

APRENDIZADO DE MÁQUINA PARA O RECONHECIMENTO FACIAL APLICADO A FECHADURAS INTELIGENTES

Victor de Rose Trunfo¹, Leandro Neckel², Merisandra Côrtes de Mattos³

Resumo: Soluções de segurança para proteção de bens, como fechaduras convencionais, inteligentes e eletrônicas estão disponíveis no mercado. Algumas fazem uso de identificação biométrica por meio de impressão digital, no entanto, o reconhecimento facial, utilizando técnicas de visão computacional por meio de *deep learning*, surge como uma alternativa promissora. O objetivo desta pesquisa consiste em analisar algoritmos de reconhecimento facial aplicados a problemática de fechaduras inteligentes, empregando-se redes neurais convolucionais e métricas de avaliação para mensurar o desempenho dos modelos gerados. A abordagem da pesquisa quanto à natureza, é aplicada e de base tecnológica, empregando algoritmos de aprendizado de máquina para o reconhecimento facial em fechaduras inteligentes. No que se refere aos objetivos é uma pesquisa descritiva, pois avalia por meio de métricas de qualidade em redes neurais artificiais o desempenho dos modelos gerados pelos métodos e arquiteturas empregadas. Os resultados evidenciaram a promissora eficácia das arquiteturas *DenseNet-121* e *DenseNet-169*, aliadas ao método de reconhecimento facial *Fisherfaces*, em termos de precisão, acurácia e F1, tanto em análises de pequena escala quanto de larga escala.

Palavras-chave: Visão computacional. *Deep learning*. Redes neurais convolucionais. *Eigenfaces*. *Fisherfaces*.

¹ Curso de Ciência da Computação, Grupo de Pesquisa em Inteligência Artificial Aplicada, Universidade do Extremo Sul Catarinense (UNESC), Criciúma - SC - Brasil. v.trunfo@hotmail.com.

² Coorientador, Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense (UNESC), Criciúma - SC - Brasil. neckel.leandro@unesc.net.

³ Orientadora, Curso de Ciência da Computação, Grupo de Pesquisa em Inteligência Artificial Aplicada, Universidade do Extremo Sul Catarinense (UNESC), Criciúma - SC - Brasil. mem@unesc.net.

ABSTRACT: *Security solutions for asset protection, such as conventional, smart and electronic locks are available in the market. Some make use of biometric identification through fingerprint. However, facial recognition, using computer vision techniques through deep learning, emerges as a promising alternative. The objective of this research is to analyze facial recognition algorithms applied to the problem of intelligent locks, using convolutional neural networks and evaluation metrics to measure the performance of the models generated. The research approach is applied and technology-based, employing machine learning algorithms for face recognition in smart locks. In terms of objectives, this is a descriptive research study, since it evaluates the performance of the models generated by the methods and architectures employed, using quality metrics in artificial neural networks. The results highlighted the promising performance of the DenseNet-121 and DenseNet-169 architectures, combined with the Fisherfaces facial recognition method, in terms of precision, accuracy, and F1 score, both in small-scale and large-scale analyses.*

Keywords: *Computer vision. Deep learning. Convolutional neural networks. Eigenfaces. Fisherfaces.*

1 INTRODUÇÃO

A segurança é uma preocupação constante em diversos setores, incluindo residências, estabelecimentos comerciais e instituições públicas. A proteção de bens e a garantia de acesso apenas as pessoas autorizadas são questões essenciais para promover a tranquilidade e prevenir incidentes indesejados. Nesse contexto, as fechaduras convencionais têm sido amplamente utilizadas, porém apresentam limitações que podem comprometer sua eficácia.

As fechaduras com chaves, por exemplo, são uma das soluções mais comuns. Entretanto, a possibilidade de duplicação das chaves e o risco de perda ou furto delas representam uma ameaça à segurança. Além disso, o processo de destrancar e trancar manualmente as portas pode ser inconveniente e limitante, especialmente em ambientes compartilhados em que várias pessoas precisam ter acesso.

Diante dessas limitações, surgiram alternativas como as fechaduras eletrônicas, que empregam tecnologias como por exemplo, de senhas ou de identificação por radiofrequência. No entanto, mesmo essas soluções não estão isentas de falhas. Senhas podem ser descobertas por pessoas mal-intencionadas, e as *tags* RFID podem ser facilmente duplicadas ou furtadas (AZNAR; PRADA, 2020). Portanto, torna-se necessário buscar métodos mais avançados e seguros para o controle de acesso.

Uma abordagem promissora é o uso de reconhecimento facial, uma técnica de biometria que permite identificar pessoas com base em características únicas de seus rostos (IGNACIO, 2021). A visão computacional, a inteligência artificial e o aprendizado de máquina desempenham um papel fundamental no desenvolvimento de algoritmos e sistemas de reconhecimento facial eficientes.

O reconhecimento facial oferece diversas vantagens em comparação com outras soluções biométricas. Ao capturar as medidas biométricas de uma pessoa a partir de uma distância específica, sem a necessidade de interação física, ele garante maior conveniência e rapidez no acesso (IGNACIO, 2021). Além disso, o reconhecimento facial pode identificar pessoas com histórico criminal ou problemas legais, fornecendo um nível adicional de segurança.

Com o avanço das tecnologias, como a inteligência artificial e o aprendizado de máquina, os computadores além de executarem tarefas repetitivas, também adquirem capacidades de aprendizado (LUDERMIR, 2021). O aprendizado profundo, também conhecido como deep learning, tem se destacado no reconhecimento facial, alcançando resultados cada vez mais próximos da capacidade humana (CHOI et al., 2020).

Nesse contexto, a presente pesquisa tem como objetivo analisar, por meio de métricas de avaliação, os modelos gerados a partir de algoritmos de aprendizado de máquina para o reconhecimento facial aplicados em fechaduras inteligentes. Para isso, são empregados nesta pesquisa dois métodos, o *Eigenfaces* e o *Fisherfaces*, e exploradas cinco arquiteturas que incluem *ResNet-50*, *VGG16*, *VGG19*, *DenseNet-121* e *DenseNet-169*. Assim, são aplicadas técnicas de visão computacional, redes neurais convolucionais e métricas de avaliação para mensurar o desempenho dos modelos de reconhecimento facial gerados.

A relevância desse estudo reside na crescente aplicação de tecnologias como inteligência artificial, aprendizado de máquina e reconhecimento facial, que

podem aprimorar a segurança e facilitar a tomada de decisões em diversos setores. A utilização de fechaduras inteligentes com reconhecimento facial apresenta benefícios como eficiência, segurança e acessibilidade, contribuindo para uma experiência mais fluida e segura para os usuários (IGNACIO, 2021). Além disso, soluções como essa podem ser aplicadas em diferentes cenários, como residências, empresas e ambientes escolares, nos quais o controle de acesso é essencial. Portanto, compreender e avaliar o desempenho dos modelos de reconhecimento facial é fundamental para o desenvolvimento e aprimoramento dessas tecnologias.

2 TRABALHOS CORRELATOS

Esta seção apresenta uma revisão dos trabalhos correlatos que abordam o tema do reconhecimento facial aplicado em fechaduras inteligentes. Dessa forma, busca-se identificar as contribuições existentes na literatura científica, as abordagens adotadas, as limitações encontradas e as lacunas de pesquisa ainda não exploradas. Esses trabalhos servem como base para o desenvolvimento deste estudo, fornecendo *insights* valiosos e direcionamentos para a pesquisa.

Um dos primeiros trabalhos relevantes no campo do reconhecimento facial foi o *Eigenfaces*, proposto por Turk e Pentland em 1991. Esse método utiliza a técnica de Análise de Componentes Principais (PCA) para extrair as principais características faciais de um conjunto de imagens de treinamento. O reconhecimento é realizado por meio da comparação dessas características com as de uma imagem de teste. O *Eigenfaces* obteve resultados promissores, porém, apresentou limitações em ambientes com variações de iluminação e pose.

Belhumeur, Hespanha e Kriegman (1997) propuseram o *Fisherfaces* com o objetivo de superar as limitações do *Eigenfaces*. Esse método utiliza a técnica de Análise Discriminante Linear (LDA) para extrair características discriminantes que maximizem a separação entre as classes. O *Fisherfaces* mostrou-se mais robusto a variações de iluminação e pose, proporcionando uma melhoria significativa no desempenho do reconhecimento facial.

Diversos estudos têm explorado diferentes abordagens de reconhecimento facial aplicadas em fechaduras inteligentes. Wang *et al.* (2016) propuseram um sistema de reconhecimento facial baseado em *deep learning* utilizando a arquitetura de rede neural convolucional *AlexNet*. O sistema foi treinado

em um grande conjunto de imagens faciais para reconhecer indivíduos autorizados. Os resultados mostraram alta taxa de acerto no reconhecimento facial e demonstraram a eficácia do uso de redes neurais convolucionais nesse contexto.

Li *et al.* (2018) exploraram o uso de redes neurais convolucionais para o reconhecimento facial em fechaduras inteligentes. Os autores utilizaram as arquiteturas *VGG16* e *ResNet-50*, treinadas em um conjunto de dados extenso. Foram realizadas análises comparativas entre as arquiteturas e avaliados parâmetros como taxa de acerto, tempo de processamento e robustez a variações de pose e iluminação. Os resultados mostraram que ambas as arquiteturas apresentaram bom desempenho, sendo a *ResNet-50* ligeiramente superior em termos de acurácia.

Lin e Wu (2020) empregaram o reconhecimento facial para o desbloqueio do sistema de portas dos veículos, para isso utilizaram análise de componente principal na condução dos experimentos e o método *Eigenfaces*. No treinamento do modelo, como usaram praticamente a totalidade de pessoas de origem asiática, o modelo se torna impreciso se utilizado no reconhecimento de pessoas ocidentais. Concluem que a abertura de uma porta por meio de reconhecimento fácil é mais eficiente e prática.

Apesar dos avanços significativos nessa área, ainda existem desafios a serem superados. Alguns desses desafios incluem a necessidade de robustez a variações de iluminação, pose e oclusões, além de lidar com problemas de privacidade e segurança relacionados ao armazenamento e uso das informações biométricas dos indivíduos.

Diante desse panorama, a presente pesquisa visa contribuir para o avanço do campo do reconhecimento facial aplicado em fechaduras inteligentes, explorando diferentes arquiteturas de redes neurais convolucionais e analisando seu desempenho em termos de acurácia, eficiência e robustez.

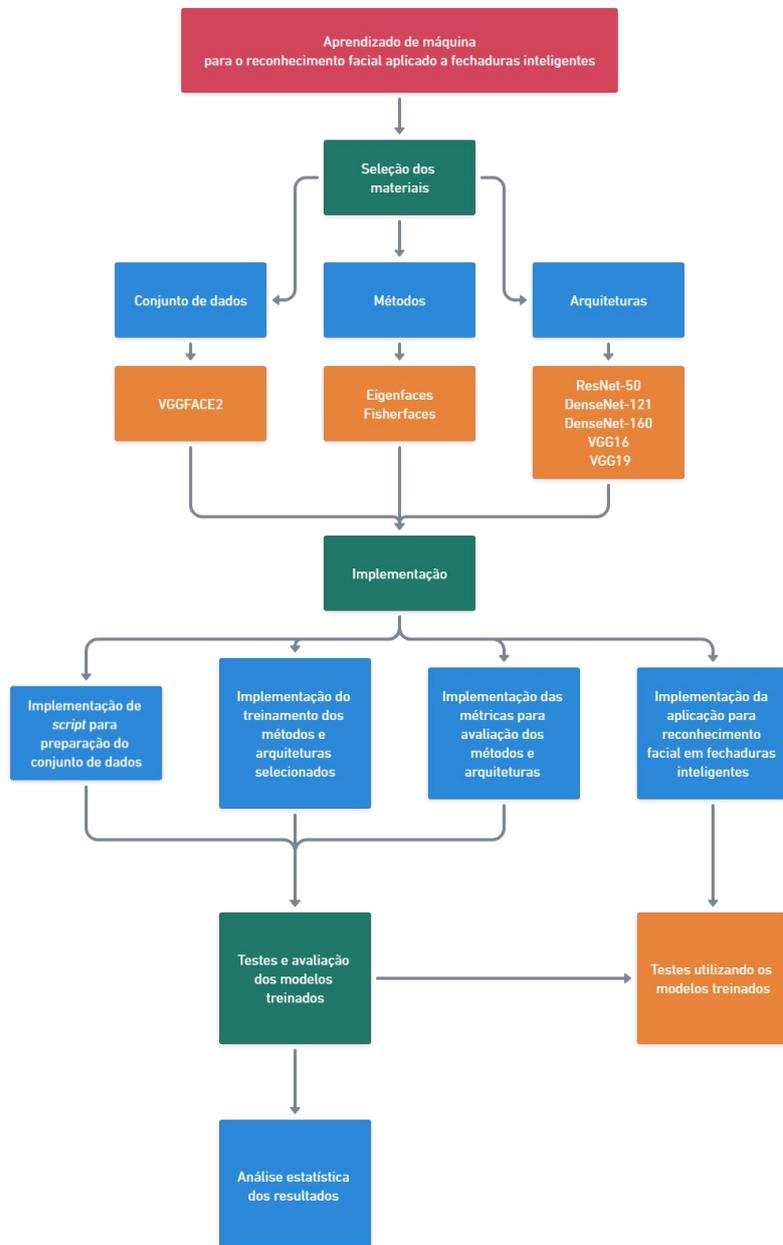
3 MATERIAIS E MÉTODOS

A abordagem da pesquisa quanto à natureza, é aplicada e de base tecnológica, empregando algoritmos de aprendizado de máquina para o reconhecimento facial em fechaduras inteligentes. No que se refere aos objetivos é uma pesquisa descritiva, pois avalia por meio de métricas de qualidade em redes

neurais artificiais o desempenho dos modelos gerados pelos métodos e arquiteturas empregadas. Quanto aos procedimentos a pesquisa é bibliográfica e experimental, compreendendo a manipulação de diferentes métodos, parâmetros e arquiteturas de redes neurais.

Na Figura 1, tem-se a estrutura geral desta pesquisa que compreendeu as etapas de seleção e preparação do conjunto de dados empregados, implementação dos métodos e arquiteturas aplicadas ao reconhecimento facial, teste e avaliação dos modelos, desenvolvimento do aplicativo de reconhecimento facial para fechaduras inteligentes.

Figura 1 - Estrutura geral da pesquisa



Fonte: Do Autor

3.1 SELEÇÃO E PREPARAÇÃO DO CONJUNTO DE DADOS

O conjunto de dados *VGGFace2* foi empregado como base de treinamento, validação e teste para os modelos de reconhecimento facial gerados. O *VGGFace2* consiste em um grande número de imagens faciais coletadas de indivíduos de diferentes origens étnicas, idades e gêneros. A seleção deste conjunto de dados deu-se em função da sua representatividade, diversidade e ampla variedade de imagens de pessoas reais, incluindo cantores, atores, políticos, entre outros. O conjunto de dados *VGGFace2*, bem como as fontes de informação, detalhes das imagens e termos de uso, está disponível publicamente e pode ser acessado por meio do *GitHub* do *VGGFace2*⁴ e do website oficial⁵.

Antes de utilizar o conjunto de dados, verificou-se a disponibilidade de diferentes versões e optou-se pelo uso de uma versão com *upscale* para garantir maior fidelidade no treinamento. A preparação do conjunto de dados incluiu o redimensionamento, a conversão para escala de cinza e outras técnicas de preparação dos dados.

3.1.1 REDIMENSIONAMENTO

O redimensionamento é realizado para ajustar todas as imagens a um tamanho fixo de 128x128 pixels usando a função `cv2.resize` da biblioteca `opencv2`.

Antes da preparação, as imagens podem ter tamanhos diferentes e variar em termos de resolução. O redimensionamento garante que todas as imagens tenham a mesma dimensão, o que é necessário para alimentar os dados em um modelo de aprendizado de máquina.

3.1.2 CONVERSÃO PARA ESCALA DE CINZA

Essa técnica envolve a conversão das imagens coloridas para tons de cinza. Isso é feito para reduzir a dimensionalidade dos dados, já que as imagens em

⁴ Disponível em: https://github.com/ox-vgg/vgg_face2

⁵ Disponível em: https://www.robots.ox.ac.uk/~vgg/data/vgg_face2

escala de cinza têm apenas um canal de cor, em oposição aos três canais (vermelho, verde e azul) das imagens coloridas. A conversão para escala de cinza simplifica o processamento das imagens e pode ajudar a reduzir a quantidade de dados necessários para treinar o modelo.

3.1.3 AUGMENTAÇÃO DE DADOS

A augmentação de dados é uma técnica utilizada para aumentar a quantidade de dados de treinamento, gerando novas amostras com base nas amostras existentes. Isso é feito aplicando transformações geométricas ou modificando as propriedades das imagens, como rotação, espelhamento, dimensão, alteração de brilho/contraste e adição de ruído. De acordo com o estudo realizado por Kakadiaris *et al.* (2007), a augmentação de dados é uma estratégia eficaz para diversificar o conjunto de treinamento, melhorando assim a capacidade do modelo de reconhecimento facial em generalizar e lidar com diferentes variações e condições de entrada.

3.2 APIs DE RECONHECIMENTO FACIAL

Realizou-se uma pesquisa de *Application Programming Interface* (API) disponíveis em busca de algumas soluções que empregavam os algoritmos *Eigenfaces* e *Fisherfaces*. Essas APIs foram testadas em bases de demonstração, a fim de analisar o seu funcionamento e a sua precisão no reconhecimento facial. Após estes testes, decidiu-se pela implementação dos próprios algoritmos de reconhecimento facial baseados em *Eigenfaces* e *Fisherfaces*, pois proporcionaria uma melhor compreensão do seu funcionamento. Para isso, realizou-se uma revisão das literaturas disponíveis e o ajuste dos parâmetros dos algoritmos de acordo com requisitos da pesquisa.

A linguagem de programação empregada para a codificação dos algoritmos *Eigenfaces* e *Fisherfaces* foi o *Python* em sua versão 3.10.9, utilizando o ambiente de desenvolvimento *JupyterLab* em sua versão 3.6.2. A escolha dessa linguagem de programação se deve à sua ampla utilização na área de aprendizado de máquina e de reconhecimento e padrões.

Na configuração do ambiente de desenvolvimento, estabeleceu-se alguns parâmetros como *seed* para o valor 42 que fora selecionado de forma aleatória, o que permite replicar os resultados em experimentos futuros. Também utilizou-se a capacidade de processamento da placa de vídeo para acelerar o processamento das imagens, que foram importadas diretamente do computador utilizado.

3.3 MÉTODOS UTILIZADOS

Dentre as abordagens utilizadas no reconhecimento facial, dois métodos clássicos se destacam, o *Fisherfaces* e o *Eigenfaces*, os quais foram utilizados nesta pesquisa. Ambos são fundamentados em técnicas de análise de componentes principais e, segundo Belhumeur *et al.* (1997) têm sido amplamente empregados em diversas aplicações de reconhecimento facial.

3.3.1 EIGENFACES

O método *Eigenfaces* baseia-se em análise de componentes principais para o reconhecimento facial, capturando as principais componentes de variação presentes nas imagens de treinamento, conhecidas como "eigenfaces". Essas eigenfaces representam as variações mais significativas encontradas nas faces do conjunto de treinamento Pentland, A. (1991).

Durante o processo de treinamento do método *Eigenfaces*, as imagens são transformadas em um espaço de baixa dimensão, sendo representadas como combinações lineares das eigenfaces. Durante a fase de teste, a face de uma nova imagem é projetada nesse subespaço e comparada com as projeções das faces de treinamento, a fim de determinar a classe de pertencimento de acordo com o artigo produzido por Belhumeur *et al.* (1997).

Segundo Pentland, A. (1991) o modelo *Eigenfaces* oferece uma abordagem computacionalmente eficiente para o reconhecimento facial, sendo capaz de lidar com um grande número de imagens e indivíduos. No entanto, ele pode ser sensível a variações na iluminação e rotação facial, o que pode afetar sua precisão em certos cenários.

3.3.2 FISHERFACES

O método *Fisherfaces*, também conhecido como *Discriminante Linear de Fisher*, é uma técnica de redução de dimensionalidade que busca maximizar a separabilidade entre diferentes classes de imagens. O objetivo deste método é encontrar as características mais discriminantes nos subespaços, com maior variabilidade interclasses e menor variabilidade intraclasses segundo BELHUMEUR *et al.* (1997).

Durante o treinamento do modelo *Fisherfaces*, as imagens são projetadas em um espaço de características de baixa dimensão, em que as informações discriminantes são maximizadas. Essas características extraídas são utilizadas para classificar novas imagens durante a fase de teste. O *Fisherfaces* oferece uma abordagem robusta e eficiente para o reconhecimento facial, sendo particularmente adequado para problemas com um número limitado de amostras por classe como explica o artigo de Yang *et al.* (2011).

3.4 ARQUITETURAS UTILIZADAS

No âmbito deste estudo de reconhecimento facial, utilizou-se um conjunto de arquiteturas de redes neurais convolucionais para o treinamento e a avaliação dos métodos. As arquiteturas empregadas foram cuidadosamente selecionadas, levando em consideração suas características distintas e o seu desempenho em tarefas de visão computacional, conforme a literatura científica na área. Na presente pesquisa foram implementadas cinco arquiteturas, as quais foram *ResNet-50*, *VGG16*, *VGG-19*, *DenseNet-121* e *DenseNet-169*.

A *ResNet-50*, de acordo com He *et al.* (2016), é reconhecida pela sua capacidade de treinar redes neurais profundas, superando o desafio conhecido como degradação de desempenho⁶ que é comum quando o número de camadas na rede aumenta. Esta arquitetura emprega o conceito de conexões residuais, permitindo que a informação flua mais facilmente por meio das camadas e promovendo uma aprendizagem mais eficaz.

⁶ Segundo Smith (2018) a degradação de desempenho ocorre quando há uma redução na capacidade ou eficiência de um sistema em relação a um estado anterior ou a um sistema de referência. Essa degradação pode ser causada por vários fatores, como alterações nas condições ambientais, desgaste de componentes ou inadequação de algoritmos.

A arquitetura *VGG16* de acordo com Simonyan *et al.* (2014) utiliza filtros convolucionais de tamanho reduzido (3x3) em várias camadas consecutivas, seguidas por camadas de *pooling*. Essa abordagem permite a extração de características detalhadas de imagens e resulta em um modelo adequado para o reconhecimento facial.

A *VGG19* também de acordo com Simonyan *et al.* (2014) é uma variação da *VGG16* que expande a capacidade de representação da rede, adicionando camadas convolucionais e de *pooling*. Com um número maior de parâmetros, a *VGG19* é capaz de capturar características mais complexas e sutis em imagens. No entanto, é importante observar que o aumento da complexidade pode implicar em um tempo de treinamento mais longo.

A arquitetura *DenseNet-121* de acordo com Huang *et al.* (2017) possui conexões densas entre as camadas, que permitem um fluxo de informação mais direto. Cada camada recebe informações das camadas anteriores, promovendo melhor comunicação entre elas. Assim, melhora o aproveitamento de características em diferentes níveis de abstração, resultando em um modelo compacto e eficiente.

A *DenseNet-169* também de acordo com Huang *et al.* (2017) é uma extensão da arquitetura anterior que apresenta maior número de camadas, tendo a capacidade de aprender representações mais complexas e capturar nuances sutis nas imagens. No entanto, é importante considerar que a *DenseNet-169* pode exigir recursos computacionais mais intensivos devido à sua maior complexidade.

3.5 TESTES E AVALIAÇÃO DOS MODELOS

Para a realização dos experimentos e treinamentos dos modelos de reconhecimento facial, utilizamos um computador com as seguintes especificações: processador AMD Ryzen 5 5600X, 32GB de memória RAM com frequência de 3600MHz, placa de vídeo RTX 3070 Ti com 8GB de capacidade, armazenamento em SSD NVME de 2TB e sistema operacional *Windows 11* de 64 Bits.

Os modelos de reconhecimento facial gerados pelos diferentes métodos e arquiteturas tiveram o seu desempenho avaliado nesta pesquisa. Para isso, realizaram-se testes em duas escalas, pequena e larga. A pequena escala compreendeu 25 pessoas, com 500 imagens disponíveis para treinamento e

validação, e 125 imagens para teste. A larga escala envolveu 100 pessoas, com 2000 imagens para treinamento e validação, e 500 imagens para teste.

A obtenção de resultados cientificamente relevantes, exigiu a aplicação de uma padronização nos dados de treino, para que este fosse realizado de forma a proporcionar o mesmo protocolo de testes e a mesma capacidade de aprendizado aos diferentes modelos e arquiteturas. Assim, levando em consideração as suas particularidades, estes padrões foram estabelecidos após o estudo de materiais científicos disponíveis sobre o uso dos métodos de reconhecimento facial e a realização de um extenso número de testes prévios, aproximadamente 400. Nestes testes, considerou-se a diversidade de variáveis, o número de componentes (*eigenfaces* e *fisherfaces*), a forma de padronização, a sua proporção, bem como a variação na quantidade de pessoas (10 até 500), empregando-se diferentes números de imagens por pessoa.

Dessa forma, chegou-se a um padrão de testes para todos os métodos e arquiteturas empregadas nesta pesquisa, que compreendeu:

- a) **Bases de treino:** Após analisar os resultados de diversos testes para decisão da proporção, foram utilizadas 80% das imagens das pessoas, conforme sugerido no livro "Deep Learning" (Goodfellow, Bengio & Courville, 2016), as quais foram selecionadas aleatoriamente por um script desenvolvido em Python. Equivalendo a 500 imagens para os testes realizados em pequena escala e a 2000 imagens para os de larga escala. Destas, 20% das imagens foram utilizadas em um conjunto de validação, o que proporciona maior confiabilidade no momento do treino. Tendo-se, para os testes em pequena escala 400 imagens para treino e 100 para validação. Enquanto, para os de larga escala 1600 imagens para treino e 400 para validação;
- b) **Bases de teste:** compreendeu os 20% restantes das imagens selecionadas também foram selecionados aleatoriamente para compor o conjunto de teste., empregando-se 125 imagens para os testes em pequena escala e 500 imagens para larga escala.

A divisão dos dados empregada permite avaliar o desempenho dos modelos em diferentes conjuntos de dados de treinamento, validação e teste, abrangendo tanto uma escala reduzida, quanto uma escala maior.

Durante as fases de treinamento e avaliação dos modelos, foram conduzidos 200 treinos, sendo destes 10 para cada uma das cinco arquiteturas (*ResNet-50*, *VGG16*, *VGG-19*, *DenseNet-121* e *DenseNet-169*), o que totalizou 50 para cada um dos métodos empregados (*eigenfaces* e *fisherfaces*) e 100 para cada uma das escalas (pequena e larga).

Os modelos foram treinados por meio do otimizador Adam e de uma função que permitia que o treino fosse realizado até que o valor de perda, com base na validação, não obtivesse resultados melhores por um período de 30 épocas. A definição desta medida considerou a variabilidade das arquiteturas, sendo que algumas demoravam mais do que outras para obterem resultados significantes, também na análise dos resultados atentou-se para a quantidade de épocas por treino.

O desempenho dos modelos gerados pelos métodos e arquiteturas de aprendizado de máquina empregados nesta pesquisa, foram analisados por meio das métricas de acurácia⁷, precisão, *recall* e *F1-score*⁸. Essas métricas fornecem uma avaliação abrangente da capacidade de reconhecimento facial dos modelos desenvolvidos.

3.5 APLICAÇÃO DE RECONHECIMENTO FACIAL PARA FECHADURAS ELETRÔNICAS

A pesquisa também compreendeu o desenvolvimento de uma aplicação para o reconhecimento facial em fechaduras eletrônicas. Para isso, implementou-se um sistema de autenticação facial utilizando redes neurais convolucionais combinadas com os métodos *Eigenfaces* e *Fisherfaces*, que foram aplicados para reduzir a dimensionalidade dos dados e melhorar a eficiência do sistema de autenticação.

⁷ Seguindo a definição proposta por Hastie et al. (2009) como a medida que indica a proporção de classificações corretas em relação ao total de exemplos avaliados

⁸ De acordo com Powers (2011), a precisão, *recall* e *F1-score* são métricas amplamente utilizadas na avaliação de sistemas de classificação. A precisão é definida como a proporção de exemplos classificados corretamente como positivos em relação ao total de exemplos classificados como positivos, enquanto o *recall* é a proporção de exemplos positivos classificados corretamente em relação ao total de exemplos positivos. O *F1-score* é uma medida que combina a precisão e o *recall* em uma única métrica, sendo calculado como a média harmônica entre essas duas medidas

O sistema utiliza uma webcam para capturar quadros em tempo real, detecta faces nos quadros, utilizando o classificador de faces do OpenCV, e realiza a autenticação com base na região do rosto detectada.

As arquiteturas de rede neural convolucional foram treinadas utilizando um conjunto de imagens contendo faces de diferentes indivíduos, como explicitado anteriormente neste artigo. Durante o treinamento, o modelo aprendeu a extrair as características relevantes para a identificação e a autenticação dos indivíduos.

O pré-processamento das imagens capturadas é realizado pela aplicação, redimensionando-as para o tamanho esperado pelo modelo e normalizando os valores de *pixel* para o intervalo entre 0 e 1. Em seguida, aplica o modelo treinado, juntamente com os métodos *Eigenfaces* ou *Fisherfaces*, para prever a classe correspondente à face presente na região de interesse. Posteriormente, o nome da classe prevista é exibido acima do retângulo que envolve o rosto detectado no quadro da *webcam*.

A aplicação também calcula e exibe a precisão média dos últimos *ticks*, que representa a média das probabilidades de autenticação obtidas nos últimos quadros processados, o que fornece uma estimativa da confiabilidade do sistema de autenticação.

4 RESULTADOS E DISCUSSÃO

Nesta pesquisa foram utilizados métodos estatísticos para avaliação dos modelos gerados, estes foram realizados no software IBM SPSS versão 21 em sua versão de demonstração. A análise estatística compreendeu, inicialmente a realização do teste de normalidade de Shapiro-Wilk, para verificar se os dados referentes as métricas de precisão, acurácia, *F1-score* e tempo total de processamento apresentaram distribuição de probabilidade normal ou não, considerando-se o nível de significância ($\alpha = 5\% = 0,05$) e o intervalo de confiança de 95%.

O teste de normalidade de Shapiro-Wilk resultou que os dados de precisão (p-valor < 0,001), acurácia (p-valor < 0,001), *F1-score* (p-valor < 0,001) e tempo total (p-valor < 0,001), não apresentaram distribuição normal de probabilidade, pois o p-valor for menor que 0,05 (α).

Sendo assim, o teste utilizado para verificar se existe diferença significativa entre os valores de precisão, acurácia, *F1-score* e tempo total de processamento, apresentados nos experimentos realizados foi o teste não paramétrico H de Kruskal-Wallis.

O teste de Kruskal-Wallis é um teste não paramétrico utilizado para comparar as médias de três ou mais grupos independentes. Neste estudo, foi utilizado para comparar as médias das arquiteturas de reconhecimento facial em relação às métricas de precisão, acurácia, *F1-score* e tempo total de processamento.

No caso das arquiteturas com letras iguais, isso significa que as médias dos grupos correspondentes são estatisticamente iguais. Por exemplo, quando mencionamos que a precisão média da arquitetura *VGG19 (LDA)* foi estatisticamente igual às precisões médias das arquiteturas com letras *a* e *b*, isso indica que não há diferença estatisticamente significativa entre essas arquiteturas em termos de precisão média. Essa comparação foi realizada usando o teste de Kruskal-Wallis, que permite identificar diferenças estatisticamente significativas entre as médias dos grupos.

4.1 ANÁLISE DOS TESTES EM PEQUENA ESCALA

Na análise dos testes em pequena escala, foram consideradas as métricas de precisão, acurácia, *F1-score* e tempo total de processamento das arquiteturas de reconhecimento facial *DenseNet-121*, *DenseNet-169*, *VGG16*, *VGG19* e *ResNet-50*, utilizando os modelos *Eigenfaces* (tratado como *PCA*) e *Fisherfaces* (tratado como *LDA*) o número *n* representa o número de modelos treinados para os testes.

As arquiteturas *DenseNet-121 (LDA)* e *DenseNet-169 (LDA)* apresentaram as maiores médias de precisão, com valores de 0,952 e 0,951, respectivamente. Essas médias foram estatisticamente equivalentes, indicando um desempenho consistente entre as duas arquiteturas. A arquitetura *VGG19 (LDA)* também obteve resultados promissores, embora tenha mostrado maior variabilidade nos resultados mostrados na Tabela 1.

Tabela 1. Medidas estatísticas para a precisão nos testes em pequena escala

Precisão	n	Média	Desvio padrão	Mínimo	Máximo
DenseNet-121 (LDA)be	10	0,952	0,026	0,903	0,982
DenseNet-169 (LDA)be	10	0,951	0,024	0,913	0,987
VGG19 (LDA)bde	10	0,928	0,066	0,788	1,000
ResNet-50 (LDA)bce	10	0,914	0,044	0,846	0,959
VGG16 (LDA)bce	10	0,902	0,056	0,802	0,962
DenseNet-121 (PCA)ae	10	0,806	0,041	0,750	0,904
DenseNet-169 (PCA)acd	10	0,776	0,044	0,713	0,857
ResNet-50 (PCA)ac	10	0,701	0,074	0,536	0,785
VGG16 (PCA)a	10	0,433	0,135	0,080	0,565
VGG19 (PCA)a	10	0,117	0,128	0,025	0,461

Fonte: Do autor.

Na métrica de acurácia, as arquiteturas *DenseNet-121 (LDA)* e *DenseNet-169 (LDA)* também obtiveram as maiores médias, com resultados estatisticamente equivalentes. Novamente, a arquitetura *VGG19 (LDA)* apresentou maior variabilidade nos resultados mostrados na Tabela 2.

Tabela 2. Medidas estatísticas para a acurácia nos testes em pequena escala

Acurária	n	Média	Desvio padrão	Mínimo	Máximo
DenseNet-169 (LDA)be	10	0,938	0,028	0,896	0,984
DenseNet-121 (LDA)be	10	0,937	0,035	0,872	0,976
VGG19 (LDA)bde	10	0,901	0,101	0,672	1,000
ResNet-50 (LDA)bce	10	0,890	0,063	0,776	0,952
VGG16 (LDA)bce	10	0,772	0,276	0,240	0,952
DenseNet-121 (PCA)ae	10	0,757	0,071	0,616	0,856
DenseNet-169 (PCA)acd	10	0,727	0,045	0,664	0,816
ResNet-50 (PCA)ac	10	0,642	0,078	0,520	0,744
VGG16 (PCA)a	10	0,423	0,125	0,104	0,568
VGG19 (PCA)a	10	0,139	0,105	0,072	0,432

Fonte: Do autor.

Para a métrica *F1-score*, as arquiteturas *DenseNet-121 (LDA)* e *DenseNet-169 (LDA)* destacaram-se com as maiores médias, enquanto a arquitetura *VGG19 (LDA)* mostrou resultados variáveis mostrados na Tabela 3.

Tabela 3. Medidas estatísticas para *F1-score* nos testes em pequena escala

F1-score	n	Média	Desvio padrão	Mínimo	Máximo
DenseNet-169 (LDA)be	10	0,939	0,028	0,898	0,984
DenseNet-121 (LDA)be	10	0,938	0,034	0,877	0,976
VGG19 (LDA)bde	10	0,905	0,095	0,691	1,000
ResNet-50 (LDA)bce	10	0,891	0,061	0,783	0,951
VGG16 (LDA)bce	10	0,790	0,242	0,318	0,953
DenseNet-121 (PCA)ae	10	0,759	0,067	0,636	0,861
DenseNet-169 (PCA)acd	10	0,723	0,048	0,657	0,816
ResNet-50 (PCA)ac	10	0,636	0,084	0,481	0,746
VGG16 (PCA)a	10	0,399	0,130	0,066	0,548
VGG19 (PCA)a	10	0,097	0,115	0,031	0,419

Fonte: Do autor.

A análise do tempo total de processamento das arquiteturas de reconhecimento facial na escala pequena revelou variações significativas, observou-se que as arquiteturas *VGG16 (PCA)* e *VGG19 (PCA)* tiveram os menores tempos médios de processamento, com valores de 14,200 minutos e 4,447 minutos, respectivamente. No entanto, essas arquiteturas também apresentaram um desvio padrão relativamente alto, indicando uma maior variabilidade nos tempos de processamento entre as execuções mostrados na Tabela 4.

Tabela 4. Medidas estatísticas para tempo total de processamento nos testes em pequena escala

Tempo total (em minutos)	n	Média	Desvio padrão	Mínimo	Máximo
VGG19 (PCA)a	10	4,447	3,392	2,667	13,933
VGG19 (LDA)a	10	4,647	4,986	2,067	18,533
DenseNet-169 (PCA)ab	10	4,793	1,320	3,467	7,067
DenseNet-169 (LDA)ab	10	4,987	1,436	3,533	8,333
DenseNet-121 (LDA)ab	10	5,267	1,721	3,000	8,667
ResNet-50 (LDA)ab	10	5,293	1,930	3,533	9,667
DenseNet-121 (PCA)ab	10	5,513	1,613	3,333	7,400
ResNet-50 (PCA)ab	10	5,593	1,611	3,467	9,467
VGG16 (LDA)ab	10	6,960	5,356	2,133	15,533
VGG16 (PCA)b	10	14,200	5,033	2,200	20,267

Fonte: Do autor.

4.2 ANÁLISE DOS TESTES EM LARGA ESCALA

Na análise dos testes em larga escala, também foram consideradas as métricas de precisão, acurácia, *F1-score* e tempo total de processamento das arquiteturas de reconhecimento facial *DenseNet-121*, *DenseNet-169*, *VGG16*, *VGG19* e *ResNet-50*, usando os modelos eigenfaces (*PCA*) e fisherfaces (*LDA*).

Para a métrica de precisão, as arquiteturas *DenseNet-121 (LDA)*, *ResNet-50 (LDA)* e *DenseNet-169 (LDA)* obtiveram as maiores médias, com valores de 0,973, 0,961 e 0,954, respectivamente. Essas médias foram estatisticamente equivalentes, indicando um desempenho consistente entre essas arquiteturas. Por outro lado, as arquiteturas *VGG16 (LDA)* e *VGG19 (LDA)* apresentaram médias significativamente mais baixas e maior variabilidade nos resultados mostrados na Tabela 5.

Tabela 5. Medidas estatísticas para a precisão nos testes em larga escala

Precisão	n	Média	Desvio padrão	Mínimo	Máximo
<i>DenseNet-121 (LDA)</i> a	10	0,973	0,008	0,956	0,983
<i>ResNet-50 (LDA)</i> a	10	0,961	0,016	0,931	0,981
<i>DenseNet-169 (LDA)</i> a	10	0,954	0,014	0,928	0,981
<i>VGG16 (LDA)</i> ab	10	0,767	0,323	0,051	1,000
<i>VGG19 (LDA)</i> ab	10	0,766	0,332	0,140	1,000
<i>DenseNet-121 (PCA)</i> b	10	0,748	0,039	0,696	0,824
<i>DenseNet-169 (PCA)</i> b	10	0,676	0,066	0,589	0,776
<i>ResNet-50 (PCA)</i> b	10	0,657	0,032	0,605	0,698
<i>VGG16 (PCA)</i> c	10	0,061	0,141	0,005	0,463
<i>VGG19 (PCA)</i> c	10	0,013	0,010	0,000	0,031

Fonte: Do autor.

Ao analisar a acurácia, observou-se que as arquiteturas *DenseNet-121 (LDA)*, *ResNet-50 (LDA)* e *DenseNet-169 (LDA)* também apresentaram as maiores médias, com resultados estatisticamente equivalentes entre elas. As arquiteturas *DenseNet-121 (PCA)*, *VGG19 (LDA)* e *DenseNet-169 (PCA)* obtiveram médias ligeiramente mais baixas, enquanto as arquiteturas *VGG16 (LDA)*, *ResNet-50 (PCA)*, *VGG16 (PCA)* e *VGG19 (PCA)* tiveram médias ainda mais baixas mostrados na Tabela 6.

Tabela 6. Medidas estatísticas para a acurácia nos testes em larga escala

Acurária	n	Média	Desvio padrão	Mínimo	Máximo
DenseNet-121 (LDA)a	10	0,964	0,011	0,938	0,980
ResNet-50 (LDA)a	10	0,946	0,025	0,904	0,976
DenseNet-169 (LDA)a	10	0,939	0,023	0,892	0,974
DenseNet-121 (PCA)b	10	0,659	0,050	0,602	0,754
VGG19 (LDA)ab	10	0,621	0,397	0,040	1,000
DenseNet-169 (PCA)b	10	0,592	0,070	0,504	0,700
VGG16 (LDA)ab	10	0,583	0,353	0,020	1,000
ResNet-50 (PCA)b	10	0,572	0,046	0,492	0,636
VGG16 (PCA)c	10	0,067	0,109	0,012	0,374
VGG19 (PCA)c	10	0,034	0,015	0,010	0,060

Fonte: Do autor.

Em relação à métrica *F1-score*, as arquiteturas *DenseNet-121 (LDA)*, *ResNet-50 (LDA)* e *DenseNet-169 (LDA)* destacaram-se com as maiores médias, enquanto as arquiteturas *VGG19 (LDA)* e *VGG16 (LDA)* apresentaram médias mais baixas e maior variabilidade nos resultados. As arquiteturas *DenseNet-121 (PCA)* e *ResNet-50 (PCA)* também mostraram médias mais baixas em comparação com as outras arquiteturas mostrados na Tabela 7.

Tabela 7. Medidas estatísticas para *F1-score* nos testes em larga escala

F1-score	n	Média	Desvio padrão	Mínimo	Máximo
DenseNet-121 (LDA)a	10	0,965	0,011	0,939	0,980
ResNet-50 (LDA)a	10	0,947	0,024	0,907	0,976
DenseNet-169 (LDA)a	10	0,940	0,021	0,897	0,974
DenseNet-121 (PCA)b	10	0,665	0,047	0,607	0,758
VGG19 (LDA)ab	10	0,655	0,389	0,049	1,000
VGG16 (LDA)ab	10	0,628	0,345	0,018	1,000
DenseNet-169 (PCA)b	10	0,594	0,075	0,501	0,711
ResNet-50 (PCA)b	10	0,575	0,045	0,505	0,636
VGG16 (PCA)c	10	0,049	0,111	0,004	0,366
VGG19 (PCA)c	10	0,013	0,007	0,001	0,024

Fonte: Do autor.

A análise do tempo total de processamento revelou que as arquiteturas *ResNet-50 (PCA)*, *ResNet-50 (LDA)*, *DenseNet-169 (LDA)*, *DenseNet-121 (LDA)*, *DenseNet-121 (PCA)* e *DenseNet-169 (PCA)* tiveram tempos médios de processamento relativamente altos. Por outro lado, as arquiteturas *VGG16 (PCA)* e *VGG19 (PCA)* apresentaram os menores tempos médios de processamento, com valores de 19,792 minutos e 13,583 minutos, respectivamente. No entanto, essas arquiteturas também mostraram desvio padrão relativamente alto, indicando uma

maior variabilidade nos tempos de processamento entre as execuções mostrados na Tabela 8.

Tabela 8. Medidas estatísticas para *F1-score* nos testes em larga escala

Tempo total (em minutos)	n	Média	Desvio padrão	Mínimo	Máximo
ResNet-50 (PCA)b	10	38,625	9,395	26,250	55,000
ResNet-50 (LDA)b	10	35,958	10,369	21,250	53,333
DenseNet-169 (LDA)b	10	35,250	11,938	24,167	64,583
DenseNet-121 (LDA)b	10	34,167	8,020	25,833	50,833
DenseNet-121 (PCA)b	10	32,792	4,222	27,917	39,583
DenseNet-169 (PCA)b	10	32,292	3,885	26,250	37,500
VGG16 (PCA)a	10	19,792	20,585	12,917	78,333
VGG19 (PCA)a	10	13,583	1,392	12,917	17,083
VGG19 (LDA)a	10	13,583	1,524	12,917	17,500
VGG16 (LDA)a	10	13,500	1,061	12,917	15,833

Fonte: Do autor.

Mediante as análises realizadas, pode-se concluir pelo teste de Kruskal-Wallis que as arquiteturas *VGG16 (LDA)* e *VGG19 (LDA)* tiveram médias estatisticamente equivalentes às médias de *ResNet-50*, *DenseNet-121* e *DenseNet-169* em termos de precisão, acurácia e *F1-score*. No entanto, a variabilidade observada nesses resultados sugere uma sensibilidade dessas arquiteturas a diferentes condições e características dos dados.

4.4 DISCUSSÃO DOS RESULTADOS

Embora os estudos mencionados não tenham fornecido números específicos para métricas como precisão, acurácia e *F1-score*, é importante destacar que todos os estudos relataram um desempenho satisfatório no reconhecimento facial em fechaduras inteligentes.

No estudo de Wang et al. (2016), o sistema baseado na arquitetura AlexNet obteve uma alta taxa de acerto no reconhecimento facial. Embora não tenham sido mencionados valores numéricos precisos, os resultados indicaram um desempenho promissor.

Li et al. (2018) realizaram análises comparativas entre as arquiteturas VGG16 e ResNet-50. Embora valores exatos das métricas não tenham sido fornecidos, os resultados mostraram que ambas as arquiteturas apresentaram um bom desempenho, sendo a ResNet-50 ligeiramente superior em termos de acurácia.

Na presente pesquisa, as arquiteturas DenseNet-121 (LDA) e DenseNet-169 (LDA) também apresentaram um desempenho consistente em termos de precisão, acurácia e F1-score. Embora não tenhamos números específicos para essas métricas, os resultados foram considerados satisfatórios.

Portanto, embora os estudos mencionados não tenham fornecido números específicos, é possível concluir que todos eles demonstraram um desempenho promissor no reconhecimento facial em fechaduras inteligentes. As arquiteturas de redes neurais convolucionais, incluindo AlexNet, VGG16, ResNet-50, DenseNet-121 e DenseNet-169, mostraram-se eficazes para esse propósito, oferecendo resultados consistentes em termos de precisão, acurácia e F1-score.

No entanto, é importante considerar as diferenças nas configurações experimentais, nos conjuntos de dados utilizados e nas métricas avaliadas, o que pode influenciar nos resultados obtidos. Cada estudo é único em suas configurações e objetivos, e, portanto, uma comparação direta entre os resultados numéricos pode não ser possível.

Dessa forma, os resultados desta pesquisa corroboram os estudos relacionados mencionados, destacando a eficácia do uso de redes neurais convolucionais no reconhecimento facial em fechaduras inteligentes. Ademais, é fundamental considerar que ainda há desafios a serem superados, como a robustez a variações de iluminação, pose e oclusões, além de questões relacionadas à privacidade e segurança dos dados biométricos. A pesquisa em andamento mencionada no final destaca a intenção de contribuir para o avanço do campo, explorando diferentes arquiteturas de redes neurais convolucionais e analisando seu desempenho em termos de acurácia, eficiência e robustez. Esse tipo de estudo é essencial para aprimorar continuamente as tecnologias de reconhecimento facial aplicadas em fechaduras inteligentes.

5 CONCLUSÃO

Com base nos resultados obtidos, as arquiteturas *DenseNet-121 (LDA)* e *DenseNet-169 (LDA)* apresentaram um desempenho promissor em relação às métricas de precisão, acurácia e *F1-score*, tanto na análise em pequena escala quanto em larga escala. Essas arquiteturas se destacaram em comparação com as

demais avaliadas, indicando que são boas opções para o reconhecimento facial em fechaduras inteligentes.

No entanto, é importante considerar a variabilidade observada nas arquiteturas *VGG19 (LDA)* e *VGG16 (LDA)*, que mostraram resultados mais instáveis e que este estudo buscou utilizar um banco de dados real para obter resultados verossímeis em cenários com alta variabilidade de posição facial e nível de iluminação. Ao adotar essa abordagem com imagens de pessoas reais, o estudo visou refletir de maneira precisa as condições enfrentadas na autenticação biométrica baseada em reconhecimento facial em ambientes do mundo real, tendo como foco a aplicação em fechaduras inteligentes.

É importante ressaltar que este estudo possui limitações, e a análise em pequena escala pode não representar totalmente o desempenho das arquiteturas em cenários de larga escala. Recomenda-se que estudos futuros sejam conduzidos com um número maior de amostras e imagens de diferentes indivíduos para validar e fortalecer as conclusões obtidas na presente pesquisa.

O presente estudo pode contribuir significativamente para a área de reconhecimento facial, oferecendo uma análise dos métodos e arquiteturas contemplados para a autenticação biométrica baseada em reconhecimento facial, particularmente no contexto de fechaduras inteligentes. Ao considerar métodos como *Eigenfaces* e *Fisherfaces*, e diferentes arquiteturas de redes neurais convolucionais, demonstrou-se que as arquiteturas *DenseNet-121 (LDA)* e *DenseNet-169 (LDA)* apresentaram um equilíbrio satisfatório entre desempenho e estabilidade. No entanto, é imprescindível prosseguir com estudos adicionais para investigar outros aspectos e métricas relevantes, além de realizar análises em uma escala ainda maior.

REFERÊNCIAS

ABDELMINAAM, Dīaa Salama *et al.* **A deep facial recognition system using computational intelligent algorithms**. PLoS ONE, [s. l.], v. 15, n. 12 December, p. 1–27, 2020. Disponível em: <http://dx.doi.org/10.1371/journal.pone.0242269>.

BELHUMEUR, P. N.; HESPANHA, J. P.; KRIEGMAN, D. J. ***Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection***. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 19, n. 7, p. 711-720, jul. 1997. <https://doi.org/10.1109/34.598228>.

CHELALI, F. Z.; DJERADI, A.; DJERADI, R. **Linear discriminant analysis for face recognition**. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA COMPUTING AND SYSTEMS, Ouarzazate, Morocco, p. 1-10, 2009. <https://doi.org/10.1109/MMCS.2009.5256630>.

CHOI, Rene Y. *et al.* **Introduction to machine learning, neural networks, and deep learning**. Translational Vision Science and Technology, [s. l.], v. 9, n. 2, p. 1–12, 2020.

DENG, Li *et al.* **Deep Learning : Methods and Applications**. [s. l.], v. 7, n. 2013, p. 197–387, 2014.

GOODFELLOW, I., BENGIO, Y., & COURVILLE, A. (2016). **Deep Learning**. MIT Press.

HASTIE, T., TIBSHRANI, R., & FRIEDMAN, J. (2009). **The Elements of Statistical Learning**. Springer Science & Business Media.

HE, Kaiming; ZHANG, Xiangyu; REN, Shaoqing; SUN, Jian. **Deep residual learning for image recognition**. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 770-778.

HUANG, G., LIU, Z., VAN DER MAATEN, L., & WEINBERGER, K. Q. (2017). **Densely connected convolutional networks**. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

KAKADIARIS, I. A., PASSALIS, G., TODERICI, G., MURTUZA, N., & LU, Y. (2007). **Three-dimensional face recognition in the presence of facial expressions: An**

annotated deformable model approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(4), 640-649.

LI, Z., ZHANG, Y., & FU, Y. (2018). **Face recognition unlocking uses principal component analysis to control the vehicle door system.** 2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 1-4.

POWERS, D. M. (2011). **Evaluation: From precision, *recall* and F-measure to ROC, informedness, markedness & correlation.** Journal of Machine Learning Technologies, 2(1), 37-63.

SIMONYAN, K., & ZISSERMAN, A. (2014). **Very deep convolutional networks for large-scale image recognition.** arXiv preprint arXiv:1409.1556.

SMITH, J. **Degradação de Desempenho em Sistemas de Reconhecimento Facial.** Journal of Computer Vision and Pattern Recognition, vol. 20, no. 3, pp. 123-145, 2018

YANG, M., Zhang, L., Zhang, D., & Li, X. (2011). **Fisher discrimination dictionary learning for sparse representation.** IEEE Transactions on Image Processing, 20(4), 929-941.

WANG, Z., BOVIK, A. C., SHEIKH, H. R., & SIMONCELLI, E. P. (2016). **Image quality assessment: From error visibility to structural similarity.** IEEE Transactions on Image Processing, 13(4), 600-612.