showed the best results.

Keywords: Optical Character Recognition, Comparison, OCR tools.

Lista de ilustrações

Figura 1 – Padrão OCR-A	27
Figura 2 – Padrão OCR-B	27
Figura 3 – Técnicas de um sistema de OCR	28
Figura 4 – Imagem DANFE digitalizada na resolução 300dpi no formato JPG . . .	48
Figura 5 – Imagem Empenho digitalizada na resolução 300dpi no formato JPG . .	49
Figura 6 – Imagem Portaria digitalizada na resolução 300dpi no formato JPG . .	50
Figura 7 – Imagem Hino digitalizada na resolução 300dpi com fonte normal tama- nho 12 no formato JPG	51
Figura 8 – Imagem Poema digitalizada na resolução 300dpi no formato JPG . . .	52
Figura 9 – Interface gráfica inicial do ambiente de comparação de ferramentas OCR	53
Figura 10 – Interface gráfica de resultados obtidos por cada OCR e cada arquivo de entrada	53
Figura 11 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem DANFE	56
Figura 12 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem DANFE	58
Figura 13 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem Empenho	59
Figura 14 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem Empenho	60
Figura 15 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem Portaria	62
Figura 16 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem Portaria	63
Figura 17 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte	64
Figura 18 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte	65
Figura 19 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte	66
Figura 20 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte	67
Figura 21 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte	68

Figura 22 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte	69
Figura 23 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 10 e com variações do estilo da fonte	70
Figura 24 – Tempo de execução das ferramentas de OCR para foto da imagem Hino com fonte tamanho 10 e com variações do estilo da fonte	71
Figura 25 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte	72
Figura 26 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte	73
Figura 27 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte	74
Figura 28 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte	75
Figura 29 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte	76
Figura 30 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte	77
Figura 31 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 12 e com variações do estilo da fonte	78
Figura 32 – Tempo de execução das ferramentas de OCR para foto da imagem Hino com fonte tamanho 12 e com variações do estilo da fonte	79
Figura 33 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte	80
Figura 34 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte	81
Figura 35 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte	82
Figura 36 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte	83
Figura 37 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte	84
Figura 38 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte	85
Figura 39 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 14 e com variações do estilo da fonte	86
Figura 40 – Tempo de execução das ferramentas de OCR para foto da imagem Hino com fonte tamanho 14 e com variações do estilo da fonte	87

Figura 41 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte . . .	88
Figura 42 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte	89
Figura 43 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte . . .	90
Figura 44 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte	91
Figura 45 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte . . .	92
Figura 46 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte	93
Figura 47 – Acurácia das ferramentas de OCR para foto da imagem Poema com fonte tamanho 10 e com variações do estilo da fonte	94
Figura 48 – Tempo de execução das ferramentas de OCR para foto da imagem Poema com fonte tamanho 10 e com variações do estilo da fonte	95
Figura 49 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte . . .	96
Figura 50 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte	97
Figura 51 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte . . .	98
Figura 52 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte	99
Figura 53 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte . . .	100
Figura 54 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte	101
Figura 55 – Acurácia das ferramentas de OCR para foto da imagem Poema com fonte tamanho 12 e com variações do estilo da fonte	102
Figura 56 – Tempo de execução das ferramentas de OCR para foto da imagem Poema com fonte tamanho 12 e com variações do estilo da fonte	103
Figura 57 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte . . .	104
Figura 58 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte	105
Figura 59 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte . . .	106

Figura 60 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte	107
Figura 61 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte . . .	108
Figura 62 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte	109
Figura 63 – Acurácia das ferramentas de OCR para foto da imagem Poema com fonte tamanho 14 e com variações do estilo da fonte	110
Figura 64 – Tempo de execução das ferramentas de OCR para foto da imagem Poema com fonte tamanho 14 e com variações do estilo da fonte	111

Lista de tabelas

Tabela 1 – Hardware e software utilizados no desenvolvimento	35
Tabela 2 – Tecnologias web de desenvolvimento utilizadas	36
Tabela 3 – Palavras-chave utilizadas nas buscas	37
Tabela 4 – Total de resultados obtidos nas buscas	39
Tabela 5 – Total de artigos analisados	39
Tabela 6 – Sumarização dos Resultados	40
Tabela 7 – Artigos selecionados	41
Tabela 8 – Comparação dos atributos das ferramentas de OCR	42
Tabela 9 – Totalização de imagens do <i>dataset</i>	46
Tabela 10 – Totalização de imagens do <i>dataset</i>	47
Tabela 11 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem DANFE	56
Tabela 12 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem DANFE	57
Tabela 13 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem Empenho	59
Tabela 14 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem Empenho	60
Tabela 15 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem Portaria	61
Tabela 16 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem Portaria	62
Tabela 17 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte	64
Tabela 18 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte	65
Tabela 19 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte	66
Tabela 20 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte	67
Tabela 21 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte	68

Tabela 22 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte	69
Tabela 23 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 10 e com variações do estilo da fonte	70
Tabela 24 – Tempo de execução das ferramentas de OCR em segundos para foto da imagem Hino com fonte tamanho 10 e com variações do estilo da fonte	71
Tabela 25 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte	72
Tabela 26 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 12 na resolução 100 dpi com variações do estilo da fonte	73
Tabela 27 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte	74
Tabela 28 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte	75
Tabela 29 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte	76
Tabela 30 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 12 na resolução 600 dpi com variações do estilo da fonte	77
Tabela 31 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 12 com variações do estilo da fonte	78
Tabela 32 – Tempo de execução das ferramentas de OCR em segundos para foto da imagem Hino com fonte tamanho 12 e com variações do estilo da fonte	79
Tabela 33 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte	80
Tabela 34 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte	81
Tabela 35 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte	82
Tabela 36 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 14 na resolução 300 dpi com variações do estilo da fonte	83
Tabela 37 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte	84

Tabela 38 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte	85
Tabela 39 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 14 e com variações do estilo da fonte	86
Tabela 40 – Tempo de execução das ferramentas de OCR em segundos para foto imagem Hino com fonte tamanho 14 e com variações do estilo da fonte	87
Tabela 41 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte	88
Tabela 42 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte	89
Tabela 43 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte	90
Tabela 44 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte	91
Tabela 45 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte	92
Tabela 46 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte	93
Tabela 47 – Acurácia das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 10 e com variações do estilo da fonte	94
Tabela 48 – Tempo de execução das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 10 e com variações do estilo da fonte	95
Tabela 49 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte	96
Tabela 50 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte	97
Tabela 51 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte	98

Tabela 52 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte	99
Tabela 53 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte	100
Tabela 54 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte	101
Tabela 55 – Acurácia das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 12 e com variações do estilo da fonte	102
Tabela 56 – Tempo de execução das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 12 e com variações do estilo da fonte	103
Tabela 57 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte	104
Tabela 58 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte	105
Tabela 59 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte	106
Tabela 60 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte	107
Tabela 61 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 600 dpi com variações do estilo da fonte	108
Tabela 62 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte	109
Tabela 63 – Acurácia das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 14 e com variações do estilo da fonte	110
Tabela 64 – Tempo de execução das ferramentas de OCR em segundos para foto da imagem Poema com fonte tamanho 14 e com variações do estilo da fonte	111
Tabela 65 – Totalização dos Resultados	112
Tabela 66 – Acurácia das ferramentas de OCR para imagem Hino	117
Tabela 67 – Tempo de execução das ferramentas de OCR para imagem Hino	119
Tabela 68 – Acurácia das ferramentas de OCR para imagem Poema	121

Tabela 69 – Tempo de execução das ferramentas de OCR para imagem Poema . . . 123

Lista de abreviaturas e siglas

API	<i>Application Programming Interface</i> (Interface de Programação de Aplicações)
CSS	<i>Cascading Style Sheets</i> (Folhas de Estilo em Cascata)
DANFE	Documento Auxiliar da Nota Fiscal Eletrônica
DDR4	<i>Double Data Rate 4</i> (Taxa de Transferência Dobrada 4)
EUA	Estados Unidos da América
GB	<i>Gigabyte</i>
HP	Hewlett Packard
HTML	<i>HyperText Markup Language</i> (Linguagem de Marcação de Hipertexto)
IDC	International Data Corporation
IFMG	Instituto Federal de Minas Gerais
JPG	<i>Joint Photographic Experts Group</i>
JSON	<i>JavaScript Object Notation</i> (Notação de Objeto JavaScript)
LSTM	<i>Long Short-term Memory</i> (Memória de Curto Prazo Longa)
OCR	<i>Optical Character Recognition</i> (Reconhecimento Óptico de Caracteres)
PBM	<i>Portable Bitmap</i> (Bitmap Portátil)
PC	<i>Personal Computer</i> (Computador Pessoal)
PDF	<i>Portable Document Format</i> (Formato de Documento Portátil)
PHP	<i>PHP: Hypertext Preprocessor</i> (PHP: Pré-processador de Hipertexto)
RAM	<i>Random Access Memory</i> (Memória de Acesso Aleatório)
RPM	<i>Rotations per Minute</i> (Rotações por Minuto)
RSL	Revisão Sistemática da Literatura
SDK	<i>Software Development Kit</i> (Kit de Desenvolvimento de Software)
SEI	Sistema Eletrônico de Informações

SGBD	Sistema Gerenciador de Bancos de Dados
TB	<i>Terabyte</i>
UTF-8	<i>8-bit Unicode Transformation Format</i> (Formato de Transformação Unicode de 8 <i>bits</i>)

Sumário

1	INTRODUÇÃO	22
1.1	Justificativa	23
1.2	Objetivos	23
1.3	Estrutura do Trabalho	24
2	FUNDAMENTAÇÃO TEÓRICA	25
2.1	Reconhecimento Óptico de Caracteres	25
2.1.1	Histórico	26
2.1.2	Técnicas de Sistemas OCR	28
2.1.2.1	Escaneamento Óptico	29
2.1.2.2	Localização e Segmentação	29
2.1.2.3	Pré-processamento	29
2.1.2.4	Extração de Características	30
2.1.2.5	Classificação	30
2.1.2.6	Pós-processamento	30
2.1.3	Aplicações	31
2.2	Trabalhos Relacionados	32
2.2.1	A Comparative Study on OCR Tools	32
2.2.2	Text recognition in multimedia documents: a study of two neural-based OCRs using and avoiding character segmentation	33
2.2.3	OCR as a service: an experimental evaluation of google docs OCR, tesseract, ABBYY finereader, and transym	33
3	METODOLOGIA	35
3.1	Materiais	35
3.1.1	Hardware e Software	35
3.2	Métodos	37
3.2.1	Revisão Sistemática da Literatura - RSL	37
3.2.1.1	Planejamento	37
3.2.1.2	Execução	38
3.2.2	Análise das Opções de OCR Disponíveis	41
3.2.2.1	Tesseract	43
3.2.2.2	Google Cloud Vision API	44
3.2.2.3	GOOCR	44
3.2.2.4	ABBYY Cloud OCR SDK	44

3.2.2.5	OCR Document Haven OnDemand	44
3.2.2.6	OCR.Space API	44
3.2.2.7	New OCR API	45
3.2.2.8	Microsoft Vision API	45
3.2.2.9	OCR Web Service	45
3.2.2.10	SemaMediaData OCR	45
3.2.2.11	Cloudmersive	45
3.2.3	<i>Dataset</i> Utilizado Nos Experimentos	45
3.2.3.1	Imagem DANFE	47
3.2.3.2	Imagem Empenho	48
3.2.3.3	Imagem Portaria	49
3.2.3.4	Imagem Hino	50
3.2.3.5	Imagem Poema	51
3.2.4	Ambiente de Experimentação	52
3.2.5	Análise Comparativa	54
4	RESULTADOS E DISCUSSÃO	55
4.1	Imagem DANFE	55
4.2	Imagem Empenho	58
4.3	Imagem Portaria	61
4.4	Imagem Hino	63
4.4.1	Fonte Tamanho 10	63
4.4.2	Fonte Tamanho 12	71
4.4.3	Fonte Tamanho 14	80
4.5	Imagem Poema	87
4.5.1	Fonte Tamanho 10	88
4.5.2	Fonte Tamanho 12	95
4.5.3	Fonte Tamanho 14	103
4.6	Totalização dos Resultados	111
5	CONSIDERAÇÕES FINAIS	113
	REFERÊNCIAS	115
	APÊNDICE A – ACURÁCIA DAS FERRAMENTAS DE OCR PARA IMAGEM HINO	117
	APÊNDICE B – TEMPO DE EXECUÇÃO DAS FERRAMENTAS DE OCR PARA IMAGEM HINO	119

APÊNDICE C – ACURÁCIA DAS FERRAMENTAS DE OCR PARA IMAGEM POEMA	121
APÊNDICE D – TEMPO DE EXECUÇÃO DAS FERRAMENTAS DE OCR PARA IMAGEM POEMA	123

1 INTRODUÇÃO

A invenção do papel procedeu uma das maiores revoluções da humanidade. Atualmente, ainda é uma das formas de mídia que mais acumula informações, embora não seja a mais eficiente. O papel traz diversas desvantagens, como o grande espaço físico necessário para armazená-lo, o que aumenta exponencialmente com a quantidade de informações (MELLO; LINS, 1999).

De acordo com Rocha (2010) em uma pesquisa realizada pelo International Data Corporation (IDC), nos EUA, os escritórios geram em torno de 1 bilhão de páginas de papel diariamente, sendo esse total formado de 600 milhões de páginas de relatórios de computador, 234 milhões de fotocópias e 24 milhões de outros tipos de documentos, referindo-se apenas aos Estados Unidos.

A invenção da digitalização tornou possível armazenar documentos em papel no formato digital, com o objetivo de preservar esses documentos do desgaste ao longo do tempo e torná-los mais acessíveis (MELLO; LINS, 1999).

De acordo com RIOS (2007) na atualidade existe um grande volume de informação disponível digitalmente e a maior parte dessa informação é composta por imagens de documentos digitalizados. Devido a essa enorme quantidade, torna-se necessário que existam ferramentas para que além dos documentos estarem disponíveis em formato digital, sejam também pesquisáveis e editáveis. Isso se tornou possível com o desenvolvimento de sistemas de Reconhecimento Óptico de Caracteres.

Reconhecimento Óptico de Caracteres, ou no inglês *Optical Character Recognition* (OCR), é uma das mais importantes áreas de Visão Computacional (MARANA; PAPA; CHIACHIA, 2009). Segundo o portal ABBYY (2018), OCR é uma tecnologia que permite converter tipos diferentes de documentos, como papéis digitalizados, arquivos em PDF e imagens capturadas com câmera digital em dados pesquisáveis.

Aires (2005) diz que os OCRs são utilizados em vários tipos de aplicações, tais como máquinas automáticas que processam formulários preenchidos manualmente, leitores automáticos de cheques bancários, máquinas automáticas de processamento de códigos postais, dentre outras.

Segundo Mello e Lins (1999), uma das principais dificuldades ao lidar com OCR, vem com a escolha da melhor ferramenta de reconhecimento. Atualmente, existem muitas ofertas de OCR disponibilizadas no mercado, como o Tesseract¹, Google Cloud Vision² e

¹ Mais informações sobre o Tesseract: <<https://github.com/tesseract-ocr>>

² Mais informações sobre o Google Cloud Vision: <<https://cloud.google.com/vision/docs/?hl=pt-br>>

ABBYY FineReader³.

1.1 Justificativa

Nos últimos anos, devido a disponibilidade e utilização de tecnologias para informação e comunicação, algumas utilidades do papel têm sido substituídas por meios digitais e eletrônicos. Os documentos em papel não atendem mais às necessidades de agilidade na circulação das informações, possuindo limitações no que se refere à simples conservação, transmissibilidade ou segurança.

Os documentos eletrônicos têm se tornado cada vez mais presentes em todas as áreas do trabalho moderno, permitindo cada vez mais agilidade na obtenção da informação. A implantação do Sistema Eletrônico de Informações (SEI)⁴ no IFMG é uma solução tecnológica que possibilita a gestão de processos e documentos eletrônicos de maneira ágil. De acordo com o portal IFMG⁵, os setores envolvidos não tiveram dificuldades em aprender a utilizar o sistema e o ganho de tempo foi enorme, pois processos que demoravam vários meses para serem concluídos, foram realizados em aproximadamente 16 dias, considerando o período de agosto a novembro de 2017.

A digitalização de documentos físicos com o propósito de transformá-los em arquivos digitais carece de alguma ferramenta que realize o Reconhecimento Óptico de Caracteres, para que seja possível que esses documentos possam se tornar pesquisáveis.

Muitas são as ofertas de OCR disponibilizadas atualmente. Porém, nenhuma dessas ferramentas domina como a melhor opção no mercado. Levando em consideração que aplicações que têm como propósito o gerenciamento de documentos necessitam de alguma forma de realizar buscas em documentos digitalizados, a extração do texto implica diretamente na qualidade desse tipo de busca. Um estudo da qualidade em função do tempo de tais ferramentas se faz necessário para a tomada de decisão de qual a melhor solução de OCR aplicar a um determinado negócio.

1.2 Objetivos

Após a contextualização do objeto de estudo deste trabalho e sua importância, sintetiza-se aqui seu objetivo primário: realizar um estudo comparativo de variadas ferramentas de OCR, tendo como foco do trabalho o reconhecimento de texto em documentos escritos em português.

São objetivos secundários e mais específicos os seguintes tópicos:

³ Mais informações sobre o ABBYY FineReader: <<https://www.abbyy.com/pt-br/finereader/>>

⁴ Mais informações sobre o SEI: <<https://www2.ifmg.edu.br/portal/pro-reitoria-de-administracao/sei>>

⁵ Portal IFMG: <<https://www2.ifmg.edu.br/portal>>

- Elencar as implementações de OCR aplicáveis;
- Apresentar um ambiente web que integra ferramentas variadas de OCR, que ofereça suporte à comparação e avaliação experimental de diversos mecanismos de Reconhecimento Óptico de Caracteres;
- Definir o *dataset* a ser utilizado nos testes;
- Executar testes individuais de cada implementação de OCR;
- Avaliar, de maneira experimental, variadas ferramentas de OCR, verificando a acurácia e eficiência das ferramentas consideradas.

1.3 Estrutura do Trabalho

Este trabalho está dividido em cinco capítulos. O capítulo dois apresenta o referencial teórico sobre Reconhecimento Óptico de Caracteres, englobando seus conceitos, história e diferentes ferramentas existentes. O capítulo três expõe a metodologia que foi utilizada no estudo e desenvolvimento do trabalho, bem como as ferramentas usadas. No capítulo quatro são apresentados os resultados obtidos nos experimentos realizados e fornece uma discussão sobre os resultados obtidos. Por fim, no capítulo cinco são feitas as considerações finais deste trabalho, seguidas pelas referências bibliográficas.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo é apresentado o Referencial Teórico sobre Reconhecimento Óptico de Caracteres, abordando sua definição, histórico, técnicas, aplicações e trabalhos relacionados.

2.1 Reconhecimento Óptico de Caracteres

A evolução tecnológica gerou um aumento na capacidade de armazenamento e uma redução no custo necessário para armazenar uma grande quantidade de arquivos. Isso fez com que consequentemente as empresas e usuários acumulassem uma grande quantidade de informações que se mantém em constante crescimento.

De acordo com RIOS (2007), algumas empresas têm adotado a prática de manter cópias digitais de alguns documentos na forma de imagem, para facilitar o compartilhamento de informações e diminuir o espaço físico necessário para armazenamento. Porém, não é vantagem apenas armazenar essas imagens, além disso deve ser possível torná-las pesquisáveis, fazendo-se necessário o uso de ferramentas de Reconhecimento Óptico de Caracteres.

Segundo Chaudhuri et al. (2017), Reconhecimento Óptico de Caracteres é um processo de classificação de padrões ópticos que estão inclusos em uma imagem digital que corresponde a caracteres alfanuméricos ou outros caracteres.

A tecnologia de Reconhecimento Óptico de Caracteres permite converter diferentes tipos de documentos, como documentos em papel digitalizados, arquivos em PDF ou imagens capturadas por uma câmera digital em dados pesquisáveis.(CHAUDHURI et al., 2017)

Chaudhuri et al. (2017) diz que os sistemas de OCR tornaram-se uma das aplicações mais bem sucedidas da tecnologia nas áreas de inteligência artificial e reconhecimento de padrões. Mas, é importante ressaltar que existem diversos problemas que impõem dificuldades na realização de um OCR e na obtenção de bons resultados. Segundo Raabe e Filho (1998), esses problemas podem ser causados pelo estado de conservação do documento, como risco e anotações, manchas, amassados ou podem estar ligados à ineficácia dos próprios OCRs para o tratamento de determinadas características, como figuras, trechos manuscritos, letras pequenas ou borradas, fórmulas matemáticas, dentre outras. O desempenho do Reconhecimento Óptico de Caracteres depende diretamente da qualidade dos documentos de entrada (EIKVIL, 1993).

2.1.1 Histórico

Os primeiros relatos sobre as origens do OCR foram encontrados em torno de 1870, com a criação do *scanner* de retina desenvolvido por CR. Carey, de Boston, Massachusetts. Esse *scanner* era um sistema de transmissão de imagem que utilizava um mosaico de fotocélulas (CHAUDHURI et al., 2017).

Duas décadas após isso, houve um enorme avanço para as máquinas de leitura e para a televisão moderna por causa da criação do *scanner* sequencial, que foi desenvolvido pelo polonês P.Nipkow. Nas primeiras décadas do século 19, através de experimentos realizados com OCR, foram feitas várias tentativas de desenvolvimento de dispositivos para auxiliar pessoas com deficiência visual. (EIKVIL, 1993).

No entanto, só com o desenvolvimento dos computadores digitais, que aconteceu por volta de 1940, é que surgiu a versão moderna de OCR Chaudhuri et al. (2017). A partir de então, as possíveis aplicações comerciais motivaram o desenvolvimento de sistemas de OCR (EIKVIL, 1993).

Em torno de 1950, a área de processamento de dados estava se transformando em um campo muito importante, devido ao grande avanço tecnológico ocorrido na época. Nesse período ocorreu uma demanda para que fosse desenvolvido algum método econômico que fosse capaz de lidar com o crescimento da quantidade de dados, sendo que a entrada desses era realizada por cartões perfurados. No meio da década de 1950, as ferramentas OCR tornaram-se disponíveis comercialmente, visto que já existia tecnologia suficiente para o desenvolvimento dessas aplicações (EIKVIL, 1993).

Conforme Chaudhuri et al. (2017), em 1954 foi instalada no Reader's Digest a primeira máquina de leitura de OCR, que era usada na conversão de relatórios de vendas datilografados em cartões perfurados que serviam como entrada para o computador.

A primeira geração de OCRs pode ser datada entre 1960 e 1965, onde as ferramentas de OCR comerciais que surgiram nessa época foram caracterizados pela leitura de um formato restrito de letras (EIKVIL, 1993). Mas, no decorrer dos anos, as ferramentas OCR tornaram-se capazes de reconhecer diversos tipos de formatos de letras, sendo a quantidade de formatos limitada pelo método de reconhecimento do padrão utilizado, que compara a imagem do caractere com uma biblioteca de protótipos de imagens de cada caractere (CHAUDHURI et al., 2017).

Os OCRs que apareceram no meio da década de 1960 e começo da década de 1970 fazem parte da segunda geração e, segundo Chaudhuri et al. (2017), eram capazes de reconhecer caracteres que tenham sido impressos por outras máquinas e também começaram a reconhecer caracteres escritos à mão. Porém, o reconhecimento de caracteres escritos à mão era limitado a poucas letras, números e símbolos. O IBM 1287, criado em 1965, foi o primeiro e mais famoso sistema desse gênero. Além disso, a Toshiba criou a

primeira máquina automática de classificação de letras para códigos postais e a Hitachi fez a primeira máquina de OCR com alto desempenho e baixo custo (EIKVIL, 1993).

Segundo Chaudhuri et al. (2017), Em 1966 foi desenvolvido um estudo dos requisitos para sistemas OCR e foi definido um padrão americano de conjunto de caracteres que foi chamado de OCR-A. Também foi desenvolvida uma fonte europeia que possuía formato mais natural do que o padrão americano. Essa fonte foi chama de OCR-B. Foram realizadas algumas tentativas para mesclar as duas fontes em um só padrão, porém, ao invés disso, foram desenvolvidas novas ferramentas que conseguiam reconhecer os dois padrões (CHAUDHURI et al., 2017).

Figura 1 – Padrão OCR-A

A	B	C	D	E	F	G	H	I	J	K	L
M	N	O	P	Q	R	S	T	U	V	W	X
Y	Z	1	2	3	4	5	6	7	8	9	0

Fonte: Chaudhuri et al. (2017)

Figura 2 – Padrão OCR-B

A	B	C	D	E	F	G	H	I	J	K	L
M	N	O	P	Q	R	S	T	U	V	W	X
Y	Z	1	2	3	4	5	6	7	8	9	0

Fonte: Chaudhuri et al. (2017)

Na década de 1970 surgiu a terceira geração de OCRs, dispendo como principais desafios os documentos de baixa qualidade e grande quantidade de conjuntos de caracteres impressos e manuscritos (CHAUDHURI et al., 2017). De acordo com Eikvil (1993), as ferramentas de OCR começaram a ser mais vendidas a partir de 1986, apesar de já estarem disponíveis comercialmente desde a década de 1950. Esse aumento nas vendas ocorreu devido ao fato de que o hardware estava se tornando mais barato. Por conseguinte, os sistemas de OCR começaram a ficar disponíveis como pacotes de software, o que fez com que aumentasse consideravelmente o número de vendas.

No entanto, o principal desenvolvimento da tecnologia de OCR se deu durante a década de 1990, com a utilização de técnicas de reconhecimento de padrões e processamento de imagem combinadas com metodologias de inteligência artificial. Atualmente, segundo Chaudhuri et al. (2017), além dos equipamentos eletrônicos mais precisos e potentes, como *scanners*, câmeras e *tablets* eletrônicos, tem-se o uso eficiente e moderno de metodologias

como redes neurais artificiais (RNAs¹), modelos ocultos de Markov (HMMs²), lógica *fuzzy*³ e processamento de linguagem natural⁴.

O sistema de OCR chamado Tesseract foi desenvolvido entre os anos de 1985 e 1994, sendo compatível com vários sistemas operacionais. Seu desenvolvimento foi apoiado pela Google e por desenvolvedores do mundo todo a partir de 2006 (FEIJÓ et al., 2017).

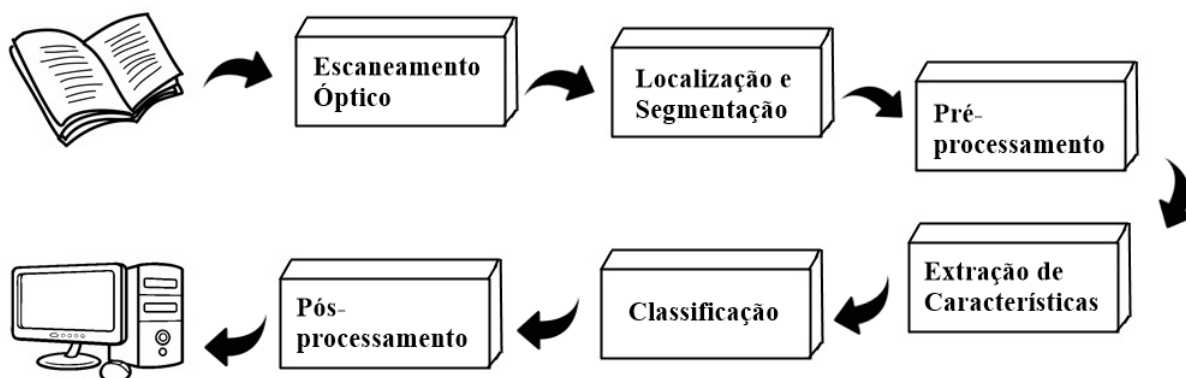
Segundo Feijó et al. (2017), atualmente, além do Tesseract existem várias aplicações modernas de OCR que conseguem desempenhar bons resultados na extração de texto de imagens, como ABBYY FineReader, Google Cloud Vision API, Microsoft Vision API, ReadIris, OmniPage, dentre outras.

2.1.2 Técnicas de Sistemas OCR

Segundo Eikvil (1993), um sistema de OCR é composto basicamente de:

- Escaneamento Óptico;
- Localização e Segmentação;
- Pré-processamento;
- Extração de Características;
- Classificação;
- Pós-processamento;

Figura 3 – Técnicas de um sistema de OCR



¹ Mais informações sobre RNAs: <<https://bit.ly/2QDDWtW>>

² Mais informações sobre HMMs: <<https://bit.ly/2REukQY>>

³ Mais informações sobre lógica *fuzzy*: <<https://bit.ly/2OM7qsi>>

⁴ Mais informações sobre processamento de linguagem natural: <<https://bit.ly/2RBr5to>>

2.1.2.1 Escaneamento Óptico

De acordo com Eikvil (1993), a imagem digital do documento original é capturada através do processo de digitalização. Atualmente, também é possível que a imagem digital seja capturada através de uma foto do documento original. Os *scanners* ópticos geralmente consistem em um mecanismo de transporte que converte a intensidade da luz em níveis de cinza, sendo esse processo conhecido como **limiarização**.

O processo de limiarização é bastante importante, porque os resultados do reconhecimento de caracteres dependem totalmente da qualidade da imagem. A limiarização feita no *scanner* geralmente é básica, sendo definido um limite fixo, onde os níveis de cinza abaixo do limiar são considerados pretos e os níveis acima são considerados brancos. Alguns tipos de documentos podem carecer de métodos mais sofisticados de limiarização para se obter bons resultados (CHAUDHURI et al., 2017).

2.1.2.2 Localização e Segmentação

Chaudhuri et al. (2017) diz que a segmentação é o isolamento de caracteres ou palavras do texto, sendo que a maioria dos algoritmos de OCR segmenta as palavras em caracteres isolados que são reconhecidos individualmente. Esse processo determina os componentes de uma imagem localizando as regiões do documento que apresentam dados impressos e diferenciando-as de figuras e gráficos.

Porém, podem surgir alguns problemas nessa fase, sendo os principais problemas, de acordo com Eikvil (1993):

- Extração de caracteres em contato e fragmentados: vários caracteres podem ser interpretados como um único caractere;
- Diferenciar o ruído de texto: acentos e pontos podem ser confundidos com ruído e vice-versa;
- Interpretação incorreta de gráficos e geometria: o texto pode não ser passado para o estágio de reconhecimento.

2.1.2.3 Pré-processamento

Segundo Islam, Islam e Noor (2017), o pré-processamento visa melhorar a qualidade da imagem, suavizando os caracteres digitalizados para tentar eliminar ruídos resultantes do processo de escaneamento, como caracteres manchados ou incompletos, que prejudicam o desempenho do OCR. Essa etapa também inclui normalização, para se obter caracteres de tamanho, inclinação e rotação uniformes (CHAUDHURI et al., 2017).

2.1.2.4 Extração de Características

O objetivo da extração de características é capturar as características de cada caractere, e é geralmente aceito que este é um dos problemas mais difíceis do reconhecimento de padrões. As técnicas para extração de tais características são frequentemente divididas em três grupos principais, onde os recursos são encontrados a partir da distribuição de pontos, das transformações e expansões de séries e de uma análise estrutural (EIKVIL, 1993).

De acordo com Chaudhuri et al. (2017), os diferentes grupos de características são avaliados de acordo com sua sensibilidade ao ruído, deformação, facilidade de implementação e uso. Os critérios utilizados nesta avaliação são:

- Robustez em termos de ruído, distorções, variação de estilo, translação e rotação;
- Uso prático em termos de velocidade de reconhecimento, complexidade de implementação e independência.

2.1.2.5 Classificação

A classificação é o processo de identificar cada caractere e atribuir a ele a uma classe de caractere correspondente. As duas categorias mais importantes de métodos de classificação para OCR são **métodos teóricos de decisão** e **métodos estruturais**, conforme visto a seguir (CHAUDHURI et al., 2017).

As principais abordagens para o reconhecimento da teoria da decisão, segundo Eikvil (1993), são:

- Distância mínima;
- Classificadores estatísticos ótimos;
- Redes neurais.

De acordo com Eikvil (1993), dentro da área de reconhecimento estrutural, os métodos sintáticos estão entre as abordagens mais utilizadas. Nos métodos sintáticos, as medidas de similaridade baseadas em relações entre componentes estruturais podem ser formuladas usando conceitos gramaticais. Existem outras técnicas, mas são menos usadas.

2.1.2.6 Pós-processamento

O resultado do reconhecimento de caracteres é um conjunto de símbolos individuais. Por isso, na fase de pós-processamento é feito o agrupamento desses caracteres, unindo palavras e números formando um texto. O agrupamento de caracteres é baseado na

localização desses no documento, sendo agrupados os caracteres que estão suficientemente próximos. (CHAUDHURI et al., 2017).

Além disso, de acordo com Eikvil (1993), nesta fase também é feita uma detecção e correção de erros. Para isso, existem duas técnicas principais, a primeira utiliza o cálculo da probabilidade de dois ou mais caracteres aparecerem juntos em sequência, usando as regras que definem a sintaxe da palavra.

A outra técnica para detecção e correção de erros que é a mais eficiente é o uso de dicionários. A desvantagem dessa técnica é que as pesquisas e comparações implícitas são demoradas (EIKVIL, 1993).

Dada uma palavra, na qual um erro pode estar presente, a palavra é pesquisada no dicionário. Se a palavra não estiver no dicionário, um erro foi detectado e pode ser corrigido alterando a palavra para a palavra mais semelhante. Se a palavra estiver presente no dicionário, isso não prova que nenhum erro ocorreu. Um erro pode ter transformado a palavra em uma outra, e tais erros são indetectáveis por esse procedimento (EIKVIL, 1993).

2.1.3 Aplicações

Existem diferentes áreas de aplicação de Reconhecimento Óptico de Caracteres, como a área de entrada de dados, entrada de texto e automação de processos Eikvil (1993).

A área de entrada de dados abrange tecnologias para que seja possível inserir grandes quantidades de dados e a princípio, essas máquinas foram utilizadas para aplicações bancárias. Esses sistemas são caracterizados pela leitura de um conjunto limitado de caracteres, normalmente números e alguns símbolos especiais, sendo projetados para ler dados como números de conta, identificação do cliente, quantias de dinheiro, números de artigo, entre outros (CHAUDHURI et al., 2017).

Já as máquinas na área de entrada de texto são utilizadas principalmente como leitores de páginas em automação de escritórios, geralmente em um ambiente de processamento de texto. O desempenho desse tipo de aplicação depende extremamente da qualidade do documento de entrada (EIKVIL, 1993).

Dentro da área de automação de processos, o principal objetivo não é ler o que é impresso, mas sim controlar determinado processo. Um exemplo dessa área é a tecnologia de leitura automática de endereços para ordenação de mensagens (EIKVIL, 1993).

Segundo Chaudhuri et al. (2017), existem muitas outras áreas de aplicação para OCR, como:

- Auxiliar deficientes visuais: Esse leitor, junto com um sistemas de síntese de fala, permite que uma pessoa cega possa entender documentos impressos;

- Leitores automáticos de placas: a entrada deve ser capturada por uma câmera muito rápida, criando problemas alguns problemas e dificuldades de reconhecimento, apesar do conjunto de caracteres ser restrito;
- Leitores de formulários: Esses sistemas são capazes de ler formas especialmente projetadas;
- Verificação e identificação de assinaturas: Nesse tipo de sistema a assinatura é considerada como padrão que é combinado com as assinaturas armazenadas no banco de dados de referência.

2.2 Trabalhos Relacionados

Nesta seção será apresentada a análise e discussão dos artigos selecionados durante a revisão sistemática da literatura, que será detalhada na seção 3.2.1.

2.2.1 A Comparative Study on OCR Tools

No trabalho de Mello e Lins (1999) é feita uma análise do desempenho de seis ferramentas de OCR comerciais mais utilizadas na época, sendo elas:

- Omnipage 9.0 (Caere Corporation);
- Corel OCR Trace 8.0 (Corel Corporation);
- SmartPage 2.1 (Recognita Corporation);
- Wordlinx (OCRON Inc.);
- TextBridge Pro 98 (Scansoft Inc.);
- TypeReader Professional 4.0 (Expervision).

Foram definidos alguns critérios para avaliação e comparação entre as ferramentas de OCR, como resolução, sensibilidade à rotação, brilho, cores de fundo e documentos históricos.

Com os testes realizados foram detectadas classes de erros comuns nas saídas dos OCRs, sendo as principais: substituição de um caractere por outro, substituição de um caractere por mais de um caractere, substituição de mais de um caractere por apenas um, perda de caracteres, mesclagem de palavras sem perda de caracteres, mesclagem de palavras com perda de caracteres, perda de linhas completas de texto, inserção de caracteres e ruído.

Pode-se observar com os testes que alguns OCRs são mais adequados para certos tipos de imagens, enquanto outros produzem melhores resultados quando se lida com outros tipos de imagens. Neste artigo, foi mostrado que o Omnipage da Caere Corp. alcançou as melhores taxas de reconhecimento em quase todos os casos analisados.

Porém, esse trabalho é de 1999 e desde então surgiram novas ferramentas de OCR, fazendo-se necessário um novo estudo comparativo sobre ferramentas OCR.

2.2.2 Text recognition in multimedia documents: a study of two neural-based OCRs using and avoiding character segmentation

O trabalho de Elagouni et al. (2014) propõe dois sistemas de reconhecimento óptico de caracteres (OCR) baseados em rede neural que lidam com o problema de reconhecimento de texto de maneiras diferentes.

A primeira é baseada em segmentação, que visa separar caracteres antes de reconhecer eles. Já a segunda evita a segmentação de caracteres usando um esquema de escala múltipla de digitalização e um modelo de gráfico.

O objetivo deste trabalho é avaliar e comparar os métodos de OCR propostos e discutir os resultados obtidos em ambas abordagens. Além disso, é feita uma revisão dos últimos métodos de reconhecimento de texto em imagens fotográficas.

Os OCRs propostos foram testados e avaliados, apresentando um resultado muito bom para Dataset I, alcançando cerca de 95% e 93% de caracteres reconhecidos corretamente para o sistema baseado em segmentação e para o sem segmentação, respectivamente.

Para o Dataset II, o sistema sem segmentação teve certa de 70% dos caracteres reconhecidos corretamente, já o sistema baseado em segmentação reconheceu 65% dos caracteres corretamente.

Os OCRs propostos também foram comparados com outros métodos de última geração e em relação ao Dataset I obteve melhores resultados no reconhecimento de caracteres do que as ferramentas ABBYY FineReader e Tesseract.

2.2.3 OCR as a service: an experimental evaluation of google docs OCR, tesseract, ABBYY finereader, and transym

O trabalho de Tafti et al. (2016) realiza avaliações experimentais qualitativas e quantitativas usando quatro serviços de OCR, o Google Docs OCR, Tesseract, ABBYY FineReader e Transym. Foram analisados a precisão e confiabilidade de cada OCR, empregando um conjunto de imagens com categorias diferentes.

Esses serviços foram escolhidos com base nas suas principais funcionalidades,

incluindo precisão de reconhecimento, desempenho, suporte multilíngue, implementação de código aberto, entrega como kit de desenvolvimento de software (SDK), alta disponibilidade e classificação na comunidade de OCR.

O OCR do Google Docs e o ABBYY FineReader produziram resultados mais promissores, e seus desvios padrão da precisão permaneceram consistentes em diferentes tipos de imagens existentes no conjunto de dados. A qualidade das imagens de entrada tem um impacto crucial nas saídas de OCR. Por exemplo, todos os sistemas de OCR examinados enfrentaram problemas com distorção e imagens com ruídos.

3 METODOLOGIA

Neste capítulo são apresentados os materiais e métodos utilizados para o desenvolvimento do presente trabalho.

3.1 Materiais

Nesta seção são apresentados os materiais que foram utilizados em todo desenvolvimento do trabalho.

3.1.1 Hardware e Software

As plataformas de hardware e de software apresentadas nas Tabelas 1 e 2, foram utilizadas no ambiente de desenvolvimento e análise experimental de todo trabalho.

Tabela 1 – Hardware e software utilizados no desenvolvimento

<i>HARDWARE</i>	
Computador PC, tipo notebook 7ª geração do Processador Intel® Core™ I5-7200U @ 2.50GHz - @ 2.70GHz 8GB de memória RAM DDR4, 2400 MHz 1TB de disco rígido 5400 RPM	
<i>SOFTWARE</i>	
ITEM	DESCRIÇÃO
Sistema Operacional Microsoft Windows 10 Home Language x64	Sistema Operacional oferecido pela Microsoft. Mais informações: https://www.microsoft.com/pt-br/windows/ .
Servidor web Apache versão 2.4.18	Projeto de servidor HyperText Transfer Protocol (HTTP) open source desenvolvido e mantido pela fundação Apache. Mais informações: https://httpd.apache.org/ .
SGBD MariaDB 10.1.32	Sistema gerenciador de bancos de dados open source. Mais informações: https://mariadb.com/ .
Editor de textos Atom versão 1.30.0	Editor de texto open source voltado para programadores. Mais informações: https://atom.io/ .

SOFTWARE	
ITEM	DESCRIÇÃO
MySQLWorkbench versão 6.3.7	Ferramenta gráfica para modelagem de banco de dados e gerenciamento dos mesmos. Mais informações: https://www.mysql.com/products/workbench/ .
Google Chrome versão 69.0.3497.100	Navegador de internet utilizado para testes do ambiente web desenvolvido. Mais informações: https://www.google.com/chrome/ .
Microsoft Excel Online	Editor de planilhas produzido pela Microsoft. Mais informações: https://office.live.com/ .

Tabela 2 – Tecnologias web de desenvolvimento utilizadas

TECNOLOGIAS WEB DE DESENVOLVIMENTO	
PHP versão 7.2.5	Linguagem de programação server-side com base em pré-processamento de scripts. Mais informações: https://www.php.net/ .
Framework Laravel versão 5.6.29	Framework de desenvolvimento PHP. Mais informações: https://laravel.com/ .
Linguagem de marcação de hipertexto 5 (HTML5)	Linguagem de marcação padrão de desenvolvimento lado client para web. Mais informações: https://www.w3.org/TR/html5/ .
Folhas de Estilo em Cascata 3 (CSS3)	Mecanismo de estilização de documentos web. Mais informações: https://www.w3.org/Style/CSS/ .
JQuery versão 3.3.1	Biblioteca Javascript. Mais informações: https://jquery.com/ .
Framework Bootstrap versão 4.0.0	Framework HTML, CSS e JavaScript grátis desenvolvimento responsivo, mobile first para a web. Mais informações: http://getbootstrap.com/
JavaScript Object Notation (JSON)	Notação em formato JavaScript para troca de objetos entre máquinas na web. Mais informações: http://www.json.org/ .

3.2 Métodos

Nesta seção são apresentados os métodos utilizados para o desenvolvimento do trabalho.

3.2.1 Revisão Sistemática da Literatura - RSL

A seguir é apresentado a aplicação do protocolo da RSL.

3.2.1.1 Planejamento

Nessa seção, são apresentados os tópicos da fase Planejamento:

- **Objetivos:** Pretende-se identificar o que é Reconhecimento Óptico de Caracteres e quais são as principais ferramentas disponíveis no mercado, para que posteriormente seja feita uma comparação entre os principais métodos disponíveis.
- **Questões de Pesquisa:** O foco da pergunta é realizar uma busca abrangente sobre Reconhecimento Óptico de Caracteres e os métodos já existentes no mercado, assim como uma comparação desses métodos.
 - Questão 1: "O que é Reconhecimento Óptico de Caracteres?"
 - Questão 2: "Como foi a evolução dos OCRs?"
 - Questão 3: "Quais os métodos utilizados para reconhecimento de caracteres?"
 - Questão 4: "Quais são as principais ferramentas de OCR disponíveis no mercado?"
- **Palavras-Chave:** As palavras-chave foram obtidas em leitura prévia de artigos relevantes, podendo se observar as escolhidas na tabela abaixo:

Tabela 3 – Palavras-chave utilizadas nas buscas

Termo 1	Termo 2
Optical Character Recognition	
OCR	
Optical Character Recognition	survey
OCR	framework
OCR	environment
OCR	Machine Learning
A Survey on Optical Character Recognition System	
Intelligent Character Recognition	
Handwritten Character Recognition	

Fonte: Elaborado pela autora.

- String de Busca: foi elaborada com base nas palavras-chave da tabela anterior, relacionado-as utilizando operadores lógicos ou logicamente entre aspas;
- Método de Busca de Fontes: Sites de bibliotecas científicas virtuais e de busca;
- Listagem de Fontes:
 1. Scopus (<<http://www.scopus.com>>);
 2. Elsevier ScienceDirect (<<http://www.sciencedirect.com>>);
 3. El Compendex (<<http://www.engineeringvillage2.org>>);
 4. SpringerLink (<<https://link.springer.com/>>);
 5. Google Acadêmico (<<https://scholar.google.com.br/>>);
 6. Outros. Artigos que foram encontrados em outras diversas fontes e que possuíam conteúdo relevante sobre o tema.
- Tipos de Artigos: Foram considerados artigos com o conteúdo sobre Reconhecimento Óptico de Caracteres e os principais métodos já disponíveis no mercado;
- Idioma dos Artigos: Foram considerados artigos em inglês e português;
- Critérios de Inclusão e Exclusão dos Artigos: Os artigos devem:
 1. Estar disponíveis na web;
 2. Oferecer conteúdo completo nas bases buscadas;
 3. Estar escritos em inglês ou português;
 4. Possuir conteúdo relevante sobre o tema abordado.

3.2.1.2 Execução

Depois da elaboração do protocolo, este é executado com a finalidade de apurar artigos que possam fazer parte da base de referencial teórico sobre o assunto.

Primeiramente foi executada uma busca as palavras-chave nas bases selecionadas. Em seguida, foi obtida uma grande quantidade de resultados, não sendo possível analisar todos os artigos encontrados, devido a limitação do número de artigos para download das bases. Os resultados dessa etapa podem ser vistos na Tabela 4.

Tabela 4 – Total de resultados obtidos nas buscas

Filtros: Ciência da Computação; Inglês; Relevância;							
Termo 1	Termo 2	Scopus	Springer	Science	Compendex	Google	Total
Optical Character Recognition		4.877	2.666	2.876	6.342	1.980.000	1.996.761
OCR		4.025	4.496	18.320	4.721	272.000	303.562
Optical Character Recognition	survey	95	840	917	98	353.000	354.950
OCR	framework	241	1.867	3.012	249	56.200	61.569
OCR	environment	179	1.824	5.156	206	71.200	78.565
OCR	Machine Learning	105	1.065	520	116	32.100	33.906
A Survey on Optical Character Recognition System		0	0	0	0	1	1
Intelligent Character Recognition		37	55	52	44	1.180.000	1.180.188
Handwritten Character Recognition		923	760	647	1.000	101.000	104.330
Total		10.482	13.573	31.500	12.776	4.045.501	4.113.832

Tabela 5 – Total de artigos analisados

Filtros: Ciência da Computação; Inglês; Relevância;							
Termo 1	Termo 2	Scopus	Springer	Science	Compendex	Google	Total
Optical Character Recognition		100	100	100	100	100	500
OCR		100	100	100	100	100	500
Optical Character Recognition	survey	95	100	100	98	100	493
OCR	framework	100	100	100	100	100	500
OCR	environment	100	100	100	100	100	500
OCR	Machine Learning	100	100	100	100	100	500
A Survey on Optical Character Recognition System		0	0	0	0	1	1
Intelligent Character Recognition		37	55	52	44	100	288
Handwritten Character Recognition		100	100	100	100	100	500
Total		732	755	752	742	801	3.782

Desse modo, foi necessário que fossem analisados um número reduzido de artigos, sendo contabilizados no máximo os 100 artigos mais relevantes de cada palavra-chave em cada base. A quantidade de artigos remanescentes pode ser analisada na Tabela 5.

A seguir são descritas cada etapa que foi realizada na RSL, podendo-se observar os resultados obtidos na Tabela 6.

- Etapa 1: Nesta etapa foi feita uma análise entre todos os artigos e foram removidos todos os artigos repetidos encontrados;
- Etapa 2: Nesta etapa foi feito uma análise de todos os títulos restantes e foram removidos os que não possuíam títulos relevantes sobre o assunto;
- Etapa 3: Nesta etapa foi realizada uma leitura do resumo de todos os artigos restantes da etapa anterior, sendo possível compreender melhor o objetivo principal de cada artigo. Após a realização dessa leitura foram removidos os artigos considerados irrelevantes;
- Etapa 4: Nesta etapa foi realizada a leitura completa dos artigos restantes, com o propósito de selecionar os artigos mais relevantes encontrados para serem base do referencial teórico do trabalho apresentado.

Tabela 6 – Sumarização dos Resultados

Etapa	Quantidade
Todos os Títulos	3840
Etapa 1 - Repetidos	1487
Etapa 2 - Leitura de Títulos	54
Etapa 3 - Leitura de Resumos	41
Etapa 4 - Leitura Total	10

Elaborado pela autora

A Tabela 7 apresenta os artigos selecionados para responder as questões da revisão sistemática e para serem utilizados como base do referencial teórico deste trabalho.

Tabela 7 – Artigos selecionados

Título	Autor	Ano	Método de Busca
Optical character recognition	Line Eikvil	1993	Google Scholar
Optical character recognition—a survey	S. IMPEDOVO, L. OTTAVIANO, and S. OCCHINEGRO	1991	Google Scholar
Optical Character Recognition Techniques: A Survey	Sukhpreet Singh	2013	Google Scholar
A Survey on Optical Character Recognition System	Noman Islam, Zeeshan Islam, Nazia Noor	2017	Google Scholar
Text recognition in multimedia documents: a study of two neural-based OCRs using and avoiding character segmentation	Khaoula Elagouni, Christophe Garcia, Franck Mamalet, Pascale Sébillot	2014	Springer
Soft computing techniques for optical character recognition systems	Chaudhuri, A., Mandaviya, K., Badelia, P., Ghosh, S.K.	2017	Scopus
OCR as a service: an experimental evaluation of google docs OCR, tesseract, ABBYY finereader, and transym	Tafti, A.P., Baghaie, A., Assefi, M., Arabnia, H.R., Yu, Z., Peissig, P.	2016	Scopus
A Comparative Study on OCR Tools	Carlos A.B de Mello, Rafael D. Lins	1999	Outros
Reconhecimento Ótico de Caracteres e Análise de Sistemas OCR Baseados em Código Aberto	Luiz Fernando Miranda Vieira Lins	2012	Outros
Aplicação do Gerenciamento Eletrônico de Documentos: Estudo de Caso de Escolha de Soluções	Sérgio Rubens Fantini	2001	Outros

3.2.2 Análise das Opções de OCR Disponíveis

Existem vários software e ferramentas disponíveis para realizar OCR atualmente no mercado. Essas ferramentas podem ser caracterizadas como *open-source*, livres ou pagas, local ou online, ter suporte para uma ou várias linguagens, dentre outras.

Mello e Lins (1999), Tafti et al. (2016) e Lins (2012) realizaram comparações entre a acurácia de várias ferramentas de OCR. Feijó et al. (2017) citou ferramentas modernas de OCR existentes atualmente. Fundamentado nesses trabalhos e acrescentando outras ferramentas, foi criada uma lista com as principais soluções de OCR disponíveis atualmente no mercado e suas principais características, que pode ser visualizada na Tabela 8.

Embora existam muitos métodos OCR e conjuntos de ferramentas disponíveis atualmente, muitas são proprietárias e não fornecem versão gratuita para teste. Dessa forma, o trabalho será limitado a um estudo comparativo de onze ferramentas de OCR que são gratuitas ou proprietárias que possuem versão para teste.

Tabela 8 – Comparação dos atributos das ferramentas de OCR

Métodos OCR	Livre	Open-Source	Disponível Online	SDK Disponível	Suporte Multilíngual	Formatos Suportados
Tesseract	Sim	Sim	Não	Sim	Sim	JPEG, PNG, GIF, BMP, TIFF
Google Cloud Vision API	Não	Não	Sim	Sim	Sim	JPEG, PNG8, PNG24, GIF, BMP, WEBP, RAW, ICO, PDF, TIFF
GOOCR	Sim	Sim	Não	Não	Sim	PBM, PCX, PGM, PNM, PPM, TGA
ABBYY FineReader	Não	Não	Sim	Sim	Sim	PDF, TIFF, JPEG, PNG, JBIG2, JPEG 2000, BMP, PCX, GIF, DjVu, XPS, HTML, RTF, TXT, ODT, ODS, ODP
ABBYY Cloud OCR SDK	Não	Não	Sim	Sim	Sim	BMP, DCX, PCX, PNG, JPEG 2000, JPEG, PDF, TIFF, GIF, DjVu, JBIG2S
OCR Document Haven OnDemand	Sim	Não	Sim	Não	Sim	TIFF, JPEG, PNG, GIF, BMP, ICO, PBM, PGM, PPM, PDF
OCR.Space	Não	Não	Sim	Sim	Sim	JPG, PNG, PDF
New OCR	Não	Não	Sim	Sim	Sim	JPEG, PNG, GIF, BMP, TIFF, PDF, DjVu
Microsoft Vision API	Não	Não	Sim	Sim	Sim	JPEG, PNG, GIF, BMP
OCR Web Service	Não	Não	Sim	Sim	Sim	PDF, TIF, JPEG, BMP, PCX, PNG, GIF
SemaMediaData API	Não	Não	Sim	Sim	Sim	JPEG, PNG, TIFF
Cloudmersive	Não	Não	Sim	Sim	Sim	PNG, JPEG
SimpleOCR SDK	Não	Não	Não	Sim	Sim	TIFF
OCR Nicomsoft	Não	Não	Não	Sim	Sim	BMP, JPEG, PNG, TIFF, GIF
OmniPage	Não	Não	Sim	Sim	Sim	PDF, PNG, JPEG, BMP, GIF
Readiris	Não	Não	Não	Sim	Sim	PDF, JPEG, DOC, RTF, JPEG, PNG, PPT
TopOCR	Não	Não	Não	Não	Sim	GIF, PNG, JPG, BMP, TIF, PDF, EPS, PCX
Adobe Acrobat Pro	Não	Não	Não	Sim	Sim	BMP, GIF, JPEG, PCX, JPEG2000, PNG, TIFF, PDF

As ferramentas consideradas neste trabalho são:

- Tesseract¹ versão 4.00
- Google Cloud Vision API² versão 1
- GOOCR³ versão 0.50

¹ Mais informações sobre o Tesseract: <<https://github.com/tesseract-ocr/tesseract>>

² Mais informações sobre Google Cloud Vision: <<https://cloud.google.com/vision/>>

³ Mais informações sobre GOOCR: <<http://jocr.sourceforge.net/>>

- ABBYY Cloud OCR SDK⁴
- OCR Document Haven OnDemand⁵ versão 1
- OCR.Space API⁶
- New OCR API⁷ versão 1
- Microsoft Vision API⁸ versão 1
- OCR Web Service REST API⁹
- SemaMediaData OCR¹⁰ versão 1.0
- Cloudmersive¹¹ versão 1

A seguir é feita uma breve descrição de cada ferramenta de OCR considerada neste estudo.

3.2.2.1 Tesseract

O Tesseract é um mecanismo que foi desenvolvido entre os anos de 1985 e 1994 pela Hewlett Packard (HP), sendo compatível com vários sistemas operacionais. Após isso, em 1996, ele se tornou portátil para Windows e seu código foi reescrito passando da linguagem C para C++ em 1998. Como seu funcionamento apresentou bons resultados, o seu código foi liberado como *software* livre em 2015 (FEIJÓ et al., 2017). Atualmente o Tesseract é oferecido e mantido pela Google sob licença Apache¹².

A versão mais recente é o Tesseract 4.00, que é baseado em rede neural *Long short-term memory*(LSTM) configurado como um reconhecedor de linha de texto. De acordo com o repositório da ferramenta a versão 4.00 oferece uma precisão significativamente maior do que as versões anteriores. O Tesseract tem suporte a unicode(UTF-8) e pode reconhecer mais de 100 idiomas, inclusive o idioma português.

⁴ Mais informações sobre ABBYY Cloud OCR SDK: <<https://www.ocrsdk.com/>>

⁵ Mais informações sobre Haven OnDemand: <<https://dev.havenondemand.com/apis/>>

⁶ Mais informações sobre OCR.Space API: <<https://ocr.space/ocrapi/>>

⁷ Mais informações sobre New OCR API: <<https://www.newocr.com/api/>>

⁸ Mais informações sobre Microsoft Vision API: <<https://azure.microsoft.com/pt-br/services/cognitive-services/computer-vision/>>

⁹ Mais informações sobre OCR Web Service: <<http://www.ocrwebservice.com/>>

¹⁰ Mais informações sobre SemaMediaData API: <<https://semamedia.com/api/documentation/>>

¹¹ Mais informações sobre Cloudmersive: <<https://api.cloudmersive.com/>>

¹² Mais informações sobre a licença Apache: <<http://www.apache.org/licenses/LICENSE-2.0>>

3.2.2.2 Google Cloud Vision API

O Google Cloud Vision é um conjunto de APIs que permite integrar recursos de detecção de imagem em aplicações, incluindo OCR, rotulagem de imagem, localização de objetos, detecção de face e ponto de referência. O recurso de reconhecimento de texto é capaz de detectar uma grande quantidade de idiomas, incluindo a língua portuguesa. O Google Cloud Vision API fornece suporte C#, GO, Java, Node.JS, PHP, Python e Ruby. Possibilita realizar OCR de até 1000 imagens por mês gratuitamente.

3.2.2.3 GOCR

GOCR é um mecanismo gratuito de OCR desenvolvido inicialmente por Jörg Schulenburg e atualmente sob licença GNU Public License. A primeira versão do GOCR foi desenvolvida em 2000. A última versão é a 0.50, que foi lançada em 2013.

3.2.2.4 ABBYY Cloud OCR SDK

ABBYY Cloud OCR SDK é um serviço *web* que oferece reconhecimento de texto através de uma API, que pode ser usada em PHP, Java, Python, C# ou qualquer outra ferramenta de desenvolvimento com suporte a comunicação pela rede. Possui suporte para mais de 200 idiomas, incluindo a língua portuguesa e uma limitação de uso de 50 páginas gratuitamente.

3.2.2.5 OCR Document Haven OnDemand

Haven OnDemand é um conjunto de APIs gratuitas que auxiliam desenvolvedores a construir aplicações no uso de dados não estruturados. Dentre as várias APIs disponíveis, existe a OCR Document API. A API retorna o texto que foi extraído e também informações sobre a localização do texto detectado no arquivo de entrada original. Possui suporte à língua portuguesa.

Segundo o portal Haven OnDemand, quando a resolução da imagem de entrada é muito baixa, o texto fica menos nítido e a imagem pixelada, levando assim a resultados piores. Porém, uma imagem de entrada com alta resolução possui mais detalhes, podendo levar o OCR a interpretar distorções de fundo como possível texto, fazendo com que não se obtenham bons resultados.

3.2.2.6 OCR.Space API

A API do OCR.Space fornece uma maneira de analisar imagens e documentos PDF e obter em um formato JSON os resultados de texto que foram extraídos. Possui suporte a língua portuguesa e uma limitação de 500 requisições gratuitas ao dia e no máximo 25.000 requisições gratuitas por mês, sendo possível se obter uma resposta mais rápida nos planos

PRO disponíveis. Também existe uma versão instalável local com o plano OCR.Space Local Enterprise.

3.2.2.7 New OCR API

A API do New OCR fornece OCR com respostas em JSON através de uma API REST que define um conjunto de funções para as quais pode-se executar solicitações e receber respostas, sendo a interação realizada através do protocolo HTTP. Possui suporte à língua portuguesa e pode reconhecer até 20 página gratuitamente.

3.2.2.8 Microsoft Vision API

Microsoft Vision API é um conjunto de APIs para processar imagens e retornar informações em JSON, dentre as APIs existe uma que fornece serviço de OCR. Possui suporte para língua portuguesa e um período para teste gratuito de 7 dias. Fornece suporte a PHP, C#, Java, JavaScript, Objective C, Python e Ruby.

3.2.2.9 OCR Web Service

OCR Web Service permite, através de uma API REST, integrar serviço OCR em *software*. Possui suporte a 46 idiomas, inclusive português e oferece uma avaliação gratuita de 30 dias, permitindo o processamento de no máximo 25 páginas diariamente. Suporta várias linguagens de programação, como Java, PHP e Python.

3.2.2.10 SemaMediaData OCR

SemaMediaData API é um conjunto de APIs baseadas em RESTful, dentre as quais possui uma API para OCR de imagens. Possui suporte à língua portuguesa e uma limitação de 50.000 requisições gratuitas por mês.

3.2.2.11 Cloudmersive

Cloudmersive é um conjunto de APIs de reconhecimento e processamento de imagem que incluem serviço de OCR para converter imagens digitalizadas em texto. Possui uma limitação de 50.000 requisições por mês e no máximo 2 requisições por segundo.

3.2.3 *Dataset* Utilizado Nos Experimentos

De acordo com Holley (2009), os principais fatores que afetam e interferem na precisão dos resultados de um OCR são a qualidade da fonte original e a resolução da digitalização do arquivo de entrada.

Um dos passos iniciais desse trabalho foi identificar um *dataset* confiável para que os experimentos pudessem ser conduzidos. Nesse passo, além de buscas na internet, foram feitos contatos com diversos pesquisadores da área. Em um desses contatos Ahmad P. Tafti, pesquisador do Marshfield Clinic Research Institute¹³, relatou que um *dataset* como esse está sendo construído, e inclusive solicitou colaboração nesse trabalho. Como não foi encontrado nenhum *dataset* público confiável, optou-se pela criação de um *dataset* próprio para o trabalho. Sendo assim, a criação e manutenção de um *dataset* público e confiável, por si só, já tem grande potencial de contribuição para a comunidade, pois permitiria a congruência dos diversos trabalhos na área de OCR. Não podemos afirmar que o *dataset* construído é o ideal. Porém, sem ele esse trabalho seria inviabilizado.

O foco deste trabalho é o OCR em idioma português brasileiro. Porém, após uma pesquisa realizada, não foram encontrados *datasets* em diferentes resoluções, tamanhos e estilos de fontes em português. Sendo assim, foi necessário a criação de um *dataset* que possui essas variações, com o intuito de analisar como essas variações impactam na acurácia obtida em cada ferramenta de OCR.

O *dataset* montado possui um total de 108 imagens, sendo que esse total possui cinco textos diferentes. Os formatos de imagens utilizados nos testes foram JPG e PBM. Cada texto possui como variação a resolução em 100, 300 e 600 Dots Per Inch(DPI), em português Pontos Por Polegada, e também uma foto do texto. Essas resoluções foram escolhidas, pois nos testes prévios realizados houve pouca diferença nos resultados entre 200 e 400, e 400 e 600 dpi. Sendo assim, convencionou-se experimentar o comportamento das ferramentas OCR para casos de baixa resolução com 100 dpi e alta resolução com 600 dpi. A resolução de 300 dpi também foi considerada por ser a resolução padrão da maioria das ferramentas OCR e também para medir o valor intermediário entre uma resolução baixa(100 dpi) e alta(600 dpi).

A Tabela 9 mostra o total de imagens presentes no *dataset* para cada tipo de imagem.

Tabela 9 – Totalização de imagens do *dataset*

Imagem	100 dpi	300 dpi	600 dpi	Foto	Total
DANFE	1	1	1	1	4
Portaria	1	1	1	1	4
Empenho	1	1	1	1	4
Hino	12	12	12	12	48
Poema	12	12	12	12	48
Total	27	27	27	27	108

¹³ <http://www.marshfieldresearch.org/>

Além da resolução, somente as imagens Hino e Poema possuem variações no tamanho(10, 12 e 14) e estilo(itálico, negrito, negrito e itálico, e normal) da fonte para verificar como essas variações influenciam no resultado do OCR, totalizando 48 imagens de cada. Não foram feitas variações na fonte das imagens DANFE, Empenho e Portaria, porque essas imagens possuem um padrão já definido. Na Tabela 10 observa-se o total de imagens para as variações de fonte do Hino e Poema, totalizando cada uma 48 imagens.

Tabela 10 – Totalização de imagens do *dataset*

Variação de Fonte	Itálico	Negrito	Negrito/Itálico	Normal	Total
10	4	4	4	4	16
12	4	4	4	4	16
14	4	4	4	4	16
Total	12	12	12	12	48

Todas as imagens utilizadas nesse *dataset* são livres e estão disponíveis *online*. O *dataset* completo pode ser acessado através do link: <https://drive.google.com/open?id=1OKerGS_Pdb4qQ_fpb6WqcRVzcJo2wx2N>.

A seguir são detalhadas cada imagem do *dataset* montado.


3.2.3.1 Imagem DANFE

O Documento Auxiliar da Nota Fiscal Eletrônica (DANFE) é usado para acompanhamento do trânsito da mercadoria, sendo uma representação simplificada da Nota Fiscal Eletrônica (BOOKS, 2018). Esse documento tem como função facilitar a consulta da Nota Fiscal Eletrônica através de uma chave de acesso, não sendo, portanto, substituto da Nota Fiscal (NFE, 2018).

DANFE é muito utilizado para a circulação de mercadorias, já que as mercadorias não podem ser transportadas sem que haja o DANFE dessas impressa (EGESTOR, 2018). Por esse motivo, é necessário que seja analisada a acurácia e a eficiência das ferramentas de OCR para esse tipo de imagem.

A imagem utilizada no *dataset* não possui valor fiscal e está disponível gratuitamente em: <<https://bit.ly/2JXaNaR>>. Na Figura 4 pode ser visualizada uma das variações para esse tipo de imagem.

Figura 4 – Imagem DANFE digitalizada na resolução 300dpi no formato JPG

RECEBEMOS DE DISTRIBUIDORA DE AGUAS MINERAIS - AV DAS FONTES, 1777 - 10 ANDAR PARQUE FONTES Sao Paulo-SP OS PRODUTOS E/OU SERVIÇOS CONSTANTES DA NOTA FISCAL ELETRÔNICA INDICADA AO LADO. EMISSÃO: 06/05/2008 VALOR TOTAL: R\$ 20.000.000,00 DESTINATÁRIO: NF-e Associação NF-e		NF-e Nº. 000.000.001 Série 001											
DATA DO RECEBIMENTO	IDENTIFICAÇÃO E ASSINATURA DO RECEBEDOR												
IDENTIFICAÇÃO DO EMITENTE NF-e Associação NF-e Rua Central,100 Fundos Distrito Industrial - 17100-171 Angatuba - SP Fone/Fax: (17) 3302-1717		DANFE Documento Auxiliar da Nota Fiscal Eletrônica 0 - ENTRADA 1 - SAÍDA 0 Nº. 000.000.001 Série 001 Folha 1/1	 CHAVE DE ACESSO 3508 0599 9990 9091 0270 5500 1000 0000 0151 8005 1273 Consulta de autenticidade no portal nacional da NF-e www.nfc.fazenda.gov.br/portal ou no site da Sefaz Autorizadora										
NATUREZA DA OPERAÇÃO Venda a vista		PROTOCOLO DE AUTORIZAÇÃO DE USO											
INSCRIÇÃO ESTADUAL 123456789012	INSCRIÇÃO ESTADUAL DO SUBST. TRIBUT.	CNPJ 99.999.090/9102-70											
DESTINATÁRIO / REMETENTE NOME / RAZÃO SOCIAL DISTRIBUIDORA DE AGUAS MINERAIS		CNPJ / CPF 00.000.000/0001-91	DATA DA EMISSÃO 06/05/2008										
ENDEREÇO AV DAS FONTES, 1777 10 ANDAR		BAIRRO / DISTRITO PARQUE FONTES	CEP 13950-000										
MUNICÍPIO Sao Paulo	UF SP	FONE / FAX (19) 3201-1234	INSCRIÇÃO ESTADUAL 171999999119										
CÁLCULO DO IMPOSTO													
BASE DE CÁLCULO DO ICMS 20.000.000,00	VALOR DO ICMS 18,00	BASE DE CÁLCULO DO ICMS S.T. 0,00	VALOR DO ICMS SUBSTITUIÇÃO 0,00										
VALOR DO PIS 130.000,00	VALOR TOTAL DOS PRODUTOS 20.000.000,00	VALOR DO FRETE 0,00	VALOR DO IPI 400.000,00										
VALOR DO SEGURO 0,00	DESCONTO 0,00	OUTRAS DESPESAS 0,00	VALOR DO COFINS 20.000.000,00										
TRANSPORTADOR / VOLUMES TRANSPORTADOS													
NOME / RAZÃO SOCIAL Distribuidora de Bebidas Fazenda de SP Ltda.	FRETE POR CONTA (0) Emitente	CODIGO ANTT 123456789	PLACA DO VEICULO BX11717										
ENDEREÇO Rua Central 100 - Fundos - Distrito Industrial	MUNICÍPIO SAO PAULO	UF SP	CNPJ / CPF 99.171.171/0001-91										
QUANTIDADE 10000	ESPECIE CAIXA	MARCA LINDOYA	NÚMERO 500										
PESO BRUTO 1.200.000.000,000		PESO LÍQUIDO 1.000.000.000,000											
DADOS DOS PRODUTOS / SERVIÇOS													
CODIGO PRODUTO	DESCRIÇÃO DO PRODUTO / SERVIÇO	NCM/SH	O/CSF	CFOP	UN	QUANT	VALOR UNIT	VALOR TOTAL	B.CALC ICMS	VALOR ICMS	VALOR IPI	ALIQ ICMS	ALIQ IPI
00001	Água Mineral		000	5101	dz	1.000,000,000	1,0000	10.000.000,00	10.000.000,00	1.800.000,00		18	
00002	Água Mineral		000	5101	pack	5.000,000,000	2,0000	10.000.000,00	10.000.000,00	1.800.000,00		18	
DADOS ADICIONAIS		RESERVADO AO FISCO											
INFORMAÇÕES COMPLEMENTARES LOCAL DA RETIRADA DA MERCADORIA : AV PAULISTA,12345 TERREO - CERQUEIRA CESAR SAO PAULO - SP LOCAL DA ENTREGA DA MERCADORIA : AV FARIA LIMA,1500 15 ANDAR - PINHEIROS SAO PAULO - SP Inf. fisco: Nota Fiscal de exemplo NF-eletronica.com													

**SEM VALOR FISCAL
AMBIENTE DE HOMOLOGAÇÃO**

3.2.3.2 Imagem Empenho

Empenhos são usados para registrar as despesas realizadas pela Administração Pública, fazendo com que o Estado seja obrigado a pagar essa despesa ao contratado e sendo esse tipo de documento utilizado por todo órgão que é gerido pelo Estado (BOSELLI, 2017).

A imagem utilizada no *dataset* está disponível gratuitamente em: <<https://bit.ly/2B1kXo4>>. Na Figura 5 pode ser visualizada uma das variações para esse tipo de imagem.

Figura 5 – Imagem Empenho digitalizada na resolução 300dpi no formato JPG

SERVICO PUBLICO FEDERAL
 SIAFI - SISTEMA INTEGRADO DE ADMINISTRACAO FINANCEIRA DO GOVERNO FEDERAL

NOTA DE EMPENHO

PAGINA: 1

EMISSAO : 21Jun18 NUMERO: 2018NE800114 ESPECIE: EMPENHO DE DESPESA
 EMITENTE : 158513/26409 - INST.FED. MINAS GERAIS/CAMPUS FORMIGA
 CNPJ : 10626896/0004-15 FONE: (37) 33228434
 ENDereco : RUA SAO LUIZ GONZAGA,S/N
 MUNICIPIO : 4521 - FORMIGA UF: MG CEP: 35570-000

CREDOR : 03145640/0001-88 - COMERCIAL BARROS GV LTDA - ME
 ENDereco : QUINTILIANO COSTA 313 LOJA: 01; LOURDES
 MUNICIPIO : 4553 - GOVERNADOR VALADARES UF: MG CEP: 35032-630
 TAXA CAMBIO:
 OBSERVACAO / FINALIDADE
 5088/2018 - INFORMAR DADOS BANCARIOS E NUMERO DO EMPENHO NA NF. SE OPTANTE PEL
 SIMPLES, APRESENTAR DECLARACAO DO ANEXO IV DA IN RFB N° 234 DE 2012. PROC OR
 IGM: 05000152017

CLASS : 1 26409 12363208020RL0031 108766 8100000000 339030 151141 LFUNCPO1FOR
 TIPO : ORDINARIO MODAL.LICIT.: PREGAO
 AMPARO: LEI10520 INCISO: PROCESSO: 23209000381201742
 UF/MUNICIPIO BENEFICIADO: MG /
 ORIGEM DO MATERIAL : NACIONAL
 REFERENCIA: ART01 LEI10520/02 NUM. ORIG.:

VALOR ORIGINAL : 171,75
 CENTO E SETENTA E UM REAIS E SETENTA E CINCO CENTAVOS*****

ESPECIFICACAO DO MATERIAL OU SERVICO

ND: 339030 SUBITEM: 25 -MATERIAL P/ MANUTENCAO DE BENS
 SEQ.: 1 QUANTIDADE: 15 VALOR UNITARIO: 11,45
 VALOR DO SEQ. : 171,75

15,0000 UNIDADE
 CÂMARA AR PNEU
 Câmara de ar para carrinho de mão - Câmara de ar 3,25 x 8" para pneu de carrin
 ho de mão
 MARCA: URIAH ITEM DO PROCESSO: 00249 ITEM DE MATERIAL: 000031305

T O T A L : 171,75

 WASHINGTON SANT.DA SILVA
 ORDENADOR

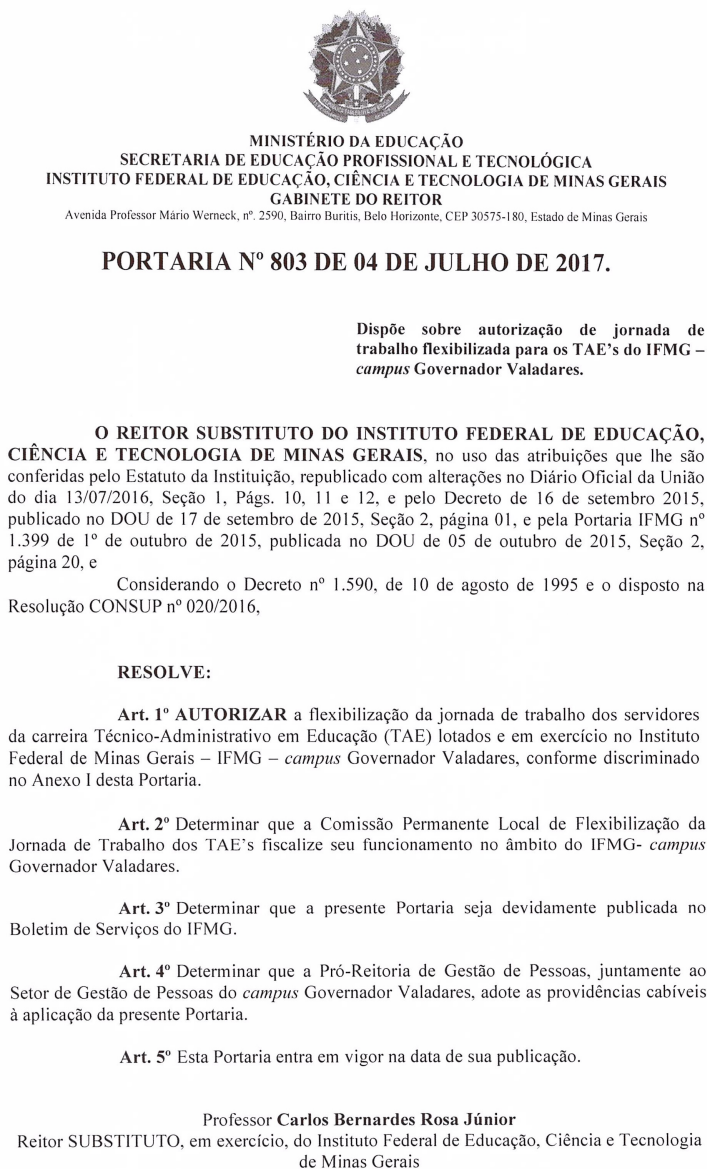
 RINALDO ALVES DE OLIVEIRA
 GESTOR FINANCEIRO

3.2.3.3 Imagem Portaria

Portaria é um documento de caráter administrativo emitido por chefes de órgãos públicos com o objetivo de estabelecer normas referentes à organização, ordem disciplinar, aplicação de decretos ou definir situações, como dispensa, remoção e lotação (DIREITO, 2018).

A imagem utilizada no *dataset* está disponível gratuitamente em: <<https://bit.ly/2RJWOYG>>. Na Figura 6 pode ser visualizada uma das variações para esse tipo de imagem.

Figura 6 – Imagem Portaria digitalizada na resolução 300dpi no formato JPG



3.2.3.4 Imagem Hino

O Hino Nacional Brasileiro é uma canção bastante complexa, seguindo o estilo parnasiano, possuindo uma linguagem rebuscada, inversões sintáticas e palavras pouco usuais atualmente (PORTUGUES, 2018).

A Figura 7 mostra uma das variações para esse tipo de imagem.

Figura 7 – Imagem Hino digitalizada na resolução 300dpi com fonte normal tamanho 12 no formato JPG

HINO NACIONAL BRASILEIRO

Ouviram do Ipiranga as margens plácidas
De um povo heroico o brado retumbante
E o sol da liberdade, em raios fúlgidos
Brillhou no céu da pátria nesse instante
Se o penhor dessa igualdade
Conseguimos conquistar com braço forte
Em teu seio, ó liberdade
Desafia o nosso peito a própria morte!
Ó Pátria amada
Idolatrada
Salve! Salve!
Brasil, um sonho intenso, um raio vívido
De amor e de esperança à terra desce
Se em teu formoso céu, risonho e límpido
A imagem do Cruzeiro resplandece
Gigante pela própria natureza
És belo, és forte, impávido colosso
E o teu futuro espelha essa grandeza
Terra adorada
Entre outras mil
És tu, Brasil
Ó Pátria amada!
Dos filhos deste solo és mãe gentil
Pátria amada
Brasil!
Deitado eternamente em berço esplêndido
Ao som do mar e à luz do céu profundo
Fulguras, ó Brasil, florão da América
Iluminado ao sol do Novo Mundo!
Do que a terra, mais garrida
Teus risonhos, lindos campos têm mais flores
Nossos bosques têm mais vida
Nossa vida no teu seio mais amores
Ó Pátria amada
Idolatrada
Salve! Salve!
Brasil, de amor eterno seja símbolo
O lábaro que ostentas estrelado
E diga o verde-louro dessa flâmula
Paz no futuro e glória no passado
Mas, se ergues da justiça a clava forte
Verás que um filho teu não foge à luta
Nem teme, quem te adora, a própria morte
Terra adorada
Entre outras mil
És tu, Brasil
Ó Pátria amada!
Dos filhos deste solo és mãe gentil
Pátria amada
Brasil!

Fonte: Elaborado pela autora

3.2.3.5 Imagem Poema

O poema Vendaval foi escrito por Fernando Pessoa na primeira fase do Modernismo em Portugal. Na obra o poeta expressa os seus sentimentos, descrevendo suas tristezas e suas inquietações (SANTOS, 2018). O poema possui palavras pouco usuais no dias de hoje, como "Minh'alma" e "pr'onde", e isso faz com que soluções OCR baseadas em aprendizagem de máquina sejam menos suscetíveis a acerto.

A imagem utilizada no *dataset* está disponível gratuitamente em: <http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=16090>. A Figura 8 mostra uma das variações para esse tipo de imagem.

Figura 8 – Imagem Poema digitalizada na resolução 300dpi no formato JPG

Vendaval (Fernando Pessoa)

Ó vento do norte, tão fundo e tão frio,
Não achas, soprando por tanta solidão,
Deserto, penhasco, coval mais vazio
Que o meu coração!
Indômita praia, que a raiva do oceano
Faz louco lugar, caverna sem fim,
Não são tão deixados do alegre e do humano
Como a alma que há em mim!
Mas dura planície, praia atra em fereza,
Só têm a tristeza que a gente lhes vê
E nisto que em mim é vácuo e tristeza
É o visto o que vê.
Ah, mágoa de ter consciência da vida!
Tu, vento do norte, teimoso, iracundo,
Que rasgas os robles — teu pulso divide
Minh'alma do mundo!
Ah, se, como levas as folhas e a areia,
A alma que tenho pudesses levar -
Fosse pr'onde fosse, pra longe da idéia
De eu ter que pensar!
Abismo da noite, da chuva, do vento,
Mar torvo do caos que parece volver -
Porque é que não entras no meu pensamento
Para ele morrer?
Horror de ser sempre com vida a consciência!
Horror de sentir a alma sempre a pensar!
Arranca-me, é vento; do chão da existência,
De ser um lugar!
E, pela alta noite que fazes mais'scura,
Pelo caos furioso que crias no mundo,
Dissolve em areia esta minha amargura,
Meu tédio profundo.
E contra as vidraças dos que há que têm lares,
Telhados daqueles que têm razão,
Atira, já pária desfeito dos ares,
O meu coração!
Meu coração triste, meu coração ermo,
Tornado a substância dispersa e negada
Do vento sem forma, da noite sem termo,
Do abismo e do nada!

Fonte: Elaborado pela autora

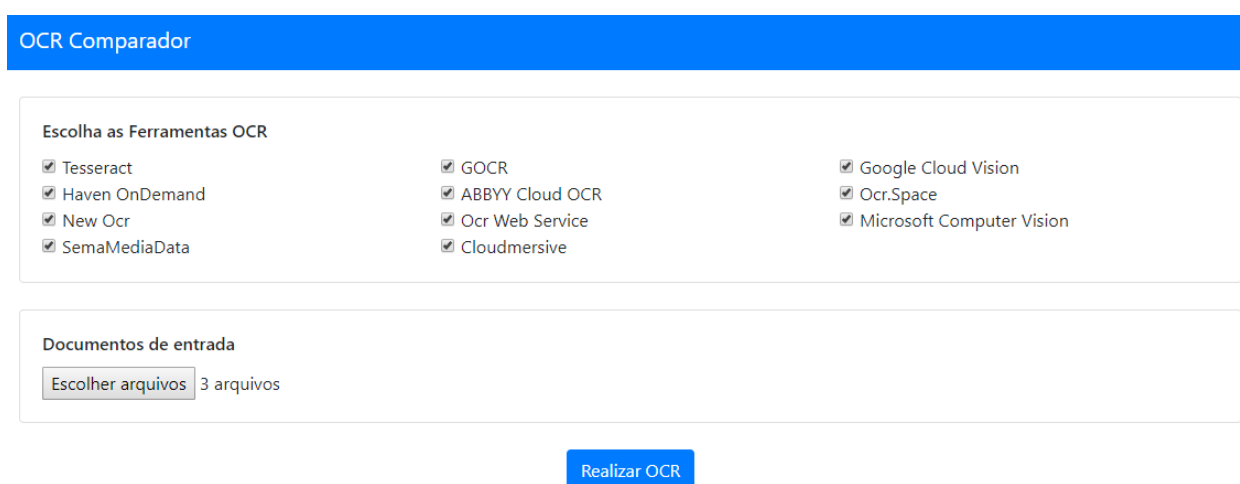
3.2.4 Ambiente de Experimentação

Para facilitar a realização dos testes das ferramentas de OCR foi desenvolvido um ambiente integrado para a extração via OCR de conteúdo de documentos digitalizados e também digitais, oferecendo através de uma plataforma *web* variadas ferramentas OCR, onde o usuário pode submeter vários arquivos de uma só vez e verificar

a qualidade obtida por cada ferramenta OCR considerada neste trabalho. O código-fonte do ambiente *web* pode ser acessado no link: <<https://drive.google.com/open?id=1mpnpklqZHOje5fY4a7IqMrbmNy3LU6ud>>.

A Figura 9 mostra a interface gráfica inicial do ambiente *web* de comparação de ferramentas OCR, onde é possível que o usuário selecione quais ferramentas deseja comparar, sendo possível selecionar todas. Além disso, o usuário pode também selecionar quantos arquivos de entrada desejar.

Figura 9 – Interface gráfica inicial do ambiente de comparação de ferramentas OCR



Ao clicar no botão Realizar OCR é mostrado ao usuário a interface gráfica com o resultado obtido por cada OCR selecionado e por cada arquivo de entrada. Pode-se observar a interface gráfica de resultados na Figura 10.

Figura 10 – Interface gráfica de resultados obtidos por cada OCR e cada arquivo de entrada



O ambiente desenvolvido facilitou a realização dos testes experimentais, pois possibilitou a execução de várias imagens de entrada de uma só vez. Além disso, a visualização dos resultados facilitou a verificação da acurácia obtida em cada arquivo de entrada por cada ferramenta OCR.

3.2.5 Análise Comparativa

Para cada imagem do *dataset* montado, foi calculada a acurácia de cada ferramenta OCR considerada no trabalho, sendo a acurácia a razão entre o número de palavras corretamente reconhecidas e o total de palavras contido na imagem.

$$Acurácia = \frac{\text{número de palavras corretamente reconhecidas}}{\text{número total de palavras da imagem}} \quad (3.1)$$

As duas alternativas levantadas para a verificação da acurácia das ferramentas de OCR foram: desenvolver um algoritmo que verificasse a acurácia de uma ferramenta para uma imagem de entrada ou verificar manualmente a saída do OCR para uma imagem e comparar com a versão original do mesmo texto. No primeiro caso, a implementação de algoritmo desse tipo consiste em um estudo extra, originando até mesmo um trabalho futuro, pois além do desenvolvimento do algoritmo seria necessário provar sua correteza, eficácia e eficiência. Para minimizar o tempo de projeto, neste trabalho optou-se por verificar manualmente a saída de cada OCR para todo o *dataset* de entrada, para então obter uma acurácia confiável de cada método.

De acordo com o portal ABBYY¹⁴ existem duas maneiras principais de verificar a acurácia de um OCR, sendo elas a acurácia em nível de caractere e acurácia em nível de palavra. Neste trabalho foi utilizada a acurácia em nível de palavra e o processo de verificação foi feito de maneira manual.

A análise comparativa realizada neste trabalho consiste na análise dos resultados obtidos por cada ferramenta de OCR para cada imagem do *dataset* e no tempo de execução gasto para cada uma delas.

A medição da acurácia de determinada ferramenta OCR é realizada pegando a saída do OCR para uma imagem e comparando-a com a versão original do mesmo texto. Após isso, é feita uma contagem de quantas palavras foram reconhecidas corretamente. Uma palavra é considerada como reconhecida se todas as suas letras foram corretamente identificadas.

¹⁴ Mais informações sobre o portal ABBYY: <<https://abbyy.technology/>>

4 RESULTADOS E DISCUSSÃO

Nesta seção são apresentados os resultados obtidos nos testes realizado com cada uma das ferramentas de OCR consideradas e para cada imagem do *dataset* montado. Serão analisadas a acurácia e o custo do tempo de execução obtidos para todas as imagens do *dataset*. Nas tabelas a seguir estão destacados em negrito os melhores resultados obtidos em cada caso.

4.1 Imagem DANFE

A Tabela 11 mostra a acurácia obtida por cada ferramenta OCR para a imagem DANFE com a variação das resoluções em 100, 300 e 600 dpi e foto. As ferramentas OCR encontraram dificuldades para reconhecer esse tipo de imagem, porque possui tabelas, textos espalhados por toda a imagem, variação de texto normal, em itálico e negrito e variação no tamanho da fonte, tendo algumas palavras uma fonte pequena e outras, muito grande.

Todas as ferramentas OCR não apresentaram um bom desempenho para a imagem DANFE com a resolução de 100 dpi, sendo que as ferramentas GOCR e SemaMediaDATA não conseguiram reconhecer nenhuma palavra corretamente para essa resolução e obtiveram uma acurácia próxima ou igual a 0 para as outras resoluções. Segundo o portal ABBYY (2018), a baixa resolução de uma imagem acarreta a uma perda da qualidade do resultado do OCR. As ferramentas OCR Web Service e Google Cloud Vision API obtiveram o melhor resultado para a imagem em 100 dpi.

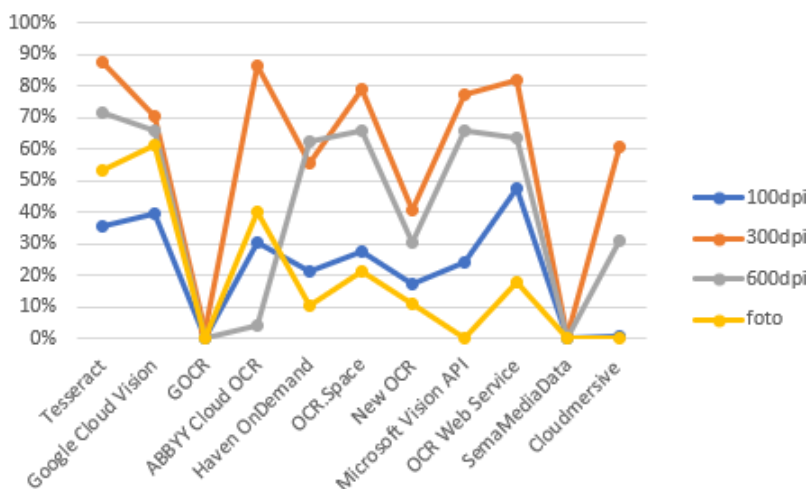
Para a imagem em 300 dpi as ferramentas que obtiveram o melhor resultado foram Tesseract, ABBYY Cloud OCR SDK e OCR Web Service, sendo que cada uma delas teve uma acurácia maior que 80%. Para a maioria das ferramentas não houve uma melhora na acurácia para a resolução da imagem em 600 dpi, isso se deve ao fato de algumas ferramentas serem ajustadas como padrão em 300 dpi, como por exemplo a ABBYY Cloud OCR SDK que possui 300 dpi como padrão de resolução. No geral as ferramentas não conseguiram reconhecer bem a foto da imagem DANFE, sendo que a ferramenta que obteve a melhor acurácia foi o Google Cloud Vision API.

Tabela 11 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem DANFE

Ferramenta OCR	Acurácia			
	100dpi	300dpi	600dpi	foto
Tesseract	35,800%	87,828%	71,599%	53,222%
Google Cloud Vision API	39,379%	70,167%	65,632%	61,098%
GOOCR	0,000%	2,148%	0,239%	0,000%
ABBYY Cloud OCR SDK	30,549%	86,396%	4,296%	40,334%
OCR Document Haven OnDemand	21,480%	55,609%	62,291%	10,501%
OCR.Space API	27,685%	78,998%	66,110%	21,480%
New OCR API	17,422%	40,573%	30,310%	10,979%
Microsoft Vision API	24,344%	77,566%	66,110%	0,000%
OCR Web Service	47,733%	82,100%	63,484%	17,900%
SemaMediaData OCR	0,000%	0,477%	0,000%	0,000%
Cloudmersive API	0,716%	60,859%	30,788%	0,239%

No gráfico da Figura 11 pode-se observar que os resultados obtidos pelas ferramentas foram melhores para a imagem DANFE com resolução de 300 dpi. As imagem com 600 dpi mesmo possuindo uma resolução maior que as de 300 dpi, não tiveram uma melhora nos resultados obtidos, sendo que na maioria das ferramentas o resultado foi pior. A acurácia obtida por todas as ferramentas para as imagens com 100 dpi e foto foram baixas, não ultrapassando 61%.

Figura 11 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem DANFE



Outra informação observada na análise é o tempo de execução de cada ferramenta para cada resolução da imagem. Pode se observar na Tabela 12 a ferramenta OCR Web Service além de ter obtido um dos melhores resultados em termos de acurácia, teve um dos menores custos no tempo de execução do OCR. O Tesseract, apesar de ter obtido bons resultados de acurácia, teve um custo no tempo de execução três vezes maior que o tempo do OCR Web Service.

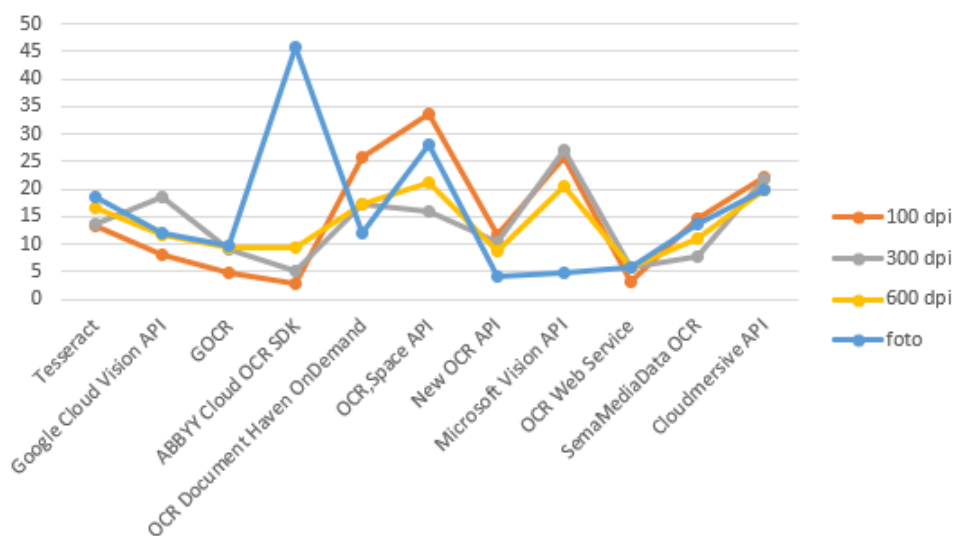
Tabela 12 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem DANFE

Ferramenta OCR	Tempo de Execução(segundos)			
	100dpi	300dpi	600dpi	foto
Tesseract	13,388	13,632	16,692	18,603
Google Cloud Vision API	8,0219	18,716	11,720	11,873
GOOCR	4,649	9,135	9,480	9,683
ABBYY Cloud OCR SDK	2,843	5,130	9,490	45,940
OCR Document Haven OnDemand	25,923	17,211	17,211	12,026
OCR.Space API	33,519	16,003	21,297	28,086
New OCR API	11,616	10,320	8,844	4,292
Microsoft Vision API	25,777	27,030	20,641	4,898
OCR Web Service	3,206	5,712	5,852	5,852
SemaMediaData OCR	14,725	7,726	11,146	13,575
Cloudmersive API	22,041	21,946	20,007	20,007

No gráfico da Figura 12 observa-se que cada ferramenta OCR tem mais facilidade em reconhecer uma imagem em uma determinada resolução. Por exemplo, o ABBYY Cloud OCR SDK obteve um custo pequeno para imagens com a resolução em 100, 300 e 600 dpi, porém, para foto teve um maior custo em relação a todas ferramentas.

A ferramenta GOOCR obteve um custo pequeno para todas as resoluções da imagem, mas não conseguiu atingir uma boa acurácia para nenhuma delas.

Figura 12 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem DANFE



4.2 Imagem Empenho

Na Tabela 13 observa-se as acurácias obtidas por cada ferramenta OCR considerada para as resoluções 100, 300 e 600 dpi e foto da imagem Empenho. Em geral, foram obtidos bons resultados para esse tipo de imagem. As ferramentas GOCR e SemaMediaData OCR não obtiveram bons resultados para esse tipo de imagem em todas as resoluções.

A ferramenta OCR Web Service obteve excelentes resultados para todas as resoluções desta imagem, com acurácia acima de 94% em todos os casos. As ferramentas Tesseract e Google Cloud Vision API também tiveram bons resultados, obtendo uma acurácia de 85,043% para a imagem com resolução de 100 dpi e acima de 90% para as demais resoluções da imagem.

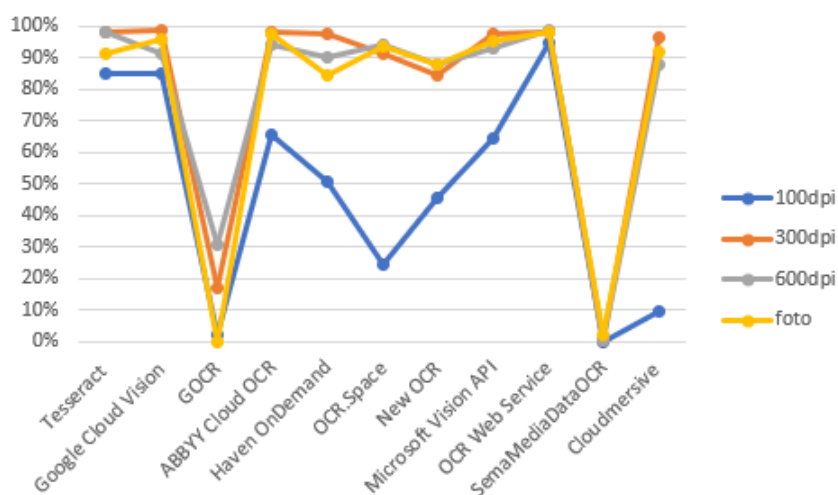
As ferramentas ABBYY Cloud OCR SDK, OCR Document Haven OnDemand, OCR.Space API, Microsoft Vision API e Cloudmersive API obtiveram bons resultados para as resoluções 300, 600 dpi e foto, porém, não tiveram uma boa acurácia para a imagem em 100 dpi.

Tabela 13 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem Empenho

Ferramenta OCR	Acurácia			
	100dpi	300dpi	600dpi	foto
Tesseract	85,043%	97,863%	98,291%	91,453%
Google Cloud Vision API	85,043%	98,718%	91,026%	95,726%
GOOCR	2,137%	17,094%	30,342%	0,000%
ABBYY Cloud OCR SDK	65,385%	98,291%	94,017%	97,436%
OCR Document Haven OnDemand	50,427%	97,436%	90,171%	84,188%
OCR.Space API	24,359%	91,026%	94,017%	93,590%
New OCR API	45,726%	84,615%	88,034%	87,607%
Microsoft Vision API	64,103%	97,436%	93,162%	95,299%
OCR Web Service	94,872%	98,291%	98,718%	98,291%
SemaMediaData OCR	0,000%	0,427%	0,427%	2,137%
Cloudmersive API	9,402%	96,154%	87,607%	91,880%

O gráfico da Figura 13 mostra que para a imagem Empenho as resoluções 300, 600 dpi e foto obtiveram uma boa acurácia. Já a imagem 100 dpi, em geral, não foi bem reconhecida pelas ferramentas.

Figura 13 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem Empenho



O GOOCR obteve o menor custo de tempo de execução para todas as resoluções de imagem, porém, teve uma acurácia baixa para todas elas. A ferramenta OCR.Space API teve um baixo custo de tempo de execução e conseguiu alcançar uma acurácia alta para a imagem em resolução 300, 600 dpi e foto.

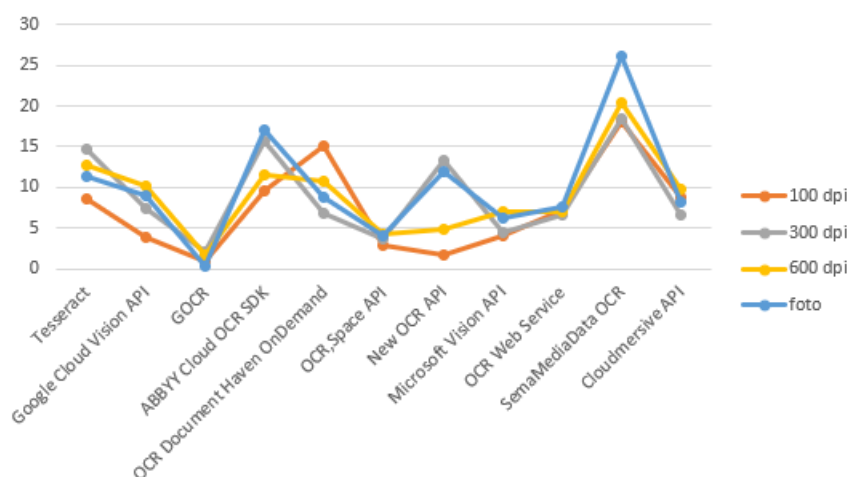
O OCR Web Service, que alcançou os melhores resultados em termos de acurácia para esta imagem, obteve um custo de tempo de execução considerável. Esses resultados podem ser observados na Tabela 14.

Tabela 14 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem Empenho

Ferramenta OCR	Tempo de Execução(segundos)			
	100dpi	300dpi	600dpi	foto
Tesseract	8,607	14,735	12,721	11,337
Google Cloud Vision API	3,917	7,392	10,181	8,916
GOOCR	0,951	2,028	1,660	0,286
ABBYY Cloud OCR SDK	9,629	15,636	11,550	17,140
OCR Document Haven OnDemand	15,117	6,876	10,679	8,691
OCR.Space API	2,910	3,576	4,332	3,976
New OCR API	1,781	13,385	4,898	11,987
Microsoft Vision API	4,003	4,502	6,956	6,184
OCR Web Service	7,280	6,561	6,943	7,554
SemaMediaData OCR	18,044	18,497	20,431	26,171
Cloudmersive API	8,859	6,573	9,748	8,150

Pode-se observar no gráfico da Figura 14 que em geral para a imagem Empenho, as ferramentas obtiveram um menor custo de execução para a imagem com uma menor resolução, aumentando o custo quanto maior a resolução da imagem.

Figura 14 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem Empenho



4.3 Imagem Portaria

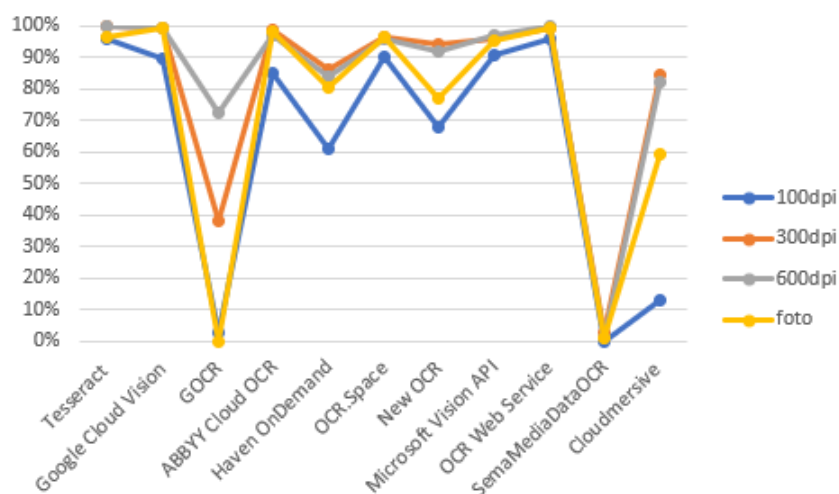
Na Tabela 15 pode-se observar que a maioria das ferramentas OCR obtiveram uma boa acurácia para a imagem Portaria, sendo grande parte dos resultados acima de 90%. Porém, destacaram-se o Tesseract e OCR Web Service, elas tiveram a mesma acurácia para a imagem em 100 e 600 dpi. Ambas tiveram uma acurácia muito próxima para a imagem em 300 dpi, entretanto, para a foto da imagem o OCR Web Service se saiu melhor.

Tabela 15 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem Portaria

Ferramenta OCR	Acurácia			
	100dpi	300dpi	600dpi	foto
Tesseract	95,625%	99,688%	99,688%	96,250%
Google Cloud Vision API	89,375%	99,375%	99,063%	99,063%
GOOCR	2,813%	37,813%	72,188%	0,000%
ABBYY Cloud OCR SDK	85,000%	98,750%	97,188%	98,125%
OCR Document Haven OnDemand	60,938%	86,250%	83,750%	80,313%
OCR.Space API	90,000%	96,563%	95,625%	96,250%
New OCR API	67,813%	94,375%	91,875%	77,188%
Microsoft Vision API	90,938%	95,938%	97,188%	95,313%
OCR Web Service	95,625%	99,375%	99,688%	99,375%
SemaMediaData OCR	0,000%	2,500%	0,625%	0,625%
Cloudmersive API	12,813%	84,375%	82,188%	59,375%

No gráfico da Figura 15 pode-se constatar que em geral as ferramentas obtiveram uma menor acurácia para a imagem em 100 dpi. Já para a imagem em 300 e 600 dpi o resultado foi semelhante, não havendo na maioria dos casos uma melhora com o aumento da resolução. Algumas ferramentas conseguem reconhecer bem a foto da imagem, tendo acurácias semelhantes as obtidas em 300 e 600 dpi.

Figura 15 – Acurácia das ferramentas de OCR para diferentes resoluções da imagem Portaria



As ferramentas OCR.Space API e Microsoft Vision API, como se pode constatar nas Tabelas 16 e 15, obtiveram um baixo tempo de execução e acurácia acima de 90% para todas as resoluções da imagem Portaria.

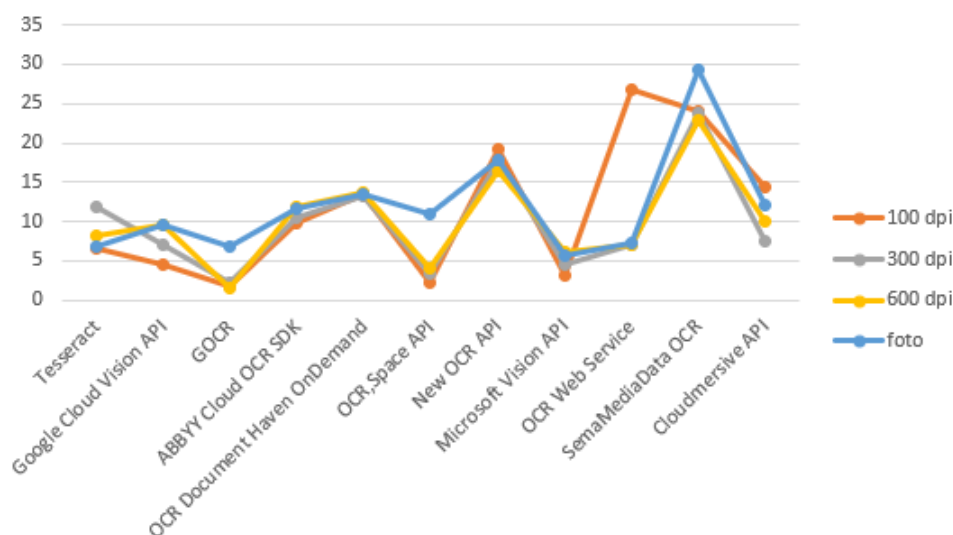
As ferramentas Tesseract e OCR Web Service também tiveram um custo de tempo de execução considerável, exceto para o OCR Web Service para a imagem em 100 dpi, que teve o custo de 26,679 segundos.

Tabela 16 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem Portaria

Ferramenta OCR	Tempo de Execução(segundos)			
	100dpi	300dpi	600dpi	foto
Tesseract	6,598	11,767	8,264	6,922
Google Cloud Vision API	4,425	7,142	9,479	9,474
GOOCR	1,757	2.191	1,506	6,907
ABBY Cloud OCR SDK	9,887	10.444	11,772	11,675
OCR Document Haven OnDemand	13,450	13,313	13,693	13,544
OCR.Space API	2,247	3,417	4,102	11,000
New OCR API	19,190	17,576	16,490	17,755
Microsoft Vision API	3,051	4,607	6,156	5,646
OCR Web Service	26,679	6,977	7,088	7,176
SemaMediaData OCR	24,049	23,823	22,787	29,190
Cloudmersive API	14,456	7,504	10,035	12,032

Para a imagem Portaria, na maior parte dos casos, as ferramentas tiveram um maior custo de execução para foto da imagem. Já para as resoluções de 100, 300 e 600 dpi, em geral, o tempo de execução foi semelhante.

Figura 16 – Tempo de execução das ferramentas de OCR para diferentes resoluções da imagem Portaria



4.4 Imagem Hino

Nessa seção são apresentados os resultados obtidos por cada ferramenta OCR para imagem Hino, contendo variações em 10, 12 e 14 no tamanho da fonte, no estilo da fonte em itálico, negrito, negrito e itálico e normal, e também nas resoluções 100, 300, 600 dpi e foto da imagem.

4.4.1 Fonte Tamanho 10

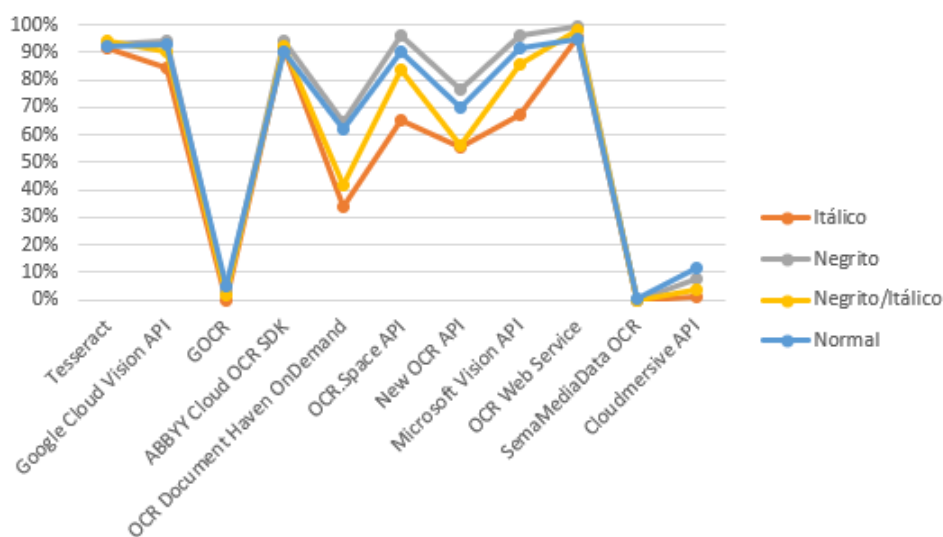
A Tabela 17 mostra que a ferramenta OCR Web Service alcançou o melhor resultado dentre todas as ferramentas analisadas para a imagem Hino com fonte de tamanho 10, resolução de 100 dpi e em todas as variações no estilo de fonte. As ferramentas Tesseract e ABBYY Cloud OCR SDK também tiveram um ótimo resultado na acurácia, tendo todas as variações na fonte uma acurácia maior que 90%.

Tabela 17 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	91,700%	92,885%	94,466%	92,095%
Google Cloud Vision API	84,190%	94,466%	90,119%	92,885%
GOOCR	0,000%	3,162%	1,581%	5,138%
ABBYY Cloud OCR SDK	91,304%	94,071%	92,490%	90,119%
OCR Document Haven OnDemand	33,992%	64,822%	41,502%	62,055%
OCR.Space API	65,613%	96,443%	83,794%	90,514%
New OCR API	55,731%	76,285%	56,126%	69,960%
Microsoft Vision API	67,589%	96,443%	85,375%	91,304%
OCR Web Service	95,257%	99,605%	98,024%	94,862%
SemaMediaDataOCR	0,000%	0,000%	0,000%	0,395%
Cloudmersive API	0,791%	7,510%	3,557%	11,858%

Analisando o gráfico da Figura 17 pode-se observar que as ferramentas tiveram um pouco mais de dificuldade em reconhecer a imagem com fonte em itálico e tiveram mais facilidade em reconhecer as imagens com fonte normal e em negrito.

Figura 17 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte



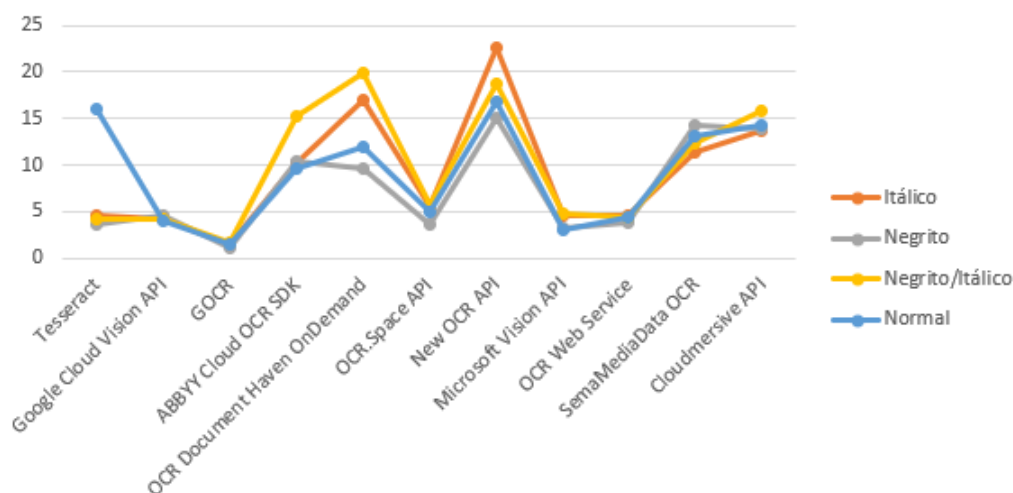
A ferramenta GOOCR o menor custo de tempo de execução para todas as variações de fonte, porém, não conseguiu reconhecer praticamente nenhuma palavra da imagem. Além de ter alcançado a melhor acurácia neste caso, o OCR Web Service também teve um dos menores custos de tempo de execução para todas as variações de fonte.

Tabela 18 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	4,623	3,709	4,193	15,951
Google Cloud Vision API	4,134	4,662	4,292	4,063
GOOCR	1,498	1,115	1,708	1,469
ABBYY Cloud OCR SDK	10,298	10,317	15,188	9,531
OCR Document Haven OnDemand	16,904	9,658	19,974	11,902
OCR.Space API	5,509	3,588	5,752	4,885
New OCR API	22,672	15,142	18,711	16,842
Microsoft Vision API	4,680	3,201	4,744	3,029
OCR Web Service	4,512	3,877	4,312	4,405
SemaMediaDataOCR	11,441	14,188	12,356	13,019
Cloudmersive API	13,698	13,912	15,899	14,207

Em geral, de acordo com o gráfico da Figura 18, as ferramentas de OCR tiveram um menor custo de tempo de execução com as fontes no estilo normal e em negrito e obtiveram um maior custo com as fontes itálico e negrito e itálico.

Figura 18 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte



De acordo com a Tabela 19, para a imagem com resolução de 300 dpi, a maioria das ferramentas tiveram um bom resultado em termos de acurácia. As ferramentas Tesseract, Google Cloud Vision API, ABBYY Cloud OCR, OCR.Space, Microsoft Vision API e OCR Web Service obtiveram uma excelente acurácia, sendo todas maiores que 94% para todas

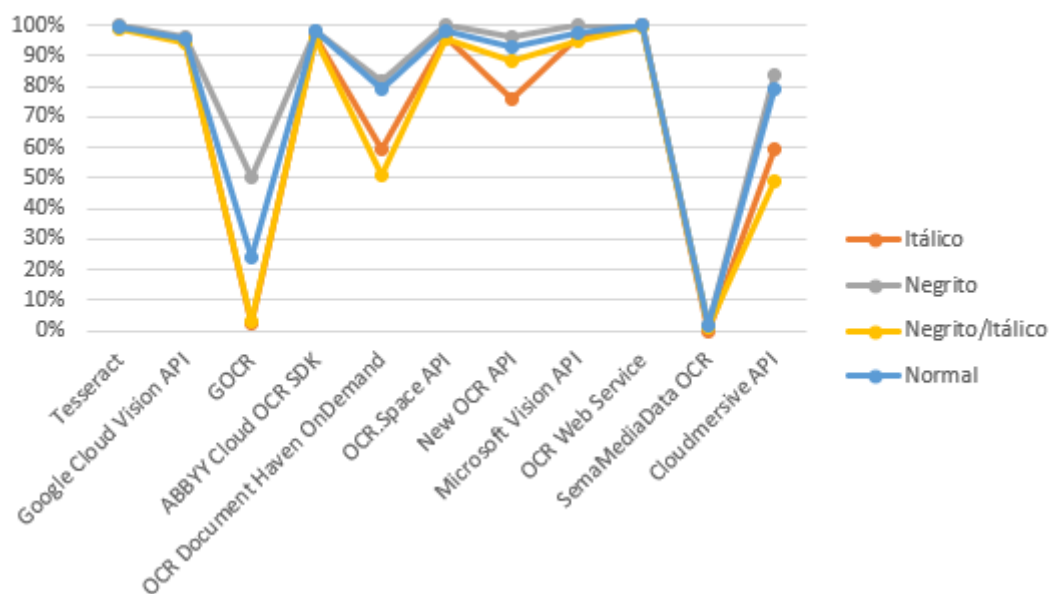
variações de fonte. O OCR Web Service foi a que obteve o melhor resultado, reconhecendo 100% das palavras para as variações itálico, negrito e normal, e próximo de 100% para negrito e itálico.

Tabela 19 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,209%	100,000%	98,814%	99,209%
Google Cloud Vision API	94,862%	96,443%	94,466%	95,652%
GOOCR	2,372%	50,593%	2,767%	24,111%
ABBYY Cloud OCR SDK	96,443%	98,419%	95,652%	98,419%
OCR Document Haven OnDemand	59,684%	81,818%	50,988%	79,051%
OCR.Space API	96,443%	100,000%	95,652%	98,024%
New OCR API	75,889%	96,047%	88,538%	92,885%
Microsoft Vision API	96,047%	100,000%	94,862%	97,628%
OCR Web Service	100,000%	100,000%	99,605%	100,000%
SemaMediaDataOCR	0,000%	1,976%	1,186%	1,581%
Cloudmersive API	59,684%	83,399%	49,012%	79,051%

O gráfico da Figura 19 mostra que para esse tipo de imagem, em geral, as ferramentas OCR tem uma maior facilidade em reconhecer o estilo de fonte negrito e normal. Já as imagens em itálico e negrito e itálico dificultam o reconhecimento.

Figura 19 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte



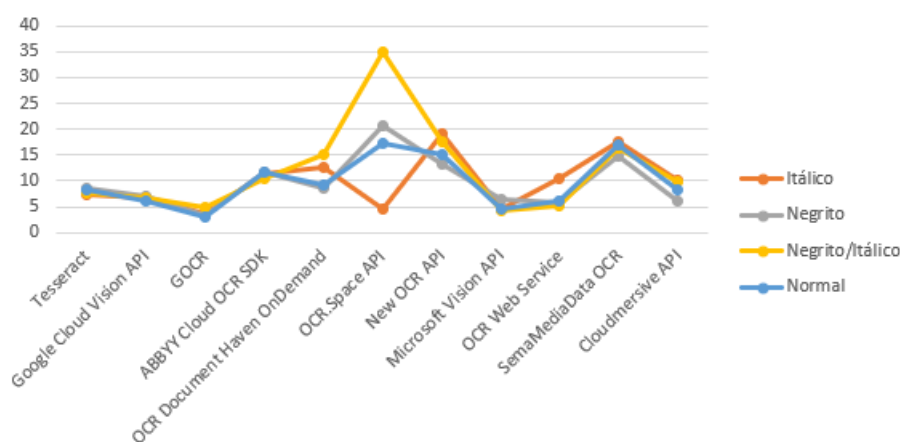
As ferramentas que apresentaram a melhor acurácia para esse tipo de imagem não tiveram um grande custo de tempo de execução, sendo a que teve um menor custo dessas foi a Microsoft Vision API.

Tabela 20 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	7,485	8,451	7,775	8,284
Google Cloud Vision API	6,583	7,137	6,663	6,055
GOOCR	3,576	3,453	4,784	2,970
ABBYY Cloud OCR SDK	11,372	11,555	10,570	11,792
OCR Document Haven OnDemand	12,710	8,715	15,085	9,192
OCR.Space API	4,427	20,799	35,041	17,326
New OCR API	19,212	13,212	17,521	15,131
Microsoft Vision API	4,635	6,389	4,388	4,456
OCR Web Service	10,340	5,692	5,241	5,944
SemaMediaDataOCR	17,438	14,867	16,483	16,881
Cloudmersive API	10,215	6,206	9,863	8,206

Observando o gráfico da Figura 20 pode-se constatar que para esse tipo e resolução de imagem, de forma geral, as ferramentas tiveram praticamente o mesmo custo de tempo de execução para as variações do estilo de fonte.

Figura 20 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte



Para a imagem em resolução de 600 dpi, de acordo com a Tabela 21, pode-se constatar que o OCR Web Service obteve uma maior acurácia para esse tipo de imagem

em praticamente todas as variações de fonte, exceto normal que foi o ABBYY Cloud OCR, porém, a diferença foi pequena.

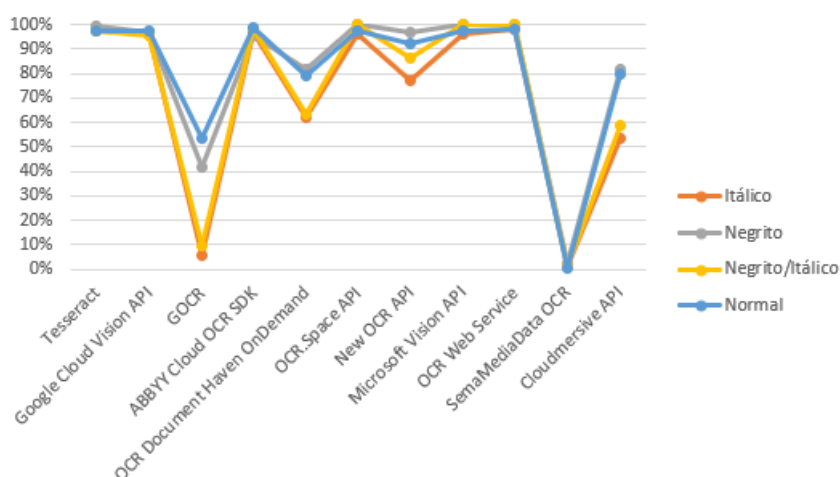
As ferramentas Tesseract, Google Cloud Vision, ABBYY Cloud OCR, OCR.Space API e Microsoft Vision API também obtiveram ótimos resultados, com acurácia acima de 95% para todas variações de fonte.

Tabela 21 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	97,233%	99,605%	97,628%	97,628%
Google Cloud Vision API	96,047%	96,838%	95,652%	97,233%
GOOCR	5,534%	41,897%	9,486%	53,360%
ABBYY Cloud OCR SDK	96,838%	95,257%	98,024%	98,814%
OCR Document Haven OnDemand	62,055%	81,818%	63,241%	79,051%
OCR.Space API	96,047%	100,000%	100,000%	97,628%
New OCR API	77,470%	96,838%	86,166%	92,095%
Microsoft Vision API	96,047%	100,000%	100,000%	97,233%
OCR Web Service	98,419%	100,000%	100,000%	98,419%
SemaMediaDataOCR	0,791%	2,372%	1,186%	0,395%
Cloudmersive API	53,360%	81,423%	58,498%	79,842%

O gráfico da Figura 21 mostra que para esse tipo de imagem, em geral, as ferramentas OCR tem uma maior dificuldade em reconhecer o estilo de fonte itálico e negrito e itálico. Já as imagens em negrito e normal tiveram um melhor reconhecimento.

Figura 21 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte



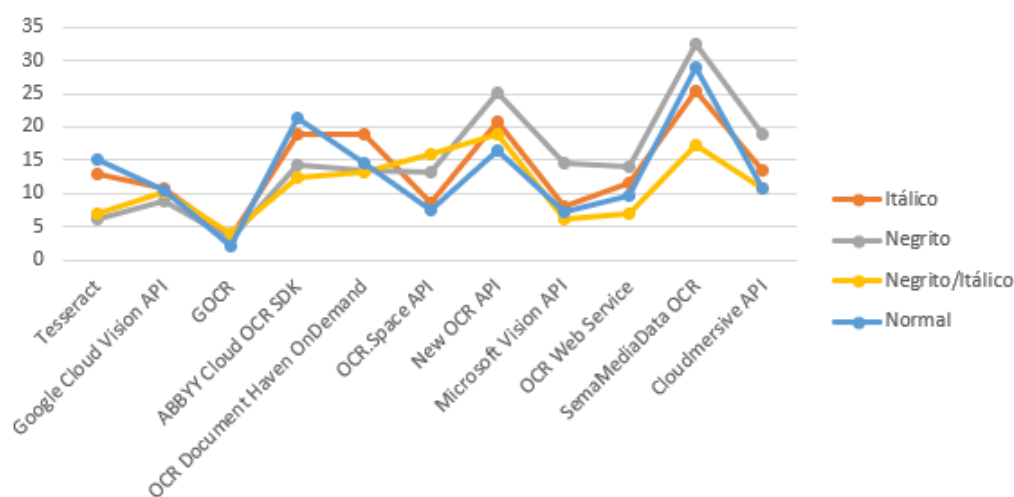
A ferramenta GOCR obteve um menor custo de tempo de execução, no entanto, não conseguiu obter uma acurácia aceitável para nenhuma das variações de fonte. Já as ferramentas que tiveram as melhores acurácias tem um custo parecido de tempo de execução.

Tabela 22 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	13,064	6,120	6,942	15,063
Google Cloud Vision API	10,752	8,761	10,226	10,558
GOCR	3,490	3,079	3,968	2,188
ABBYY Cloud OCR SDK	18,859	14,244	12,506	21,215
OCR Document Haven OnDemand	18,920	13,481	13,331	14,630
OCR.Space API	8,490	13,071	15,821	7,556
New OCR API	20,877	25,136	18,846	16,440
Microsoft Vision API	8,026	14,538	6,154	7,358
OCR Web Service	11,556	13,942	7,056	9,801
SemaMediaDataOCR	25,472	32,461	17,403	28,884
Cloudmersive API	13,603	18,879	10,728	10,763

A maioria das ferramentas obtiveram um maior custo de tempo de execução para a imagem com fonte em negrito e um menor custo para imagem com fonte em negrito e itálico.

Figura 22 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte



A Tabela 23 mostra a acurácia das ferramentas para foto da imagem Hino e pode-se observar que a ferramenta OCR.Space API obteve uma acurácia de 100% para a variação de fonte negrito, negrito e itálico e normal, porém, teve uma acurácia menor que 90% para fonte em itálico.

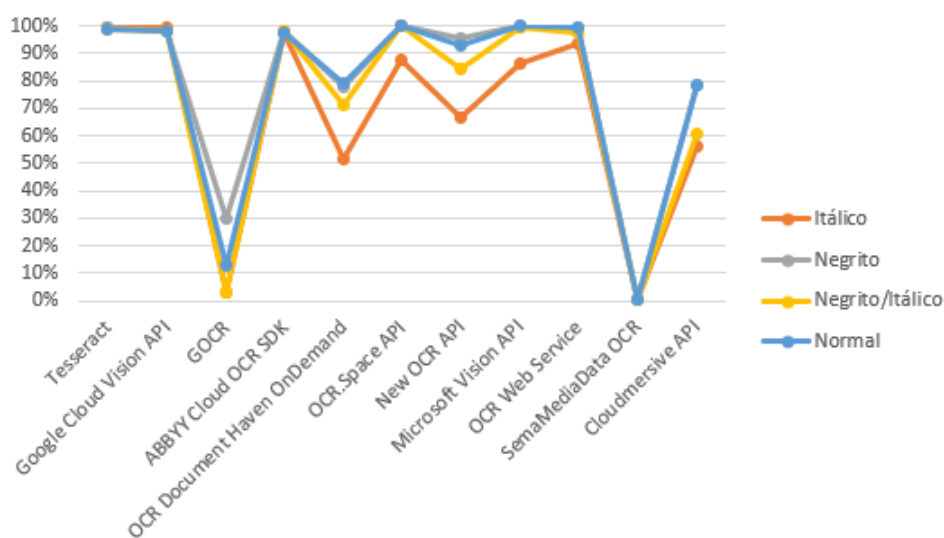
Tabela 23 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 10 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,209%	99,605%	98,814%	98,814%
Google Cloud Vision API	99,209%	98,024%	98,024%	98,419%
GOOCR	3,162%	29,644%	3,162%	12,648%
ABBYY Cloud OCR SDK	97,628%	98,024%	98,024%	97,628%
OCR Document Haven OnDemand	51,779%	77,866%	71,146%	79,051%
OCR.Space API	87,747%	100,000%	100,000%	100,000%
New OCR API	66,798%	95,257%	84,585%	92,885%
Microsoft Vision API	86,561%	100,000%	99,605%	100,000%
OCR Web Service	93,281%	99,209%	97,628%	99,209%
SemaMediaDataOCR	0,395%	0,395%	0,395%	0,395%
Cloudmersive API	56,126%	78,656%	60,474%	78,656%

Já as ferramentas Tesseract, Google Cloud Vision API, ABBYY Cloud OCR SDK e OCR Web Service tiveram uma acurácia para todas variações de fonte maior que 93%.

Novamente, as variações de fonte negrito e fonte normal tiveram uma maior acurácia comparadas as variações itálico e negrito e itálico, como mostra o gráfico da Figura 23.

Figura 23 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 10 e com variações do estilo da fonte



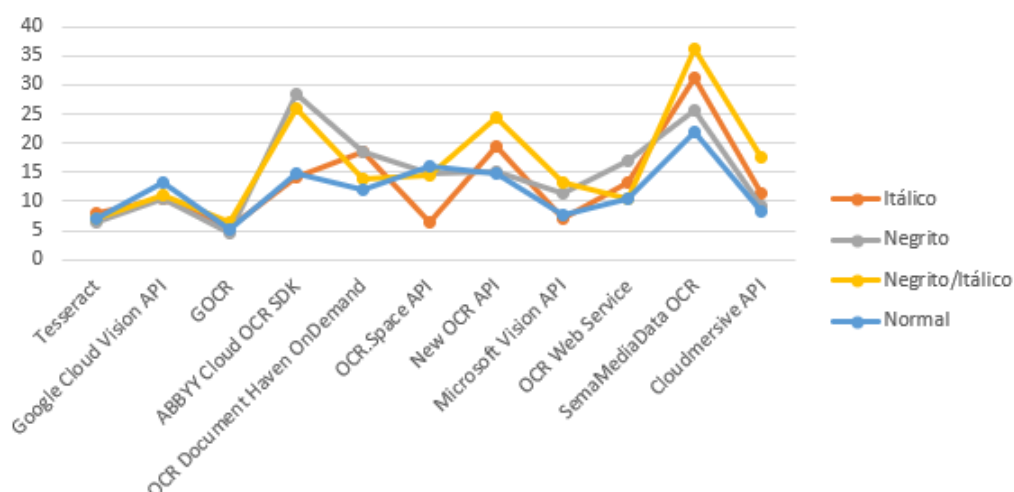
A Tabela 24 mostra que das ferramentas que tiveram uma maior acurácia para foto da imagem Hino, a que apresentou um menor custo de tempo de execução foi o Tesseract.

Tabela 24 – Tempo de execução das ferramentas de OCR em segundos para foto da imagem Hino com fonte tamanho 10 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	7,934	6,273	6,972	6,923
Google Cloud Vision API	10,518	10,374	11,182	13,379
GOOCR	5,614	4,651	6,274	5,244
ABBYY Cloud OCR SDK	14,123	28,520	25,857	14,801
OCR Document Haven OnDemand	18,454	18,535	13,974	11,880
OCR.Space API	6,516	14,864	14,443	15,977
New OCR API	19,323	15,042	24,253	14,918
Microsoft Vision API	7,009	11,421	13,313	7,510
OCR Web Service	13,359	17,000	10,461	10,308
SemaMediaDataOCR	31,222	25,756	36,189	21,788
Cloudmersive API	11,265	9,300	17,512	8,354

A variação de fonte negrito e itálico obteve um maior custo de tempo de execução para esse tipo de imagem e ainda, segundo mostra o gráfico da Figura 24, com a fonte normal as ferramentas obtiveram um menor custo de tempo de execução.

Figura 24 – Tempo de execução das ferramentas de OCR para foto da imagem Hino com fonte tamanho 10 e com variações do estilo da fonte



4.4.2 Fonte Tamanho 12

Observando-se a Tabela 25 notadamente as ferramentas Tesseract e OCR Web Service obtiveram uma acurácia maior que as demais para a imagem Hino com fonte de

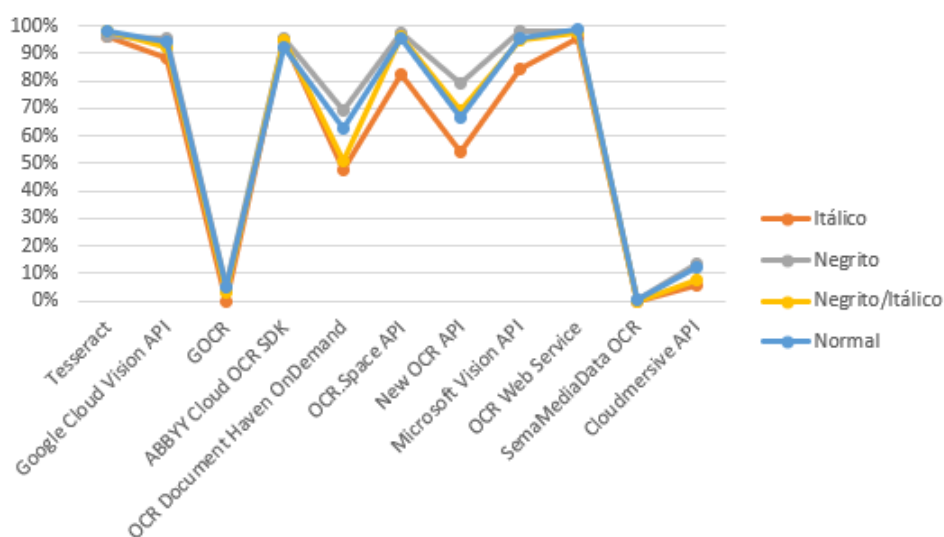
tamanho 12 e resolução de 100 dpi.

Tabela 25 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	96,443%	96,047%	98,419%	98,419%
Google Cloud Vision API	88,538%	95,652%	92,490%	94,071%
GOCR	0,000%	5,929%	2,767%	4,743%
ABBYY Cloud OCR SDK	94,862%	95,257%	94,862%	92,095%
OCR Document Haven OnDemand	47,431%	69,170%	50,988%	62,451%
OCR.Space API	82,213%	97,628%	96,047%	95,257%
New OCR API	54,150%	79,447%	69,565%	66,798%
Microsoft Vision API	84,190%	98,024%	94,862%	95,652%
OCR Web Service	95,257%	98,419%	97,628%	98,814%
SemaMediaDataOCR	0,000%	0,395%	0,000%	0,395%
Cloudmersive API	5,929%	13,834%	7,510%	12,253%

O gráfico da Figura 25 mostra que, de modo geral, as ferramentas para esse tipo de imagem obtiveram maior acurácia para fonte normal e em negrito.

Figura 25 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte



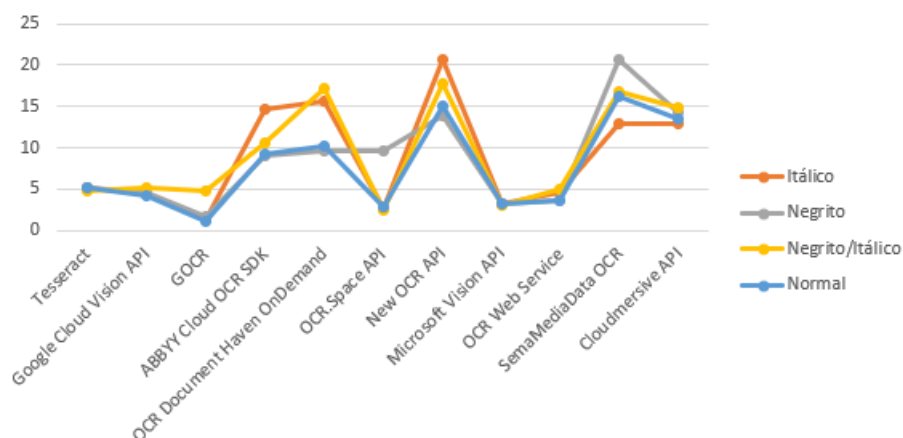
As ferramentas que tiveram uma melhor acurácia possuem custo de tempo de execução parecidos, sendo o custo do OCR WEB Service um pouco menor que do Tesseract, como mostra a Tabela 26.

Tabela 26 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 12 na resolução 100 dpi com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	5,178	5,184	4,695	5,097
Google Cloud Vision API	4,419	4,581	5,187	4,182
GOOCR	1,364	1,616	4,865	1,150
ABBYY Cloud OCR SDK	14,634	9,086	10,657	9,329
OCR Document Haven OnDemand	15,588	9,687	17,167	10,112
OCR.Space API	2,613	9,553	2,538	2,786
New OCR API	20,646	13,913	17,814	14,970
Microsoft Vision API	3,314	3,147	3,119	3,155
OCR Web Service	4,544	3,701	4,940	3,545
SemaMediaDataOCR	12,900	20,722	16,794	16,123
Cloudmersive API	12,992	14,184	14,874	13,539

Quase todas as ferramentas tiveram um menor custo de tempo de execução com a imagem em fonte normal, variando o custo para os outros estilos de fonte, como pode-se observar no gráfico da Figura 26.

Figura 26 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte



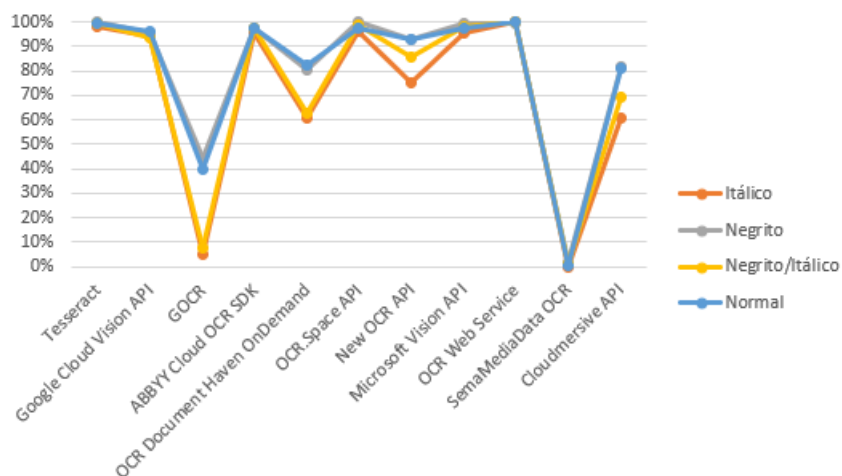
Como mostra a Tabela 27, as ferramentas Tesseract, Google Cloud Vision API, ABBYY Cloud OCR SDK, OCR.Space, Microsoft Vision API e OCR WEB Service obtiveram as melhores acurácias para imagem Hino em 300 dpi, com todas as ferramentas com a acurácia acima de 93%, sendo que o OCR Web Service obteve o melhor desempenho dessas.

Tabela 27 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	98,419%	100,000%	99,605%	99,605%
Google Cloud Vision API	94,071%	95,257%	93,676%	96,047%
GOOCR	4,743%	44,664%	7,510%	39,526%
ABBYY Cloud OCR SDK	95,257%	97,233%	97,628%	97,628%
OCR Document Haven OnDemand	60,474%	80,237%	62,451%	82,213%
OCR.Space API	96,047%	100,000%	98,814%	97,628%
New OCR API	75,494%	92,885%	85,375%	92,885%
Microsoft Vision API	95,257%	99,605%	98,419%	97,233%
OCR Web Service	100,000%	99,209%	100,000%	100,000%
SemaMediaDataOCR	0,000%	1,976%	1,186%	0,395%
Cloudmersive API	60,870%	81,818%	69,170%	81,028%

As acurácias obtidas foram maiores para a imagem com fonte normal e em negrito. Já para fontes em itálico e negrito e itálico a acurácia foi menor.

Figura 27 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte



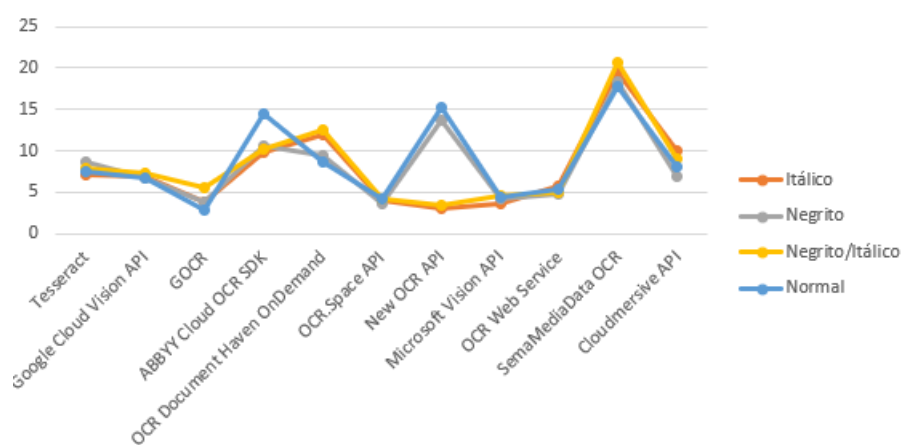
Segundo a Tabela 28, das ferramentas que apresentaram maior acurácia, as que tiveram um menor custo de execução foi Microsoft Vision API e o OCR. Space API.

Tabela 28 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	7,092	8,700	7,937	7,547
Google Cloud Vision API	6,915	6,631	7,318	6,667
GOOCR	3,821	3,807	5,600	2,841
ABBYY Cloud OCR SDK	9,827	10,629	10,284	14,551
OCR Document Haven OnDemand	11,983	9,369	12,481	8,712
OCR.Space API	4,024	3,656	4,117	4,189
New OCR API	3,111	13,607	3,404	15,308
Microsoft Vision API	3,591	4,263	4,537	4,445
OCR Web Service	5,762	4,849	5,004	5,349
SemaMediaDataOCR	19,536	18,387	20,686	17,793
Cloudmersive API	10,089	6,899	9,081	8,140

Em geral, as ferramentas de OCR tiveram custo de tempo de execução parecidos para as variações de fonte, com algumas exceções, como por exemplo o New OCR API, que teve um custo maior para reconhecer a fonte em negrito e normal.

Figura 28 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte



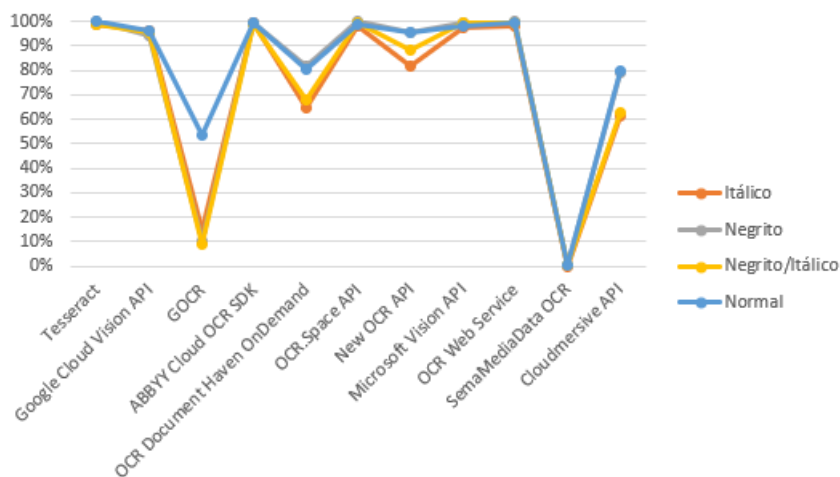
A ferramenta Tesseract obteve uma maior acurácia para essa imagem em 600 dpi, seguida do OCR Web Service, ABBYY Cloud OCR SDK e OCR.Space, que também obtiveram uma acurácia próxima de 100% em todas variações de fonte, como mostra a Tabela 29.

Tabela 29 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	98,814%	100,000%	98,814%	100,000%
Google Cloud Vision API	96,443%	94,466%	95,257%	96,443%
GOOCR	14,625%	10,277%	8,696%	53,360%
ABBYY Cloud OCR SDK	99,605%	99,209%	98,814%	99,209%
OCR Document Haven OnDemand	64,427%	81,818%	67,984%	80,237%
OCR.Space API	98,024%	100,000%	99,605%	98,814%
New OCR API	81,818%	95,652%	88,142%	95,652%
Microsoft Vision API	97,628%	99,209%	99,209%	98,419%
OCR Web Service	98,024%	100,000%	99,605%	99,605%
SemaMediaDataOCR	0,000%	0,395%	0,395%	0,395%
Cloudmersive API	61,265%	79,447%	62,451%	79,842%

Mais uma vez, as ferramentas de OCR obtiveram um melhor resultado para fontes normais e em negrito e um pior resultado para fontes em itálico e negrito e itálico.

Figura 29 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte



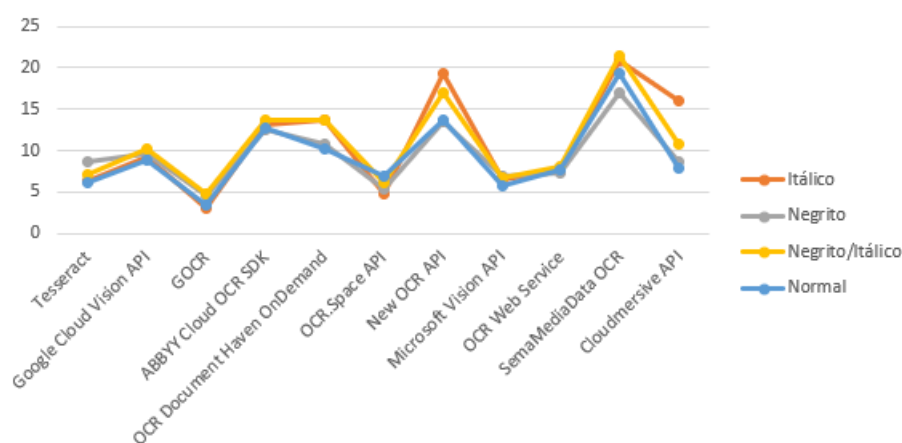
O custo de tempo de execução das ferramentas que obtiveram uma maior acurácia foi semelhante, tendo o OCR.Space API e Microsoft Vision API os menores custos.

Tabela 30 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 12 na resolução 600 dpi com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	6,335	8,640	7,017	6,167
Google Cloud Vision API	9,248	9,560	10,296	8,908
GOOCR	3,038	4,532	4,684	3,389
ABBYY Cloud OCR SDK	13,023	12,456	13,704	12,689
OCR Document Haven OnDemand	13,755	10,881	13,664	10,167
OCR.Space API	4,835	5,359	6,213	6,842
New OCR API	19,404	13,542	17,084	13,641
Microsoft Vision API	6,539	6,884	6,720	5,653
OCR Web Service	7,528	7,286	8,069	7,714
SemaMediaDataOCR	20,792	17,049	21,526	19,379
Cloudmersive API	15,964	8,743	10,866	7,828

Observando-se o gráfico da Figura 30, a maioria das ferramentas tiveram um menor custo de execução para fontes normais e em negrito. Já para as fontes itálico e negrito e itálico tiveram um maior custo.

Figura 30 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte



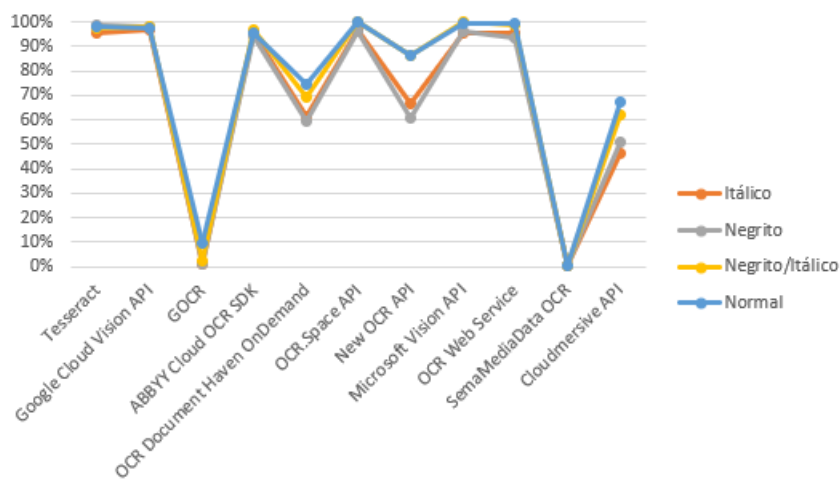
A Tabela 31 mostra que para foto da imagem Hino com fonte de tamanho 12, as ferramentas que tiveram um melhor resultado foram OCR.Space API, Tesseract, Google Cloud Vision API, ABBYY Cloud OCR SDK, Microsoft Vision API e OCR Web Service, sendo que todas tiveram uma acurácia maior que 93%.

Tabela 31 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 12 com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	95,652%	98,814%	97,233%	98,419%
Google Cloud Vision API	96,838%	98,419%	98,024%	97,628%
GOOCR	1,186%	0,791%	2,372%	9,486%
ABBYY Cloud OCR SDK	95,257%	94,466%	96,838%	95,652%
OCR Document Haven OnDemand	61,265%	59,289%	69,170%	74,308%
OCR.Space API	96,838%	96,443%	100,000%	100,000%
New OCR API	66,403%	60,474%	86,166%	86,166%
Microsoft Vision API	95,652%	96,443%	100,000%	99,605%
OCR Web Service	95,257%	93,281%	98,814%	99,209%
SemaMediaDataOCR	0,395%	0,395%	0,395%	0,395%
Cloudmersive API	46,640%	50,988%	62,055%	67,589%

O gráfico da Figura 31 mostra que as ferramentas alcançaram uma acurácia um pouco maior para fontes normais. Porém, a diferença para as outras variações de fonte não foi muita.

Figura 31 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 12 e com variações do estilo da fonte



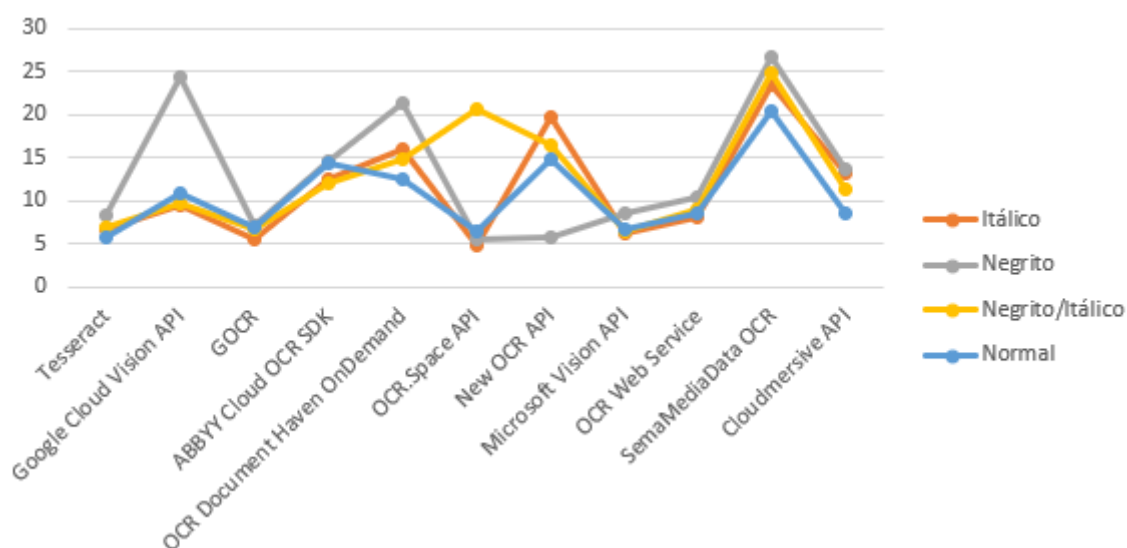
Segundo a Tabela 32, das ferramentas que tiveram um melhor resultado no OCR, o Tesseract, Microsoft Vision API e OCR Web Service obtiveram um menor custo de tempo de execução.

Tabela 32 – Tempo de execução das ferramentas de OCR em segundos para foto da imagem Hino com fonte tamanho 12 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	6,635	8,295	6,855	5,804
Google Cloud Vision API	9,346	24,405	9,705	10,780
GOCR	5,388	7,178	6,565	6,785
ABBYY Cloud OCR SDK	12,520	14,629	12,117	14,304
OCR Document Haven OnDemand	15,935	21,211	14,719	12,533
OCR.Space API	4,883	5,584	20,537	6,466
New OCR API	19,721	5,823	16,522	14,837
Microsoft Vision API	6,105	8,636	6,489	6,746
OCR Web Service	8,120	10,325	8,968	8,619
SemaMediaDataOCR	23,336	26,567	24,816	20,374
Cloudmersive API	13,223	13,632	11,204	8,467

A variação de fonte que apresentou um maior custo de tempo de execução na maioria das ferramentas foi negrito. Já a que apresentou um menor custo foi a fonte normal.

Figura 32 – Tempo de execução das ferramentas de OCR para foto da imagem Hino com fonte tamanho 12 e com variações do estilo da fonte



4.4.3 Fonte Tamanho 14

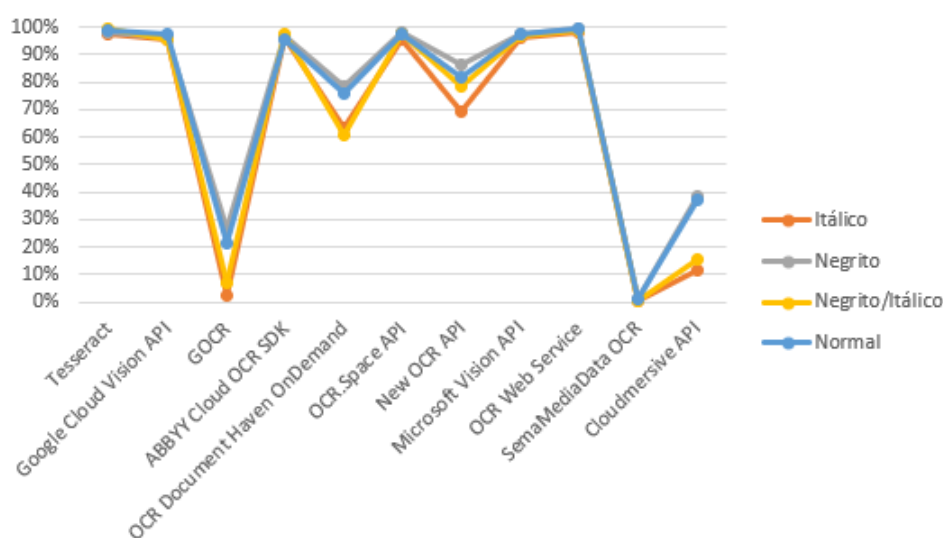
A Tabela 33 mostra que as ferramentas OCR Web Service, Tesseract, Google Cloud Vision API, ABBYY Cloud OCR SDK, OCR.Space API e Microsoft Vision API obtiveram uma acurácia maior que 95% para todas variações de fonte.

Tabela 33 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	97,233%	98,024%	99,209%	98,814%
Google Cloud Vision API	95,652%	96,047%	95,652%	97,628%
GOOCR	2,372%	26,877%	7,115%	21,739%
ABBYY Cloud OCR SDK	96,443%	96,838%	97,233%	95,652%
OCR Document Haven OnDemand	63,636%	78,656%	60,474%	75,889%
OCR.Space API	95,652%	98,024%	97,628%	97,628%
New OCR API	69,565%	86,166%	78,261%	81,818%
Microsoft Vision API	96,443%	97,233%	96,838%	97,233%
OCR Web Service	98,024%	99,209%	98,814%	99,209%
SemaMediaDataOCR	0,395%	1,186%	0,395%	1,186%
Cloudmersive API	11,462%	38,735%	15,415%	37,154%

As ferramentas de OCR alcançaram uma maior acurácia para a imagem com resolução de 100 dpi com fonte normal e em negrito, como mostra o gráfico da Figura 33.

Figura 33 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte



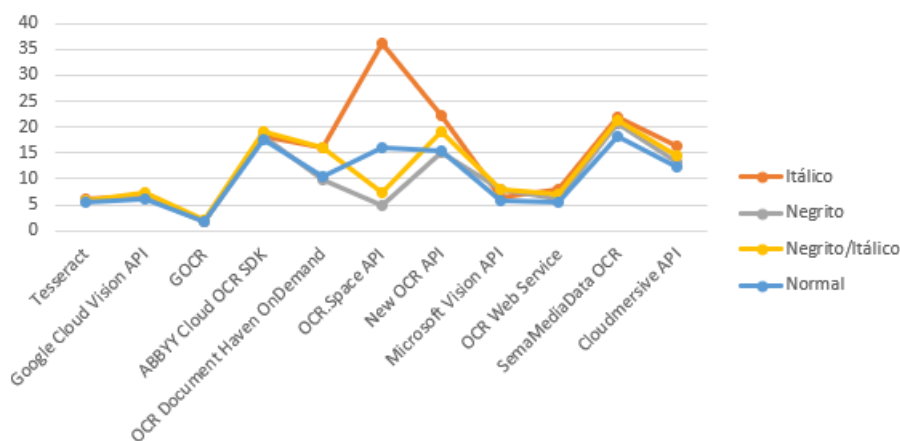
Observando-se a Tabela 34 constata-se que das ferramentas com maior acurácia para esse tipo de imagem, o Tesseract, Google Cloud Vision API, Microsoft Vision API e OCR Web Service tiveram um custo de tempo de execução baixo e semelhante.

Tabela 34 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	6,113	5,883	5,744	5,455
Google Cloud Vision API	6,800	6,422	7,196	6,097
GOOCR	1,648	1,734	2,145	1,797
ABBYY Cloud OCR SDK	18,183	18,100	18,972	17,560
OCR Document Haven OnDemand	16,161	9,712	15,977	10,440
OCR.Space API	36,067	4,824	7,319	15,968
New OCR API	22,126	15,046	19,184	15,407
Microsoft Vision API	6,258	7,868	8,070	5,637
OCR Web Service	7,936	6,034	7,085	5,361
SemaMediaDataOCR	22,034	20,559	21,139	18,281
Cloudmersive API	16,199	13,366	14,620	12,245

A maioria das ferramentas tiveram um custo de tempo de execução semelhante para as diferentes variações de fonte. Porém, a ferramenta OCR.Space API difere das demais, tendo se obtido custos diferentes para as diferentes variações de fonte.

Figura 34 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte



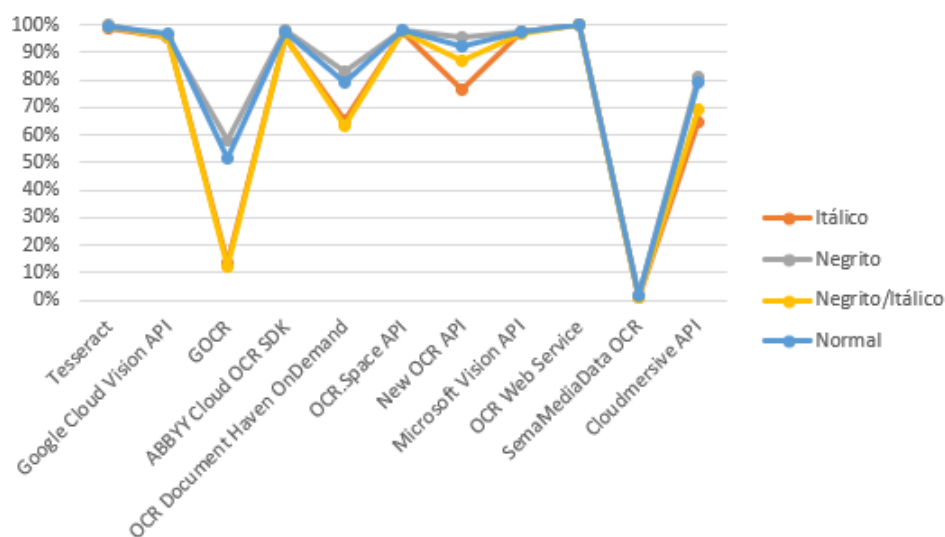
De acordo com a Tabela 35, para esse tipo de imagem com a resolução de 300 dpi, a ferramenta OCR Web Service obteve uma acurácia de 100% para todas as variações do estilo de fonte.

Tabela 35 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	98,814%	100,000%	99,209%	99,605%
Google Cloud Vision API	95,257%	95,652%	95,652%	96,838%
GOOCR	13,439%	58,103%	12,253%	51,779%
ABBYY Cloud OCR SDK	95,652%	98,024%	95,652%	97,233%
OCR Document Haven OnDemand	65,613%	83,004%	63,241%	79,447%
OCR.Space API	97,233%	98,024%	97,628%	98,024%
New OCR API	76,285%	95,257%	86,957%	92,095%
Microsoft Vision API	97,233%	97,233%	96,838%	97,628%
OCR Web Service	100,000%	100,000%	100,000%	100,000%
SemaMediaDataOCR	1,186%	1,581%	1,186%	1,581%
Cloudmersive API	64,427%	81,028%	69,170%	79,447%

A acurácia para imagens com fonte normal e negrito foi um pouco maior do que para as demais variações de fonte.

Figura 35 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte



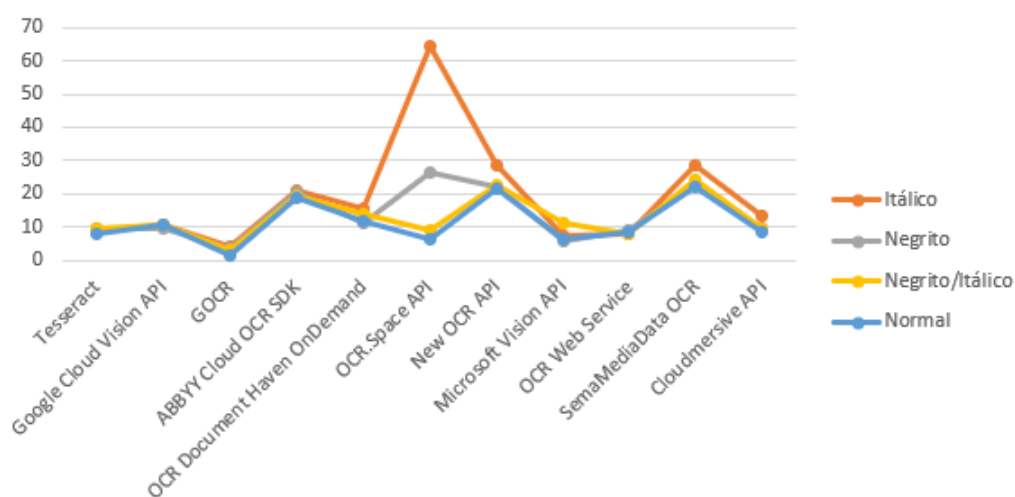
O custo de tempo de execução da ferramenta OCR Web Service foi um dos menores em relação a todas as ferramentas, como se pode observar na Tabela 36.

Tabela 36 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 14 na resolução 300 dpi com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	9,151	9,468	9,793	7,810
Google Cloud Vision API	10,757	9,850	10,768	10,574
GOOCR	4,136	3,509	2,997	1,696
ABBYY Cloud OCR SDK	21,002	20,443	19,468	18,694
OCR Document Haven OnDemand	15,333	11,121	13,747	11,891
OCR.Space API	64,373	26,509	9,071	6,099
New OCR API	28,540	21,872	22,440	21,716
Microsoft Vision API	7,424	5,712	11,055	6,432
OCR Web Service	8,134	9,202	8,179	8,646
SemaMediaDataOCR	28,563	22,696	24,067	22,201
Cloudmersive API	13,350	9,355	9,433	8,742

Novamente, a maior parte das ferramentas tiveram um custo de tempo de execução semelhante para as diferentes variações de fonte. Porém, mais uma vez, a ferramenta OCR.Space API difere das demais, tendo se obtido custos diferentes para as diferentes variações de fonte.

Figura 36 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte



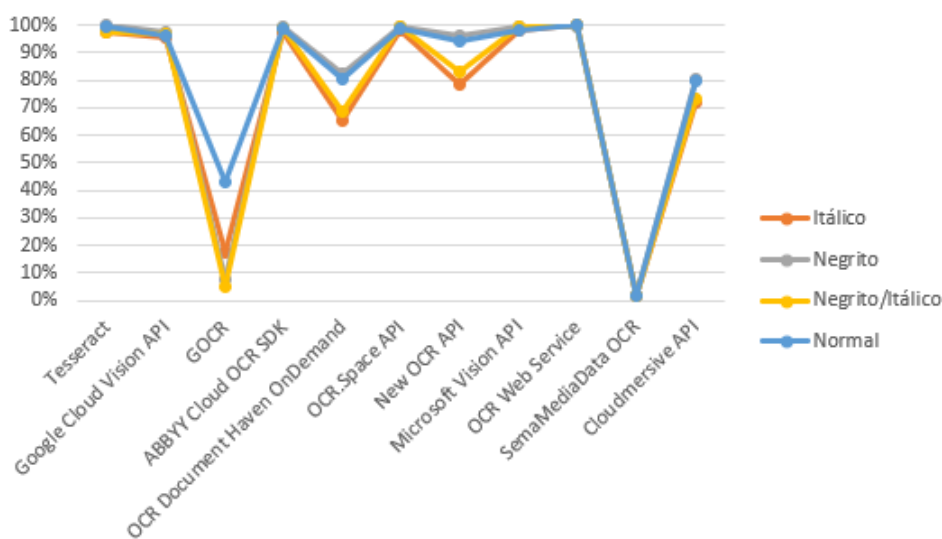
Para esse tipo de imagem em 600 dpi, novamente a ferramenta OCR Web Service obteve uma maior acurácia, seguida das ferramentas Microsoft Vision API, Tesseract, OCR.Space API e Google Cloud Vision API, que tiveram acurácias semelhantes.

Tabela 37 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	97,628%	100,000%	97,628%	99,605%
Google Cloud Vision API	95,652%	97,233%	96,838%	96,047%
GOOCR	17,391%	7,510%	4,743%	43,083%
ABBYY Cloud OCR SDK	97,628%	99,605%	98,814%	98,814%
OCR Document Haven OnDemand	65,613%	82,213%	68,775%	80,632%
OCR.Space API	98,024%	99,209%	99,605%	98,814%
New OCR API	78,656%	96,047%	83,004%	94,071%
Microsoft Vision API	98,024%	99,209%	99,209%	98,419%
OCR Web Service	100,000%	99,209%	100,000%	100,000%
SemaMediaDataOCR	1,581%	1,581%	1,581%	1,581%
Cloudmersive API	71,937%	80,237%	73,518%	79,842%

Como na maioria dos casos, as imagens com fonte normal e em negrito apresentaram uma acurácia um pouco maior do que das outras variações.

Figura 37 – Acurácia das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte



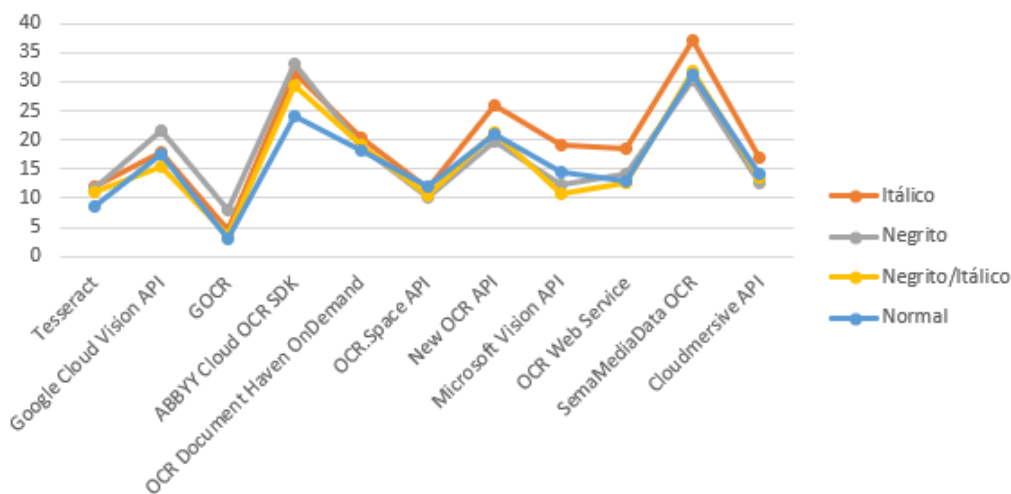
Das ferramentas com maior acurácia, o Tesseract e OCR.Space API obtiveram um custo de tempo de execução um pouco menor que as outras.

Tabela 38 – Tempo de execução das ferramentas de OCR em segundos para imagem Hino com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	11,834	11,560	11,047	8,723
Google Cloud Vision API	17,764	21,449	15,252	17,602
GOOCR	4,652	8,007	3,745	2,957
ABBYY Cloud OCR SDK	31,520	32,935	29,392	24,214
OCR Document Haven OnDemand	20,476	18,658	19,096	18,115
OCR.Space API	11,824	10,193	10,406	11,983
New OCR API	26,056	19,737	21,275	20,953
Microsoft Vision API	19,252	12,272	10,861	14,357
OCR Web Service	18,575	14,159	12,669	12,842
SemaMediaDataOCR	36,954	30,427	31,746	31,256
Cloudmersive API	16,849	12,682	13,421	14,027

O custo de tempo de execução de cada ferramenta foi semelhante para cada variação no estilo de fonte.

Figura 38 – Tempo de execução das ferramentas de OCR para imagem Hino com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte



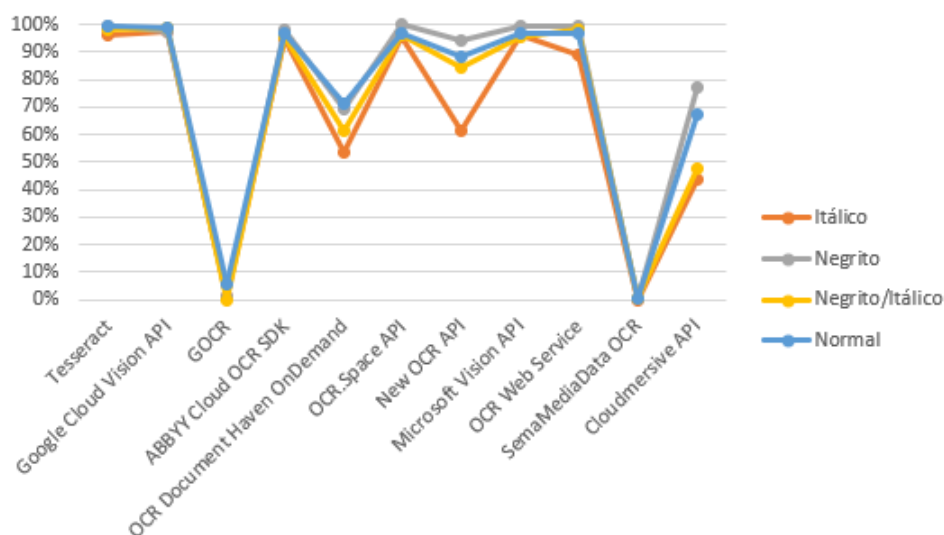
Para foto da imagem Hino com fonte tamanho 14, as ferramentas que tiveram maior acurácia, sendo para todas as variações maior que 95% foram Tesseract, Google Cloud Vision API, ABBYY Cloud OCR SDK, OCR.Space API e Microsoft Vision API.

Tabela 39 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 14 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	96,443%	99,605%	98,419%	99,605%
Google Cloud Vision API	97,233%	97,628%	98,814%	98,814%
GOOCR	1,186%	1,186%	0,000%	5,929%
ABBYY Cloud OCR SDK	94,862%	98,419%	95,257%	96,838%
OCR Document Haven OnDemand	53,360%	69,565%	61,265%	71,146%
OCR.Space API	95,652%	100,000%	96,443%	96,838%
New OCR API	61,660%	94,071%	84,585%	88,538%
Microsoft Vision API	96,047%	99,605%	95,652%	96,838%
OCR Web Service	88,933%	99,209%	98,419%	96,838%
SemaMediaDataOCR	0,000%	0,395%	0,395%	0,395%
Cloudmersive API	43,478%	77,075%	47,431%	67,194%

Como mostra o gráfico da Figura 39, a acurácia obtida pelas ferramentas é um pouco maior para fontes normais e em negrito.

Figura 39 – Acurácia das ferramentas de OCR para foto da imagem Hino com fonte tamanho 14 e com variações do estilo da fonte



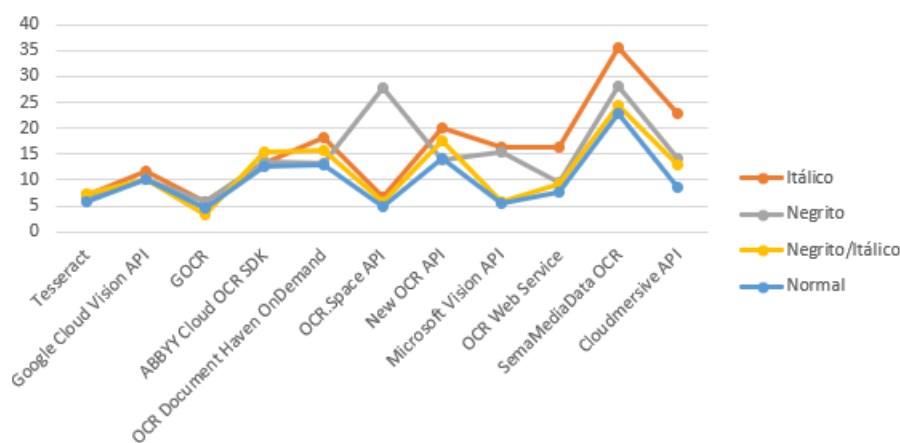
A ferramenta Tesseract, das que obtiveram maior acurácia, foi a que teve menor custo de tempo de execução.

Tabela 40 – Tempo de execução das ferramentas de OCR em segundos para foto imagem Hino com fonte tamanho 14 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	7,066	6,406	7,260	5,722
Google Cloud Vision API	11,804	10,124	10,187	10,018
GOOCR	5,748	5,876	3,249	4,681
ABBYY Cloud OCR SDK	13,295	13,645	15,498	12,667
OCR Document Haven OnDemand	18,302	13,186	15,602	12,964
OCR.Space API	6,571	27,872	5,903	4,764
New OCR API	20,169	13,733	17,646	14,166
Microsoft Vision API	16,364	15,249	5,787	5,632
OCR Web Service	16,463	9,367	9,189	7,752
SemaMediaDataOCR	35,666	28,017	24,533	22,895
Cloudmersive API	22,756	14,230	12,837	8,715

O gráfico da Figura 40 mostra que na maioria dos casos a imagem com fonte itálico obteve um maior custo de tempo de execução.

Figura 40 – Tempo de execução das ferramentas de OCR para foto da imagem Hino com fonte tamanho 14 e com variações do estilo da fonte



4.5 Imagem Poema

Nessa seção são apresentados os resultados obtidos por cada ferramenta OCR para imagem Poema, contendo variações em 10, 12 e 14 no tamanho da fonte, no estilo da fonte em itálico, negrito, negrito e itálico e normal, e também nas resoluções 100, 300, 600 dpi e foto da imagem.

4.5.1 Fonte Tamanho 10

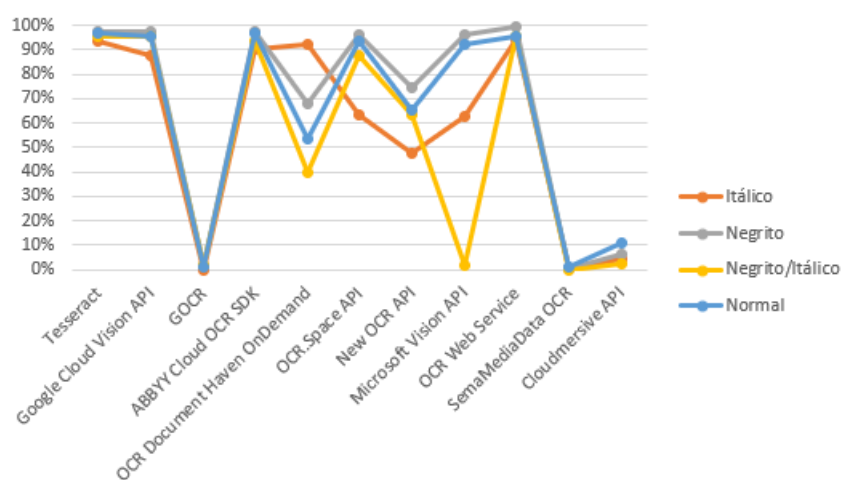
A Tabela 41 mostra que para a imagem Poema com fonte de tamanho 10 e na resolução 100 dpi, as ferramentas que obtiveram maior acurácia foram OCR Web Service e Tesseract, seguido do ABBYY Cloud OCR SDK, que também atingiu uma acurácia maior que 90% para todas as variações de fonte.

Tabela 41 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	93,870%	97,318%	95,785%	96,552%
Google Cloud Vision API	87,739%	97,318%	95,402%	95,785%
GOOCR	0,000%	1,533%	1,533%	0,766%
ABBYY Cloud OCR SDK	90,421%	97,318%	93,870%	96,935%
OCR Document Haven OnDemand	91,954%	68,199%	39,464%	53,257%
OCR.Space API	63,602%	96,169%	87,739%	93,487%
New OCR API	47,510%	74,713%	63,602%	65,134%
Microsoft Vision API	62,452%	96,169%	1,916%	92,337%
OCR Web Service	94,253%	99,234%	95,785%	95,785%
SemaMediaDataOCR	0,383%	0,383%	0,000%	1,149%
Cloudmersive API	4,215%	6,513%	2,682%	10,728%

Nota-se que, de acordo com o gráfico da Figura 41, em geral, as ferramentas tiveram um melhor resultado para imagens com fonte normal e em negrito.

Figura 41 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte



Para esse tipo de imagem, das ferramentas que apresentaram maior acurácia, o

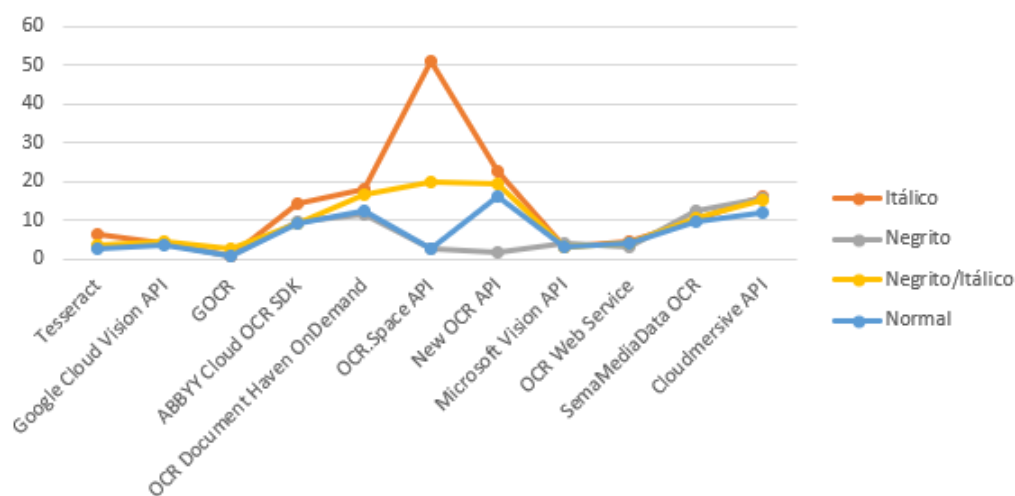
OCR Web Service e o Tesseract tiveram um baixo custo de tempo de execução, como mostra a Tabela 42.

Tabela 42 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	6,502	3,271	3,452	2,835
Google Cloud Vision API	4,107	4,257	4,604	3,651
GOOCR	1,001	0,955	2,474	0,845
ABBYY Cloud OCR SDK	14,421	9,508	8,956	9,241
OCR Document Haven OnDemand	17,758	11,621	16,722	12,437
OCR.Space API	50,798	2,563	19,944	2,447
New OCR API	22,745	1,670	19,539	16,251
Microsoft Vision API	2,959	3,924	2,914	2,930
OCR Web Service	4,569	2,984	3,941	4,057
SemaMediaDataOCR	9,873	12,241	10,615	9,455
Cloudmersive API	16,325	15,438	15,020	11,972

O gráfico da Figura 42 mostra novamente que a maioria das ferramentas possuem praticamente o mesmo custo de tempo de execução para variações do estilo de fonte. Porém, a ferramenta OCR.Space API apresenta uma grande diferença no custo para imagem com fonte em itálico, que tem um custo muito maior que as demais variações.

Figura 42 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 100 dpi e com variações do estilo da fonte



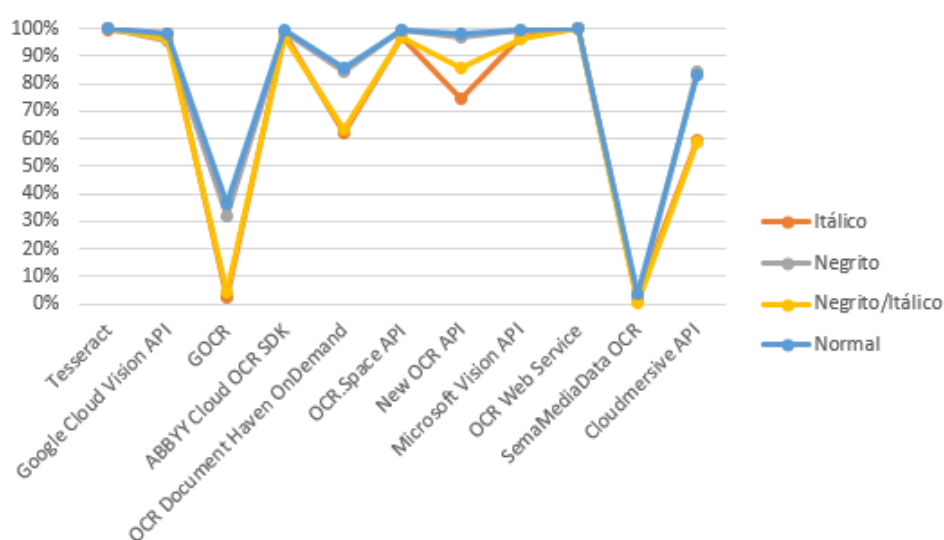
Observando-se a Tabela 43 pode-se constatar que o OCR WEB Service obteve uma maior acurácia em relação as outras ferramentas, reconhecendo 100% as palavras de todas as variações de fonte. O Tesseract também reconheceu 100% palavras para as variações negrito, negrito e itálico e normal, sendo que para itálico a porcentagem de reconhecimento ficou próxima de 100%.

Tabela 43 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,234%	100,000%	100,000%	100,000%
Google Cloud Vision API	98,467%	95,402%	96,169%	98,467%
GOOCR	2,299%	31,801%	4,215%	36,782%
ABBYY Cloud OCR SDK	98,084%	98,851%	96,935%	99,234%
OCR Document Haven OnDemand	62,069%	84,674%	63,602%	85,441%
OCR.Space API	96,552%	99,617%	96,552%	99,234%
New OCR API	74,330%	96,935%	85,441%	98,467%
Microsoft Vision API	96,552%	99,617%	96,169%	99,234%
OCR Web Service	100,000%	100,000%	100,000%	100,000%
SemaMediaDataOCR	2,682%	1,149%	0,383%	3,448%
Cloudmersive API	59,770%	84,291%	58,621%	83,142%

As ferramentas obtiveram uma maior acurácia para imagens com as variação de fonte negrito e fonte normal, como se pode observar no gráfico da Figura 43.

Figura 43 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte



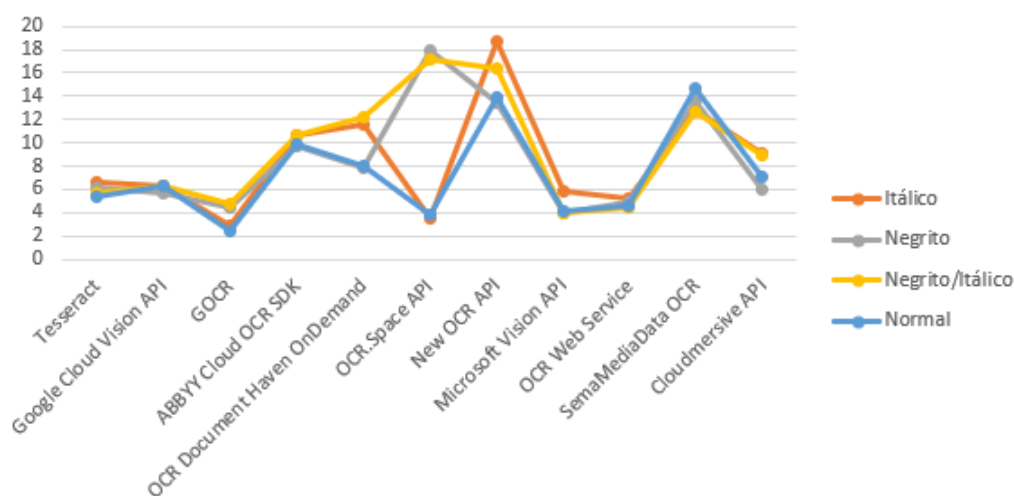
O OCR Web Service obteve para esse tipo de imagem um dos menores custos de tempo de execução em relação as outras ferramentas. O custo de tempo de execução do Tesseract é um pouco maior que do OCR Web Service, como pode-se observar na Tabela 44.

Tabela 44 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	6,612	6,198	5,619	5,436
Google Cloud Vision API	6,234	5,651	6,305	6,356
GOOCR	2,821	4,414	4,804	2,353
ABBYY Cloud OCR SDK	10,704	9,783	10,685	9,920
OCR Document Haven OnDemand	11,652	7,886	12,257	8,001
OCR.Space API	3,460	17,995	17,082	3,788
New OCR API	18,633	13,420	16,373	13,893
Microsoft Vision API	5,839	4,019	4,042	4,204
OCR Web Service	5,157	4,836	4,426	4,652
SemaMediaDataOCR	12,745	13,642	12,584	14,605
Cloudmersive API	9,046	5,995	8,896	7,044

O gráfico da Figura 44 mostra que o custo de tempo de execução para imagem com fonte normal, em geral, é menor do que para outras variações do estilo de fonte.

Figura 44 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 300 dpi e com variações do estilo da fonte



Segundo a Tabela 45, para a resolução 600 dpi desse tipo de imagem, a ferramenta Tesseract conseguiu uma maior acurácia em relação as outras ferramentas. Os resultados

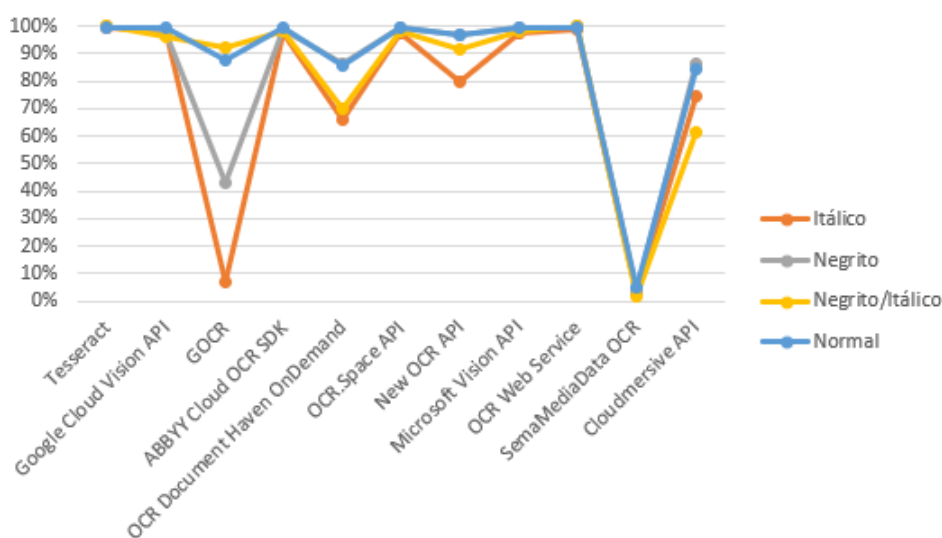
do OCR Web Service foi quase igual o do Tesseract, porém, para variação itálico a acurácia foi um pouco menor.

Tabela 45 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,234%	100,000%	100,000%	99,617%
Google Cloud Vision API	98,084%	97,318%	96,169%	99,234%
GOOCR	6,897%	43,295%	91,954%	87,739%
ABBYY Cloud OCR SDK	97,701%	98,851%	98,467%	99,617%
OCR Document Haven OnDemand	66,284%	86,207%	69,732%	85,824%
OCR.Space API	97,318%	99,617%	98,084%	99,234%
New OCR API	80,077%	96,552%	91,571%	96,552%
Microsoft Vision API	97,318%	99,234%	98,084%	99,234%
OCR Web Service	98,851%	100,000%	100,000%	99,617%
SemaMediaDataOCR	3,065%	3,448%	1,533%	4,981%
Cloudmersive API	74,330%	86,207%	61,303%	84,674%

Para a maioria das ferramentas, a acurácia obtida nos resultados para esse tipo de imagem foi maior para fonte normal e, em seguida, para variação do estilo de fonte em negrito.

Figura 45 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte



O custo do tempo de execução para o Tesseract e OCR Web Service foram semelhantes, tendo o primeiro um menor custo para fontes em negrito e negrito e itálico e o

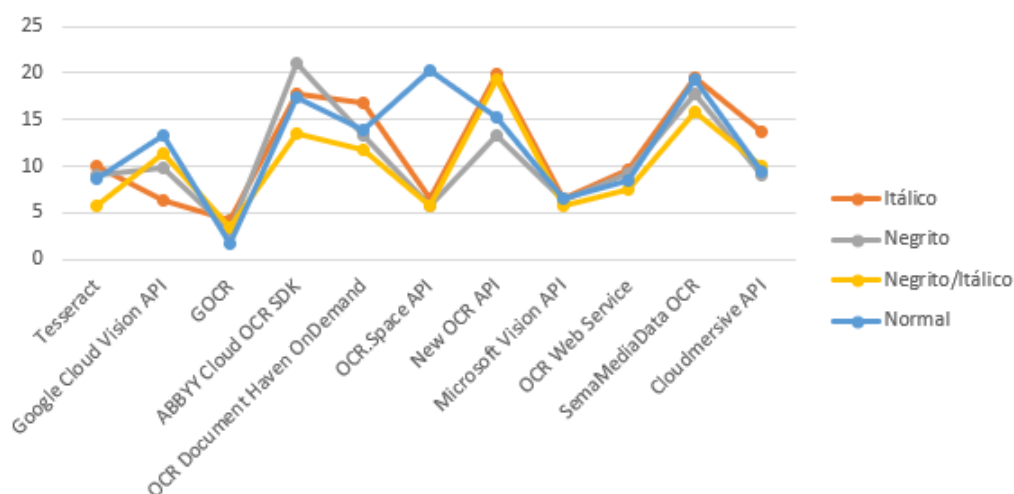
segundo um menor custo para fontes em itálico e normal, como mostra a Tabela 46.

Tabela 46 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	10,026	9,087	5,843	8,678
Google Cloud Vision API	6,234	9,837	11,380	13,389
GOOCR	4,132	3,106	3,456	1,761
ABBYY Cloud OCR SDK	17,689	21,023	13,489	17,414
OCR Document Haven OnDemand	16,744	13,215	11,743	13,944
OCR.Space API	6,595	5,833	5,820	20,375
New OCR API	19,873	13,249	19,322	15,234
Microsoft Vision API	6,466	6,342	5,778	6,619
OCR Web Service	9,567	9,184	7,545	8,422
SemaMediaDataOCR	19,596	17,787	15,844	19,241
Cloudmersive API	13,747	9,027	9,927	9,346

O gráfico da Figura 46 mostra uma diferença no custo de tempo de execução das ferramentas para cada tipo de variação do estilo de fonte.

Figura 46 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 10 na resolução 600 dpi e com variações do estilo da fonte



Para foto da imagem Poema com fonte de tamanho 10, a ferramenta OCR Web Service obteve uma maior acurácia, seguida do Tesseract, Google Cloud Vision API, ABBYY Cloud OCR SDK, OCR.Space API e Microsoft Vision API. Todas essas ferramentas

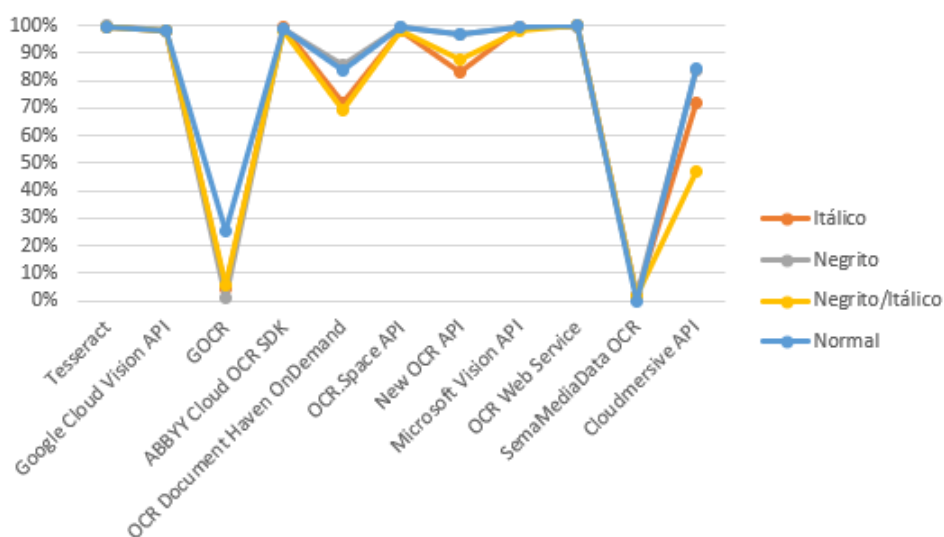
obtiveram uma acurácia maior que 98% para todas variações de estilo de fonte, como mostra a Figura 47.

Tabela 47 – Acurácia das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 10 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,234%	100,000%	99,234%	99,617%
Google Cloud Vision API	98,467%	98,467%	98,084%	98,084%
GOOCR	4,598%	1,149%	5,747%	25,670%
ABBYY Cloud OCR SDK	99,234%	98,851%	98,467%	98,851%
OCR Document Haven OnDemand	72,031%	85,441%	68,966%	83,908%
OCR.Space API	98,467%	99,617%	98,467%	99,617%
New OCR API	82,759%	96,935%	87,739%	96,552%
Microsoft Vision API	99,617%	99,234%	98,467%	99,617%
OCR Web Service	100,000%	99,617%	100,000%	100,000%
SemaMediaDataOCR	1,533%	1,916%	1,533%	0,000%
Cloudmersive API	71,648%	83,908%	47,126%	84,674%

O gráfico da Figura 47 mostra que todas as ferramentas tiveram uma maior acurácia para imagem com fonte normal.

Figura 47 – Acurácia das ferramentas de OCR para foto da imagem Poema com fonte tamanho 10 e com variações do estilo da fonte



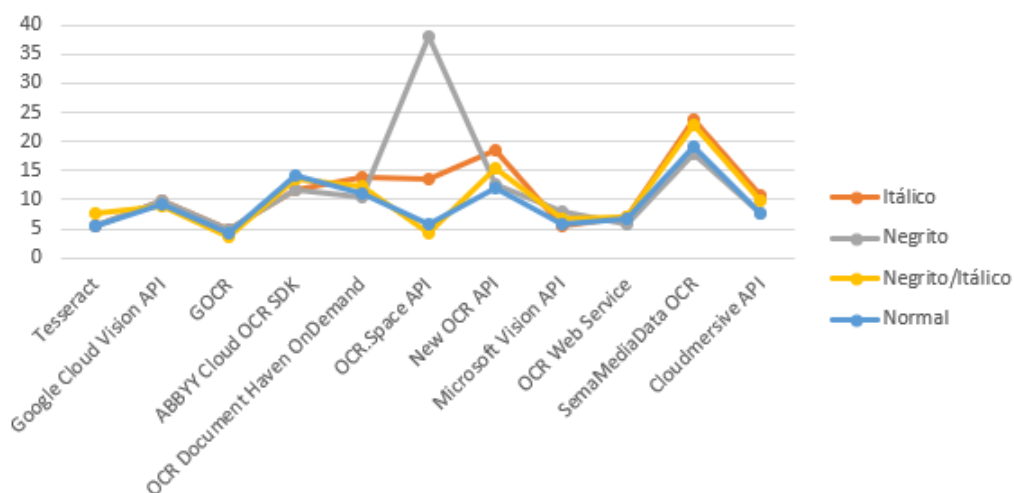
Das ferramentas que obtiveram uma maior acurácia, Microsoft Vision API, Tesseract e OCR Web Service tiveram um custo de tempo de execução semelhantes e menor que os demais, conforme pode-se observar na Figura 48.

Tabela 48 – Tempo de execução das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 10 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	5,592	5,380	7,696	5,369
Google Cloud Vision API	9,953	9,817	8,854	9,179
GOOCR	4,726	5,004	3,681	4,230
ABBYY Cloud OCR SDK	11,569	11,731	13,570	14,057
OCR Document Haven OnDemand	13,865	10,571	12,375	11,185
OCR.Space API	13,642	38,046	4,139	5,692
New OCR API	18,407	12,650	15,297	11,858
Microsoft Vision API	5,592	8,086	6,652	5,769
OCR Web Service	7,086	5,941	7,106	6,807
SemaMediaDataOCR	23,884	17,791	22,762	19,246
Cloudmersive API	10,686	7,616	9,895	7,574

De acordo com o gráfico da Figura 48, praticamente todas as ferramentas tiveram um custo semelhante para as variações do estilo de fonte, exceto o OCR.Space API, que diferiu para um custo maior para fonte itálico e bem maior para negrito.

Figura 48 – Tempo de execução das ferramentas de OCR para foto da imagem Poema com fonte tamanho 10 e com variações do estilo da fonte



4.5.2 Fonte Tamanho 12

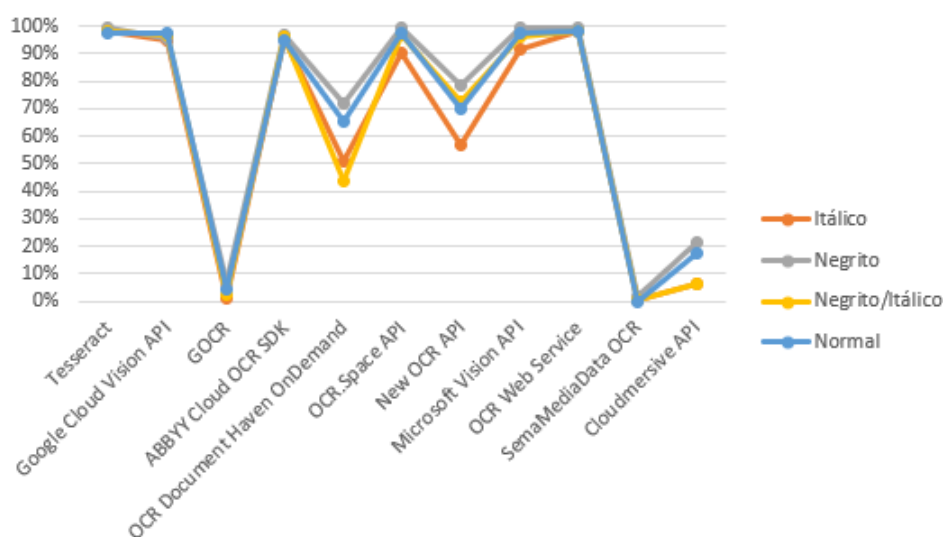
Observando-se a Tabela 49, nota-se que novamente as ferramentas que tiveram uma maior acurácia foram OCR Web Service e Tesseract. Google Cloud Vision API, ABBYY Cloud OCR SDK, OCR.Space API e Microsoft Vision API também tiveram uma acurácia maior que 90% para todas variações de fonte para imagem na resolução de 100 dpi.

Tabela 49 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	98,084%	99,234%	98,467%	97,701%
Google Cloud Vision API	95,019%	95,785%	96,935%	97,701%
GOOCR	0,766%	6,897%	2,299%	4,215%
ABBYY Cloud OCR SDK	95,785%	96,935%	96,169%	94,636%
OCR Document Haven OnDemand	50,958%	71,648%	43,678%	65,517%
OCR.Space API	90,421%	99,234%	96,935%	97,318%
New OCR API	57,088%	78,161%	72,414%	70,115%
Microsoft Vision API	91,571%	99,234%	96,169%	97,701%
OCR Web Service	98,084%	99,234%	98,467%	98,467%
SemaMediaDataOCR	0,383%	1,533%	0,383%	0,000%
Cloudmersive API	6,513%	21,456%	6,130%	17,625%

Em geral, como mostra o gráfico da Figura 49, as ferramentas obtiveram uma maior acurácia para fonte normal e depois para variação do estilo de fonte em negrito.

Figura 49 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte



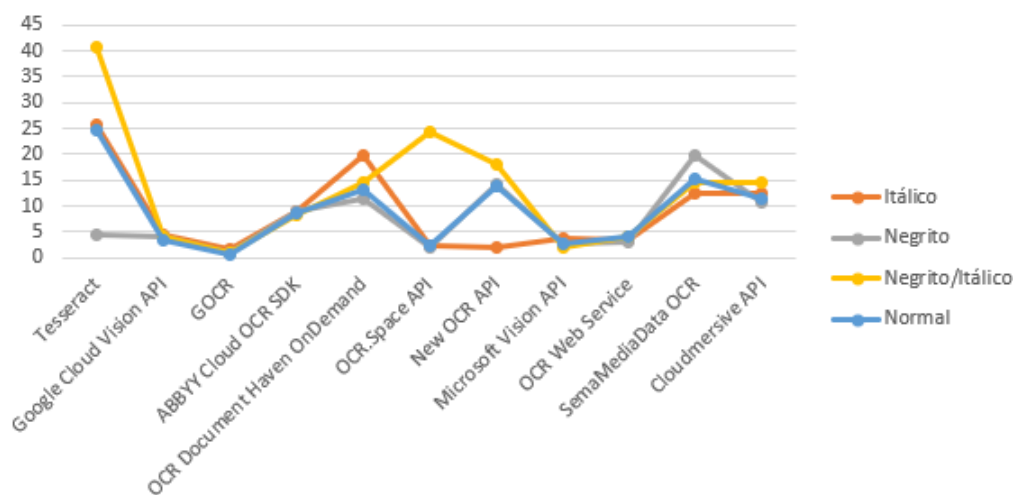
A ferramenta OCR Web Service, neste caso, obteve um baixo custo de tempo de execução. Já o Tesseract, em geral, teve um custo bem maior. Isso pode ser observado na Tabela 50.

Tabela 50 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	25,714	4,556	40,715	24,709
Google Cloud Vision API	4,541	3,907	4,015	3,511
GOOCR	1,743	1,058	1,079	0,752
ABBYY Cloud OCR SDK	8,879	8,862	8,207	8,749
OCR Document Haven OnDemand	19,787	11,278	14,365	13,119
OCR.Space API	2,397	2,025	24,155	2,355
New OCR API	1,908	14,227	18,183	13,813
Microsoft Vision API	3,625	2,824	1,812	2,768
OCR Web Service	3,377	3,105	3,978	3,918
SemaMediaDataOCR	12,531	19,740	14,694	15,363
Cloudmersive API	12,267	10,773	14,512	11,499

O gráfico da Figura 50 mostra que, em geral, a maioria das ferramentas tiveram um menor custo de tempo de execução para imagem com fonte em negrito.

Figura 50 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 100 dpi e com variações do estilo da fonte



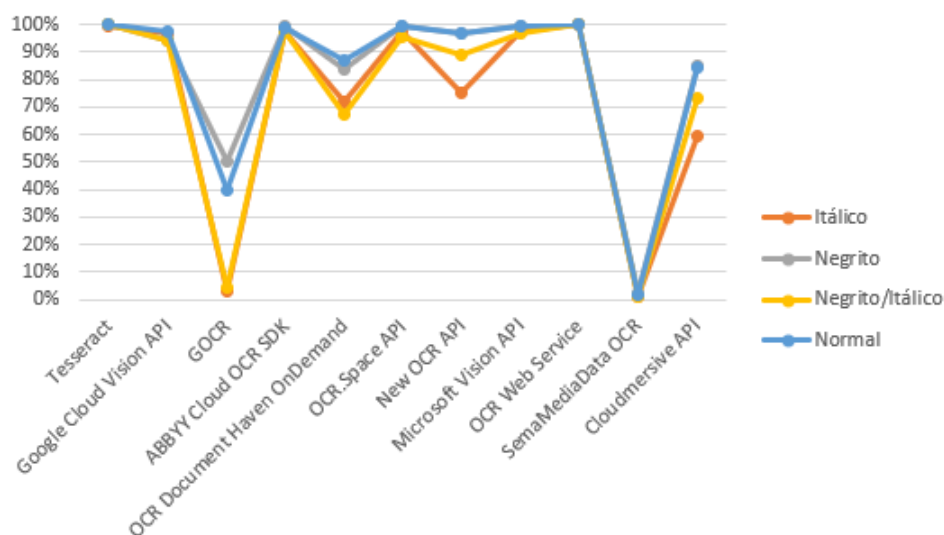
Para esse tipo de imagem com resolução em 300 dpi, como pode-se observar na Tabela 51, o OCR Web Service obteve uma acurácia de 100% para todas variações de fonte. O Tesseract também obteve 100% de acurácia para todas variações, exceto para itálico, porém, a acurácia foi próxima de 100%.

Tabela 51 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,617%	100,000%	100,000%	100,000%
Google Cloud Vision API	96,552%	94,253%	94,253%	97,701%
GOCR	3,065%	50,192%	4,598%	39,464%
ABBYY Cloud OCR SDK	97,701%	99,234%	97,318%	98,851%
OCR Document Haven OnDemand	71,648%	83,525%	67,050%	86,973%
OCR.Space API	97,701%	99,234%	95,785%	99,234%
New OCR API	75,479%	96,935%	88,889%	96,552%
Microsoft Vision API	97,701%	99,234%	96,935%	99,234%
OCR Web Service	100,000%	100,000%	100,000%	100,000%
SemaMediaDataOCR	1,149%	2,682%	1,149%	1,533%
Cloudmersive API	59,387%	85,057%	73,180%	84,674%

Observando-se o gráfico da Figura 51 nota-se que, para praticamente todas as ferramentas, as variações de fonte itálico e negrito e itálico obtiveram uma menor acurácia do que para fonte negrito e normal.

Figura 51 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte



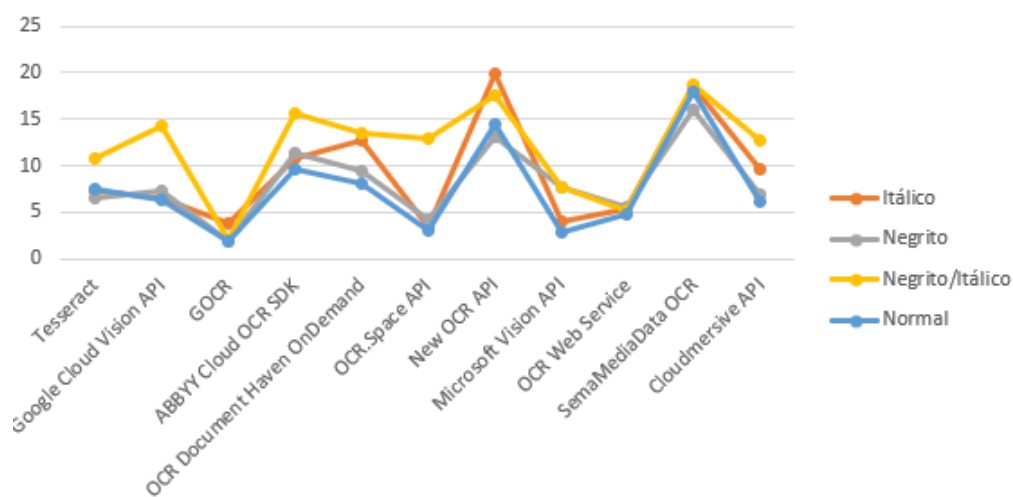
O tempo de execução da ferramenta OCR Web Service foi menor, para todas variações de fonte, do que o da ferramenta Tesseract, como mostra a Tabela 52.

Tabela 52 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	7,231	6,481	10,762	7,536
Google Cloud Vision API	6,536	7,395	14,316	6,399
GOOCR	3,829	1,936	2,161	1,828
ABBYY Cloud OCR SDK	10,779	11,339	15,579	9,652
OCR Document Haven OnDemand	12,706	9,422	13,457	8,037
OCR.Space API	3,434	4,180	12,848	3,103
New OCR API	19,959	13,090	17,633	14,538
Microsoft Vision API	3,972	7,636	7,703	2,779
OCR Web Service	5,368	5,548	5,170	4,859
SemaMediaDataOCR	18,507	15,967	18,788	17,962
Cloudmersive API	9,673	6,858	12,804	6,176

O gráfico da Figura 52 mostra que o custo do tempo de execução, para imagens com fonte em negrito e itálico, foi maior do que para as outras variações de fonte.

Figura 52 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 300 dpi e com variações do estilo da fonte



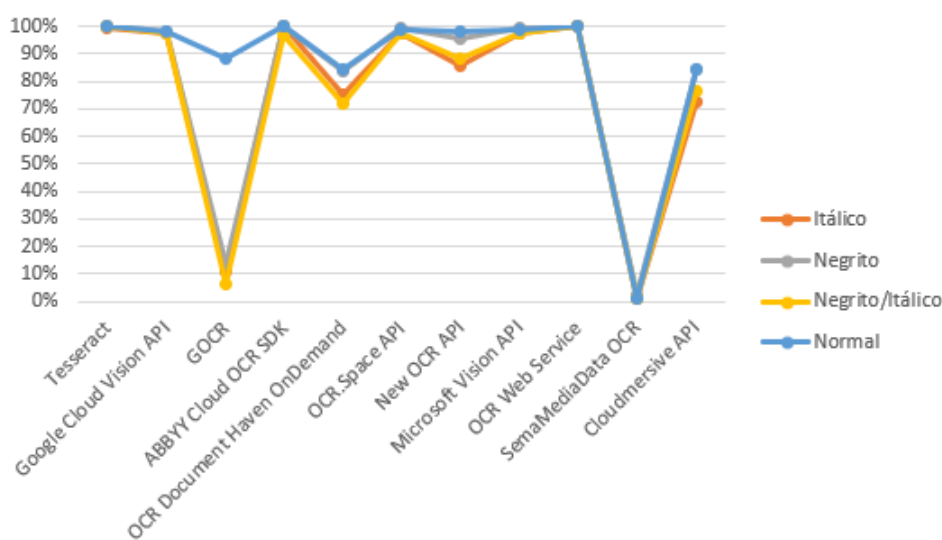
A Tabela 53 mostra que para esse tipo de imagem em 600 dpi, novamente o OCR WEB Service obteve 100% acurácia para todas variações de fonte. A acurácia do Tesseract e do ABBYY Cloud OCR SDK também foi de 100% ou próximo a 100% na maioria das variações.

Tabela 53 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,617%	100,000%	100,000%	100,000%
Google Cloud Vision API	97,318%	98,084%	97,318%	98,467%
GOOCR	10,728%	13,410%	6,513%	88,506%
ABBYY Cloud OCR SDK	100,000%	100,000%	96,935%	100,000%
OCR Document Haven OnDemand	75,479%	83,525%	72,031%	84,291%
OCR.Space API	97,701%	99,234%	97,318%	98,851%
New OCR API	85,824%	95,402%	88,123%	98,467%
Microsoft Vision API	97,318%	99,234%	97,318%	98,851%
OCR Web Service	100,000%	100,000%	100,000%	100,000%
SemaMediaDataOCR	1,149%	2,682%	0,766%	1,149%
Cloudmersive API	72,797%	84,674%	76,628%	84,291%

A acurácia das ferramentas para fonte normal se mostra maior do que para as outras variações no estilo de fonte, como mostra o gráfico da Figura 53.

Figura 53 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte



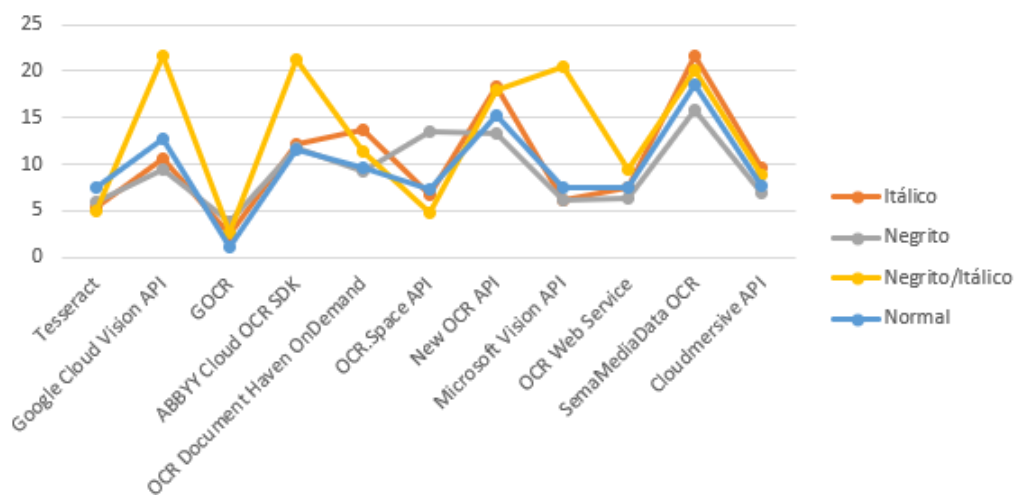
O Tesseract, das ferramentas que obtiveram uma maior acurácia, foi o que teve um menor custo de tempo de execução para essa imagem, como mostra a Tabela 54.

Tabela 54 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	5,273	5,982	5,040	7,458
Google Cloud Vision API	10,565	9,410	21,683	12,806
GOOCR	2,430	3,825	2,673	1,170
ABBYY Cloud OCR SDK	12,114	11,756	21,277	11,612
OCR Document Haven OnDemand	13,749	9,143	11,385	9,604
OCR.Space API	6,647	13,407	4,756	7,217
New OCR API	18,376	13,219	17,970	15,196
Microsoft Vision API	6,095	6,042	20,484	7,496
OCR Web Service	7,559	6,284	9,354	7,585
SemaMediaDataOCR	21,707	15,827	20,055	18,542
Cloudmersive API	9,720	6,846	8,794	7,651

As ferramentas, em geral, obtiveram um maior custo de tempo de execução para fontes em negrito e itálico.

Figura 54 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 12 na resolução 600 dpi e com variações do estilo da fonte



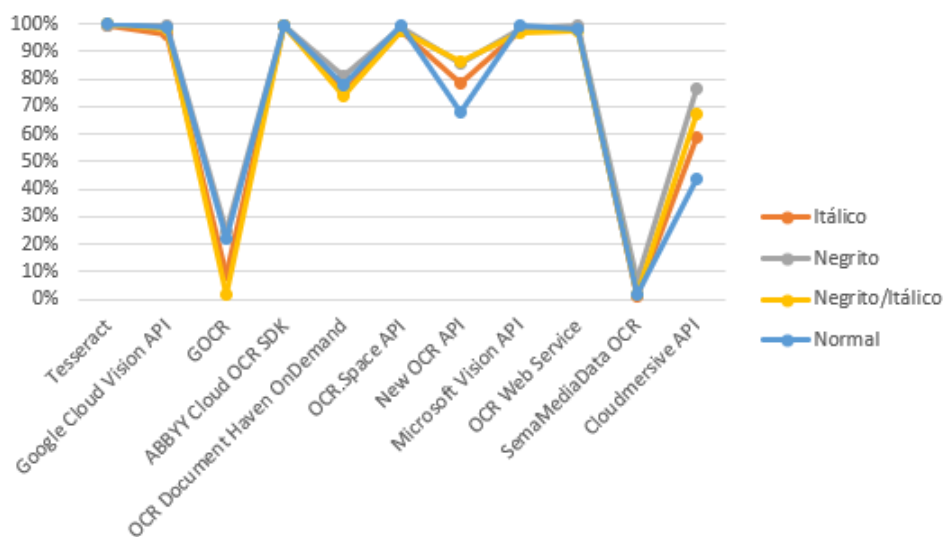
Para foto da imagem Poema com tamanho de fonte 12, como mostra a Tabela 55, a ferramenta que obteve maior acurácia foi o Tesseract. Google Cloud Vision API, ABBYY Cloud OCR SDK, OCR.Space API, Microsoft Vision API e OCR Web Service também tiveram ótimas acurácias.

Tabela 55 – Acurácia das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 12 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,617%	99,617%	100,000%	100,000%
Google Cloud Vision API	96,169%	99,617%	98,084%	98,851%
GOOCR	8,812%	24,521%	1,533%	21,839%
ABBYY Cloud OCR SDK	99,234%	99,617%	99,234%	99,234%
OCR Document Haven OnDemand	75,096%	80,843%	73,946%	77,778%
OCR.Space API	97,701%	98,851%	97,318%	99,617%
New OCR API	78,544%	85,824%	86,207%	67,816%
Microsoft Vision API	97,318%	98,467%	96,552%	99,234%
OCR Web Service	98,851%	99,617%	97,701%	98,467%
SemaMediaDataOCR	0,766%	6,897%	1,533%	1,533%
Cloudmersive API	58,621%	76,628%	67,050%	43,678%

A acurácia para as variações de fonte, em geral, foi parecida para todas as variações.

Figura 55 – Acurácia das ferramentas de OCR para foto da imagem Poema com fonte tamanho 12 e com variações do estilo da fonte



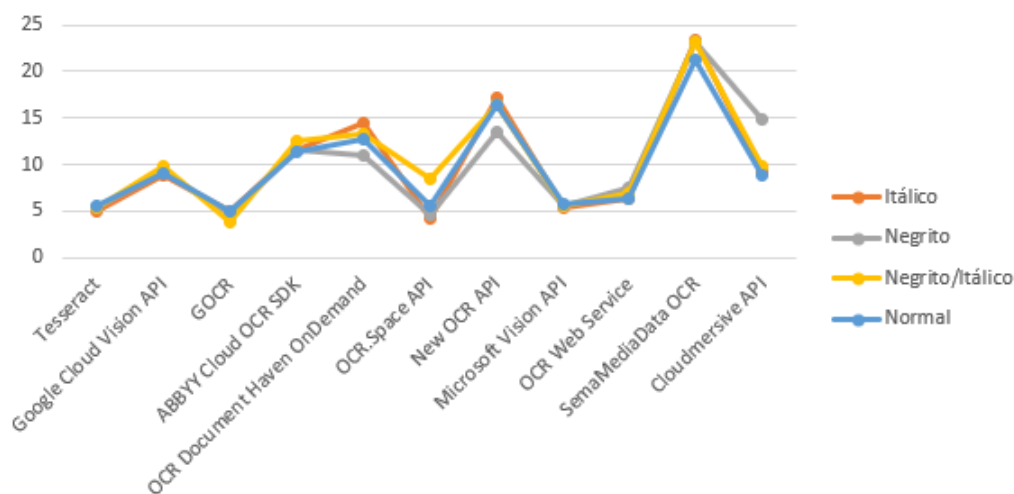
De acordo com a Tabela 56, das ferramentas que alcançaram uma maior acurácia, as que tiveram um menor custo de tempo de execução foram Tesseract, Microsoft Vision API e OCR Web Service.

Tabela 56 – Tempo de execução das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 12 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	4,959	5,306	5,411	5,609
Google Cloud Vision API	8,941	9,322	9,745	9,025
GOOCR	4,990	4,504	3,807	4,985
ABBYY Cloud OCR SDK	11,545	11,604	12,505	11,370
OCR Document Haven OnDemand	14,437	11,052	13,314	12,644
OCR.Space API	4,175	4,654	8,484	5,576
New OCR API	17,232	13,429	16,328	16,407
Microsoft Vision API	5,455	5,535	5,522	5,739
OCR Web Service	6,318	7,465	6,888	6,380
SemaMediaDataOCR	23,457	23,169	23,227	21,278
Cloudmersive API	9,321	14,933	9,742	8,919

O custo do tempo de execução para cada ferramenta se mostra parecido entre as variações de fonte existentes.

Figura 56 – Tempo de execução das ferramentas de OCR para foto da imagem Poema com fonte tamanho 12 e com variações do estilo da fonte



4.5.3 Fonte Tamanho 14

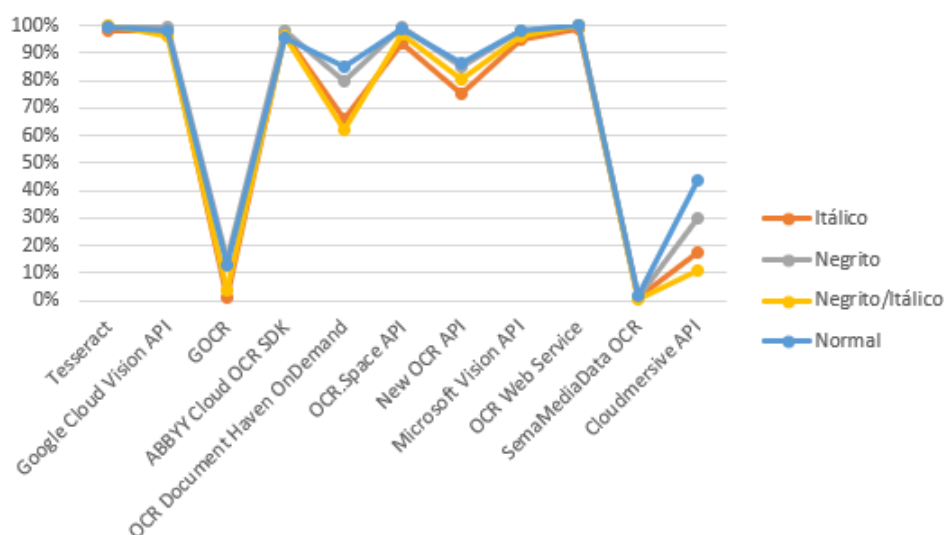
A Tabela 57 mostra que para imagem Poema com fonte de tamanho 14 em resolução de 100 dpi, a ferramenta que obteve uma maior acurácia foi o OCR Web Service, seguido por Tesseract, Google Cloud Vision API, ABBYY Cloud OCR SDK, OCR.Space API e Microsoft Vision API.

Tabela 57 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	98,084%	99,234%	100,000%	99,234%
Google Cloud Vision API	98,467%	99,234%	96,169%	98,467%
GOOCR	0,766%	14,559%	3,831%	13,027%
ABBYY Cloud OCR SDK	96,552%	98,084%	96,169%	95,785%
OCR Document Haven OnDemand	65,900%	80,077%	62,069%	85,057%
OCR.Space API	93,870%	99,617%	96,935%	98,851%
New OCR API	75,479%	85,057%	80,460%	86,590%
Microsoft Vision API	95,019%	98,084%	96,169%	98,467%
OCR Web Service	98,851%	100,000%	100,000%	100,000%
SemaMediaDataOCR	1,149%	0,766%	0,383%	1,533%
Cloudmersive API	17,241%	29,885%	10,728%	43,678%

Cada ferramenta obteve uma acurácia semelhante para as variações de fonte, tendo a fonte normal uma acurácia um pouco maior que as outras.

Figura 57 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte



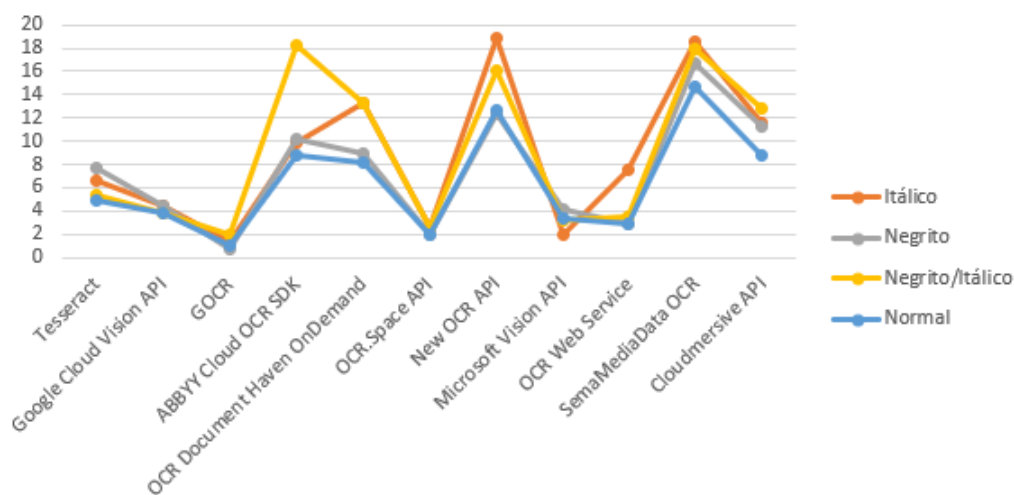
Microsoft Vision API obteve um menor custo de tempo de execução, seguido por Google Cloud Vision API e OCR Web Service.

Tabela 58 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	6,662	7,685	5,381	4,857
Google Cloud Vision API	4,403	4,395	3,889	3,892
GOOCR	1,332	0,694	1,978	1,110
ABBYY Cloud OCR SDK	9,866	10,249	18,299	8,803
OCR Document Haven OnDemand	13,239	8,914	13,246	8,192
OCR.Space API	2,564	1,994	2,579	1,911
New OCR API	18,931	12,428	16,096	12,720
Microsoft Vision API	1,994	4,112	3,141	3,392
OCR Web Service	7,518	2,896	3,456	2,836
SemaMediaDataOCR	18,522	16,660	17,911	14,668
Cloudmersive API	11,614	11,292	12,768	8,773

Em geral, as ferramentas apresentaram um maior custo de execução para imagem com fonte em negrito e itálico e itálico.

Figura 58 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 100 dpi e com variações do estilo da fonte



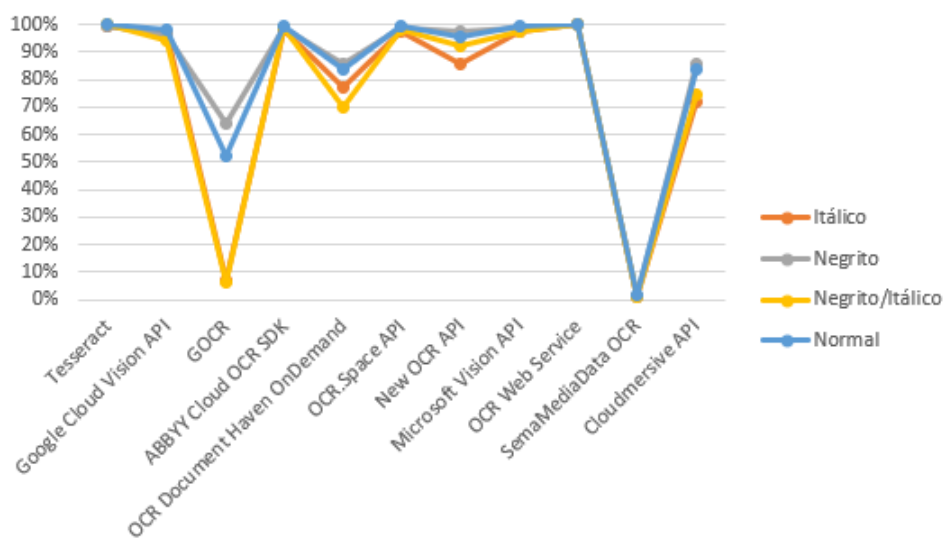
Para a imagem Poema com fonte tamanho 14 e em 300 dpi, a Tabela 59 mostra que o OCR Web Service obteve a melhor acurácia, reconhecendo 100% das palavras para todas as variações. O Tesseract também teve a acurácia próxima de 100% em todos os casos.

Tabela 59 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,234%	99,617%	100,000%	100,000%
Google Cloud Vision API	96,935%	96,169%	94,253%	98,084%
GOOCR	6,897%	64,368%	6,130%	52,107%
ABBYY Cloud OCR SDK	98,467%	98,851%	98,851%	99,234%
OCR Document Haven OnDemand	77,011%	85,824%	70,115%	83,908%
OCR.Space API	97,701%	98,851%	98,084%	99,234%
New OCR API	85,824%	97,701%	92,337%	95,402%
Microsoft Vision API	97,318%	98,851%	97,318%	99,234%
OCR Web Service	100,000%	100,000%	100,000%	100,000%
SemaMediaDataOCR	0,766%	0,766%	0,766%	1,533%
Cloudmersive API	72,031%	85,824%	74,713%	83,525%

O gráfico da Figura 59 mostra que todas as ferramentas obtiveram uma maior acurácia para fonte em negrito e fonte normal.

Figura 59 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte



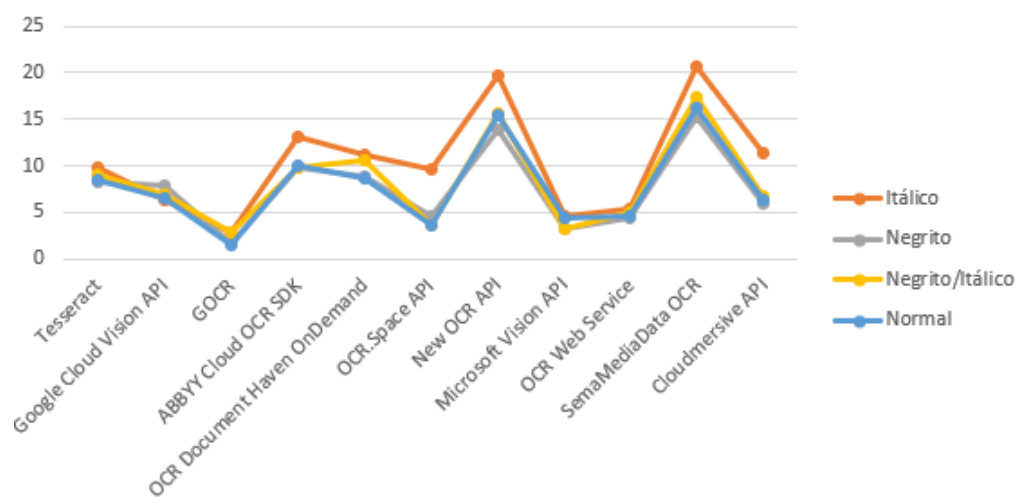
Entre OCR Web Service e Tesseract, a primeira apresentou um menor custo de tempo de execução para imagem com resolução em 300 dpi.

Tabela 60 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	9,859	8,188	9,064	8,423
Google Cloud Vision API	6,402	7,799	6,979	6,507
GOOCR	2,855	1,682	2,810	1,553
ABBYY Cloud OCR SDK	13,019	9,872	9,732	9,967
OCR Document Haven OnDemand	11,222	8,865	10,637	8,612
OCR.Space API	9,645	4,540	3,582	3,649
New OCR API	19,734	13,853	15,673	15,494
Microsoft Vision API	4,666	3,261	3,141	4,357
OCR Web Service	5,402	4,375	4,903	4,546
SemaMediaDataOCR	20,621	15,229	17,461	16,186
Cloudmersive API	11,408	5,956	6,708	6,407

Como mostra o gráfico da Figura 60, a fonte em itálico demandou um maior custo de tempo de execução para as ferramentas OCR.

Figura 60 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 300 dpi e com variações do estilo da fonte



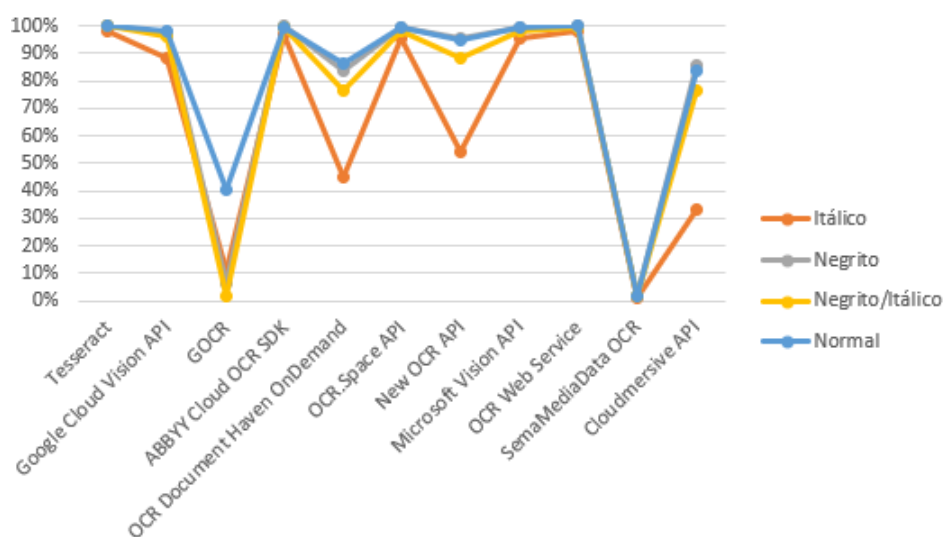
A ferramenta Tesseract, conforme a Tabela 61, apresentou uma maior acurácia para a resolução em 600 dpi para todas as variações do estilo de fonte. Outra ferramenta que apresentou acurácia próxima a 100% em todas as variações foi o OCR Web Service.

Tabela 61 – Acurácia das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 600 dpi com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	98,467%	100,000%	100,000%	100,000%
Google Cloud Vision API	88,506%	97,318%	96,169%	98,467%
GOOCR	10,728%	6,513%	1,916%	40,613%
ABBYY Cloud OCR SDK	96,552%	100,000%	99,617%	99,234%
OCR Document Haven OnDemand	45,211%	83,908%	76,628%	86,590%
OCR.Space API	95,402%	99,617%	98,084%	99,234%
New OCR API	54,406%	95,785%	88,506%	94,636%
Microsoft Vision API	95,402%	99,234%	98,467%	99,234%
OCR Web Service	98,084%	100,000%	99,617%	100,000%
SemaMediaDataOCR	1,149%	1,916%	1,916%	1,916%
Cloudmersive API	32,950%	85,824%	76,245%	83,908%

Todas as ferramentas, para essa imagem, alcançaram uma maior acurácia para a fonte normal e em negrito.

Figura 61 – Acurácia das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte



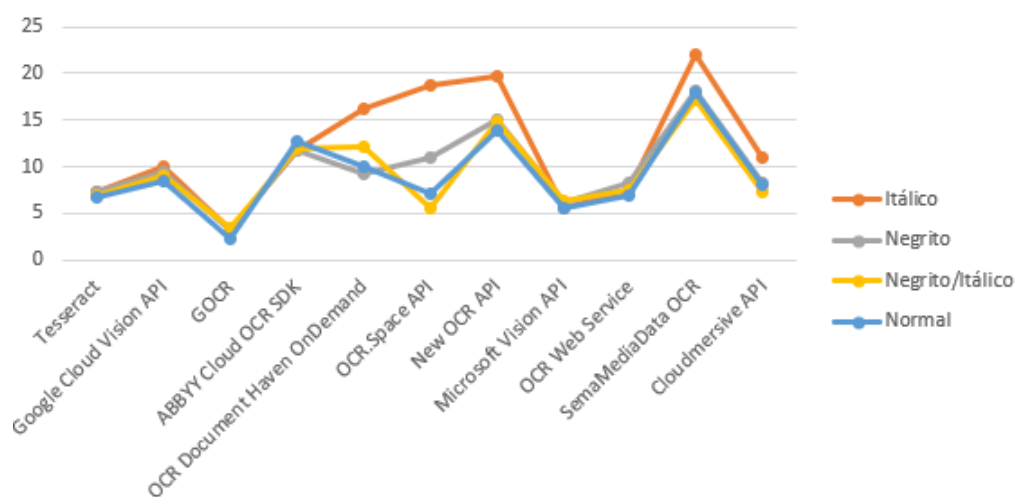
Tesseract e OCR Web Service apresentaram um custo de tempo de execução baixo e semelhante para as variações existentes.

Tabela 62 – Tempo de execução das ferramentas de OCR em segundos para imagem Poema com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	7,287	7,292	6,962	6,768
Google Cloud Vision API	10,069	9,456	9,058	8,397
GOOCR	3,204	3,218	3,514	2,196
ABBYY Cloud OCR SDK	11,816	11,769	12,023	12,705
OCR Document Haven OnDemand	16,254	9,180	12,124	10,030
OCR.Space API	18,683	10,935	5,548	7,029
New OCR API	19,734	14,973	14,776	13,923
Microsoft Vision API	5,624	6,140	6,409	5,504
OCR Web Service	7,587	8,218	7,576	6,854
SemaMediaDataOCR	22,058	18,069	17,226	17,893
Cloudmersive API	11,012	8,303	7,268	8,078

As variações de fonte negrito, negrito e itálico e normal obtiveram um custo de tempo de execução semelhante na maioria dos casos. Já para fonte em negrito, o custo de tempo de execução foi um pouco maior.

Figura 62 – Tempo de execução das ferramentas de OCR para imagem Poema com fonte tamanho 14 na resolução 600 dpi e com variações do estilo da fonte



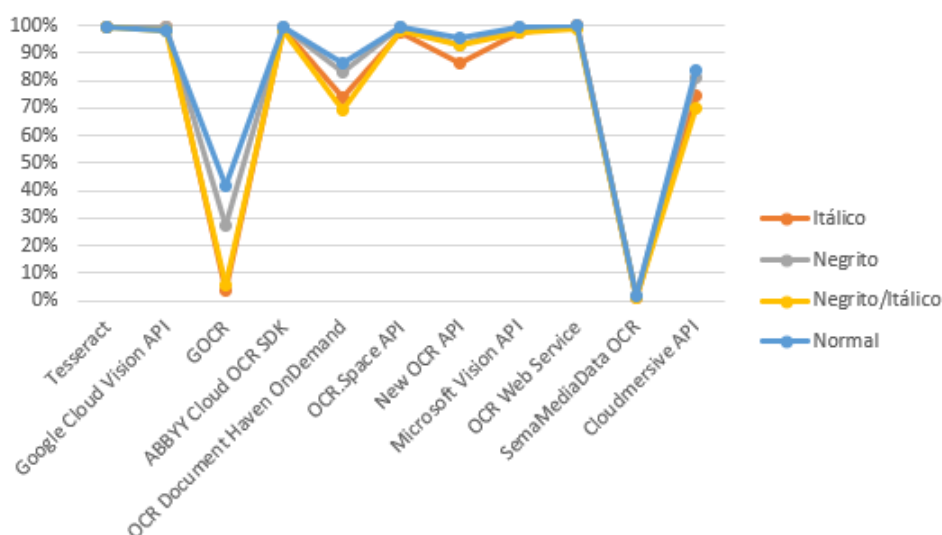
Observando a Tabela 63, pode-se constatar que para foto da imagem Poema com fonte tamanho 14 as ferramentas que obtiveram uma maioria acurácia foram OCR Web Service, Tesseract, Google Cloud Vision API e ABBYY Cloud OCR SDK, todas com acurácia acima de 98% para todas variações do estilo de fonte.

Tabela 63 – Acurácia das ferramentas de OCR em segundos para foto imagem Poema com fonte tamanho 14 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	99,617%	99,617%	99,617%	99,617%
Google Cloud Vision API	99,234%	99,617%	98,084%	98,084%
GOOCR	3,831%	27,203%	5,747%	41,762%
ABBYY Cloud OCR SDK	98,851%	99,617%	98,084%	99,234%
OCR Document Haven OnDemand	73,563%	82,759%	69,349%	86,207%
OCR.Space API	97,701%	99,617%	98,467%	99,234%
New OCR API	86,207%	92,720%	92,720%	95,785%
Microsoft Vision API	97,318%	99,234%	97,701%	99,234%
OCR Web Service	100,000%	99,617%	98,851%	100,000%
SemaMediaDataOCR	0,766%	1,916%	0,766%	1,916%
Cloudmersive API	74,713%	81,226%	69,732%	83,908%

As ferramentas tiveram uma maior acurácia para a fonte normal e depois para fonte em negrito.

Figura 63 – Acurácia das ferramentas de OCR para foto da imagem Poema com fonte tamanho 14 e com variações do estilo da fonte



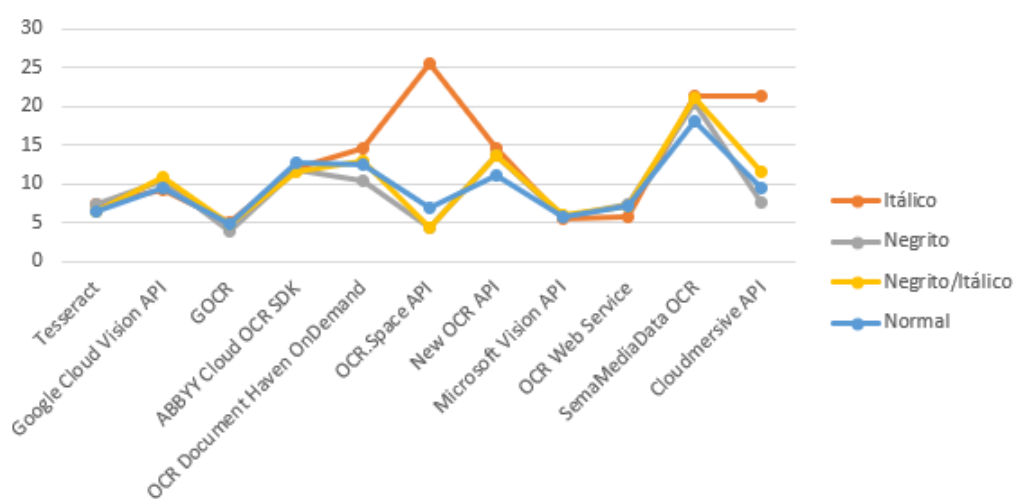
OCR Web Service e Tesseract apresentaram um custo de tempo de execução baixo e semelhante.

Tabela 64 – Tempo de execução das ferramentas de OCR em segundos para foto da imagem Poema com fonte tamanho 14 e com variações do estilo da fonte

Ferramenta OCR	Variação da Fonte			
	Itálico	Negrito	Negrito/Itálico	Normal
Tesseract	7,441	7,327	6,459	6,544
Google Cloud Vision API	9,261	10,471	10,877	9,530
GOOCR	4,950	3,774	4,903	4,788
ABBYY Cloud OCR SDK	12,081	11,892	11,593	12,638
OCR Document Haven OnDemand	14,468	10,370	12,981	12,548
OCR.Space API	25,391	4,359	4,456	6,974
New OCR API	14,619	13,734	13,664	11,124
Microsoft Vision API	5,616	5,852	5,931	5,760
OCR Web Service	5,828	7,366	7,166	7,030
SemaMediaDataOCR	21,209	20,373	21,030	17,953
Cloudmersive API	21,225	7,509	11,502	9,524

As variações de fonte negrito, negrito e itálico e fonte normal apresentaram, em geral, um custo de tempo de execução semelhante para as variações. Porém, para fonte em itálico o custo foi um pouco maior.

Figura 64 – Tempo de execução das ferramentas de OCR para foto da imagem Poema com fonte tamanho 14 e com variações do estilo da fonte



4.6 Totalização dos Resultados

Para as 108 imagens do *dataset* foi realizada uma totalização dos resultados obtidos para verificar quais ferramentas tiveram um melhor desempenho para esse *dataset* de entrada. Em caso de resultados iguais foram somados pontos para todas as ferramentas

que tiveram os melhores desempenhos. Na Tabela 65 observa-se que para 72 imagens do *dataset* a ferramenta OCR Web Service obteve a maior acurácia e o Tesseract em 48 das imagens. As ferramentas OCR.Space API, ABBYY Cloud Vision API, Google Cloud Vision API e Microsoft Vision API também conseguiram obter uma maior acurácia em algumas das imagens, obtendo respectivamente 15, 8, 8 e 4 imagens.

Tabela 65 – Totalização dos Resultados

Ferramenta OCR	Melhor Acurácia
OCR Web Service	72
Tesseract	48
OCR.Space API	15
ABBYY Cloud Vision API	8
Google Cloud Vision API	8
Microsoft Vision API	4

Então, pode-se concluir que para esse *dataset* de entrada, a ferramenta OCR Web Service obteve os melhores resultados, seguida pelo Tesseract, e mesmo nos casos em que o OCR Web Service se saiu melhor os resultados foram bem próximos aos obtidos pelo Tesseract. Além disso, vale ressaltar que o Tesseract é uma ferramenta gratuita, enquanto que o OCR Web Service é proprietário.

5 CONSIDERAÇÕES FINAIS

A utilização de ferramentas de OCR atualmente é de suma importância para tornar documentos digitalizados ou fotos pesquisáveis. Mas, como ainda não existe uma ferramenta de OCR que domina o mercado como a melhor opção para OCR, um estudo comparativo entre ferramentas de OCR é de suma importância para que essa informação possa ser utilizada como apoio a tomada de decisão ao usuário sobre qual ferramenta deve ser aplicada a um determinado negócio.

Neste trabalho foi realizada uma análise experimental de 11 ferramentas de OCR utilizando um *dataset* criado especificamente para o experimento, contendo 108 imagem de 5 diferentes categorias, a fim realizar uma comparação entre elas. Os resultados obtidos por cada ferramenta foram comparados considerando a acurácia por palavra e eficiência de cada uma. Além disso, outro fator considerado e avaliado foi o tempo de processamento, já que cada vez mais o mercado exige eficiência.

Como visto nos resultados apresentados no capítulo 4, o OCR Web Service e o Tesseract produziram resultados mais promissores quando comparados as outras ferramentas, obtendo em geral as melhores acurácias e um baixo custo de tempo de execução para diferentes tipos de imagens existentes no *dataset*.

A qualidade das imagens de entrada tem um impacto crucial no resultado do OCR. Todas as ferramentas de OCR apresentaram uma acurácia mais baixa para imagens com resolução baixa em 100 dpi. A maioria das ferramentas obtiveram uma maior acurácia nos resultados para as imagens em 300 dpi. Para a maioria das ferramentas não houve uma melhora na acurácia para a resolução da imagem em 600 dpi, sendo um aumento desnecessário do tamanho do arquivo e e que em alguns casos teve até um resultado um pouco pior que as imagens em 300 dpi. Como se sabe, o tempo de digitalização de uma imagem aumenta consideravelmente quando a resolução selecionada é maior. Os resultados deste trabalho fornecem indícios de que não é interessante pagar o custo de aumentar a resolução das imagens na busca de uma acurácia mais alta.

Para as imagens que possuíam variações no estilo de fonte pode-se se observar que em geral as ferramentas tiveram uma maior facilidade em reconhecer imagens com fonte normal e em negrito. Já para imagens com fonte em itálico e em negrito e itálico a acurácia foi menor. Outro fato importante foi que o tamanho de fonte que obteve um melhor resultado foi o 12, sendo que um aumento do tamanho da fonte não impactou no aumento da acurácia. O tamanho de fonte menor, como o tamanho 10, não diminuiu aparentemente a acurácia obtida.

Como parte de um trabalho futuro, pode-se acrescentar novas ferramentas para comparação e utilizar outro *dataset* em larga escala que possua outras linguagens além de português. Também pode-se ser feita uma análise estatística mais significativa para se obter uma maior precisão e confiabilidade das ferramentas OCR.

Referências

- ABBYY, P. *Portal ABBYY*. 2018. Disponível em <<https://www.abbyy.com/>>. Acesso em 19 set. 2018. Citado 2 vezes nas páginas 22 e 55.
- AIRES, S. Reconhecimento de caracteres manuscritos baseados em regiões perceptivas. *Pontifícia Universidade Católica do Paraná, M. Sc Thesis*, 2005. Citado na página 22.
- BOOKS, Q. *Guia DANFE: tudo o que você precisa saber*. 2018. Disponível em <<https://quickbooks.intuit.com/br/blog/nota-fiscal-eletronica/o-que-e-danfe/>>. Acesso em 19 set. 2018. Citado na página 47.
- BOSELLI, F. *A IMPORTÂNCIA DO EMPENHO NAS LICITAÇÕES*. 2017. Disponível em <<https://www.liciexpress.com.br/dicas/importancia-do-empenho-nas-licitacoes/>>. Acesso em 19 set. 2018. Citado na página 48.
- CHAUDHURI, A. et al. Soft computing techniques for optical character recognition systems. *Springer*, v. 352, p. 43–83, 2017. Citado 6 vezes nas páginas 25, 26, 27, 29, 30 e 31.
- DIREITO, D. *O que é uma Portaria? Para que serve? Conceito, Modalidades e Exemplos*. 2018. Disponível em <<https://dicionariodireito.com.br/portaria>>. Acesso em 19 set. 2018. Citado na página 49.
- EGESTOR, B. *DANFE online: como emitir, o que é, e o que representa*. 2018. Disponível em <<https://blog.egestor.com.br/o-que-e-e-como-emitir-uma-danfe-online/>>. Acesso em 19 set. 2018. Citado na página 47.
- EIKVIL, L. Optical character recognition. *citeseer. ist. psu. edu/142042.html*, 1993. Citado 7 vezes nas páginas 25, 26, 27, 28, 29, 30 e 31.
- ELAGOUNI, K. et al. Text recognition in multimedia documents: a study of two neural-based ocrs using and avoiding character segmentation. *International Journal on Document Analysis and Recognition (IJ DAR)*, Springer, v. 17, n. 1, p. 19–31, 2014. Citado na página 33.
- FEIJÓ, J. V. F. d. A. et al. *Análise e Classificação de imagens para aplicação de OCR em cupons fiscais*. Dissertação (Mestrado) — Florianópolis, SC., 2017. Citado 3 vezes nas páginas 28, 41 e 43.
- HOLLEY, R. How good can it get? analysing and improving ocr accuracy in large scale historic newspaper digitisation programs. *D-Lib Magazine*, D-Lib Magazine, v. 15, n. 3/4, 2009. Citado na página 45.
- ISLAM, N.; ISLAM, Z.; NOOR, N. A survey on optical character recognition system. *arXiv preprint arXiv:1710.05703*, 2017. Citado na página 29.
- LINS, L. F. M. V. Reconhecimento Ótico de caracteres(ocr) e análise de sistemas ocr baseados em código aberto. *Faculdade de Tecnologia de São Paul*, 2012. Citado na página 41.

MARANA, A. N.; PAPA, J. P.; CHIACHIA, G. Análise de desempenho de classificadores baseados em redes neurais, máquinas de vetores de suporte e florestas de caminhos ótimos para o reconhecimento de dígitos manuscritos. *V Workshop de Visão Computacional, Anais do V Workshop de Visão Computacional, Universidade Presbiteriana Mackenzie.-2009*, 2009. Citado na página 22.

MELLO, C. A. de; LINS, R. D. A comparative study on ocr tools. *Citeseer*, v. 99, p. 224–231, 1999. Citado 3 vezes nas páginas 22, 32 e 41.

NFE, N. F. E. *Sobre a NF-e*. 2018. Disponível em <<http://www.nfe.fazenda.gov.br/portal/sobreNFe.aspx?tipoConteudo=HaV+iXy7HdM=>>. Acesso em 19 set. 2018. Citado na página 47.

PORTUGUES, S. *Você entende o Hino Nacional Brasileiro?* 2018. Disponível em <<https://www.soportugues.com.br/secoes/curiosidades/hino.php>>. Acesso em 19 set. 2018. Citado na página 50.

RAABE, A.; FILHO, O. P. Estudo comparativo entre sistemáticas de digitalização de documentos: formatos html e pdf. *Ciência da Informação, SciELO Brasil*, v. 27, n. 3, 1998. Citado na página 25.

RIOS, I. *Busca por Palavras em Imagens de Documentos: Uma Abordagem Independente de OCR*. Tese (Doutorado) — Pontifícia Universidade Católica do Paraná, 2007. Citado 2 vezes nas páginas 22 e 25.

ROCHA, G. d. C. Ged - gerenciamento eletrônico de documentos: suas metodologias básicas de planejamentos para implantação. *Universidade Candido Mendes*, 2010. Citado na página 22.

SANTOS, E. P. A. Breve estudo semântico e estrutural do poema “vendaval”, de fernando pessoa. 2018. Citado na página 51.

TAFTI, A. P. et al. Ocr as a service: An experimental evaluation of google docs ocr, tesseract, abby finereader, and transym. *International Symposium on Visual Computing*, p. 735–746, 2016. Citado 2 vezes nas páginas 33 e 41.

APÊNDICE A – Acurácia das ferramentas de OCR para imagem Hino

A Tabela 66 apresenta os resultados completos de acurácia dos experimentos para a imagem Hino com todas as variações.

Tabela 66 – Acurácia das ferramentas de OCR para imagem Hino

Ferramenta OCR	Variação Fonte	Acurácia											
		100dpi			300dpi			600dpi			foto		
		10	12	14	10	12	14	10	12	14	10	12	14
Tesseract	Itálico	91,700%	96,443%	97,233%	99,209%	98,419%	98,814%	97,233%	98,814%	97,628%	99,209%	95,652%	96,443%
	Negrito	92,885%	96,047%	98,024%	100,000%	100,000%	100,000%	99,605%	100,000%	100,000%	99,605%	98,814%	99,605%
	Negrito e Itálico	94,466%	98,419%	99,209%	98,814%	99,605%	99,209%	97,628%	98,814%	97,628%	98,814%	97,233%	98,419%
	Normal	92,095%	98,419%	98,814%	99,209%	99,605%	99,605%	97,628%	100,000%	99,605%	98,814%	98,419%	99,605%
Google Cloud Vision API	Itálico	84,190%	88,538%	95,652%	94,862%	94,071%	95,257%	96,047%	96,443%	95,652%	99,209%	96,838%	97,233%
	Negrito	94,466%	95,652%	96,047%	96,443%	95,257%	95,652%	96,838%	94,466%	97,233%	98,024%	98,419%	97,628%
	Negrito e Itálico	90,119%	92,490%	95,652%	94,466%	93,676%	95,652%	95,652%	95,257%	96,838%	98,024%	98,024%	98,814%
	Normal	92,885%	94,071%	97,628%	95,652%	96,047%	96,838%	97,233%	96,443%	96,047%	98,419%	97,628%	98,814%
GOOCR	Itálico	0,000%	0,000%	2,372%	2,372%	4,743%	13,439%	5,534%	14,625%	17,391%	3,162%	1,186%	1,186%
	Negrito	3,162%	5,929%	26,877%	50,593%	44,664%	58,103%	41,897%	10,277%	7,510%	29,644%	0,791%	1,186%
	Negrito e Itálico	1,581%	2,767%	7,115%	2,767%	7,510%	12,253%	9,486%	8,696%	4,743%	3,162%	2,372%	0,000%
	Normal	5,138%	4,743%	21,739%	24,111%	39,526%	51,779%	53,360%	53,360%	43,083%	12,648%	9,486%	5,929%
ABBYY Cloud OCR	Itálico	91,304%	94,071%	96,443%	96,443%	95,257%	95,652%	96,838%	99,605%	97,628%	97,628%	95,257%	94,862%
	Negrito	94,071%	95,257%	96,838%	98,419%	97,233%	98,024%	95,257%	99,209%	99,605%	98,024%	94,466%	98,419%
	Negrito e Itálico	92,490%	94,862%	97,233%	95,652%	97,628%	95,652%	98,024%	98,814%	98,814%	98,024%	96,838%	95,257%
	Normal	90,119%	92,095%	95,652%	98,419%	97,628%	97,233%	98,814%	99,209%	98,814%	97,628%	95,652%	96,838%
OCR Document Haven OnDemand	Itálico	33,992%	47,431%	63,636%	59,684%	60,474%	65,613%	62,055%	64,427%	65,613%	51,779%	61,265%	53,360%
	Negrito	64,822%	69,170%	78,656%	81,818%	80,237%	83,004%	81,818%	81,818%	82,213%	77,866%	59,289%	69,565%
	Negrito e Itálico	41,502%	50,988%	60,474%	50,988%	62,451%	63,241%	63,241%	67,984%	68,775%	71,146%	69,170%	61,265%
	Normal	62,055%	62,451%	75,889%	79,051%	82,213%	79,447%	79,051%	80,237%	80,632%	79,051%	74,308%	71,146%

Ferramenta	OCR	Acurácia											
		100dpi			300dpi			600dpi			foto		
		Variação Fonte	10	12	14	10	12	14	10	12	14	10	12
OCR.Space API	Itálico	65,613%	82,213%	95,652%	96,443%	96,047%	97,233%	96,047%	98,024%	98,024%	87,747%	96,838%	95,652%
	Negrito	96,443%	97,628%	98,024%	100,000%	100,000%	98,024%	100,000%	100,000%	99,209%	100,000%	96,443%	100,000%
	Negrito e Itálico	83,794%	96,047%	97,628%	95,652%	98,814%	97,628%	100,000%	99,605%	99,605%	100,000%	100,000%	96,443%
	Normal	90,514%	95,257%	97,628%	98,024%	97,628%	98,024%	97,628%	98,814%	98,814%	100,000%	100,000%	96,838%
New OCR API	Itálico	55,731%	54,150%	69,565%	75,889%	75,494%	76,285%	77,470%	81,818%	78,656%	66,798%	66,403%	61,660%
	Negrito	76,285%	79,447%	86,166%	96,047%	92,885%	95,257%	96,838%	95,652%	96,047%	95,257%	60,474%	94,071%
	Negrito e Itálico	56,126%	69,565%	78,261%	88,538%	85,375%	86,957%	86,166%	88,142%	83,004%	84,585%	86,166%	84,585%
	Normal	69,960%	66,798%	81,818%	92,885%	92,885%	92,095%	92,095%	95,652%	94,071%	92,885%	86,166%	88,538%
Microsoft Vision API	Itálico	67,589%	84,190%	96,443%	96,047%	95,257%	97,233%	96,047%	97,628%	98,024%	86,561%	95,652%	96,047%
	Negrito	96,443%	98,024%	97,233%	100,000%	99,605%	97,233%	100,000%	99,209%	99,209%	100,000%	96,443%	99,605%
	Negrito e Itálico	85,375%	94,862%	96,838%	94,862%	98,419%	96,838%	100,000%	99,209%	99,209%	99,605%	100,000%	95,652%
	Normal	91,304%	95,652%	97,233%	97,628%	97,233%	97,628%	97,233%	98,419%	98,419%	100,000%	99,605%	96,838%
OCR Web Service	Itálico	95,257%	95,257%	98,024%	100,000%	100,000%	100,000%	98,419%	98,024%	100,000%	93,281%	95,257%	88,933%
	Negrito	99,605%	98,419%	99,209%	100,000%	99,209%	100,000%	100,000%	100,000%	99,209%	99,209%	93,281%	99,209%
	Negrito e Itálico	98,024%	97,628%	98,814%	99,605%	100,000%	100,000%	100,000%	99,605%	100,000%	97,628%	98,814%	98,419%
	Normal	94,862%	98,814%	99,209%	100,000%	100,000%	100,000%	98,419%	99,605%	100,000%	99,209%	99,209%	96,838%
SemaMediaData OCR	Itálico	0,000%	0,000%	0,395%	0,000%	0,000%	1,186%	0,791%	0,000%	1,581%	0,395%	0,395%	0,000%
	Negrito	0,000%	0,395%	1,186%	1,976%	1,976%	1,581%	2,372%	0,395%	1,581%	0,395%	0,395%	0,395%
	Negrito e Itálico	0,000%	0,000%	0,395%	1,186%	1,186%	1,186%	1,186%	0,395%	1,581%	0,395%	0,395%	0,395%
	Normal	0,395%	0,395%	1,186%	1,581%	0,395%	1,581%	0,395%	0,395%	1,581%	0,395%	0,395%	0,395%
Cloudmersive API	Itálico	0,791%	5,929%	11,462%	59,684%	60,870%	64,427%	53,360%	61,265%	71,937%	56,126%	46,640%	43,478%
	Negrito	7,510%	13,834%	38,735%	83,399%	81,818%	81,028%	81,423%	79,447%	80,237%	78,656%	50,988%	77,075%
	Negrito e Itálico	3,557%	7,510%	15,415%	49,012%	69,170%	69,170%	58,498%	62,451%	73,518%	60,474%	62,055%	47,431%
	Normal	11,858%	12,253%	37,154%	79,051%	81,028%	79,447%	79,842%	79,842%	79,842%	78,656%	67,589%	67,194%

APÊNDICE B – Tempo de execução das ferramentas de OCR para imagem Hino

A Tabela 67 apresenta os resultados completos de tempo de execução dos experimentos para a imagem Hino com todas as variações.

Tabela 67 – Tempo de execução das ferramentas de OCR para imagem Hino

Ferramenta OCR	Variação Fonte	Acurácia											
		100dpi			300dpi			600dpi			foto		
		10	12	14	10	12	14	10	12	14	10	12	14
Tesseract	Itálico	4,623	5,178	6,113	7,485	7,092	9,151	13,064	6,335	11,834	7,934	6,635	7,066
	Negrito	3,709	5,184	5,883	8,451	8,700	9,468	6,120	8,640	11,560	6,273	8,295	6,406
	Negrito e Itálico	4,193	4,695	5,744	7,775	7,937	9,793	6,942	7,017	11,047	6,972	6,855	7,260
	Normal	15,951	5,097	5,455	8,284	7,547	7,810	15,063	6,167	8,723	6,923	5,804	5,722
Google Cloud Vision API	Itálico	4,134	4,419	6,800	6,583	6,915	10,757	10,752	9,248	17,764	10,518	9,346	11,804
	Negrito	4,662	4,581	6,422	7,137	6,631	9,850	8,761	9,560	21,449	10,374	24,405	10,124
	Negrito e Itálico	4,292	5,187	7,196	6,663	7,318	10,768	10,226	10,296	15,252	11,182	9,705	10,187
	Normal	4,063	4,182	6,097	6,055	6,667	10,574	10,558	8,908	17,602	13,379	10,780	10,018
GOOCR	Itálico	1,498	1,364	1,648	3,576	3,821	4,136	3,490	3,038	4,652	5,614	5,388	5,748
	Negrito	1,115	1,616	1,734	3,453	3,807	3,509	3,079	4,532	8,007	4,651	7,178	5,876
	Negrito e Itálico	1,708	4,865	2,145	4,784	5,600	2,997	3,968	4,684	3,745	6,274	6,565	3,249
	Normal	1,469	1,150	1,797	2,970	2,841	1,696	2,188	3,389	2,957	5,244	6,785	4,681
ABBYY Cloud OCR	Itálico	10,298	14,634	18,183	11,372	9,827	21,002	18,859	13,023	31,520	14,123	12,520	13,295
	Negrito	10,317	9,086	18,100	11,555	10,629	20,443	14,244	12,456	32,935	28,520	14,629	13,645
	Negrito e Itálico	15,188	10,657	18,972	10,570	10,284	19,468	12,506	13,704	29,392	25,857	12,117	15,498
	Normal	9,531	9,329	17,560	11,792	14,551	18,694	21,215	12,689	24,214	14,801	14,304	12,667

Ferramenta	OCR	Acurácia											
		100dpi			300dpi			600dpi			foto		
		Variação	10	12	14	10	12	14	10	12	14	10	12
Fonte													
OCR Document Haven OnDemand	Itálico	16,904	15,588	16,161	12,710	11,983	15,333	18,920	13,755	20,476	18,454	15,935	18,302
	Negrito	9,658	9,687	9,712	8,715	9,369	11,121	13,481	10,881	18,658	18,535	21,211	13,186
	Negrito e Itálico	19,974	17,167	15,977	15,085	12,481	13,747	13,331	13,664	19,096	13,974	14,719	15,602
	Normal	11,902	10,112	10,440	9,192	8,712	11,891	14,630	10,167	18,115	11,880	12,533	12,964
OCR.Space API	Itálico	5,509	2,613	36,067	4,427	4,024	64,373	8,490	4,835	11,824	6,516	4,883	6,571
	Negrito	3,588	9,553	4,824	20,799	3,656	26,509	13,071	5,359	10,193	14,864	5,584	279,872
	Negrito e Itálico	5,752	2,538	7,319	35,041	4,117	9,071	15,821	6,213	10,406	14,443	20,537	5,903
	Normal	4,885	2,786	15,968	17,326	4,189	6,099	7,556	6,842	11,983	15,977	6,466	4,764
New OCR API	Itálico	22,672	20,646	22,126	19,212	3,111	28,540	20,877	19,404	26,056	19,323	19,721	20,169
	Negrito	15,142	13,913	15,046	13,212	13,607	21,872	25,136	13,542	19,737	15,042	5,823	13,733
	Negrito e Itálico	18,711	17,814	19,184	17,521	3,404	22,440	18,846	17,084	21,275	24,253	16,522	17,646
	Normal	16,842	14,970	15,407	15,131	15,308	21,716	16,440	13,641	20,953	14,918	14,837	14,166
Microsoft Vision API	Itálico	4,680	3,314	6,258	4,635	3,591	7,424	8,026	6,539	19,252	7,009	6,105	16,364
	Negrito	3,201	3,147	7,868	6,389	4,263	5,712	14,538	6,884	12,272	11,421	8,636	15,249
	Negrito e Itálico	4,744	3,119	8,070	4,388	4,537	11,055	6,154	6,720	10,861	13,313	6,489	5,787
	Normal	3,029	3,155	5,637	4,456	4,445	6,432	7,358	5,653	14,357	7,510	6,746	5,632
OCR Web Service	Itálico	4,512	4,544	7,936	10,340	5,762	8,134	11,556	7,528	18,575	13,359	8,120	16,463
	Negrito	3,877	3,701	6,034	5,692	4,849	9,202	13,942	7,286	14,159	17,000	10,325	9,367
	Negrito e Itálico	4,312	4,940	7,085	5,241	5,004	8,179	7,056	8,069	12,669	10,461	8,968	9,189
	Normal	4,405	3,545	5,361	5,944	5,349	8,646	9,801	7,714	12,842	10,308	8,619	7,752
SemaMediaData OCR	Itálico	11,441	12,900	22,034	17,438	19,536	28,563	25,472	20,792	36,954	31,222	23,336	35,666
	Negrito	14,188	20,722	20,559	14,867	18,387	22,696	32,461	17,049	30,427	25,756	26,567	28,017
	Negrito e Itálico	12,356	16,794	21,139	16,483	20,686	24,067	17,403	21,526	31,746	36,189	24,816	24,533
	Normal	13,019	16,123	18,281	16,881	17,793	22,201	28,884	19,379	31,256	21,788	20,374	22,895
Cloudmersive API	Itálico	13,698	12,992	16,199	10,215	10,089	13,350	13,603	15,964	16,849	11,265	13,223	22,756
	Negrito	13,912	14,184	13,366	6,206	6,899	9,355	18,879	8,743	12,682	9,300	13,632	14,230
	Negrito e Itálico	15,899	14,874	14,620	9,863	9,081	9,433	10,728	10,866	13,421	17,512	11,204	12,837
	Normal	14,207	13,539	12,245	8,206	8,140	8,742	10,763	7,828	14,027	8,354	8,467	8,715

APÊNDICE C – Acurácia das ferramentas de OCR para imagem Poema

A Tabela 68 apresenta os resultados completos de acurácia dos experimentos para a imagem Poema com todas as variações.

Tabela 68 – Acurácia das ferramentas de OCR para imagem Poema

Ferramenta OCR	Variação Fonte	Acurácia											
		100dpi			300dpi			600dpi			foto		
		10	12	14	10	12	14	10	12	14	10	12	14
Tesseract	Itálico	93,870%	98,084%	98,084%	99,234%	99,617%	99,234%	99,234%	99,617%	98,467%	99,234%	99,617%	99,617%
	Negrito	97,318%	99,234%	99,234%	100,000%	100,000%	99,617%	100,000%	100,000%	100,000%	100,000%	99,617%	99,617%
	Negrito e Itálico	95,785%	98,467%	100,000%	100,000%	100,000%	100,000%	100,000%	100,000%	100,000%	99,234%	100,000%	99,617%
	Normal	96,552%	97,701%	99,234%	100,000%	100,000%	100,000%	99,617%	100,000%	100,000%	99,617%	100,000%	99,617%
Google Cloud Vision API	Itálico	87,739%	95,019%	98,467%	98,467%	96,552%	96,935%	98,084%	97,318%	88,506%	98,467%	96,169%	99,234%
	Negrito	97,318%	95,785%	99,234%	95,402%	94,253%	96,169%	97,318%	98,084%	97,318%	98,467%	99,617%	99,617%
	Negrito e Itálico	95,402%	96,935%	96,169%	96,169%	94,253%	94,253%	96,169%	97,318%	96,169%	98,084%	98,084%	98,084%
	Normal	95,785%	97,701%	98,467%	98,467%	97,701%	98,084%	99,234%	98,467%	98,467%	98,084%	98,851%	98,084%
GOOCR	Itálico	0,000%	0,766%	0,766%	2,299%	3,065%	6,897%	6,897%	10,728%	10,728%	4,598%	8,812%	3,831%
	Negrito	1,533%	6,897%	14,559%	31,801%	50,192%	64,368%	43,295%	13,410%	6,513%	1,149%	24,521%	27,203%
	Negrito e Itálico	1,533%	2,299%	3,831%	4,215%	4,598%	6,130%	91,954%	6,513%	1,916%	5,747%	1,533%	5,747%
	Normal	0,766%	4,215%	13,027%	36,782%	39,464%	52,107%	87,739%	88,506%	40,613%	25,670%	21,839%	41,762%
ABBYY Cloud OCR	Itálico	90,421%	95,785%	96,552%	98,084%	97,701%	98,467%	97,701%	100,000%	96,552%	99,234%	99,234%	98,851%
	Negrito	97,318%	96,935%	98,084%	98,851%	99,234%	98,851%	98,851%	100,000%	100,000%	98,851%	99,617%	99,617%
	Negrito e Itálico	93,870%	96,169%	96,169%	96,935%	97,318%	98,851%	98,467%	96,935%	99,617%	98,467%	99,234%	98,084%
	Normal	96,935%	94,636%	95,785%	99,234%	98,851%	99,234%	99,617%	100,000%	99,234%	98,851%	99,234%	99,234%
OCR Document Haven OnDemand	Itálico	91,954%	50,958%	65,900%	62,069%	71,648%	77,011%	66,284%	75,479%	45,211%	72,031%	75,096%	73,563%
	Negrito	68,199%	71,648%	80,077%	84,674%	83,525%	85,824%	86,207%	83,525%	83,908%	85,441%	80,843%	82,759%
	Negrito e Itálico	39,464%	43,678%	62,069%	63,602%	67,050%	70,115%	69,732%	72,031%	76,628%	68,966%	73,946%	69,349%
	Normal	53,257%	65,517%	85,057%	85,441%	86,973%	83,908%	85,824%	84,291%	86,590%	83,908%	77,778%	86,207%

Ferramenta	OCR	Acurácia											
		100dpi			300dpi			600dpi			foto		
		Variação Fonte	10	12	14	10	12	14	10	12	14	10	12
OCR.Space API	Itálico	63,602%	90,421%	93,870%	96,552%	97,701%	97,701%	97,318%	97,701%	95,402%	98,467%	97,701%	97,701%
	Negrito	96,169%	99,234%	99,617%	99,617%	99,234%	98,851%	99,617%	99,234%	99,617%	99,617%	98,851%	99,617%
	Negrito e Itálico	87,739%	96,935%	96,935%	96,552%	95,785%	98,084%	98,084%	97,318%	98,084%	98,467%	97,318%	98,467%
	Normal	93,487%	97,318%	98,851%	99,234%	99,234%	99,234%	99,234%	98,851%	99,234%	99,617%	99,617%	99,234%
New OCR API	Itálico	47,510%	57,088%	75,479%	74,330%	75,479%	85,824%	80,077%	85,824%	54,406%	82,759%	78,544%	86,207%
	Negrito	74,713%	78,161%	85,057%	96,935%	96,935%	97,701%	96,552%	95,402%	95,785%	96,935%	85,824%	92,720%
	Negrito e Itálico	63,602%	72,414%	80,460%	85,441%	88,889%	92,337%	91,571%	88,123%	88,506%	87,739%	86,207%	92,720%
	Normal	65,134%	70,115%	86,590%	98,467%	96,552%	95,402%	96,552%	98,467%	94,636%	96,552%	67,816%	95,785%
Microsoft Vision API	Itálico	62,452%	91,571%	95,019%	96,552%	97,701%	97,318%	97,318%	97,318%	95,402%	99,617%	97,318%	97,318%
	Negrito	96,169%	99,234%	98,084%	99,617%	99,234%	98,851%	99,234%	99,234%	99,234%	99,234%	98,467%	99,234%
	Negrito e Itálico	1,916%	96,169%	96,169%	96,169%	96,935%	97,318%	98,084%	97,318%	98,467%	98,467%	96,552%	97,701%
	Normal	92,337%	97,701%	98,467%	99,234%	99,234%	99,234%	99,234%	98,851%	99,234%	99,617%	99,234%	99,234%
OCR Web Service	Itálico	94,253%	98,084%	98,851%	100,000%	100,000%	100,000%	98,851%	100,000%	98,084%	100,000%	98,851%	100,000%
	Negrito	99,234%	99,234%	100,000%	100,000%	100,000%	100,000%	100,000%	100,000%	100,000%	99,617%	99,617%	99,617%
	Negrito e Itálico	95,785%	98,467%	100,000%	100,000%	100,000%	100,000%	100,000%	100,000%	99,617%	100,000%	97,701%	98,851%
	Normal	95,785%	98,467%	100,000%	100,000%	100,000%	100,000%	99,617%	100,000%	100,000%	100,000%	98,467%	100,000%
SemaMediaData OCR	Itálico	0,383%	0,383%	1,149%	2,682%	1,149%	0,766%	3,065%	1,149%	1,149%	1,533%	0,766%	0,766%
	Negrito	0,383%	1,533%	0,766%	1,149%	2,682%	0,766%	3,448%	2,682%	1,916%	1,916%	6,897%	1,916%
	Negrito e Itálico	0,000%	0,383%	0,383%	0,383%	1,149%	0,766%	1,533%	0,766%	1,916%	1,533%	1,533%	0,766%
	Normal	1,149%	0,000%	1,533%	3,448%	1,533%	1,533%	4,981%	1,149%	1,916%	0,000%	1,533%	1,916%
Cloudmersive API	Itálico	4,215%	6,513%	17,241%	59,770%	59,387%	72,031%	74,330%	72,797%	32,950%	71,648%	58,621%	74,713%
	Negrito	6,513%	21,456%	29,885%	84,291%	85,057%	85,824%	86,207%	84,674%	85,824%	83,908%	76,628%	81,226%
	Negrito e Itálico	2,682%	6,130%	10,728%	58,621%	73,180%	74,713%	61,303%	76,628%	76,245%	47,126%	67,050%	69,732%
	Normal	10,728%	17,625%	43,678%	83,142%	84,674%	83,525%	84,674%	84,291%	83,908%	84,674%	43,678%	83,908%

APÊNDICE D – Tempo de execução das ferramentas de OCR para imagem Poema

A Tabela 69 apresenta os resultados completos de tempo de execução dos experimentos para a imagem Poema com todas as variações.

Tabela 69 – Tempo de execução das ferramentas de OCR para imagem Poema

Ferramenta OCR	Variação Fonte	Acurácia											
		100dpi			300dpi			600dpi			foto		
		10	12	14	10	12	14	10	12	14	10	12	14
Tesseract	Itálico	4,623	5,178	6,113	7,485	7,092	9,151	13,064	6,335	11,834	7,934	6,635	7,066
	Negrito	3,709	5,184	5,883	8,451	8,700	9,468	6,120	8,640	11,560	6,273	8,295	6,406
	Negrito e Itálico	4,193	4,695	5,744	7,775	7,937	9,793	6,942	7,017	11,047	6,972	6,855	7,260
	Normal	15,951	5,097	5,455	8,284	7,547	7,810	15,063	6,167	8,723	6,923	5,804	5,722
Google Cloud Vision API	Itálico	4,134	4,419	6,800	6,583	6,915	10,757	10,752	9,248	17,764	10,518	9,346	11,804
	Negrito	4,662	4,581	6,422	7,137	6,631	9,850	8,761	9,560	21,449	10,374	24,405	10,124
	Negrito e Itálico	4,292	5,187	7,196	6,663	7,318	10,768	10,226	10,296	15,252	11,182	9,705	10,187
	Normal	4,063	4,182	6,097	6,055	6,667	10,574	10,558	8,908	17,602	13,379	10,780	10,018
GOOCR	Itálico	1,498	1,364	1,648	3,576	3,821	4,136	3,490	3,038	4,652	5,614	5,388	5,748
	Negrito	1,115	1,616	1,734	3,453	3,807	3,509	3,079	4,532	8,007	4,651	7,178	5,876
	Negrito e Itálico	1,708	4,865	2,145	4,784	5,600	2,997	3,968	4,684	3,745	6,274	6,565	3,249
	Normal	1,469	1,150	1,797	2,970	2,841	1,696	2,188	3,389	2,957	5,244	6,785	4,681
ABBY Cloud OCR	Itálico	10,298	14,634	18,183	11,372	9,827	21,002	18,859	13,023	31,520	14,123	12,520	13,295
	Negrito	10,317	9,086	18,100	11,555	10,629	20,443	14,244	12,456	32,935	28,520	14,629	13,645
	Negrito e Itálico	15,188	10,657	18,972	10,570	10,284	19,468	12,506	13,704	29,392	25,857	12,117	15,498
	Normal	9,531	9,329	17,560	11,792	14,551	18,694	21,215	12,689	24,214	14,801	14,304	12,667

Ferramenta	OCR	Acurácia											
		100dpi			300dpi			600dpi			foto		
		Variação	10	12	14	10	12	14	10	12	14	10	12
OCR Document Haven OnDemand	Itálico	16,904	15,588	16,161	12,710	11,983	15,333	18,920	13,755	20,476	18,454	15,935	18,302
	Negrito	9,658	9,687	9,712	8,715	9,369	11,121	13,481	10,881	18,658	18,535	21,211	13,186
	Negrito e Itálico	19,974	17,167	15,977	15,085	12,481	13,747	13,331	13,664	19,096	13,974	14,719	15,602
	Normal	11,902	10,112	10,440	9,192	8,712	11,891	14,630	10,167	18,115	11,880	12,533	12,964
OCR.Space API	Itálico	5,509	2,613	36,067	4,427	4,024	64,373	8,490	4,835	11,824	6,516	4,883	6,571
	Negrito	3,588	9,553	4,824	20,799	3,656	26,509	13,071	5,359	10,193	14,864	5,584	279,872
	Negrito e Itálico	5,752	2,538	7,319	35,041	4,117	9,071	15,821	6,213	10,406	14,443	20,537	5,903
	Normal	4,885	2,786	15,968	17,326	4,189	6,099	7,556	6,842	11,983	15,977	6,466	4,764
New OCR API	Itálico	22,672	20,646	22,126	19,212	3,111	28,540	20,877	19,404	26,056	19,323	19,721	20,169
	Negrito	15,142	13,913	15,046	13,212	13,607	21,872	25,136	13,542	19,737	15,042	5,823	13,733
	Negrito e Itálico	18,711	17,814	19,184	17,521	3,404	22,440	18,846	17,084	21,275	24,253	16,522	17,646
	Normal	16,842	14,970	15,407	15,131	15,308	21,716	16,440	13,641	20,953	14,918	14,837	14,166
Microsoft Vision API	Itálico	4,680	3,314	6,258	4,635	3,591	7,424	8,026	6,539	19,252	7,009	6,105	16,364
	Negrito	3,201	3,147	7,868	6,389	4,263	5,712	14,538	6,884	12,272	11,421	8,636	15,249
	Negrito e Itálico	4,744	3,119	8,070	4,388	4,537	11,055	6,154	6,720	10,861	13,313	6,489	5,787
	Normal	3,029	3,155	5,637	4,456	4,445	6,432	7,358	5,653	14,357	7,510	6,746	5,632
OCR Web Service	Itálico	4,512	4,544	7,936	10,340	5,762	8,134	11,556	7,528	18,575	13,359	8,120	16,463
	Negrito	3,877	3,701	6,034	5,692	4,849	9,202	13,942	7,286	14,159	17,000	10,325	9,367
	Negrito e Itálico	4,312	4,940	7,085	5,241	5,004	8,179	7,056	8,069	12,669	10,461	8,968	9,189
	Normal	4,405	3,545	5,361	5,944	5,349	8,646	9,801	7,714	12,842	10,308	8,619	7,752
SemaMediaData OCR	Itálico	11,441	12,900	22,034	17,438	19,536	28,563	25,472	20,792	36,954	31,222	23,336	35,666
	Negrito	14,188	20,722	20,559	14,867	18,387	22,696	32,461	17,049	30,427	25,756	26,567	28,017
	Negrito e Itálico	12,356	16,794	21,139	16,483	20,686	24,067	17,403	21,526	31,746	36,189	24,816	24,533
	Normal	13,019	16,123	18,281	16,881	17,793	22,201	28,884	19,379	31,256	21,788	20,374	22,895
Cloudmersive API	Itálico	13,698	12,992	16,199	10,215	10,089	13,350	13,603	15,964	16,849	11,265	13,223	22,756
	Negrito	13,912	14,184	13,366	6,206	6,899	9,355	18,879	8,743	12,682	9,300	13,632	14,230
	Negrito e Itálico	15,899	14,874	14,620	9,863	9,081	9,433	10,728	10,866	13,421	17,512	11,204	12,837
	Normal	14,207	13,539	12,245	8,206	8,140	8,742	10,763	7,828	14,027	8,354	8,467	8,715