Reconhecimento Facial Para Controle de Acesso em Ambientes Fechados

Caio César Sabino Soares
caio.soares@edu.univali.br
Universidade do Vale do Itajaí
Itajaí, Santa Catarina

Anita Maria da Rocha Fernandes
anita.fernandes@univali.br
Universidade do Vale do Itajaí
Itajaí, Santa Catarina

Eduardo Alves da Silva
eas@univali.br
Universidade do Vale do Itajaí
Itajaí, Santa Catarina

Rudimar Luís Scaranto Dazzi
rudimar@univali.br
Universidade do Vale do Itajaí
Itajaí, Santa Catarina

Wemerson Delcio Parreira
wemerson.delcio@puc-campinas.edu.br
Pontifícia Universidade Católica de Campinas
Campinas, São Paulo

Resumo

A set of Computer Vision algorithms was validated through a facial recognition-based access control system designed for indoor environments, motivated by the necessity for contactless authentication and the imperative to surmount the limitations inherent to traditional cards and passwords. The research, characterized as applied and quantitative and utilizing the hypothetical-deductive method, performed performance evaluations based on objective metrics within controlled scenarios. The methodology included a Systematic Literature Review (SLR) in which nine databases were consulted, resulting in 381 initial records, which were subsequently reduced to 45 eligible primary studies following the application of exclusion criteria. Within the literature, precision ranged from 78.4% to 100% (mean: 96%) and accuracy from 80% to 100% (mean: 94%), signaling maturity in controlled environments, despite persistent gaps regarding standardized reporting. Regarding the prototype, implemented in Python with facial detectors and extractors, trials conducted with 6 cameras of distinct configurations demonstrated a strong dependence on capture conditions. specifically, average accuracy increased from 65.7% at 110 lux to 82.1% at 515 lux. Conversely, long-distance capture (350 cm) caused a drop in the average from 92.7% to 32.2%, except when utilizing dedicated optical zoom, thereby reinforcing the importance of facial scale on the sensor. Furthermore, the utilization of a pre-processing technique enhanced accuracy, ensuring greater reliability in facial recognition tasks. Multiple additional scenarios were evaluated, such as spoofing attempts, user motion, the use of accessories, partial facial occlusion, and night vision capture. This work presents a set of techniques operationalized through a web-based system and attests to the applicability of this type of solution in both controlled and real-world scenarios. By considering a range of acquisition devices under distinct conditions, this study provides the necessary tools to facilitate replicability.

Keywords

Access control, artificial intelligence, computer vision, face recognition;

1 Introdução

A Tecnologia de Reconhecimento Facial (TRF) consolidou-se como um campo proeminente dentro da IA, impulsionada pelos avanços contínuos em Aprendizado de Máquina (Machine Learning – ML) e Aprendizado Profundo (Deep Learning – DL) [1–3]. Esse crescimento tem favorecido sua adoção em diferentes setores, especialmente em aplicações que demandam identificação rápida, confiável e automatizada de indivíduos.

O reconhecimento facial consiste em identificar ou verificar pessoas a partir de imagens ou fluxos de vídeo [4]. Apesar da aparente simplicidade, o processo envolve desafios significativos decorrentes de variações de iluminação, pose, oclusões parciais e expressões faciais [5]. Para mitigar esses efeitos, técnicas baseadas em *landmarks* realizam a detecção de pontos estruturais do rosto, permitindo alinhar e normalizar a imagem antes das etapas posteriores [6], conforme ilustrado na Figura 1.

Após essa normalização, o *pipeline* de reconhecimento facial organiza-se em três etapas principais: detecção, extração de características e classificação, nas quais vetores gerados pela Inteligência Artificial (IA) são comparados a representações previamente armazenadas em um banco de dados [5]. Esse modelo tem possibilitado aplicações diversas, como autenticação em serviços digitais, verificação de identidade em redes sociais, análise comportamental, recomendação de conteúdo e sistemas de pagamento [5, 7, 8].

Entre essas aplicações, destaca-se o uso de TRF para controle de acesso e registro de presença, área em que a demanda por plataformas robustas de autenticação tem crescido substancialmente [1]. A biometria facial também tem sido incorporada a sistemas de gestão fronteiriça e verificação supranacional de identidade, substituindo métodos tradicionais baseados em cartões, senhas ou registro manual [9–11]. Tais soluções oferecem benefícios como automação, redução de erros e maior eficiência operacional.

O principal desafio técnico permanece na capacidade de reconhecer rostos mesmo diante de transformações inevitáveis, como variações de escala, rotação, envelhecimento e uso de acessórios. Diante desse cenário, este trabalho propõe e valida um modelo de reconhecimento facial para controle de acesso, com as seguintes contribuições:

- Proposição de uma metodologia para seleção e aplicação de diferentes TRF presentes no estado da arte;
- Demonstração da eficiência de técnicas específicas para detecção, extração de características e classificação facial; e
- Disponibilização de uma solução alternativa aos métodos tradicionais de controle de acesso, visando otimização administrativa.

Figura 1: *Landmarks* no processo de reconhecimento facial de indivíduo [6].

1.1 Problematização

A dependência contínua de métodos tradicionais de controle de acesso, como crachás, cartões por aproximação, códigos de barras e listas de presença, expõe organizações a fragilidades operacionais, riscos de fraude e limitações decorrentes de falhas humanas [12]. Esses sistemas, amplamente utilizados em escolas, hospitais, repartições públicas, empresas e ambientes de grande circulação, mostram-se pouco eficientes em cenários onde a identificação rápida, precisa e automatizada é essencial.

Eventos críticos registrados nas últimas décadas, como os ataques em Columbine (1999) e na Escola Estadual Professor Raul Brasil, em Suzano (2019), evidenciaram que mecanismos de acesso baseados em dispositivos físicos ou verificação manual são insuficientes frente a situações de risco elevado, revelando falhas estruturais na prevenção e no monitoramento de indivíduos não autorizados [13, 14]. Da mesma forma, em hospitais e ambientes corporativos, a ausência de sistemas robustos pode comprometer a segurança de pacientes, colaboradores e informações sensíveis, especialmente em locais onde o fluxo diário de pessoas é alto e heterogêneo [15, 16]. Em fronteiras e aeroportos, a falta de biometria confiável dificulta a verificação eficiente de identidades e a detecção de múltiplos perfis associados ao mesmo indivíduo [9].

Tecnologias de reconhecimento facial baseadas em IA surgem como alternativa promissora a esse cenário, por permitir identificação em tempo real sem contato físico e com maior agilidade [17]. Contudo, sua adoção prática é limitada pela ausência de consenso sobre quais técnicas oferecem melhor desempenho em condições operacionais adversas, como iluminação irregular, variação de ângulo, distâncias inconsistentes e oclusões parciais. A literatura apresenta soluções heterogêneas, resultados fragmentados e avaliações realizadas sob métricas distintas, dificultando a compreensão do estado da arte e a tomada de decisão para implementação em escala institucional.

Essas lacunas tornam necessário investigar: (i) quais técnicas de reconhecimento facial têm sido mais empregadas em sistemas de controle de acesso e frequência; (ii) quais desafios operacionais persistem na aplicação prática dessas tecnologias; e (iii) quais métricas e *benchmarks* sustentam a validação dos modelos disponíveis. A inexistência de uma síntese sistematizada desses elementos compromete tanto a avaliação crítica do desempenho dos sistemas quanto a projeção de sua viabilidade em ambientes reais.

Diante disso, este estudo busca avaliar abordagens de reconhecimento facial aplicadas ao controle de acesso, contribuindo com comparações para o avanço técnico do campo e fornecendo subsídios quantitativos para decisões de adoção tecnológica e de dispositivos de aquisição em diferentes instituições.

2 Fundamentação Teórica

2.1 Reconhecimento Facial

A Visão Computacional (Computer Vision - CV) constitui um dos pilares da IA, dedicando-se ao desenvolvimento de algoritmos capazes de interpretar informações visuais de forma automatizada e eficiente [18, 19]. Seu objetivo central é extrair e analisar padrões presentes em imagens e vídeos, provenientes de diferentes sensores e cenários, de modo a permitir que sistemas computacionais compreendam elementos relevantes do ambiente [20].

Inserido nesse campo, o reconhecimento facial destaca-se como uma tecnologia biométrica amplamente consolidada, capaz de identificar ou verificar indivíduos por meio das características únicas de seus rostos [21–23]. O avanço recente do poder computacional e da qualidade das imagens tem ampliado significativamente suas aplicações, que abrangem segurança e vigilância, autenticação, controle de acesso, redes sociais, e-commerce e apoio a pessoas com deficiência visual [23, 24].

2.2 Aprendizado de Máquina

O ML é uma subárea central da IA, dedicada ao desenvolvimento de algoritmos capazes de aprender a partir de dados e aprimorar seu desempenho com a experiência, sem depender de programação explícita para cada tarefa [25, 26]. Seu objetivo principal é explorar grandes volumes de dados para extrair padrões e informações relevantes, sobretudo em cenários em que abordagens tradicionais se tornam inviáveis devido à complexidade ou variabilidade dos problemas [26, 27].

2.3 Rede Neural Convolucional

As Redes Neurais Convolucionais (Convolutional Neural Network - CNN) são um tipo especializado de Rede Neural Artificial (Artificial Neural Network - ANN) que se consolidou como componente central em sistemas de visão computacional, obtendo desempenho de destaque em tarefas visuais complexas e figurando entre as técnicas mais influentes do DL [28, 29]. Inspiradas em princípios biológicos, as CNNs utilizam estruturas em camadas para aprender representações hierárquicas dos dados, sendo amplamente empregadas em aplicações que exigem interpretação automatizada de imagens e vídeos [30, 31].

2.3.1 Redes Neurais Convolucionais Multitarefa em Cascata. As Redes Neurais Convolucionais Multitarefa em Cascata (Multi-task Cascaded Convolutional Networks - MTCNN) é uma arquitetura em cascata composta por três estágios de CNN, desenhada para detecção de faces e localização de pontos faciais de forma progressiva, do nível mais grosseiro ao mais refinado [32]. No primeiro estágio, a Proposal Network (P-Net) opera em múltiplas escalas gerando

janelas candidatas e vetores de regressão de caixas delimitadoras. Em seguida, a Refine Network (R-Net) filtra candidatos falsos e realiza calibração com supressão não máxima. Por fim, a Output Network (O-Net) refina as detecções e estima a posição de cinco pontos faciais: olhos esquerdo e direito, nariz e cantos esquerdo e direito da boca [33, 34].

3 Trabalhos Relacionados

Para contextualizar e fundamentar esta pesquisa, conduziu-se uma Revisão Sistemática da Literatura (RSL) com o objetivo de identificar como o reconhecimento facial tem sido aplicado em sistemas de controle de acesso. As perguntas de pesquisa consideradas foram:

- Quais técnicas de inteligência artificial são utilizadas no reconhecimento facial aplicado ao controle de acesso?
- Quais algoritmos são adotados nos estudos?
- Quais *datasets* são empregados?
- Quais desafios ou lacunas são relatados nos trabalhos analisados?

3.1 Métodos e Resultados

A revisão foi conduzida com base na lista de verificação Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) e planejada por meio da plataforma Parsifal[1], onde se definiu o protocolo inicial utilizando o modelo População, Intervenção, Contexto, Outcome/Saída e Comparação (PICOC) para orientar as decisões metodológicas.

Com base no protocolo, definiram-se palavras-chave, sinônimos e a *string* de busca aplicada nas bases selecionadas:

> *(("AI"OR "artificial intelligence"OR "computational vision"OR "computer vision") AND ("access control"OR "face recognition"OR "facial recognition"))*

A busca contemplou artigos publicados entre janeiro de 2020 e abril de 2025, nos idiomas português, inglês ou espanhol em 9 bases de periódicos: ACM Digital Library, Embase, Engineering Village, IEEE Xplore, PubMed, ScienceDirect, Scopus, Springer e Web of Science.

Os critérios de inclusão foram: (i) uso de técnicas de reconhecimento facial com IA; e (ii) métricas quantitativas e experimentos reportados. Os critérios de exclusão foram: estudos fora do escopo, inacessíveis, teóricos sem validação experimental, duplicados, retratados ou que representassem apenas partes de sistemas maiores.

A partir da aplicação dos critérios e da triagem baseada no fluxo PRISMA, foram selecionados 381 estudos primários empregados para análise qualitativa e quantitativa. Após verificação automática, 208 periódicos foram eliminados por duplicação. Após a aplicação dos critérios, uma listagem preliminar de 45 artigos aceitos foi considerada para leitura.

Entre a totalidade dos 45 artigos analisados, a técnica de IA mais utilizada[2] foi a CNN, empregada em 24 estudos (53,33%). Já no que tange às técnicas algorítmicas relacionadas ao reconhecimento facial, a mais frequentemente empregada foi o Haar Cascades Classifier (HCC), utilizada em 15 estudos (33,33%). Por fim. o *dataset* mais utilizado foram as bases próprias ou personalizadas (BPP), empregadas em 18 estudos (40,00%).

A síntese dos resultados da RSL subsidiou a escolha das técnicas experimentais deste trabalho.

4 Desenvolvimento

A aplicação de controle de acesso desenvolvida tem como finalidade demonstrar o potencial dos algoritmos selecionados para autenticação baseada em reconhecimento facial em ambientes controlados. A solução foi projetada para laboratórios e espaços institucionais, eliminando a necessidade de credenciais físicas ou senhas ao realizar a identificação de usuários por meio de suas características faciais. O sistema processa imagens capturadas por câmeras locais ou remotas, permitindo sua instalação em diferentes configurações e ampliando sua aplicabilidade prática. Além disso, a aplicação integra ferramentas de gestão centralizada de usuários, registrando data e hora de cada acesso e disponibilizando uma interface web para administração das informações.

4.1 Implementação

Os testes para validação foram realizados em laboratórios, com ambientes controlado, com as câmeras posicionadas de modo que o usuário fosse capturado em trânsito normal. Indivíduos previamente cadastrados foram identificados ao olhar diretamente para o dispositivo.

A implementação foi desenvolvida em Python com reconhecimento e processamento facial a partir das bibliotecas DeepFace, TensorFlow e Keras. O DeepFace fornece uma interface de alto nível para modelos de reconhecimento facial, incluindo o FaceNet, uma arquitetura baseada em Inception-ResNet-v1 que mapeia rostos em um espaço vetorial euclidiano por meio de *embeddings*, conforme ilustrado na Figura 2.

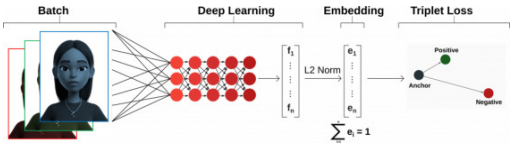

Figura 2: Fluxo de funcionamento do FaceNet [35].

Para detecção facial, além da implementação interna do DeepFace, foram utilizados dois detectores adicionais: a MTCNN, com estrutura em cascata multitarefa, e o RetinaFace, modelo baseado em redes neurais profundas de alta precisão. Essas abordagens complementares foram empregadas para melhorar o desempenho diante de variações de iluminação, ângulo e distância.

As bibliotecas OpenCV e Pillow foram utilizadas para o processamento de imagens, captura em tempo real, manipulação de formatos e pré-processamento. A interface web e a API REST foram desenvolvidas utilizando o *framework* Flask, com suporte a CORS via *flask-cors*. Para execução em produção, adotou-se o servidor WSGI Gunicorn. O armazenamento dos dados foi realizado no Firebase, um banco de dados NoSQL em nuvem estruturado em formato JSON.

[1] Disponível em https://parsif.al
[2] Mais utilizado em quantidade de artigos e porcentagem de artigos em relação ao todo (45 artigos).

4.1.1 Dispositivos de Aquisição de Imagens. Foram utilizados diferentes dispositivos de captura de vídeo, com o objetivo de avaliar o comportamento do sistema sob distintas condições de resolução, qualidade óptica e iluminação. No total, seis câmeras foram empregadas: (C1) a câmera integrada do notebook (sensor CMOS frontal, resoluções HD e *Full HD* a 30 FPS), adotada como dispositivo principal de teste; (C2) uma webcam USB *Full HD* genérica, de baixo custo e configuração *plug & play*; (C3) uma webcam Razer Kiyo, com resolução *Full HD* e anel de luz integrado (*ring light*), projetada para uso em ambientes internos com iluminação controlável; (C4) uma webcam 4K (TW-350), com resolução 3840 × 2160 *pixels* e foco automático; (C5) uma câmera DSLR Canon EOS Rebel T6, utilizada como referência óptica de maior qualidade; e (C6) uma câmera IP Orbitronic ORB776, capaz de transmitir vídeo em *Full HD* via Wi-Fi ou conexão cabeada.

4.1.2 Treinamento do Modelo. O sistema utiliza o modelo FaceNet pré-treinado, disponibilizado pela biblioteca DeepFace, não sendo necessário realizar um treinamento do zero para esta aplicação. O modelo permanece congelado durante toda a operação, sendo utilizado exclusivamente para extração de *embeddings* faciais e comparação entre vetores de características. As fotos usadas para cadastrar cada pessoa no sistema são totalmente diferentes das fotos usadas para testar o reconhecimento depois.

4.2 Pré-processamento

A aplicação implementa um *pipeline* de pré-processamento de imagens adaptativo ao nível de zoom utilizado. Quando o zoom digital é ativado, intensifica-se o tratamento da imagem para mitigar a perda de detalhes decorrente da interpolação de *pixels*. O processamento é realizado no espaço de cores YCrCb, separando luminância e crominância, o que permite atuar de forma mais eficaz sobre contraste e iluminação.

O fluxo inclui: (i) equalização adaptativa de histograma (CLAHE), para realce de contraste local em regiões com diferentes níveis de iluminação; (ii) filtragem bilateral, para suavização de ruído preservando bordas relevantes para o reconhecimento; (iii) operações de aguçamento (*sharpening*) em um ou dois estágios, dependendo do nível de zoom, a fim de enfatizar contornos e texturas faciais; e (iv) normalização dos valores de *pixel* em um intervalo restrito, reduzindo saturação em tons muito escuros ou muito claros [36–38].

4.3 Classificação

A etapa de classificação facial é baseada em reconhecimento por características (*feature-based recognition*), utilizando o FaceNet pré-treinado para gerar *embeddings* vetoriais das faces. No momento do cadastro, uma imagem do rosto do usuário é capturada, processada e convertida em vetor de características, armazenado no Firebase associado ao identificador do usuário.

Durante o reconhecimento, uma nova captura facial é submetida ao mesmo *pipeline* de pré-processamento e extração de *embeddings*. Em seguida, o vetor resultante é comparado com os vetores de referência armazenados no banco de dados por meio da distância cosseno, definida pelo escore de similaridade:

$$\text{Similaridade} = 1 - \frac{\varphi(x) \cdot \varphi(y_i)}{\|\varphi(x)\| \, \|\varphi(y_i)\|}. \tag{1}$$

Onde $\varphi(x)$ representa o embedding vetorial extraído da face capturada na etapa de consulta, e $\varphi(y_i)$ representa o embedding vetorial da i-ésima face de referência armazenada no banco de dados. Valores de similaridade inferiores ao limiar estabelecido a partir de testes realizados e comparações com o estado da arte da Seção 3 (0,6 para zoom inativo e 0,7 para zoom ativo) são interpretados como correspondência positiva entre a face capturada e o vetor de referência. O modelo não sofre atualização de pesos em tempo de execução; a classificação consiste apenas em *matching* entre *embeddings*.

5 Resultados e discussões

A prezada seção visa apresentar os resultados dos testes realizados com os algoritmos selecionados na literatura em seis dispositivos de aquisição distintos, com ênfase em métricas quantitativas como acurácia e precisão.

5.1 Configuração Geral e Procedimento Experimental

Os testes foram conduzidos com seis câmeras (*C1* a *C6*) em três ambientes distintos (residência do autor, LIA e IoTec Lab), variando condições de iluminação, distância entre usuário e dispositivo, altura da câmera em relação ao solo, presença de movimento e aplicação de filtros de pré-processamento. Em cada sessão, foram definidos:

- lista de rostos alvo (10 usuários cadastrados, incluindo o autor)[3];
- modo de captura contínua, com início e término por botão único;
- meta de aproximadamente 50 imagens por sessão, registrando-se o total efetivo de amostras;
- duração de até três dias por sessão, dentro de um período total de 20 dias de experimentação;
- variação moderada de pose (cerca de 5 cm em todas as direções) e expressões faciais distintas (sorriso, raiva, medo, tristeza e neutra);
- limiar padrão de decisão de 0,6 para o score de similaridade.

O procedimento experimental seguiu três etapas principais: (i) criação da sessão com metadados (dispositivo, zoom, *threshold*, distância, altura, iluminância, local, pré-processamento, cenário e data) e seleção da lista de rostos a reconhecer; (ii) captura contínua de imagens com auditoria manual quadro a quadro, rotulando cada registro como verdadeiro positivo, verdadeiro negativo, falso positivo ou falso negativo; e (iii) envio desses rótulos a um módulo de análise responsável pelo cálculo das métricas e exportação de relatórios consolidados.

5.2 Instrumentação e Condições de Captura

A iluminância (lux) foi aferida com um luxímetro digital Minipa MLM-1332, utilizando-se fita métrica para distâncias e alturas em relação ao solo. Embora se tenha buscado precisão em todas as etapas, reconhece-se margem de erro inerente aos instrumentos e variações ambientais.

[3] Os outros 9 rostos foram obtidos a partir de figuras públicas com imagens disponíveis na internet.

A iluminação foi sempre ambiente, com luminárias posicionadas no teto e sem foco direto no rosto, e as câmeras foram ajustadas para ficar aproximadamente na altura da face do indivíduo. As sessões foram agrupadas em cenários específicos (distância, iluminação, oclusão, acessórios, fraude, zoom, pré-processamento e movimento), permitindo avaliar separadamente o impacto de cada fator sobre o desempenho do sistema.

5.3 Métricas Quantitativas

O desempenho foi avaliado a partir da matriz de confusão por sessão, com os rótulos Verdadeiro Positivo (True Positive - TP), Verdadeiro Negativo (True Negative - TN, Falso Positivo (False Positive - FP) e Falso Negativo (False Negative - FN), e das seguintes métricas derivadas:

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$\text{Precisão} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall (Sens.)} = \frac{TP}{TP + FN} \tag{4}$$

$$\text{F1-Score} = \frac{2 \cdot \text{Precisão} \cdot \text{Recall}}{\text{Precisão} + \text{Recall}} \tag{5}$$

$$\text{FMR (False Match Rate)} = \frac{FP}{FP + TN} \tag{6}$$

$$\text{FNMR (False Non-Match Rate)} = \frac{FN}{FN + TP} \tag{7}$$

$$\text{FDR (Fail Detection Rate)} = \frac{\text{Faces não identificadas}}{\text{Total de registros}} \tag{8}$$

EER (Equal Error Rate). Definida no ponto de limiar τ^* em que $\text{FMR}^{(\tau^*)} = \text{FNMR}^{(\tau^*)}$ nas curvas ROC/DET.

Além disso, foram medidos tempos médios de detecção (TMD), reconhecimento (TMR) e tempo médio total (TMT), bem como reportado o escore de similaridade por cosseno.

Essas métricas permitem avaliar simultaneamente a capacidade global de acerto (acurácia), o equilíbrio entre rejeição de impostores (FMR) e aceitação de genuínos (FNMR, *recall*, F1), e a eficiência temporal do *pipeline*. No contexto de controle de acesso, FMR e FNMR assumem papel central: a primeira está associada à segurança (evitar TN) e a segunda, à usabilidade (evitar FN).

5.4 Desempenho Agregado por Câmera

A Tabela 1 apresenta o resumo por câmera, em que cada dispositivo agrega médias e desvios padrão das sessões correspondentes. Considerando todas as condições de teste, a câmera integrada (C1) concentrou o maior número de sessões e imagens (616 registros em 12 sessões), sendo adotada como referência prática por representar o cenário mais comum em notebooks de uso geral.

De forma geral:

- a C1 apresentou acurácia média de $\approx$ 60% ms, com variação elevada entre cenários (desvio-padrão de 38,3%), refletindo sensibilidade marcada à distância, iluminação e oclusões;
- a C2 obteve a melhor combinação entre acurácia, *recall* e F1 em distâncias maiores, com FMR praticamente nulo e FNMR inferior ao da C1, o que a destacou em cenários desafiadores;
- a C3 exibiu excelente precisão (próxima a 100%), mas com decréscimo de *recall* em condições adversas, o que indica postura mais conservadora (poucos falsos positivos, porém mais falsos negativos);
- a C4 e a C6 apresentaram desempenho mais modesto, com acurácia média inferior e maior FNMR, sobretudo em longas distâncias e iluminação desfavorável;
- a C5 alcançou métricas próximas de 100% nas poucas sessões válidas, mas os problemas de integração e instabilidade inviabilizaram uma avaliação representativa de seu potencial.

Em síntese, os resultados sugerem que, embora webcams integradas sejam suficientes para curtas distâncias com boa iluminação, câmeras com ótica superior ou zoom óptico tendem a preservar melhor as métricas quando o cenário se torna mais exigente.

5.4.1 Distância. A Tabela 2 sintetiza o impacto da distância (100, 200 e 350 cm) na acurácia, sob iluminação alta (515 lux) e indivíduo parado. A média entre dispositivos caiu de 92,7% a 100 cm para 68,4% a 200 cm e 32,2% a 350 cm, evidenciando degradação acentuada à medida que o rosto ocupa porção menor do quadro.

Tabela 2: Comparação de acurácia entre câmeras por distância.

Cenário	C1	C2	C3	C4	C5	C6	Méd.
100 cm	*98%*	***100%***	*82%*	*98%*	*98%*	*80.4%*	*92.7%*
200 cm	*88.2%*	***82%***	*76%*	*68%*	*n/a*	*28%*	*68.4%*
350 cm	*1.90%*	***96.1%***	*6%*	*2%*	*n/a*	*0%*	*32.2%*

A C2 (Razer), com zoom óptico, manteve acurácia de 96,1% a 350 cm, em contraste com a queda abrupta das demais câmeras, algumas chegando próximo de 0%. Isso indica que a combinação de lente e campo de visão mais estreito, ao concentrar *pixels* na região facial, é determinante para viabilidade do reconhecimento em longas distâncias.

5.4.2 Iluminação. A Tabela 3 compara o desempenho em níveis baixo (20 lux), médio (110 lux) e alto (515 lux), com distância de 200 cm e indivíduo parado. Em 20 lux, a C1 não obteve acertos, demonstrando que, abaixo desse patamar, a textura facial torna-se insuficiente para detecção e alinhamento. Em 110 lux, a acurácia passa a ser viável, porém com forte heterogeneidade entre dispositivos.

Tabela 3: Comparação de acurácia entre câmeras por iluminação.

Cenário (Lux)	C1	C2	C3	C4	C5	C6	Méd.
Baixo (20 lux)	***0%***[*]	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*0%*
Médio (110 lux)	***72%***	*52.6%*	*44%*	*0%*	*n/a*	*67.3%*	*65.7%*
Alto (515 lux)	***88.2%***	*82%*	*76%*	*68%*	*n/a*	*28%*	*82.1%*

[*] Sessão capturada a 100 cm para testar a aderência do algoritmo a iluminações extremamente baixas.

Ao elevar a iluminância para 515 lux, observa-se ganho consistente na maioria das câmeras: aumentos de até 30 pontos percentuais em acurácia, com destaque para a C3 e a C4. A C6 (IP) foi

Tabela 1: Resumo por câmera (média ± desvio-padrão entre sessões).

Câmera	Sessões	Imgs.	Acur.	Prec.	Recall	F1	FMR	FNMR	EER
INTEGRADA (C1)*	12	616	59.9% ± 38.3%	75.0% ± 43.3%	53.3% ± 40.2%	58.7% ± 41.8%	0.3% ± 1.1%	39.8% ± 37.5%	27.57%
RAZER (C2)	6	311	69.2% ± 34.5%	83.3% ± 37.3%	71.9% ± 35.1%	76.5% ± 35.4%	0.0% ± 0.0%	28.2% ± 35.1%	14.10%
CONV. 4K (C3)	6	250	57.6% ± 29.3%	99.1% ± 1.8%	58.9% ± 29.1%	68.5% ± 30.0%	10.0% ± 20.0%	41.1% ± 29.1%	25.54%
CONV. HD (C4)	6	250	43.2% ± 38.0%	80.0% ± 40.0%	43.6% ± 38.0%	50.1% ± 40.7%	0.0% ± 0.0%	36.4% ± 36.1%	18.20%
CANON (C5)**	**2**	**58**	**99.0% ± 1.0%**	**100.0% ± 0.0%**	**100.0% ± 0.0%**	**100.0% ± 0.0%**	**0.0% ± 0.0%**	**0.0% ± 0.0%**	**0.00%**
IP (C6)	6	306	40.3% ± 32.7%	80.0% ± 40.0%	41.7% ± 34.5%	49.8% ± 38.4%	0.0% ± 0.0%	58.3% ± 34.5%	29.12%
Total	**38**	**1791**	-	-	-	-	-	-	-

* Câmera escolhida como padrão, com maior quantidade de testes.
** Câmera abandonada após múltiplos erros. Dados incluídos para análise e comparação. A primeira sessão teve 51 capturas, enquanto a segunda obteve-se 7.

exceção, piorando sob luz alta, provavelmente por comportamento específico do *pipeline* de exposição e compressão do dispositivo. Os ensaios em visão noturna (0 lux, modo infravermelho) não resultaram em reconhecimentos válidos, o que é coerente com a mudança de domínio (RGB para NIR) e com a ausência de treinamento do modelo nessa faixa espectral.

5.4.3 Movimentação. A Tabela 4 resume o desempenho em cenário de movimentação, com altura, distância e iluminação variáveis e indivíduo caminhando no campo de visão. Nessa condição mais próxima do uso real, a C2 e a C3 mantiveram acurácia acima de 80% e F1 superior a 89%, enquanto a C1, a C4 e a C6 apresentaram resultados mais modestos. Ainda assim, em todas as câmeras a precisão permaneceu em 100%, indicando que, quando o sistema decide positivamente, tende a não aceitar impostores, embora rejeite mais vezes usuários legítimos em movimento.

Tabela 4: Comparação entre dispositivos em cenário de movimentação (colunas selecionadas).

Disp.	Acur.	Prec.	Rec.	F1	FMR	FNMR	EER	FDR	Score
C1	35.3%	100.0%	37.5%	54.5%	0.0%	62.5%	31.2%	5.9%	50.0%
C2	**84.6%**	**100.0%**	**91.7%**	**95.7%**	**0.0%**	**8.3%**	**4.2%**	**7.7%**	**78.0%**
C3	80.0%	100.0%	81.6%	89.9%	0.0%	18.4%	9.2%	2.0%	78.0%
C4	48.0%	100.0%	50.0%	66.7%	0.0%	50.0%	25.0%	4.0%	56.0%
C5*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*
C6	66.0%	100.0%	66.0%	79.9%	0.0%	34.0%	17.0%	0.0%	83.0%

* Câmera Canon, cujos testes não foram realizados para este cenário por problemas técnicos.

5.4.4 Oclusão, acessórios, fraude e pré-processamento. Os ensaios de oclusão parcial e de uso de acessórios mostram um padrão claro: quando a região das sobrancelhas é encoberta por mãos ou pela aba de bonés, a acurácia, o *recall* e o F1 caem de forma significativa (até cerca de 30%), enquanto a precisão e o FMR se mantêm estáveis. Em oclusões leves (dedos no queixo, na bochecha ou na testa e pequenos toques no nariz) o algoritmo tolerou razoavelmente bem as variações, com menor impacto nas métricas.

Esses achados reforçam a importância da região periocular (olhos e sobrancelhas) para detecção e alinhamento, sugerindo que políticas de uso de EPIs ou acessórios devem considerar esse efeito em aplicações práticas.

Nos testes de fraude, em que o autor tentou se passar por outros usuários cadastrados, a acurácia geral manteve-se alta (96%), com FMR de apenas 4%. O sistema tendeu a rejeitar sistematicamente tais tentativas, o que se refletiu em *recall* e F1 nulos no cenário de fraude (ausência de acertos positivos esperados para o alvo genuíno, dado o arranjo experimental). Em termos de segurança, o comportamento é desejável: o algoritmo raramente aceita um impostor, ainda que isso implique em maior cautela nas decisões positivas.

A Tabela 5 compara sessões com e sem pré-processamento (equalização adaptativa, filtragem bilateral, *sharpening* e normalização de iluminação). Com filtros ativos, a acurácia subiu de 72% para 98%, o *recall* de 72% para 98% e o F1 de 83,7% para 99%, enquanto o FNMR despencou de 28% para 2%. O FMR permaneceu em 0% em ambos os casos.

Tabela 5: C1 (integrada) — comparação com e sem pré-processamento (filtros).

Pré-proc.	Acur.	Prec.	Recall	F1	FMR	FNMR	EER	FDR	Score
Sim	98,0%	100,0%	98,0%	99,0%	0%	2,0%	1,0%	0%	0,9
Não	72,0%	100,0%	72,0%	83,7%	0%	28,0%	14,0%	0%	0,7

Mesmo considerando diferenças de ambiente entre as sessões, o ganho observado no cenário mais desafiador (iluminância menor) indica que o pré-processamento contribuiu decisivamente para estabilizar a detecção, o alinhamento e a qualidade dos *embeddings*, especialmente quando se recorre a zoom digital.

5.5 Tempos de processamento

A Tabela 6 consolida os tempos médios por câmera. Entre as webcams USB (C1, C3 e C4), TMD e TMR foram baixos e semelhantes, levando a TMT em torno de 2,3 ms, o que evidencia que o *pipeline* é suficientemente rápido para aplicações em tempo quase real. A C2 apresentou TMT um pouco maior (≈3,5 ms), possivelmente devido a formatos de captura mais custosos (por exemplo, MJPEG/H.264) que exigem conversão adicional.

Tabela 6: Tempos de processamento médio por câmera (ms).

Câm.	TMD	TMR	TMT
C1	*0,2 ± 0,24*	*0,44 ± 0,52*	*2,43 ± 0,87*
C2	*0,16 ± 0,04*	*0,36 ± 0,08*	*3,53 ± 1,18*
C3	*0,1 ± 0,01*	*0,23 ± 0,03*	*2,26 ± 0,03*
C4	*0,12 ± 0,03*	*0,27 ± 0,06*	*2,31 ± 0,10*
C5	*5,54 ± 3,61*	*12,92 ± 8,44*	*20,56 ± 12,06*
C6	***0,16 ± 0,03***	***0,34 ± 0,09***	***1,01 ± 0,30***

A câmera IP (C6) obteve o menor TMT agregado (≈ 1 ms), sugerindo captura eficiente, mas seu TMR não foi o mais baixo, provavelmente por sobrecarga de decodificação e *buffering* do fluxo de rede. Já a Canon (C5) apresentou TMT muito superior às demais,

coerente com a necessidade de software proprietário intermediário e maior propensão a travamentos, o que aumenta a variância e compromete a experiência.

Em termos práticos, os resultados indicam que a escolha do dispositivo deve considerar não apenas a ótica, mas também o caminho de captura (UVC direto, IP, software intermediário), que pode introduzir *overheads* relevantes e instabilidades difíceis de mitigar.

5.6 Spoofing e cenário multiusuário

Nos testes de *spoofing*, com fotos exibidas em celular e imagens impressas, a integração de um módulo de *anti-spoofing* baseado em detecção de objetos (celular/foto) e de mãos com modelos YOLO mostrou-se promissora, reduzindo reconhecimentos indevidos ao rejeitar quadros em que um objeto plano era segurado à frente da câmera (Figura 3). A heurística de rejeitar decisões quando a face ocupava mais de 40% da área do quadro mitigou ataques por aproximação extrema, mas trouxe custo de usabilidade e aumento de latência, além de não eliminar totalmente os falsos aceites. Por limitações de cronograma e complexidade de integração, esse módulo foi desativado na versão final, sendo apontado como trabalho futuro.

Figura 3: Captura com anti-*spoofing*, rotulando objeto restringido.

Em cenário multiusuário, o sistema foi capaz de detectar e reconhecer simultaneamente múltiplas faces em uma cena (Figura 4). Entretanto, não foram conduzidos testes quantitativos em larga escala nesse contexto; assim, os resultados devem ser interpretados como demonstração de capacidade técnica, e não como avaliação estatisticamente robusta.

6 Considerações Finais

O trabalho cumpriu o objetivo de investigar e comparar um conjunto de técnicas computacionais para reconhecimento facial para controle de acesso em ambiente fechado, integrando revisão sistemática e validação prática em laboratório através de um sistema mediador. A investigação demonstrou viabilidade técnica para autenticação em tempo próximo a 1 segundo, com registro automatizado de acessos e operação com conexão *wireless*.

Figura 4: Captura com múltiplos usuários em cena.

Os resultados mostraram que o desempenho depende fortemente das condições de captura. A acurácia foi elevada em distâncias curtas (100 cm), mas caiu com o aumento da distância e em iluminação insuficiente. Níveis intermediários e altos de iluminância melhoraram substancialmente a detecção e o alinhamento. Oclusões e acessórios também prejudicaram o reconhecimento, ainda que a precisão e a taxa de falsos positivos tenham permanecido próximas de 100% e 0%, respectivamente, indicando boa especificidade do *pipeline*.

O pré-processamento proposto destacou-se como contribuição prática, elevando a acurácia em cenários adversos (de 72% para 98%) e reduzindo o erro. Ensaios de ataque por apresentação mostraram tendência de rejeição de impostores, embora com vulnerabilidades residuais, justificando a desativação do módulo de *anti-spoofing* e sua indicação como trabalho futuro. Os tempos médios de processamento ficaram na ordem de milissegundos, atendendo ao requisito de tempo de resposta.

Considera-se que o objetivo geral foi atingido. A hipótese de viabilidade técnica foi confirmada, enquanto a hipótese de acurácia acima de 90% foi atendida apenas em condições favoráveis de captura. As limitações do conjunto de testes incluem tamanho amostral reduzido, diversidade demográfica reduzida, variações instrumentais de iluminação e ausência de avaliação quantitativa de múltiplos usuários simultaneamente.

Como perspectivas, recomenda-se ampliar os cenários e perfis avaliados, incorporar técnicas de *anti-spoofing* e explorar sensores de profundidade (RGB-D ou câmeras NIR), amplamente utilizados em sistemas comerciais de controle de acesso. Em síntese, os resultados confirmam que o *pipeline* baseado em *embeddings* de FaceNet é consistente e pode alcançar alto desempenho quando a captura preserva qualidade, fornecendo base sólida para evoluções futuras rumo a sistemas mais robustos e escaláveis.

Referências

[1] M. K. Alomari et al. Systematic analysis of artificial intelligence-based platforms for identifying governance and access control. *Security and Communication Networks*, 2021:1–10, 2021. doi: 10.1155/2021/8686469. URL https://doi.org/10.1155/2021/8686469.

[2] H. S. Lee et al. Efficient defect identification via oxide memristive crossbar array based morphological image processing. *Advanced Intelligent Systems*, 3(2), 2021. doi: 10.1002/aisy.202000202. URL https://doi.org/10.1002/aisy.202000202. Article ID 2000202.

[3] S. I. Shafiq et al. Designing intelligent factory: Conceptual framework and empirical validation. *Procedia Computer Science*, 96:1801–1808, 2016. doi: 10.1016/j.procs.2016.09.351. URL https://doi.org/10.1016/j.procs.2016.09.351.

[4] A. S. Rafika et al. Face Recognition based Artificial Intelligence With AttendX Technology for Student Attendance. In *2022 International Conference on Science and Technology (ICOSTECH)*, 2 2022. doi: 10.1109/icostech54296.2022.9829122. URL https://doi.org/10.1109/icostech54296.2022.9829122.

[5] M. Kumar and T. Hussaini. Face Recognition Algorithm based on Traditional and Artificial Intelligence: A Systematic Review. *2021 International Conference on Intelligent Technologies (CONIT)*, pages 1–5, 6 2021. doi: 10.1109/conit51480.2021.9498476. URL https://doi.org/10.1109/conit51480.2021.9498476.

[6] E. N. Korte et al. Reconhecimento facial para controle de acesso de estudantes ao instituto federal do paraná – campus cascavel. In *Anais do 20º Congresso Latino-Americano de Software Livre e Tecnologias Abertas (Latinoware 2023)*, pages 162–165, Foz do Iguaçu, PR, 2023. Sociedade Brasileira de Computação. doi: 10.5753/latinoware.2023.236527. URL https://doi.org/10.5753/latinoware.2023.236527.

[7] M. Mortensen. Sneaking ai through the back door: Constructing the identity of capitol hill rioters through social media images and facial recognition technologies. *Information, Communication & Society*, 28(2):278–294, 2024. doi: 10.1080/1369118X.2024.2358164. URL https://doi.org/10.1080/1369118X.2024.2358164.

[8] N. Nawaz. Artificial intelligence applications for face recognition in recruitment process. SSRN Working Paper #3883916, jul 2021. URL https://ssrn.com/abstract=3883916. Posted 2 ago. 2021.

[9] N. Vavoula. Artificial intelligence (AI) at Schengen Borders: automated processing, algorithmic profiling and facial recognition in the era of Techno-Solutionism. *European Journal of Migration and Law*, 23(4):457–484, 12 2021. doi: 10.1163/15718166-12340114. URL https://doi.org/10.1163/15718166-12340114.

[10] J. Srikanth et al. Artificial Intelligence based Multi-Face Recognition and Attendance Marking System. In *2024 5th International Conference on Image Processing and Capsule Networks (ICIPCN)*, volume 45, pages 47–51. 2024 5th International Conference on Image Processing and Capsule Networks (ICIPCN), 7 2024. doi: 10.1109/icipcn63822.2024.00017. URL https://doi.org/10.1109/icipcn63822.2024.00017.

[11] G. A. Senthil, S. Geerthik, R. Karthikeyan, and G. Keerthana. Face Recognition based Automated Smart Attendance using Hybrid Machine Learning Algorithms and Computer Vision. *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pages 606–611, 6 2024. doi: 10.1109/icaaic60222.2024.10574896. URL https://doi.org/10.1109/icaaic60222.2024.10574896.

[12] J. de S. Santos. Reconhecimento facial para a identificação de pessoas em ambientes restritos: um estudo avaliativo, 2023. URL https://repositorio.ufrn.br/handle/123456789/54046.

[13] S. A. da Fonseca et al. Análise das políticas públicas na prevenção dos massacres nas escolas à luz do direito. *Revista Ibero-Americana de Humanidades, Ciências e Educação*, 9(9):3530–3547, out 2023. doi: 10.51891/rease.v9i9.11318. URL https://periodicorease.pro.br/rease/article/view/11318.

[14] T. da S. Seixas et al. A efetividade das medidas socioeducativas diante dos massacres em escolas no brasil. *Revista Multidisciplinar do Nordeste Mineiro*, 8(1):1–13, ago 2023. doi: 10.61164/rmnm.v8i1.1500. URL https://revista.unipacto.com.br/index.php/multidisciplinar/article/view/1500.

[15] J. Al-Nabulsi et al. IoT solutions and AI-Based frameworks for Masked-Face and face recognition to fight the COVID-19 pandemic. *Sensors*, 23(16):7193, 8 2023. doi: 10.3390/s23167193. URL https://doi.org/10.3390/s23167193.

[16] T. V. Aravinda et al. Implementation of facial recognition (AI) and its impact on the service sector. In *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pages 74–80, 5 2022. doi: 10.1109/icaaic53929.2022.9793161. URL https://doi.org/10.1109/icaaic53929.2022.9793161.

[17] A. Krishnamurthi et al. Real time face recognition system based on yolo and insightface. *Multimedia Tools and Applications*, 83(11):31893–31910, set 2023. doi: 10.1007/s11042-023-16831-7. URL https://doi.org/10.1007/s11042-023-16831-7.

[18] M. Klaus Scheuerman, Alex Hanna, and Emily Denton. Do Datasets have politics? Disciplinary values in Computer Vision Dataset development. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–37, 10 2021. doi: 10.1145/3476058. URL https://doi.org/10.1145/3476058.

[19] T Huang. Computer Vision: Evolution And Promise. *1996 CERN School of Computing*, 1 1996. doi: 10.5170/cern-1996-008.21. URL https://cds.cern.ch/record/400313/files/p21.pdf.

[20] Paris Amerikanos and Ilias Maglogiannis. Image analysis in digital pathology utilizing machine learning and deep neural networks. *Journal of Personalized Medicine*, 12(9):1444, 9 2022. doi: 10.3390/jpm12091444. URL https://doi.org/10.3390/jpm12091444.

[21] M. Smith and S. Miller. The ethical application of biometric facial recognition technology. *AI Society*, 37(1):167–175, 4 2021. doi: 10.1007/s00146-021-01199-9. URL https://doi.org/10.1007/s00146-021-01199-9.

[22] X. Wang and Y. Li. Facial recognition system based on genetic algorithm improved ROI-KNN Convolutional Neural Network. *Applied Bionics and Biomechanics*, 2022:1–11, 10 2022. doi: 10.1155/2022/7976856. URL https://doi.org/10.1155/2022/7976856.

[23] P. Peng et al. A face recognition software framework based on principal component analysis. *PLoS ONE*, 16(7):e0254965, 7 2021. doi: 10.1371/journal.pone.0254965. URL https://doi.org/10.1371/journal.pone.0254965.

[24] W. A. Mahmoud, A. I. Abbas, and N. A. S. Alwan. FACE IDENTIFICATION USING BACK-PROPAGATION ADAPTIVE MULTIWAVENET. *Journal of Engineering*, 18(03):392–402, 7 2023. doi: 10.31026/j.eng.2012.03.12. URL https://doi.org/10.31026/j.eng.2012.03.12.

[25] M. Dirik. Application of machine learning techniques for obesity prediction: a comparative study. *Journal of Complexity in Health Sciences*, 6(2):16–34, 10 2023. doi: 10.21595/chs.2023.23193. URL https://doi.org/10.21595/chs.2023.23193.

[26] P. Razmi et al. Topology identification in distribution system via machine learning algorithms. *PLoS ONE*, 16(6):e0252436, 6 2021. doi: 10.1371/journal.pone.0252436. URL https://doi.org/10.1371/journal.pone.0252436.

[27] A. Barkhordari et al. Machine learning approach for predicting electrical features of Schottky structures with graphene and ZnTiO3 nanostructures doped in PVP interfacial layer. *Scientific Reports*, 13(1), 8 2023. doi: 10.1038/s41598-023-41000-z. URL https://doi.org/10.1038/s41598-023-41000-z.

[28] C. Szegedy et al. Rethinking the inception architecture for computer vision. *arXiv (Cornell University)*, 1 2015. doi: 10.48550/arxiv.1512.00567. URL https://arxiv.org/abs/1512.00567.

[29] David Opeoluwa Oyewola, Emmanuel Gbenga Dada, Sanjay Misra, and Robertas Damaševičius. Detecting cassava mosaic disease using a deep residual convolutional neural network with distinct block processing. *PeerJ Computer Science*, 7:e352, 3 2021. doi: 10.7717/peerj-cs.352. URL https://doi.org/10.7717/peerj-cs.352.

[30] A. Krizhevsky, I Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 5 2017. doi: 10.1145/3065386. URL https://doi.org/10.1145/3065386.

[31] Z. Li et al. A Survey of Convolutional Neural Networks: Analysis, applications, and Prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12):6999–7019, 6 2021. doi: 10.1109/tnnls.2021.3084827. URL https://doi.org/10.1109/tnnls.2021.3084827.

[32] Dinesh Acharya, Zhiwu Huang, Danda Paudel, and Van Gool Luc. Covariance pooling for facial expression recognition. *arXiv (Cornell University)*, 1 2018. doi: 10.48550/arxiv.1805.04855. URL https://arxiv.org/abs/1805.04855.

[33] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, 8 2016. doi: 10.1109/lsp.2016.2603342. URL https://doi.org/10.1109/lsp.2016.2603342.

[34] Kostiantyn Khabarlak and Larysa Koriashkina. Fast Facial Landmark Detection and Applications: A survey. *Journal of Computer Science and Technology*, 22(1):e02, 4 2022. doi: 10.24215/16666038.22.e02. URL https://doi.org/10.24215/16666038.22.e02.

[35] J. A. Mensah et al. FaceNet recognition algorithm subject to multiple constraints: Assessment of the performance. *Scientific African*, 23:e02007, 12 2023. doi: 10.1016/j.sciaf.2023.e02007. URL https://doi.org/10.1016/j.sciaf.2023.e02007.

[36] R. Sarki et al. Image preprocessing in classification and identification of diabetic eye diseases. *Data Science and Engineering*, 6(4):455–471, 8 2021. doi: 10.1007/s41019-021-00167-z. URL https://doi.org/10.1007/s41019-021-00167-z.

[37] M. Mohapatra et al. Botanical leaf disease detection and classification using convolutional neural network: a hybrid metaheuristic enabled approach. *Computers*, 11(5):82, 5 2022. doi: 10.3390/computers11050082. URL https://doi.org/10.3390/computers11050082.

[38] M. Elavarasu and K. Govindaraju. Effectiveness of filtering methods in enhancing pulmonary carcinoma image quality: A comparative analysis. *International Journal of Power Electronics and Drive Systems/International Journal of Electrical and Computer Engineering*, 14(1):358, 11 2023. doi: 10.11591/ijece.v14i1.pp358-365. URL https://doi.org/10.11591/ijece.v14i1.pp358-365.