

ESTUDO DAS EXPRESSÕES FACIAIS COM USO DA INTELIGÊNCIA ARTIFICIAL

César Silva Pedro¹, Gabriel Gomes Machioni¹,

Rodrigo de Oliveira Plotze¹, Anna Patricia Zakem China¹

¹Faculdade de Tecnologia de FATEC Ribeirão Preto (FATEC)

Ribeirão Preto, SP – Brasil

cesar.pedro@fatec.sp.gov.br,

gabriel.machioni@fatec.sp.gov.br,

rodrigo.plotze@fatec.sp.gov.br, anna.china@fatec.sp.gov.br

Resumo. *O objetivo desse trabalho é apresentar uma pesquisa sobre a detecção de expressões faciais usando algoritmos de inteligência artificial e classificá-las nas principais emoções às quais essas expressões pertencem, empregando as técnicas-chave para pré-processamento, extração de características e classificação.*

Abstract. *The objective of this work is to present a research study on the detection of facial expressions using artificial intelligence algorithms and classify them into the main emotions to which these expressions belong, employing key techniques for pre-processing, feature extraction, and classification.*

1. Introdução

Desde a Revolução Industrial, a sociedade tem acompanhado um constante avanço tecnológico em inúmeros campos do conhecimento. Entre esses campos, destaca-se a inteligência artificial, que tem sido cada vez mais utilizada para automatizar tarefas repetitivas e melhorar a eficiência em diversos setores. Uma das áreas promissoras da inteligência artificial é o reconhecimento de expressões faciais, que permite a identificação e classificação de emoções a partir das expressões de uma pessoa.

O reconhecimento de expressões faciais tem atraído a atenção de pesquisadores devido ao seu potencial de aplicação em diversas áreas, como na medicina, psicologia, publicidade e segurança pública, entre outras. Essa tecnologia permite a detecção e análise de emoções em tempo real, o que pode ser útil para monitorar o bem-estar emocional de indivíduos e detectar sinais de problemas de saúde mental. Além disso, o reconhecimento de expressões faciais pode ser usado para melhorar a interação humano-computador, possibilitando a criação de interfaces mais intuitivas e adaptáveis às emoções do usuário (CYBIS; BETION; FAUST, 2010).

Este trabalho tem como objetivo explorar as possibilidades do reconhecimento de expressões faciais por meio de algoritmos de inteligência artificial. São abordados temas como a concepção das emoções, a forma como as expressões faciais são detectadas e classificadas em diferentes tipos de sentimentos, bem como as possibilidades de aplicação dessa tecnologia em numerosos contextos. Espera-se contribuir para um melhor entendimento do funcionamento da inteligência artificial e suas aplicações no campo do reconhecimento de emoções, ampliando assim o conhecimento sobre essa área promissora.

A aplicação da detecção de expressões faciais por IA apresenta uma enorme potencialidade para agregar valor em diversos campos. Esse trabalho permite uma compreensão das emoções humanas, auxiliando em pesquisas de mercado, saúde mental, segurança e em muitas outras aplicações. Ao capacitar as máquinas com a capacidade de interpretar as expressões faciais, abrem-se possibilidades para uma interação mais rica e inteligente entre humanos e sistemas tecnológicos.

A estrutura do artigo é organizada da seguinte forma: na introdução, é apresentada o contexto histórico do avanço tecnológico desde a Revolução Industrial e a importância da inteligência artificial, com foco no reconhecimento de expressões faciais. Em seguida, são explorados os fundamentos teóricos relacionados à concepção das emoções e aos métodos utilizados para a detecção e classificação das expressões faciais. Na seção de materiais e métodos, são detalhados os principais algoritmos, ferramentas e técnicas de inteligência artificial utilizados no reconhecimento de expressões faciais. Além disso, são apresentados estudos de caso e experimentos realizados para validar a eficácia dos modelos propostos. Já os resultados apresentam as taxas de acurácia alcançadas pelos modelos de inteligência artificial, assim como as principais descobertas e insights obtidos. São destacados padrões e características identificados nas expressões faciais relacionados a cada emoção. Por fim, na conclusão, são sumarizados os principais pontos abordados no trabalho, ressaltando a importância do reconhecimento de expressões faciais por meio da inteligência artificial em diferentes áreas de aplicação.

2. Referencial Teórico

2.1. Reconhecimento de expressões faciais

De acordo com Zhao e Zhang (2016) O Reconhecimento de Expressão Facial (FER) é um procedimento fundamental em várias áreas, como psicologia, segurança e interação humano-computador. Durante o processo de FER, é essencial extrair e classificar as características faciais, que podem ser baseadas em características geométricas ou de aparência. A extração de recursos geométricos inclui a localização e forma dos olhos, boca, nariz, sobrancelhas e outros componentes faciais. Enquanto isso, a extração de recursos baseada em aparência envolve a análise da seção exata do rosto.

O rosto apresenta geralmente três tipos distintos de sinais: sinais estáticos, sinais lentos e sinais rápidos. Os sinais estáticos englobam características duradouras como a cor da pele, incluindo aspectos relacionados à pigmentação facial, depósitos de gordura, formato do rosto, estrutura óssea e cartilaginosa, além da localização e tamanho de características faciais como sobrancelhas, olhos, nariz e boca. Já os sinais lentos referem-se a rugas permanentes, abrangendo alterações no tônus muscular e mudanças graduais na textura da pele que ocorrem ao longo do tempo (ZHAO, X.; ZHANG, 2016).

Por fim, Ekman e Friesen (2003) afirmam que os sinais rápidos envolvem mudanças na aparência facial que acontecem em questão de segundos, como o movimento dos músculos da face que levantam as sobrancelhas. Esses sinais são altamente variáveis entre os indivíduos e podem transmitir informações sobre suas emoções e intenções de forma rápida e eficaz.

2.2. Pré-Processamento de imagens

As técnicas de pré-processamento de imagens são fundamentais para melhorar a qualidade das imagens de entrada e facilitar a aplicação de outras técnicas de processamento de imagens. Segundo Gonzalez e Woods (2018), a maioria das técnicas de processamento de imagem requer que a imagem de entrada seja pré-processada antes que as etapas principais de processamento possam ser aplicadas. Algumas dessas técnicas incluem:

Segmentação e redimensionamento de imagens são processos cruciais no pré-processamento de imagens. A segmentação envolve isolar a região de interesse (ROI), permitindo análises específicas nessa área. O redimensionamento padroniza as imagens para uma resolução ou tamanho específico, facilitando a comparação e a aplicação de algoritmos consistentes. Essas técnicas são essenciais para obter resultados confiáveis em várias aplicações de processamento de imagens. (CHEN et al., 2003).

A amostragem descendente de Bessel é uma técnica de amostragem que preserva as informações de alta frequência em uma imagem, reduzindo a taxa de amostragem da imagem em uma proporção específica. Isso pode reduzir a resolução da imagem, mas mantém as características importantes. (SZELISK, 2010)

Filtro Gaussiano é um filtro de suavização que reduz o ruído, realiza suavização de imagens, substituindo cada pixel pela média ponderada dos pixels vizinhos. O grau de suavização é determinado pelo tamanho da máscara utilizada no filtro, é um dos filtros mais comuns e eficazes para remover ruído gaussiano em imagens. (SANCHES, 2015).

Segundo Gonzales e Woods (2018) a Normalização envolve a normalização da escala de cinza em uma imagem para um intervalo padrão. Este processo auxilia a garantir que a imagem tenha uma faixa dinâmica consistente e facilita a comparação de imagens.

Já o filtro mediano é um filtro de suavização não linear que remove ruído de impulso (ou ruído *salt and pepper*) em uma imagem. Ele substitui o valor de um pixel pelo valor mediano dos pixels em sua vizinhança. Segundo Gonzalez e Woods (2018), o filtro mediano é eficaz para remover ruído de impulso em imagens.

Além dessas técnicas, existem outras que podem ser úteis para melhorar a qualidade das imagens, como a localização usando o algoritmo Viola-Jones, o alinhamento facial com o algoritmo de fluxo *Scale-Invariant Feature Transform* (SIFT), a segmentação de ROI e a equalização de histograma. (VIOLA e JONES, 2001).

2.3. Extração de Recursos das imagens

O processo de extração de recursos é uma etapa importante no sistema de Reconhecimento de Expressão Facial (FER). Tal processo envolve encontrar e descrever características relevantes em uma imagem para processamento posterior. Existem diferentes métodos de extração de recursos categorizados em cinco tipos: baseados em textura, baseados em borda, baseados em recursos globais e locais, baseados em características geométricas e baseados em patches. (ZHAO, X.; ZHANG, 2016)

Segundo Revina (2018) os métodos baseados em recursos de textura utilizam descritores como o filtro Gabor, que extrai informações de magnitude e fase para representar a organização da imagem facial. O *Local Binary Pattern* (LBP) é outro

descritor utilizado para extrair características de textura por meio de códigos binários. Outros descritores, como *Gaussian Laguerre* (GL), *Vertical Time Backward* (VTB) e Descritor Local de Weber (DLV), também são empregados para extrair características de textura das imagens faciais.

Por sua vez, os métodos baseados em borda utilizam descritores como *Line Edge Map* (LEM), *Histogram of Oriented Gradients* (HOG) e *Graphics-processing unit based Active Shape Model* (GASM) para extrair características relacionadas à estrutura geométrica e informações de borda das imagens faciais. (REVINA, 2018)

Um outro método que realiza extração de características é o PCA. Este tem por finalidade básica a análise dos dados usados visando sua redução, eliminação de sobreposições e a escolha das formas mais representativas de dados a partir de combinações lineares das variáveis originais. (ALCOFORADO, 2016, p. 382).

Asthana, Zafeiriou e Pantic (2013) afirmam que os métodos baseados em recursos globais e locais empregam técnicas como *Principal Component Analysis* (PCA), *Independent Component Analysis* (ICA) e *Stepwise Linear Discriminant Analysis* (SWLDA) para extrair características globais e locais das imagens faciais. Já os métodos baseados em características geométricas utilizam descritores como *Local Curvelet Transform* (LCT) para extrair características geométricas, como média, entropia e desvio padrão.

De acordo com Revina (2018) os métodos baseados em patches extraem características faciais de acordo com as características de distância. Isso é feito por meio da extração e correspondência de patches, traduzindo os patches extraídos em características de distância.

Esses descritores de recursos podem produzir vetores de características de alta dimensão, e, geralmente, algoritmos de redução de dimensionalidade, como PCA, são aplicados para reduzir a dimensão dos vetores. Além disso, algoritmos de seleção de características, como *Adaboost* e pontuações de similaridade, também são utilizados para selecionar as características mais relevantes (VIOLA E JONES, 2001).

2.4. Classificação

Como exemplificado no diagrama da figura 1 a etapa de classificação está próxima de concluir o objetivo que é detectar a emoção.

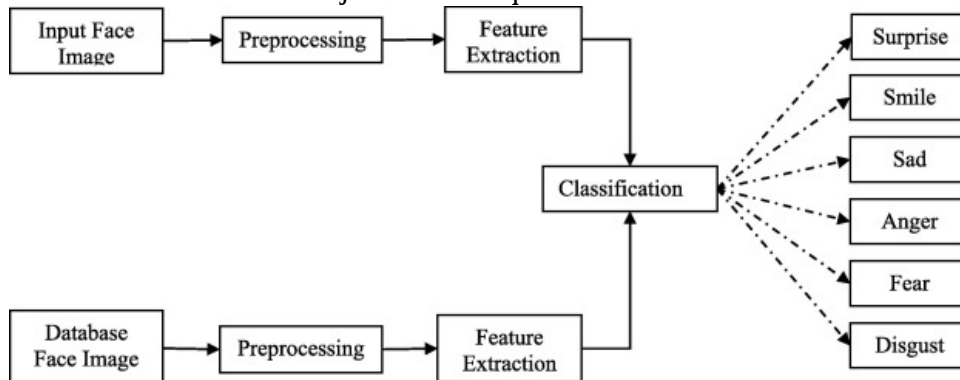


Figura 1. Processo para reconhecimento de expressões faciais e classificação

Fonte: (REVINA, 2018)

Segundo (GUDIPATI et al.,2016, p. 1) “O termo Reconhecimento de Expressão Facial significa a classificação dos elementos faciais em uma das seis emoções básicas: alegria, tristeza, repulsa, medo, raiva e surpresa”. Como demonstra na figura 2.



Figura 2. Seis expressões facial básicas
Fonte: (GUDIPATI et al.,2016)

De acordo com GUDIPATI et al. (2016) extraída as características faciais, algoritmos são aplicados para identificar a emoção expressa na face, podendo ser baseados em *machine learning* como o algoritmo *AdaBoost*, que combina vários classificadores fracos para formar um classificador forte. Continuamente utilizando características baseadas em Haar, trazendo padrões retangulares de intensidade de pixel, para acelerar a detecção de rosto e a extração de características.

Outro exemplo é a técnica de detecção a partir de caricaturas faciais. Tal técnica utiliza de detector de bordas extraíndo características estruturais e geométricas de um modelo de expressão esboçado pelo usuário, combinando-o com o descritor de mapa de borda de linhas (LEM) da imagem de rosto de entrada, como exemplifica na figura 3. A partir desse passo é utilizado a distância proposta de Hausdorff de segmentos de linhas direcionadas (dLHD) medindo a similaridade entre o LEM do rosto de consulta e os modelos de caricatura de expressões. Essa medida de disparidade é robusta a variações de expressão, o que é importante para o reconhecimento automático (Gao et al., 2003).

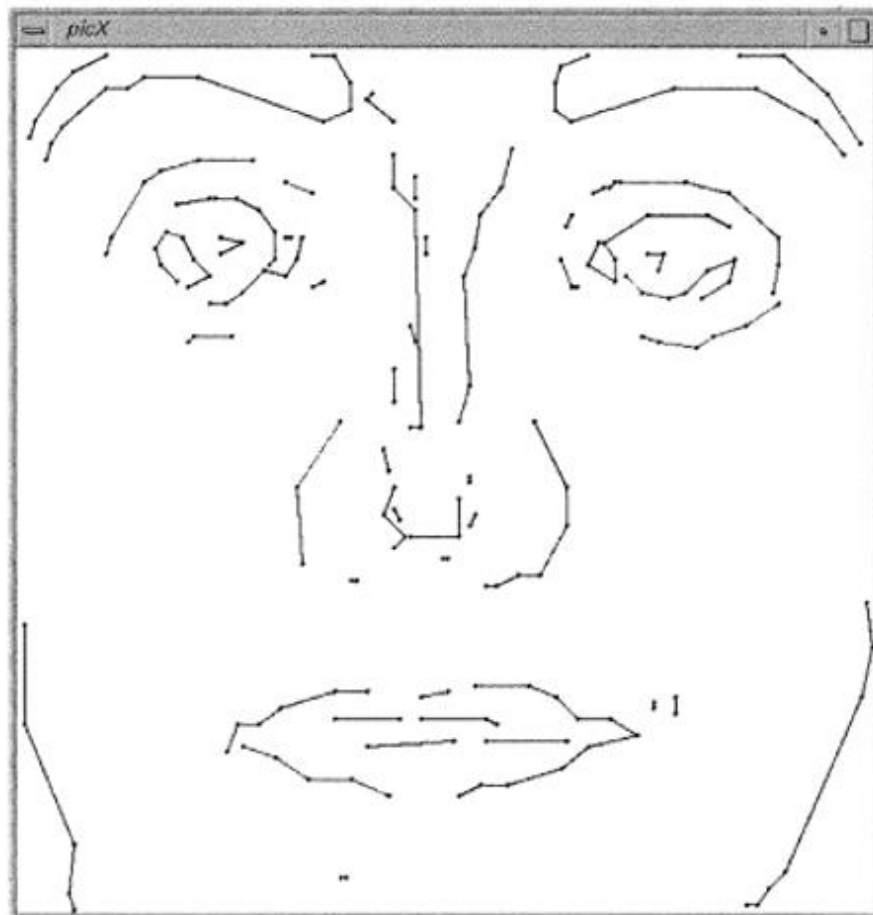


Figura 3. Ilustração do LEM de um rosto

Fonte: (Gao et al., 2003)

3. Materiais e Métodos

3.1 Pipeline software

O processo para realizar o reconhecimento das expressões faciais estão representadas na Figura 4. Para esse estudo foi utilizado um script Python disponibilizado no *Github*. Todo o software foi desenvolvido utilizando bibliotecas livres ou padrões do Python, no caso foram utilizadas as bibliotecas *opencv*, *numpy* e *tensorflow* e todos os passos para testar o código estão descritos no *readme.md* do repositório. Para reconhecimento das expressões faciais foram convidadas vinte pessoas.

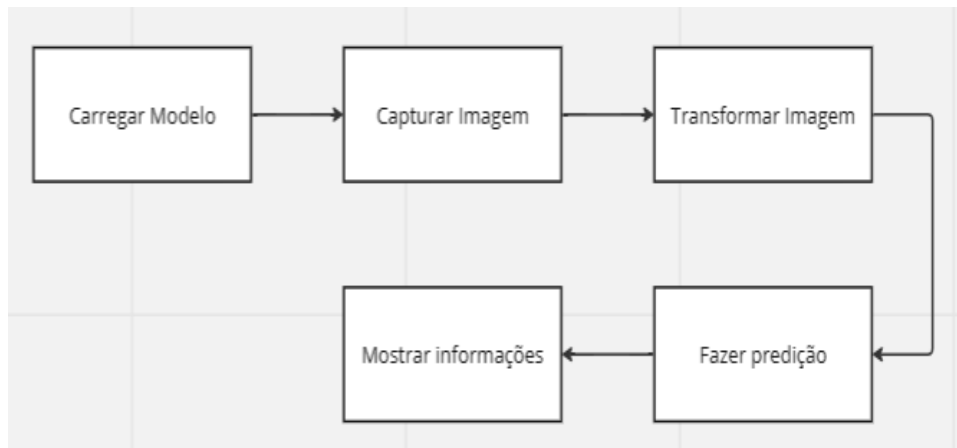


Figura 4. Diagrama de blocos do software
Fonte: (Autoria própria)

3.2 Tecnologias de desenvolvimento

Python é uma linguagem de programação de alto nível, interpretada e de propósito geral, criada em 1991 por Guido van Rossum. Muito conhecida por sua sintaxe simples e clara, facilidade de leitura e escrita, além de sua vasta biblioteca padrão e grande comunidade de desenvolvedores. Suas principais características incluem tipagem dinâmica, linguagem orientada a objeto, há uma grande quantidade de módulos e bibliotecas padrão, assim como é uma linguagem interpretada o que significa que o código é compilado e executado em tempo de execução. Além disso, a linguagem é amplamente utilizada em áreas como desenvolvimento web, ciência de dados, inteligência artificial, automação de tarefas, entre outras. É uma linguagem de código aberto e gratuita, disponível para Windows, Mac e Linux. (PYTHON, 2023)

A OpenCV (*Open Source Computer Vision Library*) é uma biblioteca de código aberto, esta oferece suporte para o processamento de imagens e visão computacional em tempo real. Ela foi desenvolvida originalmente pela Intel em 1999, também é utilizada em várias áreas, como robótica, automação industrial, segurança, entre outras. Esta biblioteca suporta várias linguagens de programação, incluindo Python. Algumas das suas principais funcionalidades incluem captura e processamento de imagens em tempo real, detecção de objetos e reconhecimento de padrões, processamento de imagem e análise de vídeo (OPENCV, 2022).

TensorFlow é uma biblioteca de software de código aberto desenvolvida pelo Google para aprendizado de máquina e inteligência artificial. Tal biblioteca foi lançada em 2015 e tornou-se uma das bibliotecas mais populares para construção de modelos de aprendizado profundo. Esta biblioteca é altamente flexível e pode ser utilizada em várias plataformas, incluindo CPUs, GPUs e dispositivos móveis. (TENSORFLOW)

De acordo com a documentação *Pandas Documentation* (2023) e a *Numpy Documentation* (2022), Pandas e Numpy são bibliotecas de software de código aberto para Python, que são frequentemente utilizadas em ciência de dados, análise financeira e outras áreas relacionadas à matemática e estatística. NumPy é uma biblioteca que fornece suporte para trabalhar com matrizes e vetores multidimensionais, além de funções matemáticas que podem ser aplicadas em grandes quantidades de dados. Já a biblioteca Pandas é para análise de dados que oferece suporte para a manipulação, limpeza e análise

de dados tabulares.

4. Resultados

A partir da revisão bibliográfica e da análise de estudos de caso, foi possível identificar as principais técnicas utilizadas na detecção e classificação de expressões faciais, bem como os desafios e limitações dessa tecnologia.

O estudo também investigou o reconhecimento de expressões faciais em diferentes sujeitos utilizando um algoritmo com um modelo já treinado. Os dados adquiridos indicam que há uma certa confiança acima de 50% para expressões de neutro e feliz, sendo as demais emoções com uma taxa de confiança muito menor, como a de medo, que em uma automação poderia ser confundida facilmente com outra expressão. Como demonstrado na Tabela 1.

Tabela 1. Análise das imagens de teste

	César	Gabriel	Edna	Leticia	Roberto
Bravo	18%	21%	17%	23%	25%
Medo	12%	9%	13%	10%	8%
Surpreso	13%	15%	15%	14%	12%
Triste	37%	35%	39%	36%	33%
Feliz	90%	63%	81%	75%	84%
Nojo	38%	26%	29%	35%	28%
neutro	55%	69%	76%	59%	58%

Fonte: Autoria própria

5. Considerações finais

O uso da inteligência artificial e suas técnicas para reconhecimento de expressões faciais é altamente eficaz em diversas áreas de atuação. No entanto, é importante destacar que a obtenção de resultados precisos requer um modelo de inteligência artificial bem treinado. A qualidade do modelo utilizado desempenha um papel crucial na detecção das expressões faciais.

Para alcançar uma boa precisão, é fundamental contar com um grande volume de dados para o treinamento do modelo, bem como realizar ajustes para evitar viés e garantir sua eficácia em diferentes contextos. Entretanto, durante a implementação desse projeto, dois fatores limitadores surgiram, a restrição de tempo e o número limitado de pessoas disponíveis para testar a implementação.

A limitação de tempo impôs restrições ao tempo dedicado para o desenvolvimento do projeto. Essa limitação pode ter impactado a extensão em que os objetivos propostos foram totalmente alcançados.

Além disso, o número limitado de pessoas disponíveis para testar a implementação também representa um desafio. A análise e validação dos resultados obtidos exigem a participação de um grupo diversificado de indivíduos, a fim de garantir a eficácia do modelo em diferentes contextos e situações.

Apesar desses fatores limitadores, o projeto em sua versão final conseguiu explorar ao máximo o potencial do reconhecimento de expressões faciais por meio de

algoritmos de inteligência artificial. Porém, é importante ressaltar que a falta de tempo e o número reduzido de pessoas para testar a implementação podem ter influenciado parcialmente a conclusão dos objetivos propostos.

A partir desse trabalho será dada continuidade com aplicações na área da saúde mental, uma vez que o reconhecimento de expressões faciais pode ser utilizado para auxiliar no diagnóstico e tratamento de distúrbios emocionais. Com o aprimoramento dos algoritmos, é possível identificar padrões específicos de expressões faciais associados a diferentes condições de saúde mental, permitindo intervenções precoces e personalizadas. O trabalho subsequente será para otimizar esta aplicação para *real time*.

6. REFERÊNCIAS

- ASTHANA, A.; ZAFEIRIOU, S.; PANTIC, M. Incremental Face Alignment in the Wild. In: IEEE Conference on Computer Vision and Pattern Recognition, 2013. p. 1859-1866. DOI: 10.1109/CVPR.2013.242.
- ALCOFORADO, M. Redes Neurais Autoassociativas de Arquitetura Profundo para Reconhecimento de Expressões Faciais. Revista de Engenharia e Pesquisa Aplicada, v. 2, n. 1, 8 nov. 2016.
- CHEN, H.-M.; VARSHNEY, P. K.; SLAMANI, M.-A. On registration of regions of interest (ROI) in video sequences. In: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, 2003. IEEE, 2003. p. 313-318.
- CYBIS, W.; BETION, A. H.; FAUST, R. Ergonomia e Usabilidade: Conhecimentos, métodos e técnicas. [São Paulo]: Novatec Editora, 2010
- EKMAN, P.; FRIESEN, W. V. Unmasking the face: A guide to recognizing emotions from facial expressions. [S.l.]: Malor Books, 2003.
- GAO, Y. et al. Facial expression recognition from line-based caricatures. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, v. 33, n. 3, p. 407-412, 2003.
- GONZALEZ, Rafael C.; WOODS, Richard E. Processamento de Imagens Digitais. 4. ed. São Paulo: Editora Blucher, 2018.
- GUDIPATI, V. K. et al. Efficient facial expression recognition using AdaBoost and Haar cascade classifiers. In: 2016 Annual Connecticut Conference on Industrial Electronics, Technology & Automation (CT-IETA). IEEE, 2016. p. 1-4.
- NUMPY Documentation. Disponível em: <<https://numpy.org/doc/>>. Último acesso em 17/05/2023.
- OPENCV. Open Source Computer Vision. Disponível em: <<https://docs.opencv.org/4.x/index.html>> Último acesso em 16/05/2023.
- PANDAS Documentation. Disponível em: <<https://pandas.pydata.org/pandas-docs/stable/>>. Último acesso em 17/05/2023.
- PYTHON. Beginner's Guide to Python. Disponível em: <<https://www.python.org/doc/>>. Último acesso em 17/05/2023.
- REVINA, I. M.; EMMANUEL, W. R. S. A survey on human face expression

- recognition techniques. *Journal of King Saud University - Computer and Information Sciences*, v. 33, n. 6, p. 619–628, 2018.
- SANCHES, C. H. et al. Técnicas de suavização de imagens e eliminação de ruídos. *Anais do EATI–Encontro Anual de Tecnologia da Informação*. Frederico Westphalen–RS, p. 21-30, 2015.
- SZELISKI, R. *Computer Vision: Algorithms and Applications*. London: Springer-Verlag, 2010.
- TENSORFLOW. TensorFlow. Disponível em: < <https://www.tensorflow.org/?hl=pt-br> > Último acesso em 17/05/2023
- VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001. p. I-511-I-518.
- YUNG KONG, T.; ROSENFELD, A. *Topological algorithms for digital image processing*. [S.l.]: North-Holland, 1996.
- ZHAO, X.; ZHANG, S. A Review on Facial Expression Recognition: Feature Extraction and Classification. *IETE Technical Review*, v. 33, n. 5, p. 505–517, 2016.