

# Condução de experimentos para o treinamento de modelos de Rede Neurais Artificiais em tarefas de inferência de expressões faciais em um ambiente Testbed

Ítalo J. da S. Melo<sup>1</sup>, Edson M. da C. Filho<sup>1</sup>,  
Alana M. Medeiros<sup>4</sup>, Yomara P. Pires<sup>3</sup>, Saulo W. da S. Costa<sup>2</sup>, Marcos C. da R. Seruffo<sup>3</sup>

<sup>1</sup>Faculdade de Computação, Universidade Federal do Pará (UFPA) -  
Campus Castanhal, Pará, 68746-630, Brasil

<sup>2</sup>Programa de Pós Graduação em Engenharia Elétrica (PPGEE) -  
Universidade Federal do Pará(UFPA), Belém, Pará, 66075-110, Brasil.

<sup>3</sup>Programa de Pós Graduação em Estudos Antrópicos na Amazônia (PPGEAA) –  
Universidade Federal do Pará (UFPA), Belém, Pará, 66075-110, Brasil.

<sup>4</sup>Programa de Pós Graduação em Ciência da Computação (UFPA) –  
Universidade Federal do Pará (UFPA), Belém, Pará, 66075-110, Brasil

italo.melo@castanhal.ufpa.com, edson.filho@castanhal.ufpa.com,

saulo.costa@ifpa.edu.br, yomara@ufpa.br, seruffo@ufpa.br

**Abstract.** *Advances in the field of Artificial Intelligence (AI) in recent times have allowed the development of technologies and research present in our daily lives, an example of the area of affective computing in the recognition of advanced expressions. However, carrying out experiments in this area requires allocation of resources and a controllable test environment for training neural models capable of inferring emotions. This article reports the use of Testbeds for training an Artificial Neural Network (ANN) model for emotion inference tasks through facial expression recognition. For this, the Teachable Machine tool was used to compose the neural network model. The test environment used was Testbeds from the National Teaching and Research Network (RNP). The obtained results indicate that the use of the Teachable Machine in a TestBeds enables the creation of accurate and efficient facial recognition models.*

**Resumo.** *Os avanços da área de Inteligência Artificial (IA) nos últimos tempos, têm permitido o desenvolvimento de tecnologias e pesquisas presentes em nosso cotidiano, a exemplo a área de computação afetiva no reconhecimento de expressões faciais. No entanto, a realização de experimentos nesta área, requer alocação de recursos e um ambiente de testes controlável para treinamento dos modelos neurais capazes de realizar a inferência de emoções. Este artigo relata o uso de Testbeds para o treinamento de um modelo de Redes Neurais Artificiais (RNA) para tarefas de inferência de emoções através do reconhecimento de expressões faciais. Para tal, foi utilizada a ferramenta Teachable Machine para composição do modelo de rede neural. O ambiente de testes utilizado foi Testbeds da Rede Nacional de Ensino e Pesquisa (RNP). Os resultados obtidos indicam que o uso Teachable Machine em um TestBeds possibilita a criação de modelos de reconhecimento facial precisos e eficientes.*

## 1. Introdução

No desenvolvimento de Inteligência Artificial (IA), a subárea *Machine Learning* (ML) se destaca como uma das tecnologias fundamentais para a criação de modelos computacionais capazes de realizar o reconhecimento de emoções, por meio da detecção de expressões faciais.

Porém, mesmo diante de tantos avanços, detectar e reconhecer emoções computacionalmente e com a mesma facilidade que humanos reconhecem ainda é um problema relevante a ser explorado [Heck et al. 2020]. A alocação de recursos e um ambiente de testes controlável com baixo custo computacional para o treinamento dos modelos de IA, configuram-se como um dos principais desafios.

Neste contexto destacam-se os *TestBeds*, que são plataformas comumente utilizados para realização de experimentos, oferecem um ambiente controlável e escalável, permitindo assim a construção e validação de modelos de IA a baixo custo computacional.

Este trabalho visa a utilização de ML para realização da tarefa de reconhecimento de emoções, por meio de expressões faciais, em um ambiente *Testbeds*. Para tal, foi utilizado o serviço de *Testbeds* da Rede Nacional de Ensino e Pesquisa (RNP) juntamente com *Teachable Machine*, uma importante ferramenta que pode ser aplicada para o tarefas de reconhecimento e classificação de imagens [Wong and Fadzly 2022].

O principal objetivo consiste em relatar o uso do *Testbeds* RNP para treinar um modelo de ML capaz de realizar inferência de emoções, por meio do reconhecimento de expressões faciais, com o intuito de aprimorar uma ferramenta de reconhecimento de expressões faciais que auxilia na construção de perfis psicocomportamentais de mulheres vítimas de violência doméstica e familiar.[Costa et al. 2021].

O restante do trabalho está organizado da seguinte forma: Na seção 2 são descritos trabalhos correlatos a esta pesquisa. A seção 3 descreve a metodologia adotada. Os resultados obtidos são descritos na seção 4. Por fim, as considerações finais e perspectivas de trabalhos futuros são mostrados na seção 5.

## 2. Correlatos

Nesta seção, revisaremos alguns trabalhos sobre o uso de algoritmos de *Deep Learning* (DL) para reconhecimento de expressões facias.

Em [Oguine et al. 2022] foi utilizado um Modelo Híbrido de Reconhecimento de Expressão Facial para Classificação e Predição de Emoções em Tempo Real. A base de dados utilizada foi a Facial Expression Recognition 2013 (FER-2013). O modelo DL utilizado foi o Rede Neural Convolutacional (CNN) com o objetivo de classificar as imagens faciais digitais em tempo real em uma das sete categorias de emoções faciais consideradas na base FER-2013. Os resultados obtidos indicaram 70% de precisão, mostrando um desempenho de classificação 6% superior aos modelos convencionais.

O trabalho apresentado por [Domingos, 2022] apresentou uma abordagem meta-heurística com a implementação de arquiteturas CNN e algoritmos genéticos (AG) para classificação de emoções por meio de imagens de expressões faciais. O uso de algoritmos genéticos permitiu encontrar a melhor arquitetura (hiperparâmetros) da CNN. Foram

utilizados dois conjuntos de dados distintos. O conjunto de dados FER-2013, com desempenho de 63,84% no conjunto de dados de treino e 62,39% no conjunto de dados de validação. O segundo conjunto de dados utilizado foi o *Cohn-Kanade AU*, versão 2 (CK+). Os experimentos apresentaram o desempenho de 99,32% no conjunto de dados de validação. Os autores destacam que, esse conjunto de dados possui imagens criadas especificamente para o problema proposto, não apresentando os mesmos problemas da área de visão computacional já conhecidos no conjunto FER-2013 (oclusão, perspectiva e ruídos).

Já em [Bukhari et al. 2022], foi usado o modelo de CNN como um extrator de características para detecção de emoções por meio de expressões faciais. Neste estudo, foram utilizados três modelos pré-treinados VGG-16, ResNet-50 e Inception-V3, tendo como bases de dados utilizadas Fer-2013 e CK+. Com o conjunto de dados FER-2013, as taxas de precisão encontradas para CNN, ResNet-50, VGG-16 e Inception-V3 foram de 76,74%, 85,71%, 85,7%, 97,93%, respectivamente. Os resultados obtidos com o conjunto de dados CK+ indicaram taxas de precisão para CNN, ResNet-50, VGG-16 e Inception-V3 de 84,18%, 92,91%, 91,07% e 73,16%, respectivamente. Os resultados experimentais mostraram resultados excepcionais para Inception-V3, com 97,93% usando o conjunto de dados FER-2013, e ResNet-50, com 91,92% usando o conjunto de dados CK+.

Em [Bukhari et al. 2022], o autor destaca o FER-2013 como um dos bancos de dados mais utilizados, para o reconhecimento de emoções. Vale destacar que, apesar de sua ampla utilização, é importante notar que a base de dados FER-2013 possui algumas limitações que podem afetar a precisão e confiabilidade dos modelos de reconhecimento facial desenvolvidos a partir dela [MEDEIROS et al. 2022], já que a base de dados consiste em imagens de baixa resolução e qualidade, o que pode afetar a precisão do reconhecimento facial. Todavia, ao ser usado em projetos em sua maioria prevalece uma precisão acima de 50% , assim como abordado por [Domingos et al. 2022] em seu estudo sobre redes neurais convolucionais utilizando *Transfer Learning* no conjunto de dados *Japanese Female Facial Expression* (JAFFE) e FER-2013, para detecção de emoções.

Cabe destacar que, os trabalhos descritos acima, apesar de alguns assemelharem-se na utilização de algoritmos de DL, CNN e na base de dados utilizada para reconhecimento de expressões faciais (FER-2013), não mencionam o uso de ferramentas como o *Teachable Machine*, tão pouco o uso de um ambiente de experimentos do tipo *Testbeds*.

Outro ponto a ser observado, diz respeito ao uso da base de dados FER-2013. Diante da relevância desta base para projetos de reconhecimento facial, torna-se importante o estabelecimento de uma metodologia eficiente para a construção de modelos utilizando essa base. Nesse sentido, é fundamental considerar algumas etapas importantes do processo, como o pré-processamento das imagens, o treinamento do modelo e a avaliação dos resultados obtidos, bem como o uso de um ambiente para experimentos remotos. É sobre essas etapas que iremos discutir a seguir.

### **3. Metodologia**

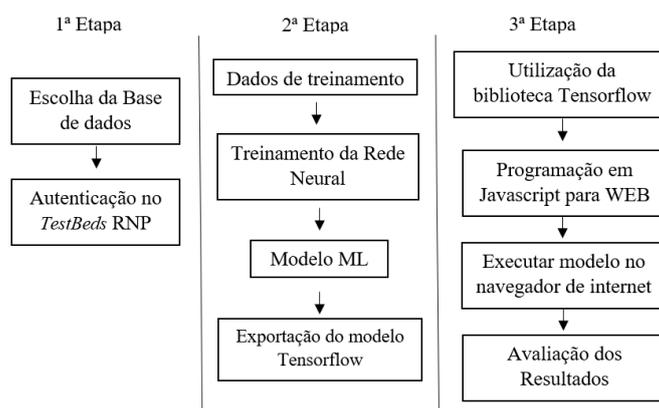
A metodologia adotada nesta pesquisa consistiu em três etapas, que foram executadas, por meio de algoritmos de processamento de dados.

A primeira etapa consistiu na autenticação e configuração da plataforma de *Testbeds* RNP. Essa plataforma permite a construção de um ambiente de testes controlável e

escalável, o que possibilita a realização de experimentos de forma rigorosa, transparente e reaplicável. Atividades estas que são fundamentais para garantir a confiabilidade dos resultados obtidos neste trabalho. Quanto a aquisição e pré-processamento das imagens, a seção 3.1 descreve a base utilizada.

A segunda fase da metodologia utilizou a plataforma do *Teachable Machine*, uma ferramenta baseada na web que permite o desenvolvimento de modelos de forma ágil e simples, utilizando técnicas de aprendizado de máquina para reconhecimento e classificação de imagens. Nesta fase, foi realizada a criação e treinamento do modelo de reconhecimento facial.

Na terceira fase, o modelo treinado foi importado para o *TensorFlow*<sup>1</sup>, uma biblioteca de código aberto para aprendizado de máquina que possibilita a criação de modelos de redes neurais profundas para reconhecimento de padrões em dados. Em seguida, o modelo foi executado automaticamente em um navegador de internet para validação dos resultados obtidos. Todos os resultados apresentados neste estudo foram gerados automaticamente, garantindo a objetividade e precisão dos resultados obtidos. A Figura 1 ilustra as três etapas da metodologia adotada na pesquisa:



**Figura 1. Estrutura de funcionamento do Teachable Machine. Fonte: Autores**

### 3.1. A base de dados FER-2013

Os bancos de dados são componentes essenciais para o desenvolvimento de projetos de reconhecimento facial. Dentre as diversas bases de dados disponíveis, cada uma apresenta suas características e vantagens. Um exemplo de base de dados utilizada para o reconhecimento facial é FER-2013, disponibilizada na plataforma *Kaggle*<sup>2</sup>. Essa base de dados contém imagens de rostos humanos em diferentes expressões emocionais, como raiva, alegria, tristeza, medo, surpresa, nojo e neutro, que são usadas para a construção de modelos que utilizam reconhecimento facial.

Logo, vale-se destacar que cada banco de dados tem suas particularidades e o banco de dados FER-2013 foi o que melhor se encaixou no projeto, por conter uma grande quantidade de imagens em escala de cinza de rostos humanos em 7 tipos de expressões emocionais, sobretudo, as imagens têm dimensões de 48 x 48 pixels, o que as torna adequadas para o treinamento de modelos de aprendizado de máquina.

<sup>1</sup><https://www.tensorflow.org/?hl=pt-br>

<sup>2</sup><https://www.kaggle.com>

De acordo com [MEDEIROS et al. 2022], a base de dados FER-2013 tem se destacado em projetos de desenvolvimento de redes neurais devido à sua gratuidade, fácil acesso e amplo conjunto de imagens rotuladas com as sete expressões faciais mencionadas anteriormente. Em uma revisão sistemática realizada pelo autor, constatou-se que 15% dos artigos analisados utilizaram essa base de dados em seus projetos, reforçando sua relevância e popularidade na comunidade científica.

Essa base de dados é composta por 35.881, sendo 28.709 imagens para treinamento e 7.178 para teste. As imagens são rotuladas nas sete emoções e podem apresentar algumas divergências com o cenários realísticos, devido a presença de mãos, óculos e no formato de barbas e cabelos, diminuindo a precisão. A Figura 2 a seguir demonstra um exemplo das expressões faciais utilizadas pela FER-2013.



**Figura 2. Amostra de Imagens do banco de Dados FER-2013. Fonte: Base de dados FER-2013**

Dada a importância do banco de dados FER-2013 e sua ampla utilização na comunidade científica, é importante ressaltar que este pode ser utilizado com diversas ferramentas tecnológicas, como a plataforma *Teachable Machine*. Diante disso, será apresentado a seguir esta plataforma e, como ela foi utilizada para criar um modelo que possa identificar expressões faciais a partir desse banco de dados.

### 3.2. Ferramenta *Teachable Machine*

O desenvolvimento de ML exige uma quantidade de conhecimentos técnicos muito elevada. Todavia, o uso de ferramentas como o *Teachable Machine*<sup>3</sup>, com sua interface simples e intuitiva, permite que usuários sem habilidades técnicas possam criar e treinar modelos de ML, usando dados de entrada como imagens, sons ou gestos. Essa facilidade de uso torna a plataforma uma ferramenta útil para pesquisadores e desenvolvedores de diferentes áreas, possibilitando a criação de sistemas de ML com aplicações em diversas áreas, incluindo a análise de imagens, reconhecimento de fala e identificação de objetos.

O modelo de rede neural, utilizado neste trabalho, foi criado usando a plataforma *Teachable Machine*, que utiliza a biblioteca *TensorFlow* para treinar e executar modelos de ML em um navegador da internet. O *TensorFlow.js* é uma biblioteca que permite que modelos treinados em *Python* sejam facilmente exportados para uso em *JavaScript*.

<sup>3</sup>Ferramenta desenvolvida pela Google que permite criar modelo de ML. <https://teachablemachine.withgoogle.com>

O modelo é carregado em memória durante a inicialização do site e pode ser usado para realizar previsões em tempo real. Por exemplo, se o computador do usuário tiver uma câmera embutida, o modelo pode ser usado para detectar a emoção em tempo real com base na expressão facial do usuário. A rede neural da ferramenta adota o método de *Transfer Learning*, ou seja, é um método de aprendizado de máquina em que um modelo desenvolvido para uma tarefa é utilizado como ponto de partida para um modelo em uma segunda tarefa. A principal vantagem do uso da ferramenta *Teachable Machine* e da biblioteca *TensorFlow*, é que esta apresenta uma enorme facilidade em seu uso já que é um código aberto para desenvolvimento de pesquisa, aprendizado de máquina e de redes neurais, por meio do reconhecimento de imagens.

Além disso, é uma abordagem popular em aprendizado profundo, onde modelos pré-treinados são usados como ponto de partida em tarefas de visão computacional e processamento de linguagem natural, dada a grande quantidade de computação e tempo necessários para desenvolver modelos do zero. O uso de modelos pré-treinados reduzem os custos computacionais em tarefas desta natureza, uma vez que as redes neurais artificiais podem ser estruturas complexas que necessitam de uma grande quantidade de dados e, conseqüentemente, do processamento de milhares de parâmetros.

O modelo de RNA criado reconhece 7 classes, cada uma representando uma das emoções. Após finalizar o *upload* das imagens na plataforma, é realizado o treinamento da neural com a técnica *Transfer Learning*. Uma vez que treinado, o modelo é exportado para posteriormente ser usado na biblioteca *TensorFlow* e poder ser utilizado na *web* através de *JavaScript*. Dessa forma, é possível fornecer uma experiência mais personalizada para os usuários do site com base em suas expressões faciais e emoções detectadas.

### **3.3. O Serviço de *Testbeds* RNP**

O uso de *Testbeds* tem se mostrado uma solução promissora para a realização de experimentos remotos em diferentes aplicações relacionadas a Tecnologia de Informação e Comunicação (TIC). Durante o desenvolvimento deste trabalho, procuramos diversos ambientes de testes que pudessem auxiliar nossa pesquisa. No entanto, limitações de integração da *Interface* de Linha de Comando (CLI) com a plataforma *Teachable Machine*, impossibilitaram o uso das mesmas.

Além disso, buscamos explorar serviços de nuvem pública como alternativa. Todavia, nos deparamos com limitações nas configurações disponíveis, especialmente em relação à memória RAM, importante componente para o desenvolvimento deste trabalho. Diante dessas dificuldades, encontramos uma solução mais adequada no Testebed da RNP, o qual apresentou as funcionalidades e recursos necessários para dar continuidade à nossa pesquisa.

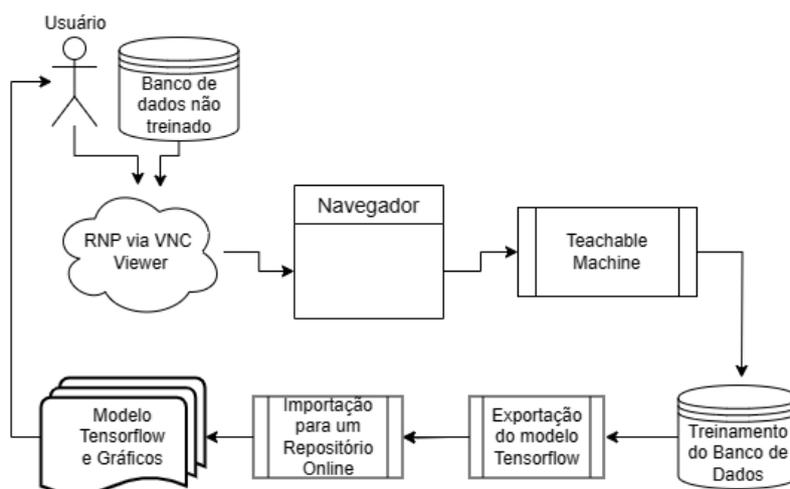
A seguir descreveremos como utilizamos o serviço de *Testbeds* da RNP, juntamente com o *Teachable Machine*, para inferência de emoções por meio de um banco de dados de expressões faciais.

Nosso objetivo foi treinar a rede neural para reconhecer sete emoções básicas: raiva, alegria, tristeza, medo, surpresa, desgosto e neutro. Para isso, foi preciso aumentar a memória RAM e a memória interna disponíveis, isso se deve ao fato de que a ferramenta utilizada para o treinamento exigia uma quantidade significativa de memória, a fim de evitar possíveis limitações durante o processo de treinamento do modelo. Dessa forma,

garantimos uma execução mais fluida e eficiente do treinamento, evitando problemas relacionados à falta de recursos de memória. Somado a isso, optamos por utilizar o sistema operacional Linux Ubuntu no *TestBed*, pois ele oferece suporte a uma ampla variedade de softwares de aprendizado de máquina e possui recursos que nos permite otimizar o uso de recursos de *hardware*.

Para acessar o *TestBed*, utilizamos o software VNC *Viewer*<sup>4</sup>, que nos permitiu interagir com a interface gráfica da máquina virtual e realizar o experimento. A interface gráfica é importante para facilitar a interação com a máquina virtual, permitindo que os usuários visualizem e controlem a execução de softwares e aplicativos de maneira mais intuitiva.

Um aspecto relevante é que a ferramenta *Teachable Machine* não foi integrada diretamente. Em vez disso, optamos por utilizar o formato remoto, no qual o usuário inicia a página *web* do *Teachable Machine* para acessar e utilizar a ferramenta. Essa abordagem nos permitiu aproveitar todas as funcionalidades e etapas mencionadas na seção anterior, porém, sem a necessidade de uma integração completa da ferramenta, conforme ilustra a Figura 3.



**Figura 3. Estrutura de funcionamento utilizando o Testbed RNP**

Com o *TestBed*, conseguimos aumentar significativamente a capacidade de processamento disponível para nosso projeto de RNA. Utilizamos uma máquina virtual com configuração de 64GB de RAM, 2.00GHz, 4 discos de 1TB, dois processadores Intel Xeon e Sistema Operacional Ubuntu. O uso de *TestBeds* permite que pesquisadores acessem recursos computacionais sem a necessidade de investir em *hardware* ou infraestrutura local. Isso possibilita a realização de projetos com requisitos mais avançados e permite que pesquisadores de diferentes locais possam colaborar em projetos conjuntos.

Em resumo, o uso de *Testbeds* pode ser uma solução interessante para projetos de pesquisa que exigem recursos computacionais avançados e colaboração remota.

<sup>4</sup><https://www.realvnc.com/pt/connect/download/viewer/>

## 4. Resultados

Os resultados obtidos indicam que a utilização da ferramenta *Teachable Machine*, no desenvolvimento de um modelo de reconhecimento facial baseado em ML em um ambiente *Testbeds*, são promissores. A rede neural treinada com essa ferramenta foi capaz de identificar com uma média de precisão de 62,85% as sete emoções presentes na base de dados FER-2013. No que se refere a precisão de cada emoção, o pior resultado foi a categoria tristeza (Sad) que obteve 29% de precisão e a melhor categoria foi surpresa (Surprise) obtendo 87% conforme Figura 4.

Accuracy per class ?

CLASS	ACCURACY	# SAMPLES
Angry	0.32	144
Disgust	0.79	34
Fear	0.77	308
Happy	0.59	267
Neutral	0.77	370
Sad	0.29	188
Surprise	0.87	250

Figura 4. Acuracidade por classes de emoções. Fonte: Teachable Machine.

Essa precisão é baseada na matriz de confusão que pode ser analisada, por meio de uma matriz  $N \times N$  usada para avaliar o desempenho do modelo de classificação, onde  $N$  é o número de classes do atributo alvo. A matriz compara o valor real com o valor previsto pelo modelo [Bhandari 2020]. A matriz de confusão é gerada automaticamente pelo *Teachable Machine* e a Figura 5 mostra essa matriz do modelo ML.

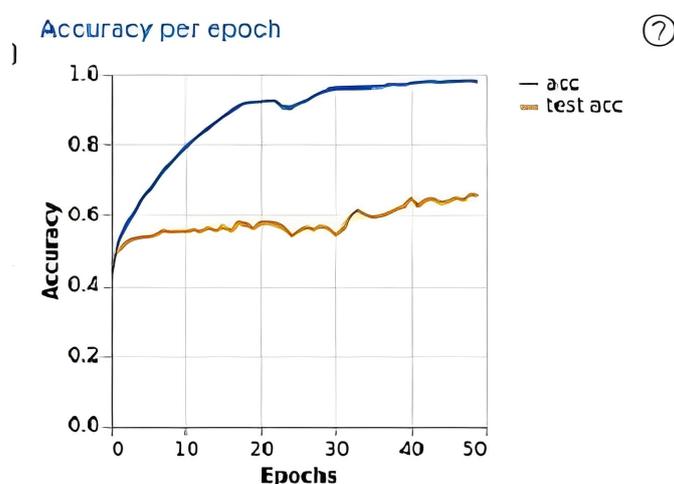
Confusion Matrix ?

Class	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	46	1	33	14	21	21	8
Disgust	1	27	3	1	1	0	1
Fear	12	0	236	4	21	12	23
Happy	22	1	25	157	28	15	19
Neutral	18	0	27	21	285	7	12
Sad	37	1	47	15	27	54	7
Surprise	5	0	18	4	5	1	217

Figura 5. Matriz de confusão do modelo ML. Fonte: Teachable Machine

Essa matriz 7 x 7, é composta pelas 7 emoções e cada célula representa uma informação de erro ou acerto no modelo, a sua construção é feita utilizando a Figura de acuracidade por classes e a diagonal principal da matriz demonstra o resultado esperado.

Com base nos dados fornecidos pelo *Teachable Machine*, é possível gerar um gráfico que destaca a acurácia (acc) em azul e o teste da acurácia em amarelo por época, conforme demonstra a Figura 6. Dessa forma, é possível gerar um gráfico de precisão por época, que fornece informações mais detalhadas sobre o desempenho do modelo ao longo do tempo. Esse gráfico é útil para ajudar a identificar padrões e tendências no desempenho do modelo ao longo do tempo, permitindo uma análise mais precisa sobre como ajustar o modelo para melhorar sua precisão. Como pode ser visualizado a seguir na Figura 6 de precisão por época, que nos indica os possíveis ajustes no modelo.



**Figura 6. Precisão por época. Fonte: Teachable Machine**

A Figura 6 indica que a precisão do modelo aumentou significativamente após a primeira época de treinamento, mas estabilizou nas épocas subsequentes. Esse tipo de informação pode ajudar a identificar o momento em que o modelo começa a apresentar diminuição no desempenho e a necessidade de realizar ajustes no treinamento, como adicionar mais dados ao conjunto de treinamento.

Além do gráfico de precisão por época, é possível gerar um gráfico que mostra a perda (*loss*) do modelo ao longo das épocas. A perda é uma medida de quão distantes as previsões do modelo estão dos valores reais. Em geral, quanto menor a perda, melhor é o desempenho do modelo. Abaixo, apresentamos a Figura 7 que aborda a perda por época, onde podemos perceber que a precisão por época e a perda por época são medidas importantes para avaliar o desempenho de modelos de aprendizado de máquina ao longo do tempo, para que venha-se a analisar os parâmetros do modelo.

A precisão por época indica a proporção de exemplos de treinamento que foram classificados corretamente pelo modelo em cada época. Já a perda por época mede a diferença entre as previsões do modelo e os valores reais dos exemplos de treinamento.

A análise dessas métricas é fundamental para identificar padrões e tendências no desempenho do modelo ao longo do tempo. A Figura de precisão por época, por exemplo,

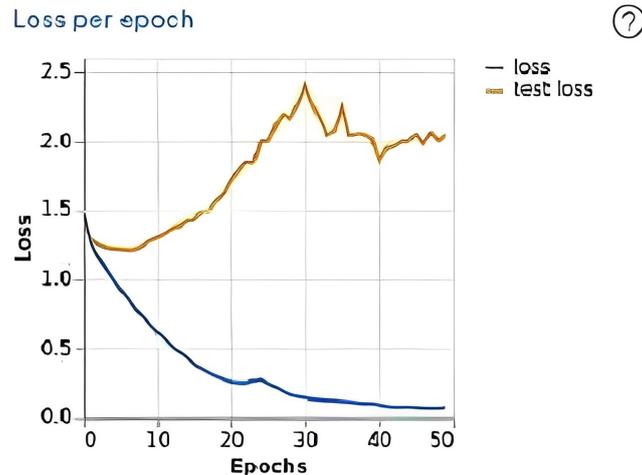


Figura 7. Perda por época. Fonte: Teachable Machine

pode revelar que o modelo está tendo dificuldades para classificar determinadas classes de exemplos, o que pode indicar a necessidade de ajustar os parâmetros do modelo ou adicionar mais dados de treinamento.

Já o gráfico de perda por época é útil para avaliar o quanto o modelo está errando em suas previsões, permitindo identificar casos em que a perda está aumentando ao longo das épocas, o que pode indicar *overfitting* do modelo ou a necessidade de aplicação de técnicas de regularização [Passos 2022].

No geral, os gráficos gerados pelo *Teachable Machine* são ferramentas úteis para a análise e ajuste de modelos de ML. Eles fornecem informações precisas sobre o desempenho do modelo ao longo do tempo, permitindo que os usuários identifiquem tendências e padrões em suas bases de dados.

Com base na perspectiva acima, foi possível encontrar uma precisão (acurácia) média de 62,85 % entre as emoções, utilizando a base de teste do banco de dados FER-2013. Essa precisão pode ser considerada razoável, mas ainda há espaço para melhorias, como o aumento da diversidade na base de dados e a utilização de técnicas mais avançadas de pré-processamento de imagens.

Os resultados obtidos demonstram a utilidade do *Teachable Machine* como uma ferramenta para o desenvolvimento de modelos de aprendizado de máquina. Com sua *interface* e recursos gráficos intuitivos, é possível analisar o desempenho do modelo e ajustá-lo de maneira eficaz para obter melhores resultados.

Assim, podemos concluir que o *Teachable Machine* é uma ferramenta poderosa e de fácil uso para o desenvolvimento de modelos de aprendizado de máquina, com grande potencial para ser utilizada em diversas áreas, desde reconhecimento de emoções até detecção de objetos e classificação de dados.

## 5. Conclusão

O desenvolvimento de modelos de reconhecimento facial baseados em redes neurais tem se mostrado uma área promissora em diversos campos. Apesar da acurácia obtida em

nostros testes seja de 62,85 %, essa pode ser melhorada com a utilização de outras técnicas, é importante ressaltar que o uso de bases de dados adequadas e o monitoramento constante do desempenho do modelo são fundamentais para que o modelo de reconhecimento facial seja ainda mais preciso e eficaz.

Durante o estudo, foi constatado que o treinamento da rede neural exigiu um alto poder computacional, sendo esta uma das principais dificuldades encontradas. Além disso, a base de dados utilizada no treinamento do modelo apresentou algumas limitações, como a falta de representatividade de grupos específicos, como pessoas negras, com barba e com acessórios (óculos, chapéu, fone de ouvido). Tais limitações podem afetar a precisão do modelo, o qual segundo [Kuruvayil and Palaniswamy 2022] pode ser sanado utilizando técnicas de *meta-learning*.

Outro ponto relevante é que fatores externos, como a qualidade das imagens e a variação na expressão facial de cada indivíduo, podem afetar a acurácia do modelo. Vale destacar que o modelo desenvolvido pode não ser completamente preciso em todas as situações, sendo necessários ajustes durante a fase de treinamento do modelo.

Dessa forma, é importante realizar testes rigorosos e avaliações de desempenho em diferentes cenários e condições de uso para garantir que o modelo tenha um desempenho satisfatório e confiável. Portanto, o desenvolvimento de um modelo de reconhecimento facial precisa levar em consideração não apenas questões técnicas, mas também fatores externos que podem afetar sua acurácia e precisão em diferentes situações.

Apesar das limitações encontradas, a utilização da *Teachable Machine* mostrou-se uma estratégia eficaz para o treinamento de modelos de reconhecimento facial baseados em aprendizado de máquina. No entanto, é fundamental ressaltar que o sucesso do modelo depende da qualidade e diversidade dos dados utilizados no treinamento, bem como do poder computacional disponível para conduzir o treinamento da rede neural.

Assim, é possível concluir que o uso de técnicas de *Transfer Learning*, aliado a bases de dados mais diversas, pode resultar em modelos de reconhecimento facial mais precisos e eficientes. Além disso, a utilização de técnicas de pré-processamento de imagem pode ser investigada para aprimorar a acurácia do modelo em diferentes condições de iluminação e qualidade de câmera.

O uso do ambiente de *Testbed* da RNP destacou-se como uma solução eficiente para pesquisadores que precisam utilizar recursos computacionais avançados para seus projetos. Neste artigo, destacamos a facilidade de acesso ao *Testbed* e como ele nos permitiu otimizar o uso de recursos de *hardware* para melhorar a aplicação do *Teachable Machine* em nosso estudo de ML para o reconhecimento de emoções por expressões faciais. No entanto, é importante mencionar que as limitações do navegador afetaram o desenvolvimento do trabalho, impedindo que alguns procedimentos de testes e adições de imagens no banco de dados fossem realizados de forma recorrente.

Uma possível melhoria para o ambiente de teste é a disponibilização de serviços com Unidades de Processamento Gráfico (GPUs). Durante os testes realizados localmente, observamos que a ferramenta *Teachable Machine* demandava demasiadamente de recursos da GPU para treinar o modelo.

Apesar das dificuldades relatadas, a utilização do *Testbed* se apresentou como uma

solução valiosa e acessível para os interessados em explorar essa temática, possibilitando a realização de experimentos remotos com a colaboração entre pesquisadores de diferentes locais, além de tornar possível a execução de projetos com requisitos mais avançados.

## Referências

- Bhandari, A. (2020). Everything you should know about confusion matrix for machine learning. analytics vidhya, 2020. Disponível em: <https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/>. Acesso em 2023-03-20.
- Bukhari, N., Hussain, S., Ayoub, M., Yu, Y., and Khan, A. (2022). Deep learning based framework for emotion recognition using facial expression. *Pakistan Journal of Engineering and Technology*, 5(3):51–57.
- Costa, S. W. D. S., Pires, Y. P., De Sousa, A. L., Costa, F. A. R., De Oliveira, E., Araújo, F. P., and Seruffo, M. C. D. R. (2021). Whot, a novel tool to assist women victims of violence: a case study in the brazilian amazon. *IEEE Access*, 9:95046–95060.
- Domingos, D. A., Cortes, O. A. C., and Lobato, F. M. F. (2022). Evoluindo redes neurais convolucionais na detecção de emoções usando micro ags. *Anais do Computer on the Beach*, 13:266–273.
- Heck, L., Künas, C., and Padoin, E. (2020). Análise de desempenho e efetividade de redes neurais convolucionais em plataformas de gpu e cpu aplicadas ao reconhecimento de emoções através de expressões faciais em seres humanos. In *Anais Estendidos do XXI Simpósio em Sistemas Computacionais de Alto Desempenho*, pages 8–13, Porto Alegre, RS, Brasil. SBC.
- Kuruvayil, S. and Palaniswamy, S. (2022). Emotion recognition from facial images with simultaneous occlusion, pose and illumination variations using meta-learning. *Journal of King Saud University-Computer and Information Sciences*, 34(9):7271–7282.
- MEDEIROS, A. M. et al. (2022). Uma revisão sistemática sobre estudos de detecção de expressões faciais a partir do emprego de modelos de inteligência artificial.
- Oguine, O. C., Oguine, K. J., Bisallah, H. I., and Ofuani, D. (2022). Hybrid facial expression recognition (fer2013) model for real-time emotion classification and prediction. *arXiv preprint arXiv:2206.09509*.
- Passos, L. F. C. (2022). Métodos de regularização no aprendizado de máquinas: Ridge e lasso.
- Wong, J. J. N. and Fadzly, N. (2022). Development of species recognition models using google teachable machine on shorebirds and waterbirds. *Journal of Taibah University for Science*, 16(1):1096–1111.