

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E  
TECNOLOGIA DE SÃO PAULO  
CÂMPUS GUARULHOS

THIAGO PINTO MARTINS

**ENERGIA VISUAL EMOCIONAL (E.V.E. VERSÃO 3.0) – PLATAFORMA DE  
ACESSIBILIDADE COMUNICATIVA QUE UTILIZA RECONHECIMENTO  
FACIAL E MENSAGENS EM ÁUDIO PARA INCLUSÃO DE PESSOAS COM  
DEFICIÊNCIA VISUAL OU CEGUEIRA**

GUARULHOS, SP

2021

THIAGO PINTO MARTINS

**ENERGIA VISUAL EMOCIONAL (E.V.E. VERSÃO 3.0) – PLATAFORMA DE  
ACESSIBILIDADE COMUNICATIVA QUE UTILIZA RECONHECIMENTO  
FACIAL E MENSAGENS EM ÁUDIO PARA INCLUSÃO DE PESSOAS COM  
DEFICIÊNCIA VISUAL OU CEGUEIRA**

Relatório de projeto apresentado como exigência para participar da 11ª Feira de Ciência e Tecnologia da região de Bragantina (BRAGRANTEC), desenvolvido por alunos do Curso Técnico em Informática para Internet Integrado ao Ensino Médio, do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Câmpus Guarulhos.

Orientador: Prof. Robson Ferreira Lopes

Coorientador: Nathalia Assunção das Chagas

GUARULHOS, SP

2021

*Dedicamos este trabalho primeiramente a Deus, e aos  
nossos pais e familiares, pois sem eles não chegaríamos  
até aqui.*

## **AGRADECIMENTOS**

Gostaria de agradecer ao Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP), Câmpus Guarulhos, que me proporcionou essa oportunidade de aprendizado, além da disponibilização de recursos que foram utilizados no projeto até o momento e, também, ao apoio da direção do IFSP, mediado pelo Diretor-Geral Professor Ricardo Agostinho Rezende.

Agradeço principalmente ao meu orientador, professor Me. Robson Ferreira Lopes, que sempre motivou a continuar o projeto com paciência e sabedoria, e a professora Gema Bezerra, pelas correções e incentivos.

Agradeço também a minha coorientadora Nathalia Assunção das Chagas, por ter iniciado o projeto e acompanhando cada etapa, seguindo para nova função coorientar o projeto para suas novas fases.

Não menos importante, agradeço a todos que direta ou indiretamente nos ajudaram e tornaram possível essa realização: a todos vocês, em nome de toda a equipe, nossos agradecimentos.

## RESUMO

Na contemporaneidade, a comunicação cinésica é considerada um aspecto facilitador nas relações interpessoais, isso porque transmite informações que muitas vezes passam despercebidas no discurso verbal. Desse modo, pessoas com deficiência visual ou cegueira acabam perdendo uma parte das informações transmitidas no decorrer de uma conversa, o que pode resultar em má compreensão durante as interações sociais e na reclusão desses indivíduos. Em virtude disso, a plataforma Energia Visual Emocional (E.V.E. versão 3.0) tem como objetivo fomentar a inclusão social das pessoas com deficiência visual a partir de um protótipo de software capaz de reconhecer as faces de pessoas presentes em uma imagem e de traduzir em mensagens de áudio sua identificação e traços visuais que indiquem suas emoções. Utilizando o método de engenharia a partir da prototipação, a plataforma utiliza as técnicas de Aprendizagem Profunda (*Deep Learning*) em conjunto com a biblioteca *Face Recognition* para realizar o reconhecimento facial. Os próximos passos consistem em aprimorar a eficiência do reconhecimento de emoções, além de aprofundar os estudos para adaptar o projeto para o dispositivo móvel. Os processos de desenvolvimento, bem como os fatores levados em consideração na etapa inicial do projeto são detalhados no decorrer dessa pesquisa.

**Palavras-chave:** Inclusão. Deficiência visual ou cegueira. Inteligência artificial. Reconhecimento de faces.

## **ABSTRACT**

Nowadays, kinesic communication is considered a facilitating aspect in interpersonal relationships, because it transmits information that often goes unnoticed in verbal discourse. Thus, people with visual impairment or blindness end up losing part of the information transmitted during a conversation, which can result in poor understanding during social interactions and in the seclusion of these individuals. As a result, the Visual Emotional Energy platform (EVE version 3.0) aims to promote the social inclusion of people with visual impairments through a software prototype capable of recognizing the faces of people present in an image and translating them into messages. audio your identification and visual traits that indicate your emotions. Using the engineering method based on prototyping, the platform uses Deep Learning techniques together with the Face Recognition library to perform facial recognition. The next steps are to improve the efficiency of emotion recognition, in addition to further studies to adapt the project to the mobile device. The development processes, as well as the factors taken into consideration in the initial stage of the project, are detailed during the course of this research.

**Keywords:** Inclusion. Visual impairment or blindness. Artificial intelligence. Face recognition.

## SUMÁRIO

|                                              |           |
|----------------------------------------------|-----------|
| <b>1. INTRODUÇÃO</b>                         | <b>8</b>  |
| <b>2. OBJETIVOS E RELEVÂNCIA DO TRABALHO</b> | <b>9</b>  |
| <b>3. FUNDAMENTAÇÃO TEÓRICA</b>              | <b>10</b> |
| <b>3.1 Deep Learning</b>                     | <b>10</b> |
| <b>3.2 Python e as bibliotecas</b>           | <b>10</b> |
| <b>3.3 Rede Neural Convolucional</b>         | <b>12</b> |
| <b>4. DESENVOLVIMENTO DO PROTÓTIPO</b>       | <b>12</b> |
| <b>5. RESULTADOS</b>                         | <b>17</b> |
| <b>6. CONSIDERAÇÕES FINAIS</b>               | <b>18</b> |
| <b>7. REFERÊNCIAS</b>                        | <b>19</b> |

## 1. INTRODUÇÃO

Na contemporaneidade, a sociedade torna-se cada vez mais orientada visualmente. Segundo Viana (2014) as expressões faciais são um dos meios mais importantes nas relações interpessoais, isso porque funcionam como uma confirmação de ideias, reforçando o discurso, além de serem a forma mais comum de expressar as emoções dos indivíduos. Ainda segundo Viana (2014), que analisa o estudo do antropólogo Ray Birdwhistell (1971)<sup>1</sup>, a comunicação cinésica, como gestos e expressões faciais, transmitem entre 65% e 70% das informações durante uma interação social, mostrando que apenas cerca de 30% das informações são transmitidas pela comunicação verbal. Desse modo, pessoas com deficiência visual ou cegueira (pessoas doravante mencionadas neste trabalho como deficientes visuais ou cegas) acabam perdendo uma parte das informações transmitidas durante uma conversa, o que pode resultar em má compreensão durante as interações sociais. De acordo com a psicóloga Maria C.C. Barczinski (2001), que comenta o estudo do psiquiatra Luis Cholden (1984), a falta de visão se torna um grande gatilho para a perda da identidade, falta de vontade de viver e depressão, reações semelhantes às do luto. Em vista disso, contornar esse obstáculo na comunicação é um grande passo para reduzir as lacunas de uma vida sem visão, incluindo os deficientes visuais na sociedade e diminuindo os casos de isolamento e depressão.

Em virtude do que foi mencionado, o projeto tem como objetivo a construção de uma plataforma capaz de identificar as faces e emoções humanas, primeiramente em imagens, e, depois, em interações reais, utilizando a inteligência artificial, e descrevê-las por meio de mensagens de áudio, melhorando a qualidade de vida das pessoas com deficiência visual, facilitando seu convívio e interações sociais. Nesse trabalho serão descritos os processos de levantamento de necessidades e material teórico, bem como o processo de prototipação da plataforma que utiliza elementos do *Deep Learning* com a ciência social aplicada.

## 2. OBJETIVOS E RELEVÂNCIA DO TRABALHO

O protótipo da plataforma em desenvolvimento conta com o reconhecimento facial por meio de uma biblioteca denominada *face\_recognition*, que usa conceitos de rede neural para identificar quantas e quem são as pessoas presentes em uma imagem, e a biblioteca *Espeak*, que emite um áudio com o nome das pessoas encontradas na imagem.

O objetivo é melhorar a autonomia das pessoas com deficiência visual usando inteligência artificial para identificar faces humanas em uma imagem e reconhecê-las por meio de mensagens de áudio, melhorando a sua qualidade de vida e facilitando o seu convívio com pessoas videntes. O projeto, que aplica elementos do *Deep Learning* com a ciência social aplicada, partiu de uma pesquisa bibliográfica e vem utilizando o método de engenharia a partir da prototipação.

Objetivo Geral:

- Desenvolver uma plataforma, denominada Energia Visual Emocional (E.V.E), capaz de realizar o reconhecimento facial em imagens, extraindo suas informações e transmitindo-as em formato de mensagens de áudio.

Objetivos específicos:

- Realizar uma pesquisa aprofundada acerca das ações fundamentais dos músculos individuais (*Action Units*) e investigar o melhor meio para a sua implementação no software.
- Realizar testes de funcionalidade e validação do E.V.E.

### 3. FUNDAMENTAÇÃO TEÓRICA

Essa seção deste trabalho tem como objetivo apresentar os conceitos utilizados para a fomentação dessa pesquisa.

#### 3.1 *Deep Learning*

Segundo Goodfellow, Bengio e Courville (2016), *Deep Learning* ou aprendizagem profunda, tecnologia que se encontra como uma subcategoria da inteligência artificial, é um ramo da ciência da computação que propõe modelos de dispositivos que simulem a capacidade de raciocínio humano em termos de hierarquia de conceitos, produzindo decisões e resultados confiáveis, com a menor intervenção humana possível.

Dessa tecnologia resulta uma série de camadas de processamento: a máquina realiza etapas para tentar descobrir determinada informação. Apresentando a imagem de um objeto para um computador – um *input* (dados fornecidos) –, ele tentaria entendê-lo dividindo a análise desse objeto em diversas camadas de processamento chamadas de *Hidden Layers*. Usando várias camadas de processamento de dados não lineares, obtém-se uma representação complexa e abstrata dos dados de forma hierárquica, sem a intervenção humana no processo.

Dados sensoriais (pixels de imagens, por exemplo) são alimentados em uma primeira camada, sendo que a saída de cada camada se torna a entrada da camada seguinte. O aprendizado profundo é baseado no conceito de redes neurais artificiais, ou sistemas computacionais que imitam a maneira como o cérebro humano funciona.

#### 3.2 Python e as bibliotecas

Segundo a [python.org](https://www.python.org), Python é uma linguagem fácil de aprender e poderosa, pois possui estruturas de dados de alto nível, o que a torna mais eficiente com uma abordagem simples e efetiva de programação orientada a objetos.

Neste projeto os módulos usados mais importantes são: A *Face Recognition*, a biblioteca Dlib, Pyttx4, um sintetizador de voz compacto de software de código aberto para inglês e outros idiomas, e a API Keras em conjunto com o dataset FER2013 (Facial Expression Recognition 2013 Dataset) para identificar as emoções.

A Face Recognition foi criada por Adam Geitgey, uma API (*Application Programming Interface*) que possibilita o reconhecimento facial, baseada no estado da arte da biblioteca Dlib. É desenvolvida na linguagem Python e disponibilizada gratuitamente no GitHub do próprio autor. ([https://github.com/ageitgey/face\\_recognition](https://github.com/ageitgey/face_recognition)).

O Dlib é uma biblioteca gratuita, desenvolvida em C ++, projetada para ser utilizada como uma coleção de componentes de framework que podem ser utilizados em qualquer tipo de aplicação, seja comercial ou acadêmica. No presente, ele fornece módulos que suportam álgebra linear, processamento de imagens, otimização e outras funções, e todos fornecem uma documentação rica e de simples acesso.

O Keras é um framework, biblioteca Python voltada a aprendizado de máquina profundo. Esta biblioteca implementa funcionalidades de frameworks de deep learning, já consolidadas, entre elas o TensorFlow.

Para desenvolvimento do trabalho foi utilizado a base de dados Facial Expression Recognition 2013 (FER-2013). Esta base de dados foi desenvolvida por Pierre Luc Carrier e Aaron Courville, e possui um total 35887 imagens faciais RGB de diferentes expressões, com dimensões de 48x48 pixels (GOODFELLOW et al., 2013). Os dados possuem as seis expressões faciais de Ekman, adicionadas à expressão neutra. As amostras desta base de dados estão distribuídas em: 7215 imagens de alegria; 436 de desgosto; 4097 de medo; 4965 de neutra; 3995 de raiva; 3171 de surpresa; e 4830 de tristeza.

Para o desenvolvimento da interface foi utilizado o pacote (GUI - Graphic User Interface) Tkinter, que é o pacote GUI padrão do Python. É uma fina camada orientada a objetos em cima de Tcl/Tk. Tkinter não é o único kit de ferramentas gui-programming para Python. No entanto, é o mais usado.

### 3.3 Rede Neural Convolucional

De acordo com Vargas, Carvalho e Vasconcelos (2016), uma CNN - *Convolutional Neural Network*, cuja tradução literal é “Rede Neural Convolucional”, é um algoritmo do *Deep Learning* inspirado no “processo biológico de processamentos de dados visuais”.

Enquanto a abordagem tradicional de computação utiliza uma série de blocos lógicos para executar uma determinada tarefa, Redes Neurais utilizam redes compostas por nós (os quais atuam como neurônios) e arestas (que por sua vez atuam como sinapses) para processar dados. As entradas então percorrem o sistema e uma série de saídas é gerada. Essas saídas são então comparadas com dados conhecidos.

Segundo Goodfellow, Bengio e Courville (2016) a CNN, na sua execução, utiliza uma construção matemática que é tipicamente composta por três tipos de camadas (ou blocos de construção): convolução, agrupamento e camadas totalmente conectadas. As duas primeiras, camadas de convolução e agrupamento, realizam extração de recursos, enquanto a terceira, uma camada totalmente conectada, mapeia as características extraídas para a saída final, como classificação.

## 4. DESENVOLVIMENTO DO PROTÓTIPO

Para programar o protótipo, foi necessária uma pesquisa que resultou na utilização de uma programação em python em conjunto com a biblioteca *Face Recognition* (escrita em linguagem python), que identifica características do rosto e é uma das ferramentas fundamentais usadas para o reconhecimento facial.

Para viabilizar o uso da biblioteca, foram instalados os pacotes python3-numpy==1.16.4, python3-scipy, python3-matplotlib ipython3, python3-pandas, python3-sympy, python3-nose, python3-keras, python3-opencv, python3-pyttsx3, python3-imutils em uma máquina com sistema operacional Windows e, depois, com o pip (ferramenta utilizada para baixar pacotes do repositório Python) foi feita a instalação da *face\_recognition*. Além da biblioteca, foram utilizados módulos como dlib, que é um moderno conjunto de algoritmos escritos em C++ para aprendizado de máquina, PIL.image e Pyttx4 (sintetizador de voz compacto de software de código aberto para inglês e outros idiomas).

Vários testes da biblioteca foram realizados até chegar ao protótipo usado neste projeto. Na primeira versão foi desenvolvido o código `facesample01.py` e os resultados foram as coordenadas de cada rosto em qualquer foto indicada no programa. No segundo teste foi possível identificar as faces e fazer o software indicar a quantidade de pessoas na foto, e assim as atualizações foram acontecendo a cada nova versão.

Neste protótipo atual utilizamos alguns pacotes conhecidos no mundo da deep learning voltada para reconhecimento facial. A nossa aplicação utiliza da (API - Application Programming Interface) `face-recognition`, que dentro dela estão compactadas algumas bibliotecas, tais como o `Open-CV` que utilizamos para a parte visual do projeto como apresentação do funcionamento do algoritmo.

Na primeira etapa temos um arquivo nomeado de (*reconhecimento treinamento.py*), ele irá receber as imagens que serão armazenadas no software e treinar a rede neural principal para identificar as faces cadastradas. Para cada inserção de imagem, é necessário treinar uma foto .png ou .jpeg com apenas uma face contida. Assim, o algoritmo é capaz de salvar duas faces ao mesmo tempo em uma única imagem. Caso o usuário insira uma imagem que contenha duas faces, o algoritmo irá interromper o processo e irá dizer que a mais de uma face contida na imagem.

Ao anexar as imagens de teste no treinamento conforme o anexo 01. Obtemos o seguinte resultado:

Anexo 01 - Arquivos de imagem usados como base para comparação pelo programa do projeto E.V.E.



Foto: (Reprodução/Cinema com Rapadura;  
Reprodução/AdoroCinema; Reprodução/Legião dos Heróis).

fotos/treinamento\Ben Affleck.jpg

```
[[[-0.11437102 0.16535544 0.06166501 -0.0385839 -0.13644788 0.03504861  
0.02407607 -0.0626087 0.14703475 -0.0028598 0.28948197 0.02588779  
-0.26941833 -0.09501561 0.01938695 0.06426555 -0.20150705 -0.06051556  
....]]
```

fotos/treinamento\Gal Gadot.jpg

```
[[[-1.31862029e-01 6.77738041e-02 1.38327137e-01 -5.27683385e-02  
-1.79630592e-01 -9.18008573e-03 -9.12394524e-02 -7.59068877e-02  
1.74739018e-01 -1.35839880e-01 1.98038772e-01 -8.60403925e-02  
-2.71009326e-01...]]
```

fotos/treinamento\Henry Cavill.jpg

```
[[[-0.17476112 0.15562367 -0.01584224 0.03230717 -0.06836422 -0.06813291  
-0.05605397 -0.05459433 0.06428429 0.0446728 0.24186163 0.02503302  
-0.24766943 -0.03824538 -0.05291041 -0.01219529 -0.1510424 -0.0446354  
-0.06458185 0.00499556 0.12470792.....]]
```

Process finished with exit code 0

Cada imagem foi lida individualmente pelo *dlib.get\_frontal\_face\_detector()* - aplicação interna do pacote dlib - que detecta as faces no projeto. Em seguida, a imagem passa pelo *shape\_predictor\_68\_face\_landmarks.dat*, que marca os 68 pontos faciais na imagem.

Na terceira etapa do treinamento a imagem passa pelo *dlib\_face\_recognition\_resnet\_model\_v1.dat*. Este modelo é uma rede ResNet com 29 camadas convolucionais. É essencialmente uma versão da rede ResNet-34.

O ResNet-34 é uma rede neural convolucional de 34 camadas que pode ser utilizada como um modelo de classificação de imagem de última geração. Este é um modelo que foi pré-treinado no conjunto de dados ImageNet — um conjunto de dados que tem mais de 100.000 imagens em mais de 200 classes diferentes. Diferente das redes neurais, o ResNet-34 pega resíduos de cada camada e os usa nas camadas conectadas subsequentes (semelhante às redes neurais residuais usadas para previsão de texto). Após esse processo os índices de cada imagem é salvo em um arquivo numpy como uma matriz.

Na etapa seguinte do algoritmo, (*reconhecimnto\_facial.py*) é carregado um arquivo numpy que foi gerado pelo mesmo. Assim, a base de dados, junto com a rede, fica apta para identificar as imagens treinadas (Ben Affleck, Henry Cavill e Gal Gadot).

Em seguida, é aberta uma aba para encontrar a imagem desejada a partir do tkinter, conforme o anexo 02.

Anexo 02 - Primeira implementação gráfica com o Tkinter.

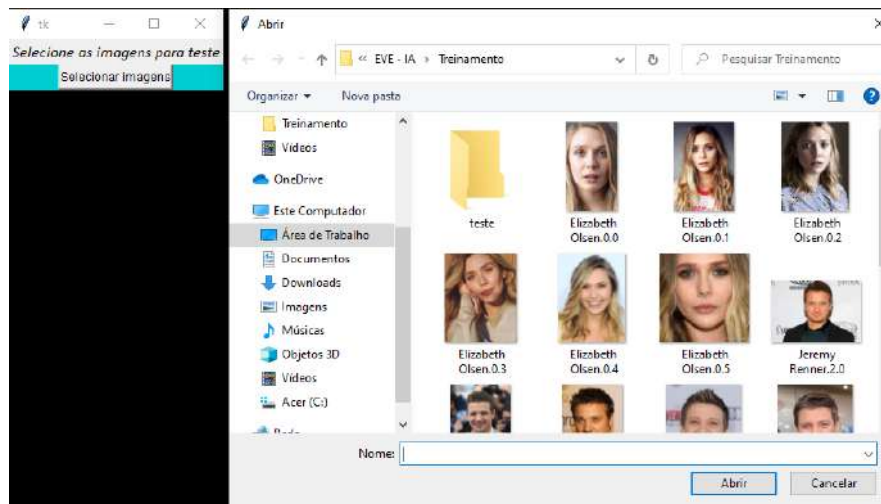
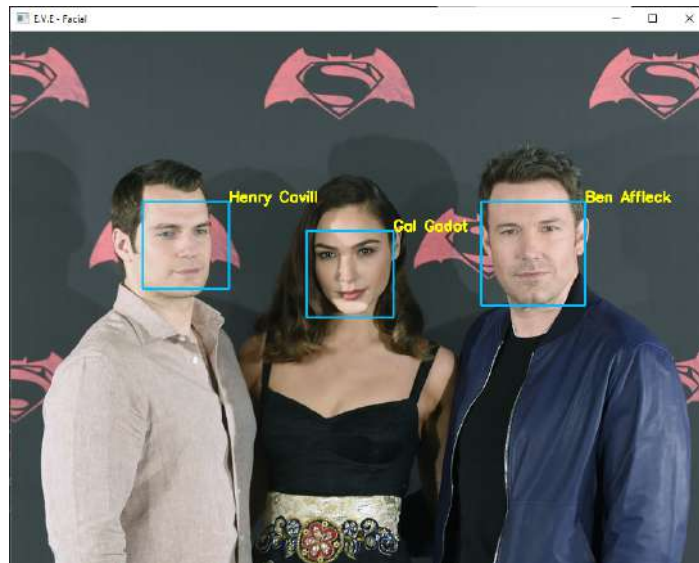


Foto: Arquivo pessoal do projeto

Tendo carregado a imagem selecionada, o algoritmo filtra os dados da imagem de entrada e compara com a base de dados já treinada. Caso a imagem de entrada tenha os dados semelhantes às imagens treinadas, ele irá dizer o nome do autor e logo em seguida irá reproduzir o nome de quem está presente na imagem. No caso da imagem não tenha sido treinada inicialmente no *reconhecimento\_treinamento.py*, o algoritmo irá reproduzir a face como (desconhecido), pois ela não foi treinada pela rede.

Anexo 03 - Imagem com os três rostos identificados pelo programa EVE versão beta.



(Foto: Reprodução/CinePop)

Na versão gama, assim como na versão beta, o programa precisa ter as imagens treinadas e receber uma imagem de entrada e comparar com as imagens pré-registradas no banco de dados, criando uma caixa de texto ao redor da face com o nome do indivíduo identificado e em seguida reproduzir o nome em voz sintetizada. Em seguida, é realizada uma extração dos pontos faciais das imagens de treino. Esses pontos são arquivados em um arquivo .npy e lançados na rede neural, que compara os descritores faciais da nova imagem de entrada com os descritores faciais já salvos no treinamento, mas diferentemente da versão beta, a versão gama vem com o propósito de identificar emoções e reproduzi-las em áudio, ex.: “Jeremy Renner está bravo”, conforme exemplificado no anexo 04.

Anexo 04 - Imagem utilizada para validação da do programa E.V.E. versão gama.

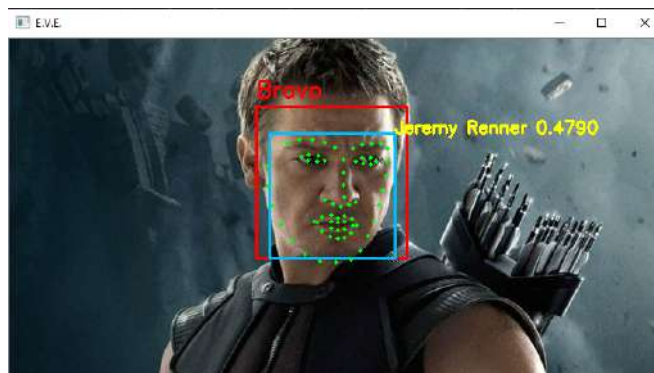


Foto: Reprodução/Geek Blog.

Com a comparação finalizada e os pontos identificados, a imagem passa pela biblioteca Keras, que cria uma matriz com os descritores faciais, reduzindo a imagem em pixels de 48x48px, realizando uma comparação com os pontos salvos no dataset FER2013.

O programa então pega os descritores faciais da imagem de teste e compara com as imagens presentes no dataset, classificando com qual emoção a imagem mais se encaixa.

Anexo 05 - Imagem com o rosto e emoção identificados pelo programa EVE versão gama.



Foto: Reprodução/Geek Blog.

Para validar o protótipo, o programa foi executado novamente com outro arquivo de teste.

Anexo 06 - Imagem utilizada para validação da do programa E.V.E. versão gama.

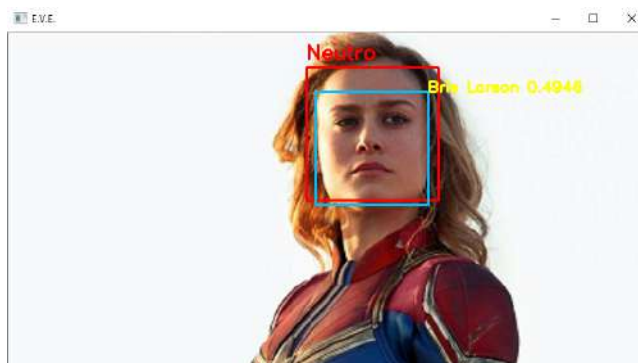


Foto: Reprodução/Ei Nerd. Disponível em:

<<https://www.einerd.com.br/vingadores-ultimato-estreia-2-anos/>>

## 5. RESULTADOS

Até o presente momento, o programa conseguiu identificar a quantidade de pessoas na foto e exibi-las separadamente, identificando cada uma delas por meio de áudio.

Foto: Reprodução/Zinema.

No entanto, o protótipo consegue identificar apenas uma emoção por vez caso haja mais de um indivíduo presente na imagem.

O protótipo mostrou-se eficiente quanto às técnicas de programação e uso da biblioteca, mas cabe salientar que a sua validação ainda não foi realizada devido à pandemia de Covid-19. Nas fases seguintes, tem-se o objetivo de realizar novas simulações para que seja possível a identificação das emoções em todos os indivíduos presentes na imagem, além de aprofundar as pesquisas do reconhecimento de emoções a partir da utilização das ações fundamentais dos músculos individuais (*Action Units*), o que permitirá uma eficiência acima de 70% nos resultados do programa. Vale ressaltar que o projeto tem como prioridade testar o programa com pessoas deficientes visuais ou cegas, validando-o devidamente.

Por fim, com a utilização da biblioteca Tkinter montamos uma interface simples (anexo 02) que permitirá uma maneira mais dinâmica de realizar os testes. Futuramente, pretende-se aprimorar a interface e torná-la totalmente acessível. Após alguns testes com as bibliotecas GUI, o Tkinter foi a melhor opção entre o Kivy, devido a sua simplicidade, facilidade e bons resultados.

## 6. CONSIDERAÇÕES FINAIS

O objetivo desta pesquisa foi realizar um estudo que demonstre a importância da consolidação da autonomia comunicativa da pessoa com deficiência visual ou cegueira por meio da criação de novos meios de comunicação, o que pode aumentar a qualidade de vida desses indivíduos para uma melhor socialização no mundo contemporâneo. Dada a importância do assunto, o projeto E.V.E tem como principais objetivos o reconhecimento

facial em imagens, tendo como seu diferencial a utilização da tecnologia *Deep Learning*, em conjunto com o reconhecimento facial, como uma ferramenta de inclusão. O desenvolvimento do presente estudo possibilitou o aprendizado de conceitos como *Deep Learning* e redes neurais, que são conhecimentos avançados na área de inteligência artificial, além de permitir um melhor conhecimento sobre programação utilizando a linguagem python.

A versão 3.0 do projeto é capaz de identificar a quantidade de pessoas na foto e exibi-las separadamente com apoio da biblioteca *face recognition*, identificando cada uma das faces presentes por meio da reprodução de áudio, além de identificar as emoções dos mesmos.

Para o futuro, a intenção é realizar testes com software, permitindo uma validação do sistema em testes com pessoas com deficiência visual ou cegueira.

## 7. REFERÊNCIAS

BARCZINSKI, Maria Cristina de Castro. Reações psicológicas à perda da visão. **Benjamin Constant**, n. 18, 2001. Disponível em: <<http://revista.ibc.gov.br/index.php/BC/article/view/585>> Acesso em: 13 abril. 2021.

DLIB C++ Library. **Documentation**. Disponível em: <http://dlib.net/>. Acesso em: 17 nov. 2020.

GEITGEY, Adam. **Face Recognition documentation**. 2017. Disponível em: <<https://face-recognition.readthedocs.io/en/latest/index.html>>. Acesso em: 17 nov. 2020.

GOODFELLOW, Ian *et al.* **Deep Learning**. Cambridge: MIT press, 2016. Disponível em: <[https://www.academia.edu/38223830/Adaptive\\_Computation\\_and\\_Machine\\_Learning\\_series\\_Deep\\_learning\\_The\\_MIT\\_Press\\_2016\\_pdf](https://www.academia.edu/38223830/Adaptive_Computation_and_Machine_Learning_series_Deep_learning_The_MIT_Press_2016_pdf)> Acesso em: 17 nov. 2020

PYTHON Software Foundation. **Documentação Python 3.8.2** Disponível em: <<https://docs.python.org/pt-br/3/index.html>>. Acesso em: 17 nov. 2020.

VARGAS, Ana Caroline Gomes; PAES, Aline; VASCONCELOS, Cristina Nader. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: **Proceedings of the XXIX Conference on Graphics, Patterns and Images**. 2016. Disponível em: <<http://gibis.unifesp.br/sibgrapi16/eproceedings/wuw/7.pdf>>. Acesso em: 19 nov. 2020

VIANA, Isabel. Comunicação não verbal e expressões faciais das emoções básicas. **Revista de Letras**, v. 13, p. 165-181, 2014. Disponível em: <[https://www.researchgate.net/profile/Bel-Viana/publication/330729356\\_Comunicacao\\_nao\\_](https://www.researchgate.net/profile/Bel-Viana/publication/330729356_Comunicacao_nao_)

verbal\_e\_expressoes\_faciais\_das\_emocoes\_basicas/links/5c518ffca6fdccd6b5d4112b/Comunicacao-nao-verbal-e-expressoes-faciais-das-emocoes-basicas.pdf>. Acesso em: 29 abril. 2021.