

SISTEMA DE RECONHECIMENTO DE MÁSCARAS FACIAIS COM TEACHABLE MACHINE E PYTHON

DAVID DE MORAIS SILVA¹, VITOR HUGO THOMAZ LEAL², MARIO T SHIMANUKI³

¹ Graduando em Tecnologia em Análise e Desenvolvimento de Sistemas, IFSP, Câmpus Caraguatatuba, david.morais@ifspcaragua.net

² Graduando em Tecnologia em Análise e Desenvolvimento de Sistemas, IFSP, Câmpus Caraguatatuba, vitor.thomaz@aluno.ifsp.edu.br

³ Professor do curso de Tecnologia em Análise e Desenvolvimento de Sistemas, IFSP, Câmpus Caraguatatuba, mario@ifspcaragua.net

Área de conhecimento (Tabela CNPq): Ciência da Computação – 1.03.00.00-7

RESUMO: Este estudo com Teachable Machine e Python desenvolveu um sistema automatizado para detecção de máscaras faciais, com foco principal em reconhecimento em tempo real. O sistema se mostra eficiente quando em condições favoráveis, onde apresentou uma precisão de 100%, mas seu desempenho diminuiu significativamente em situações adversas, tais como iluminação inadequada e ângulos diferentes, chegando a atingir apenas 20% de precisão em ângulos onde o rosto se coloca em 60° e chegando a falhar completamente em 90°. Como proposta para a melhoria do software, sugere-se a utilização de uma inteligência artificial com maior robustez e um dataset mais diversificado e preciso para lidar com os diferentes contextos. Para estudos futuros, mira-se uma expansão do sistema para a detecção de outros tipos de EPIs.

PALAVRAS-CHAVE: reconhecimento automático; padrões faciais; machine learning; detecção de EPIs; classificação em tempo real.

1 INTRODUÇÃO

O reconhecimento automático de objetos e padrões faciais é um campo amplamente pesquisado na computação, que vem sendo utilizado em diversas áreas, como as de monitoramento, controle de acesso e segurança. Entre os desafios, está a detecção de equipamentos de proteção individual (EPIs), especialmente a de máscaras faciais, em variados cenários. Automatizando esse processo por via de técnicas de machine learning que oferece uma solução robusta e eficaz, aplicável em ambientes industriais, comerciais e na segurança pública, onde o uso de EPIs é exigido por regulamentações internas de segurança.

O estudo foi motivado pela crescente demanda por sistemas que possam identificar, com precisão e em tempo real, o uso de EPIs em locais monitorados. Essa necessidade é ainda mais relevante à medida que se intensifica a busca por tecnologias que assegurem o cumprimento das normas de segurança, minimizando a dependência de monitoramento humano e imprecisão do mesmo. O desenvolvimento de soluções tecnológicas acessíveis, alicerçadas em bibliotecas de código aberto e ferramentas de aprendizado de máquina, ajuda na criação desses sistemas de baixo custo.

O principal objetivo deste artigo é construir um sistema automatizado capaz de detectar, com alta precisão, o uso de máscaras faciais. Utilizando técnicas de aprendizado supervisionado e alimentado por, principalmente, uma base de dados pública, o sistema faz uso de plataformas como o Teachable Machine e Python para garantir a detecção em

tempo real. A integração com sistemas de monitoramento existentes pode aprimorar significativamente a segurança em locais onde o uso de máscaras faciais é obrigatório.

O problema investigado neste trabalho envolve o desenvolvimento de uma solução para detectar o uso de máscaras faciais em diferentes circunstâncias, considerando as diferentes variáveis, tais como condições de iluminação e ângulos de captura. A relevância desse estudo está na necessidade de garantir o cumprimento de normas de segurança e fornecer uma ferramenta eficaz para o controle automatizado do uso de máscaras.

2 TEORIA

A identificação automática de objetos e traços faciais é um campo de estudo extenso na computação, com usos em vários setores, como monitoramento, controle de acesso e segurança. Dentre os obstáculos, ressalta-se a identificação de equipamentos de proteção individual (EPIs), particularmente as máscaras faciais, em diversas situações. A automação deste procedimento, por meio de métodos de aprendizado de máquina, proporciona uma solução sólida e eficiente, adequada para uso em contextos industriais, comerciais e na segurança pública, onde a utilização de Equipamentos de Proteção Individual é requerida por normas internas de proteção.

A contínua procura por um sistema que possa garantir a observação de EPIs de maneira assertiva e em tempo real nos locais restritos despertou o interesse no assunto e me motivou a realizar o estudo. A tecnologia que garante o cumprimento das normas de segurança em locais restritos, ao mesmo tempo, minimiza a necessidade de uma constante vigilância humana e torna-se cada vez mais necessária para a humanidade. A utilização de tecnologias acessíveis construídas em torno de bibliotecas de código aberto e recursos de aprendizado de máquina ajuda com uma maior acessibilidade para construir sistemas baratos sem afetar sua eficiência.

Nos últimos anos, o desenvolvimento e avanço das CNNs tem sido uma prática, especialmente de solução, que requer limitação de classificação em tempo real. Alguns estudos com dados recentes têm mostrado o desempenho das redes para casos práticos de detecção de máscara facial em condições controladas. Além disso, os prestadores utilizaram uma abordagem do artigo recente de Nagrath et al. Material, que integraram o modelo MobileNetV2 com o detector SSD para classificação rápida e precisa em tempo real. Além disso, o estudo de Ullah et al. (2022) introduziu o modelo DeepMaskNet, que demonstrou robustez frente a variações de ângulo e condições de iluminação, alcançando alta precisão mesmo em cenários complexos.

No entanto, modelos mais tradicionais, como o InceptionV3, conforme explorado por Chowdary et al. (2020), enfrentam desafios em ambientes com variações significativas de iluminação e ângulos, como demonstrado por Addagarla et al. (2020). A utilização dessas redes mais profundas requer um maior poder computacional e um nível técnico mais elevado, o que limita sua aplicação em cenários práticos e em tempo real.

Recentemente, novos avanços têm sido feitos para melhorar a precisão em condições adversas. Estudos como o de Zhang et al. (2023), por exemplo, exploraram o uso de redes generativas adversariais (GANs) para aumentar a diversidade de dados e melhorar a generalização dos modelos. Outros pesquisadores, como Li et al. (2024), implementaram técnicas de normalização dinâmica de imagens para lidar com diferentes fontes de variação na qualidade, obtendo resultados promissores para sistemas de detecção em tempo real.

A abordagem de deep learning continua a ser uma fronteira para o aprimoramento da detecção de máscaras faciais, com esforços focados em superar limitações de precisão em ambientes não controlados. A combinação de técnicas de aprendizado supervisionado

com dados mais diversificados e métodos de normalização mais avançados tem mostrado grande potencial, como demonstrado nas pesquisas mais recentes.

3 MATERIAL E MÉTODOS

Para o desenvolvimento desse sistema, utilizou-se o Teachable Machine para treinar um modelo de reconhecimento facial com o foco, principalmente, na detecção de máscaras. O principal motivo para a escolha desse framework foi a facilidade de uso, integração com as diferentes plataformas de machine learning e acessibilidade para a fácil e rápida implementação em ambientes de produção. O banco de dados de imagens escolhido para o treinamento do modelo foi o Kaggle, uma plataforma conhecida por sua qualidade e diversidade de datasets. As imagens foram escolhidas de acordo com a disponibilidade do site, considerando as condições de iluminação, ângulos e máscaras, para assegurar uma melhor abrangência.

O principal dataset utilizado no nosso projeto foi o Face Mask Lite Dataset: <<https://www.kaggle.com/datasets/prasoonkottarathil/face-mask-lite-dataset>>

O processo de preparação do dataset envolveu as seguintes etapas:

Seleção das Imagens Foram escolhidas imagens de alta qualidade disponíveis no Kaggle. Prioritariamente, selecionamos imagens com o rosto em vista frontal, pois os datasets de qualidade disponíveis eram, em sua maioria, compostos por imagens de rostos nessa posição. O dataset foi dividido em dados de treinamento (95%) e dados de teste (5%).

Pré-processamento As imagens passaram por um processo de normalização para ajuste de contraste e brilho, utilizando a biblioteca OpenCV. Também foram utilizadas técnicas de aumento de dados para simular condições e ângulos variados.

Critérios de Treinamento e Validação Durante o treinamento, o modelo foi alimentado com uma combinação equilibrada de imagens de rostos com e sem máscaras, buscando uma distribuição representativa de condições.

Em relação a linguagens e bibliotecas, o projeto foi desenvolvido em Python, devido à ampla adoção na comunidade de ciência de dados e machine learning, além de sua compatibilidade com as ferramentas. As bibliotecas utilizadas encontram-se na Tabela 1.

TABELA 1. BIBLIOTECAS

Bibliotecas	Descrição
OpenCV	Utilizado para o processamento de imagens e detecção facial.
TensorFlow	Utilizado para integrar o modelo treinado no Teachable Machine
MediaPipe	Utilizado para a detecção de pontos-chave no rosto, facilitando a análise visual.
Tkinter	Utilizado para a criação de uma interface gráfica que permite visualizar e interagir com o sistema de uma forma mais eficiente, eliminando a necessidade de executar diferentes arquivos para diferentes funções do sistema.

A escolha de utilizar o Teachable Machine foi feita pela sua acessibilidade e capacidade de criar modelos eficientes para a classificação de imagens de forma fácil e

rápida. O foco do estudo é a detecção de máscaras faciais; a abordagem baseada em redes neurais convolucionais, facilitada pelo Teachable Machine e TensorFlow, mostrou-se a melhor. O uso do banco de dados da Kaggle permitiu o treinamento com um conjunto amplo de imagens, proporcionando robustez e precisão ao modelo.

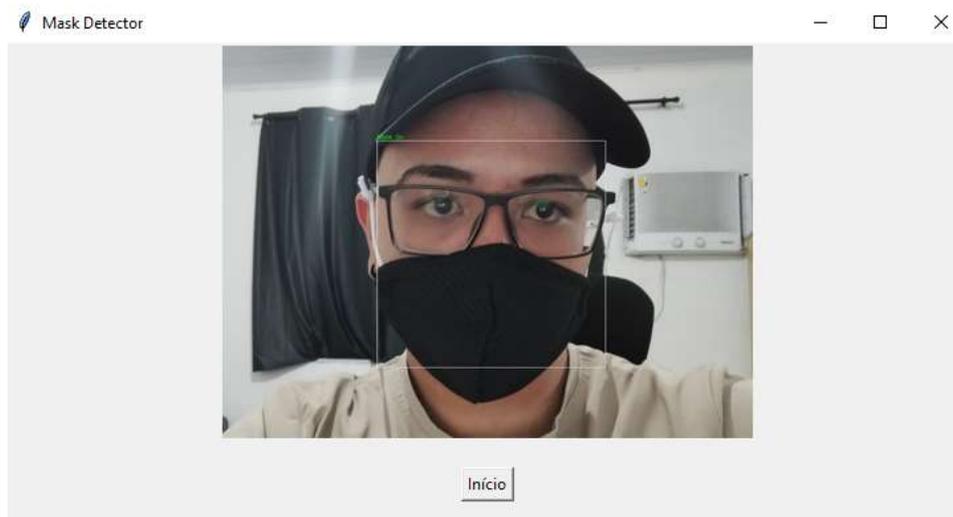
A arquitetura do sistema pode ser dividida em três componentes principais:

Captura e Pré-processamento de Imagens As imagens são capturadas em tempo real através de uma webcam utilizando a biblioteca OpenCV. As imagens passam por um processo de detecção facial e de pontos-chave do rosto utilizando o MediaPipe, para que então possam ser redimensionadas e normalizadas para o formato necessário do modelo.

Classificação e Reconhecimento O modelo treinado no Teachable Machine é carregado e utilizado para classificar as imagens processadas, identificando se o indivíduo está ou não usando uma máscara facial.

Interface Gráfica e Feedback Visual A interface gráfica foi desenvolvida a partir do Tkinter, permitindo que os usuários visualizem as imagens escolhidas ou então as imagens da webcam e verifiquem o status da detecção em tempo real, assim como mostrado na Figura 1. Essa abordagem proporciona uma interação mais lúdica com o sistema, eliminando a necessidade de executar scripts individuais manualmente.

FIGURA 1. Imagem do aplicativo mostrando a detecção em tempo real.



Confira os principais componentes, módulos e funcionalidades do sistema na Tabela 2.

TABELA 2. COMPONENTES, MÓDULOS E FUNCIONALIDADES

Módulos	Descrição
mask_detector.py	Módulo que é responsável pela detecção e classificação de rosto com ou sem as máscaras através de uma webcam utilizando o modelo treinado.
mask_detector_img.py	Parecido com o anterior, porém nesse modulo a análise é feita apenas com imagens estáticas.
main.py	Esse é o arquivo principal, o qual é responsável por integrar os dois módulos anteriores e fornecer uma interface visual para uma melhor utilização do sistema.

A implementação do software começou utilizando apenas a webcam. No entanto, à medida que o sistema se tornou mais complexo e surgiram dificuldades, como a inconsistência na detecção, decidimos dividir o processo em duas partes: a detecção pela webcam e a detecção por imagem estática. Posteriormente, adicionamos um módulo principal para permitir o uso simultâneo dos dois módulos, tornando a utilização da aplicação mais fácil e dinâmica.

Durante o desenvolvimento do sistema, um dos principais desafios foi a detecção precisa do rosto, sendo a iluminação e os ângulos escolhidos os principais fatores. Esses problemas foram mais evidentes durante a análise com a webcam.

Para solucionar essas questões, utilizamos um conjunto de imagens mais diversificado, com diferentes ângulos e condições de iluminação. Além disso, aplicamos técnicas de normalização de imagem e ajuste dinâmico de brilho utilizando OpenCV.

O sistema adotou algoritmos de aprendizado supervisionado, com ênfase no uso de redes neurais convolucionais (CNNs). O design escolhido seguiu o padrão Model-View-Controller (MVC), que separa de forma clara a lógica de detecção (modelo), o processamento das imagens, e a interface visual (view), garantindo uma organização eficiente do código.

4 RESULTADOS E DISCUSSÃO

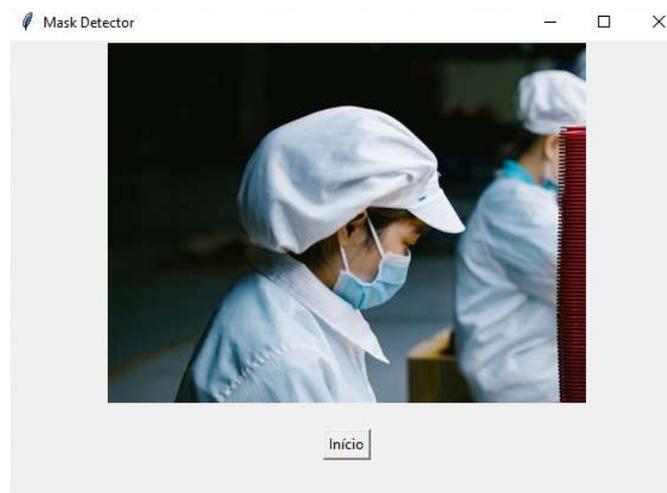
A Figura 2 ilustra os resultados dos testes que, sob condições controladas, com o rosto em uma posição estática e iluminação constante, o sistema alcançou uma precisão de 100%. Este resultado demonstra a efetividade do modelo treinado em reconhecer corretamente as máscaras faciais quando as variáveis ambientais estão sob controle. A Figura 2 ilustra o aplicativo efetuando com êxito a identificação facial, ressaltando o reconhecimento exato em tempo real.

FIGURA 2. Imagem do aplicativo, mostrando a detecção facial com resultado correto.



Contudo, ao submeter o sistema a condições de ângulos e iluminação inadequadas, seja em excesso ou em intensidade insuficiente, a precisão do sistema foi consideravelmente prejudicada, conforme demonstrado na Figura 3. Os erros detectados em situações com iluminação insuficiente demonstram a sensibilidade do sistema à qualidade da imagem, um elemento fundamental para a exatidão. Pesquisas similares, como a realizada por Addagarla et al. (2020), que empregaram imagens de alta resolução, notaram uma queda considerável no rendimento do sistema quando a qualidade da imagem era diminuída. A Figura 3 representa a situação em que o sistema não funcionou por causa de uma iluminação insuficiente.

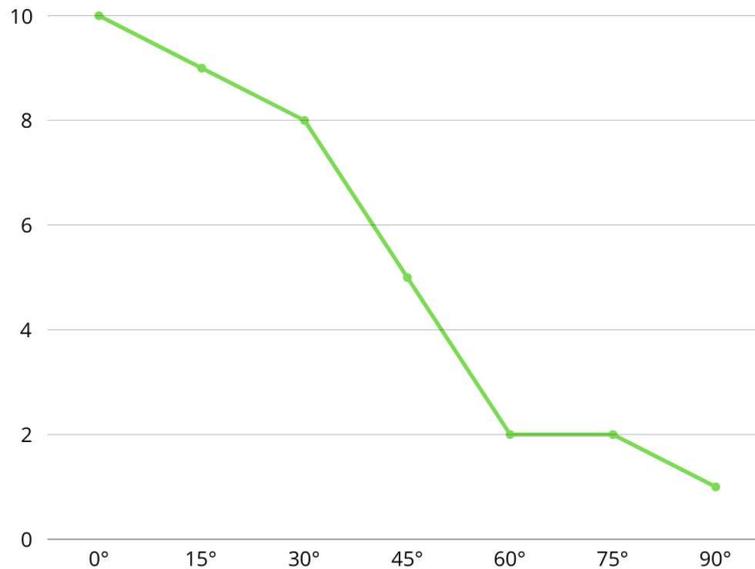
FIGURA 3. Imagem do aplicativo, mostrando a detecção facial com resultado incorreto.



A precisão do sistema também foi influenciada pelo ângulo de captura, além da iluminação. A Figura 4 apresenta os resultados da identificação facial em variados ângulos de registro, variando de 0° a 90°, sendo 0° o rosto visto de forma frontal e 90° visto de perfil. O sistema manteve uma exatidão de 100% quando o rosto estava perpendicular à câmera (0°). Contudo, conforme o ângulo aumentava, a taxa de acertos caía: com um ângulo de 30°, o número de acertos diminuiu para 8 em 10 tentativas, enquanto com um ângulo de 60°, o número de acertos diminuiu para 2. O sistema não foi capaz de identificar a máscara adequadamente em ângulos de 90°. Esses resultados estão alinhados com as

descobertas de Chowdary et al. (2020), que também notaram uma diminuição no rendimento em circunstâncias com ângulos extremos.

FIGURA 4. Imagem de um gráfico mostrando os resultados em determinados ângulos.



Embora o modelo proposto tenha mostrado resultados satisfatórios sob condições controladas, sua performance foi inferior à de outros sistemas mais avançados descritos na literatura, como o DeepMaskNet de Ullah et al. (2022). O modelo DeepMaskNet se demonstrou mais robusto frente a variações de iluminação e ângulos, alcançando uma precisão de 100% em uma ampla gama de cenários. Essa performance superior foi obtida pela utilização de redes neurais mais profundas e técnicas de normalização avançadas, que permitiram uma adaptação mais eficiente às condições variáveis. A implementação de métodos semelhantes no nosso sistema poderia melhorar consideravelmente os resultados, especialmente em cenários com variações angulares e de iluminação.

Além disso, a utilização de um conjunto de dados mais diversificado, com maior variação nos ângulos e nas condições de iluminação, como sugerido por Zhang et al. (2023), poderia proporcionar ao sistema uma maior robustez. No nosso estudo, o aumento da diversidade de dados foi uma tentativa de superar essas limitações, mas uma maior variedade ainda é necessária para garantir a precisão do modelo em condições do mundo real. A normalização dinâmica de imagens, como sugerido por Li et al. (2024), também pode ser uma estratégia vantajosa para melhorar o desempenho do sistema.

A Tabela 3 fornece um resumo completo dos resultados principais alcançados pelo sistema, enfatizando sua exatidão em diversas situações de teste. Esta tabela mostra como o rendimento do modelo variou com base em variáveis ambientais, como a iluminação e os ângulos de captura, proporcionando uma perspectiva nítida da efetividade do sistema em variados contextos. A análise dos resultados alcançados em situações ideais e desfavoráveis proporciona percepções valiosas sobre as restrições do modelo, bem como aponta áreas com potencial para aprimoramento.

Tabela 3. Desempenho do sistema em diferentes condições de teste.

Condição	Descrição
Iluminação controlada	100
Iluminação inadequada	50
Ângulo 30°	80
Ângulo 60°	20
Ângulo 90°	0

Estes achados destacam a exigência de um modelo mais sólido, apto a gerir mudanças nas condições locais. Ademais, a comparação com pesquisas anteriores sugere que, mesmo que o modelo sugerido seja eficiente em circunstâncias ideais, ele pode ser melhorado por meio do uso de métodos mais sofisticados de aprendizado profundo e conjuntos de dados mais representativos.

5 CONSIDERAÇÕES FINAIS

O sistema de detecção de máscaras faciais desenvolvido neste estudo apresentou bons resultados em condições controladas de iluminação e posição fixa do rosto. Contudo, algumas limitações importantes foram observadas. O desempenho foi impactado negativamente por variações de iluminação e ângulos de captura, além da necessidade de uma maior diversidade de imagens para o treinamento.

Diante disso, uma melhoria no processo de treinamento com uma IA mais robusta e um dataset com imagens de maior qualidade e variedade são sugeridas como aprimoramentos. A inclusão de mais exemplos, especialmente em ângulos diferentes, deve aumentar a precisão e a generalização do modelo.

Para estudos futuros, planeja-se expandir o sistema para detectar outros EPIs, como capacetes e óculos de proteção, aumentando a aplicabilidade em diversos ambientes, como indústrias e comércios. Além disso, a integração com sistemas de monitoramento pode tornar a solução ainda mais eficaz no controle automatizado de EPIs.

O código-fonte completo do sistema pode ser acessado em: <<https://github.com/david-uzkk/maskDetector>>

REFERÊNCIAS

ADDAGARLA, S. K. Real Time Multi-Scale Facial Mask Detection and Classification Using Deep Transfer Learning Techniques. ago. 2020.

HAJIRA SIDDIQUA, S. C. A study on post-operative change in intraocular pressure and anterior chamber depth after retinal detachment surgery. 2020.

KOTTARATHIL, P. Face Mask Lite Dataset. Disponível em: <<https://www.kaggle.com/datasets/prasoonkottarathil/face-mask-lite-dataset>>. Acesso em: 15 out. 2024.

LARXEL. Face Mask Detection. Disponível em: <<https://www.kaggle.com/datasets/andrewmvd/face-mask-detection>>. Acesso em: 15 out. 2024.

SICLN 2024

Seminário de Iniciação Científica e Pesquisa do Litoral Norte

NAEEM ULLAH A , ALI JAVED B , MUSTANSAR ALI GHAZANFAR C , ABDULMAJEED ALSUFYANI D , SAMI BOUROUIS E. A novel DeepMaskNet model for face mask detection and masked facial recognition. nov. 2022.

PREETI NAGRATH A , RACHNA JAIN A , AGAM MADAN A , ROHAN ARORA A , PIYUSH KATARIA A , JUDE HEMANTH B. SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. mar. 2021.

SITI MURNI HANUM MD ARIFF NORLELA ISHAK MAZIDAH TAJJUDIN. IoT-Based and Teachable Machine Platform for Covid-19 Prevention and Control. Semarak Ilmu, 8 out. 2024.