



Ministério da Educação
Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Biológicas
Departamento de Estatística



MÉTODOS ESTATÍSTICOS ADEQUADOS AO RECONHECIMENTO FACIAL

PAULO VITOR DE SOUZA E SILVA

Ouro Preto MG
2023

PAULO VITOR DE SOUZA E SILVA

MÉTODOS ESTATÍSTICOS ADEQUADOS AO RECONHECIMENTO FACIAL

Monografia apresentada ao Curso de Estatística da Universidade Federal de Ouro Preto como parte dos requisitos necessários para a obtenção do grau em Bacharelado em Estatística.

Orientador: Prof. Dr. Marcelo Carlos Ribeiro

Ouro Preto - MG
30 de março de 2023

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S586m Silva, Paulo Vitor de Souza e.
Métodos estatísticos adequados ao reconhecimento facial.
[manuscrito] / Paulo Vitor de Souza e Silva. - 2023.
59 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Marcelo Carlos Ribeiro.
Monografia (Bacharelado). Universidade Federal de Ouro Preto.
Instituto de Ciências Exatas e Biológicas. Graduação em Estatística .

1. Biometria. 2. Inteligência artificial. 3. Reconhecimento facial. I.
Ribeiro, Marcelo Carlos. II. Universidade Federal de Ouro Preto. III. Título.

CDU 31

Bibliotecário(a) Responsável: Luciana de Oliveira - SIAPE: 1.937.800



FOLHA DE APROVAÇÃO

Paulo Vitor de Souza e Silva

Estatísticos adequados ao reconhecimento facial

Monografia apresentada ao Curso de Estatística da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Estatística

Aprovada em 30 de março de 2023

Membros da banca

Dr. Marcelo Carlos Ribeiro - Orientador - Universidade Federal de Ouro Preto
Dr. Eduardo Bearzoti - Universidade Federal de Ouro Preto
Dr. Tiago Martins Pereira - Universidade Federal de Ouro Preto

Professor Dr. Marcelo Carlos Ribeiro, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 30/03/2023



Documento assinado eletronicamente por **Marcelo Carlos Ribeiro, PROFESSOR DE MAGISTERIO SUPERIOR**, em 04/04/2023, às 16:45, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0499422** e o código CRC **C9359AA5**.

Dedico este trabalho aos meus pais, que sempre estiveram ao meu lado em cada passo da minha jornada acadêmica. Foram eles que me deram amor, apoio e os recursos necessários para que eu pudesse alcançar meus objetivos e me tornar a pessoa que sou hoje.

Mãe e pai, este trabalho é uma pequena homenagem ao que vocês têm feito por mim desde o meu nascimento. Sempre me inspirando a ser a melhor versão de mim mesmo, vocês me ensinaram valores como dedicação, perseverança e humildade que carregarei para sempre.

Não tenho palavras para expressar minha gratidão por tudo que vocês fizeram e fazem por mim. Obrigado por serem os melhores pais que alguém poderia pedir.

Agradecimentos

Agradeço, em primeiro lugar, a Deus por me dar a vida e me acompanhar ao longo desta jornada. Sua força e perseverança me ajudaram a superar os obstáculos e a alcançar meus objetivos.

Agradeço aos meus pais, Geraldo e Gilsirene, por me ensinarem a viver com dignidade, pelo incentivo a perseguir meus sonhos e pelo amor incondicional. Seu apoio e presença foram fundamentais para que eu chegasse até aqui.

Também quero agradecer à minha irmã, Ana Beatriz, aos meus tios, primos e avós, por me acompanharem ao longo do caminho, oferecendo sua companhia e apoio sempre que necessário.

À minha namorada, Lais, agradeço pelo companheirismo, carinho e pelos momentos inesquecíveis que compartilhamos juntos. Sua presença em minha vida tem sido uma bênção.

Ao Prof.Dr.Marcelo Carlos Ribeiro, pelo empenho, paciência e por todo o conhecimento transmitido na realização desse trabalho.

Aos demais professores e o departamento de Estatística da UFOP, por todo o conhecimento e apoio oferecidos ao longo desta jornada. Serei eternamente grato por tudo o que aprendi com eles.

Por fim, agradeço a todos os amigos e familiares que acompanharam minha trajetória, torceram por mim e compreenderam minha ausência em alguns momentos. Sem o apoio de vocês, eu não teria chegado até aqui.

Tudo o que vivi durante este período - os medos, as decisões, as dúvidas, as alegrias e as descobertas - contribuíram para meu crescimento pessoal e profissional. Sei que há muito trabalho pela frente, mas estou confiante, pois tenho o suporte das pessoas que amo e a proteção de Deus.

Mais uma vez, a todos que contribuíram de maneira direta ou indireta para que eu chegasse até aqui, o meu muito obrigado!

Eu sei que o meu trabalho é uma gota no oceano, mas sem ele o oceano seria menor.

Madre Teresa de Calcutá

Resumo

O presente trabalho teve como objetivo implementar um sistema de reconhecimento facial baseado em expressões faciais básicas, utilizando a linguagem de programação Python e algumas funções da biblioteca OpenCV. Foi realizada a conversão das imagens em matrizes de pixels, o pré-processamento do banco de imagens JAFFE e o treinamento dos algoritmos Eigenfaces e Fisherfaces. Foram avaliadas seis classificações diferentes com a utilização de 152 imagens de teste contendo diversas expressões faciais, sendo realizada a avaliação da eficiência dos algoritmos através da matriz de confusão, acurácia, sensibilidade, especificidade, precisão e F1-Score. Os resultados mostraram que imagens expressivas aumentam a probabilidade de reconhecer os indivíduos com sucesso, melhorando assim a eficiência dos métodos utilizados. O sistema desenvolvido possui grande potencial para aplicação em diferentes áreas, como segurança e reconhecimento de pessoas desaparecidas.

Palavras-chave: Biometria, Classificação, Componentes principais, Discriminante linear, Eigenfaces, Expressões faciais, Fisherfaces, Inteligência artificial, OpenCV, Reconhecimento facial.

Abstract

The present work aimed to implement a facial recognition system based on basic facial expressions, using the Python programming language and some functions from the OpenCV library. The images were converted into pixel matrices, the JAFFE image database was preprocessed, and the Eigenfaces and Fisherfaces algorithms were trained. Six different classifications were evaluated using 152 test images containing various facial expressions, and the efficiency of the algorithms was evaluated using the confusion matrix, accuracy, sensitivity, specificity, precision, and F1-Score. The results showed that expressive images increase the probability of successfully recognizing individuals, thus improving the efficiency of the methods used. The developed system has great potential for application in different areas, such as security and recognition of missing persons.

Keywords: Biometrics, Classification, Principal Components, Linear Discriminant, Eigenfaces, Facial Expressions, Fisherfaces, Artificial Intelligence, OpenCV, Facial Recognition.

Lista de figuras

Figura 1 – Exemplo de expressões faciais em pessoas e animais. Na parte Superior temos varias fotografias de uma mulher, enquanto na parte inferior vemos fotografias de cachorros e um gato	4
Figura 2 – Componentes principais obtidos do banco de dados Iris, muito usados em exemplos de aplicações de técnicas de redução de dimensionalidade.	8
Figura 3 – Porcentagem de variância explicada pelos componentes principais em ordem decrescente.	9
Figura 4 – Redução de características em uma imagem após aplicação do <i>Eigenfaces</i> . Imagem adaptada de Dinalankara (2017).	14
Figura 5 – Diagrama detalhando os passos do algoritmo <i>Eigenfaces</i> . Imagem adaptada de Lima e Neto (2017).	15
Figura 6 – Transformação de imagem em vetor, o que possibilita o calculo das componentes principais. Imagem adaptada de Kortli et al. (2020).	16
Figura 7 – Conjunto de exemplos de <i>Eigenfaces</i> redimensionadas para o tamanho das imagens do conjunto de treinamento. Imagem adaptada de Silva (2016).	17
Figura 8 – Fluxograma do sistema de reconhecimento facial proposto.	22
Figura 9 – Imagens de indivíduos do banco <i>JAFFE</i> em diferentes expressões faciais.	23
Figura 10 – Exemplo da imagem de um indivíduo da base <i>JAFFE</i> com sua respectiva identificação.	24
Figura 11 – Imagens com a expressão de felicidade de todas as mulheres da base <i>JAFFE</i> e suas respectivas iniciais do nome.	25
Figura 12 – Transformação da imagem em matriz de pixels. Cada <i>pixel</i> na imagem pode ser representado na escala de cinza com valores que variam entre 0 (Preto) e 255 (Branco).	26
Figura 13 – <i>Eigenfaces</i> geradas a partir das imagens neutras do dataset <i>JAFFE</i>	33
Figura 14 – Variância cumulativa explicada para o treinamento com as expressões neutras.	34
Figura 15 – <i>Fisherfaces</i> geradas a partir das imagens neutras do dataset <i>JAFFE</i>	38
Figura 16 – Médias ponderadas dos resultados obtidos para o treinamento utilizando imagens neutras.	43
Figura 17 – Médias ponderadas dos resultados obtidos para o treinamento utilizando imagens expressando emoção felicidade.	43
Figura 18 – Médias ponderadas dos resultados obtidos para o treinamento utilizando imagens expressando emoções neutra e felicidade.	44

Lista de tabelas

Tabela 1	– Distribuição das imagens do banco de imagens JAFFE. Abreviaturas: NE - Neutro, HA - feliz, SA - Tristeza, SU- Surpresa, AN - Raiva, DI - Nojo.	24
Tabela 2	– <i>Dataframe</i> contendo os dados usados para treinamento após processo de pré-processamento e organização do banco de dados JAFFE	27
Tabela 3	– Quantidade de imagens utilizadas na base de teste	30
Tabela 4	– Matriz de Confusão	30
Tabela 5	– Variância acumulada para o treinamento do <i>Eigenfaces</i> com as expressões neutras.	35
Tabela 6	– Resultado do modelo <i>Eigenfaces</i> treinado com as imagens neutras.	35
Tabela 7	– Resultado do modelo <i>Eigenfaces</i> treinado com as imagens expressando felicidade.	36
Tabela 8	– Resultado do modelo <i>Eigenfaces</i> treinado com as imagens expressando emoções neutra e felicidade.	37
Tabela 9	– Resultado do modelo <i>Fisherfaces</i> treinado com as imagens neutras.	39
Tabela 10	– Resultado do modelo <i>Fisherfaces</i> treinado com imagens expressando felicidade.	40
Tabela 11	– Resultado do modelo <i>Fisherfaces</i> treinado com imagens expressando emoções neutra e felicidade.	41
Tabela 12	– Matriz de confusão <i>Eigenfaces</i> Neutra	53
Tabela 13	– Matriz de confusão <i>Fisherfaces</i> Neutra	53
Tabela 14	– Matriz de confusão <i>Eigenfaces</i> Feliz	54
Tabela 15	– Matriz de confusão <i>Fisherfaces</i> Feliz	54
Tabela 16	– Matriz de confusão <i>Eigenfaces</i> Neutro e Feliz	54
Tabela 17	– Matriz de confusão <i>Fisherfaces</i> Neutro e Feliz	55
Tabela 18	– Médias ponderadas dos resultados obtidos para o treinamento utilizando as imagens neutras.	56
Tabela 19	– Médias ponderadas dos resultados obtidos para o treinamento utilizando as imagens Felizes	56
Tabela 20	– Médias ponderadas dos resultados obtidos para o treinamento utilizando as imagens neutras e Feliz juntas.	56

Lista de abreviaturas e siglas

AM	aprendizado de máquina
CNNs	Redes neurais convolucionais
DEEST	Departamento de Estatística da Universidade Federal de Ouro Preto
IA	Inteligência artificial
IDE	Integrated Development Environment
JAFFE	Japanese Female Facial Expression
LDA	Linear Discriminant Analysis
OpenCV	Open Source Computer Vision
PCA	Principal Component Analysis
UFOP	Universidade Federal de Ouro Preto

Sumário

1	INTRODUÇÃO	1
2	REVISÃO BIBLIOGRÁFICA	2
2.1	Visão computacional	2
2.2	Expressões faciais	3
2.3	Reconhecimento de faces	4
2.4	Pré-Processamento de Imagens	5
2.5	Extração de características	6
2.5.1	Análise de componentes principais - PCA	7
2.5.1.1	<i>Eigenfaces</i>	14
2.5.2	Análise Discriminante Linear - LDA	18
2.5.2.1	<i>Fisherfaces</i>	19
3	MATERIAL E MÉTODOS	22
3.1	Banco de dados	22
3.2	Python	24
3.2.1	Manipulação das figuras com Pyplot	25
3.3	Transformação da imagem em matriz de pixels	26
3.4	Pré-processamento da base de dados	26
3.5	OpenCV	26
3.6	Criação das bases de Treinamento e Teste	28
3.7	Avaliação dos modelos	30
4	RESULTADOS E DISCUSSÕES	33
4.1	<i>Eigenfaces</i>	33
4.1.1	EXPERIMENTO 1: Treinamento - Neutra	34
4.1.2	EXPERIMENTO 2: Treinamento - Felicidade	36
4.1.3	EXPERIMENTO 3: Treinamento - Neutra + Felicidade	37
4.2	<i>Fisherfaces</i>	38
4.2.1	EXPERIMENTO 1: Treinamento - Neutra	39
4.2.2	EXPERIMENTO 2: Treinamento - Felicidade	40
4.2.3	EXPERIMENTO 3: Treinamento - Neutra + Felicidade	41
4.3	<i>Eigenfaces vs Fisherfaces</i>	42
4.3.1	EXPERIMENTO 1: Treinamento - Neutra	42
4.3.2	EXPERIMENTO 2: Treinamento - Felicidade	43
4.3.3	EXPERIMENTO 3: Treinamento - Neutra + Felicidade	44

4.4	Discussão dos resultados	45
5	CONSIDERAÇÕES FINAIS	47
5.1	Conclusão	47
5.2	Trabalhos Futuros	47
5.3	Materiais Elaborados	48
	Referências	49
	Apêndices	52
	APÊNDICE A – MATRIZES DE CONFUSÃO	53
	APÊNDICE B – MÉDIAS PONDERADAS DOS RESULTADOS	56

1 Introdução

A tecnologia de reconhecimento facial tem suas raízes no início dos anos 1960, quando Woodrow Wilson Bledsoe, um cientista da computação, começou a desenvolver um sistema que identificava pessoas automaticamente em fotografias. No entanto, não foi até a década de 1990 que a tecnologia começou a ganhar atenção generalizada. Uma das primeiras aplicações comerciais foi no Japão, onde a *NEC Corporation* introduziu um sistema de reconhecimento facial para fins de segurança.

O reconhecimento facial tem avançado rapidamente nos últimos anos, oferecendo novas oportunidades e desafios para uma variedade de setores, incluindo segurança pública, aplicação da lei, marketing e saúde. Na aplicação da lei, é usado para identificar suspeitos e rastrear criminosos. Também é usado em sistemas de segurança pública, como a identificação em espaços públicos de pessoas desaparecidas ou na segurança de aeroportos, com o intuito de detectar ameaças e prevenir o terrorismo. No marketing, é usado para personalizar anúncios e melhorar a experiência do cliente. Na área da saúde, é usado para diagnosticar certas condições médicas, como a síndrome de *Down* e distúrbios genéticos raros. Essa tecnologia usa algoritmos para detectar e analisar as características únicas do rosto de uma pessoa, como a distância entre os olhos, o formato do nariz e os contornos das bochechas, para identificar e verificar sua identidade.

O intuito deste trabalho é realizar uma análise comparativa entre dois algoritmos de reconhecimento facial, sob a hipótese de que a eficiência desses algoritmos pode ser melhorada quando utilizados imagens de pessoas expressando emoções neutra ou de felicidade. Para tanto, serão estudados algoritmos que fazem uso de duas técnicas bem conhecidas e difundidas pela literatura, o PCA (*Principal Component Analysis*) ou do português, Análise de componentes principais e o LDA (*Linear Discriminant Analysis*) em português Análise Discriminante Linear. Tais técnicas consistem em reduzir a redundância ou mesmo para detecção de padrões (faces, objetos, etc) em um conjunto de dados utilizando procedimentos da álgebra linear e estatística aplicada.

Os dois algoritmos escolhidos são intitulados como *Eigenface* e *Fisherface*, ambos fornecidos pela biblioteca *OpenCV* e implementados em linguagem *Python*. Os métodos foram aplicados ao banco de imagens *JAFFE*, obtido através de um requerimento disponível no endereço <<https://zenodo.org/record/3451524#.ZCsVZnbMLIU>>. O *Eigenface* e *Fisherface* foram avaliados através das medidas de acurácia, precisão, sensibilidade e especificidade. Os resultados obtidos mostraram que imagens expressivas e neutras afetam a probabilidade de reconhecimento dos algoritmos.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo apresentamos alguns dos principais conceitos de técnicas e métodos fundamentados em outros trabalhos de pesquisa publicados na literatura, ao qual nos baseamos para as análises presentes neste trabalho.

2.1 Visão computacional

A Visão Computacional é um campo interdisciplinar que tem ganhado significativa atenção e desenvolvimento nos últimos anos, devido aos avanços da tecnologia e à necessidade de sistemas inteligentes para processar informações visuais. Segundo Marengoni e Stringhini (2010), a Visão Computacional permite aos computadores interpretar, entender e analisar imagens e vídeos. Desde seu surgimento na década de 1960, a Visão Computacional avançou significativamente, com o desenvolvimento de algoritmos mais complexos que podem reconhecer rostos, objetos e cenas.

Uma das principais áreas da Visão Computacional é o uso de técnicas de inteligência artificial (IA) e aprendizado de máquina (ML), que permitem que os computadores aprendam com os dados e melhorem seu desempenho ao longo do tempo. Os algoritmos de IA e ML permitem que os computadores analisem grandes quantidades de dados visuais e extraiam informações significativas deles. Na área da saúde, por exemplo, algoritmos de IA e ML foram desenvolvidos para diagnosticar doenças e fornecer recomendações de tratamento personalizadas aos pacientes.

Outra área importante na Visão Computacional é a biometria, que envolve o uso de dados visuais para identificar indivíduos. Os sistemas biométricos tornaram-se cada vez mais populares nos últimos anos e são usados para fins de segurança, como controle de acesso e vigilância. Os sistemas de reconhecimento facial são um exemplo comum de sistemas biométricos e são amplamente utilizados em aplicações da lei e segurança.

As técnicas estatísticas também desempenham um papel significativo na Visão Computacional, particularmente na área de extração de recursos. A extração de recursos envolve a identificação de recursos ou padrões relevantes em imagens que podem ser usados para tarefas de reconhecimento ou classificação. De acordo com Belhumeur, Hespanha e Kriegman (1997), análise de componentes principais (PCA) e análise discriminante linear (LDA) são duas técnicas estatísticas comumente usadas na extração de características. O PCA é usado para reduzir a dimensionalidade dos dados identificando os recursos mais significativos, enquanto o LDA é usado para identificar os recursos que podem discriminar entre diferentes classes de dados.

Nos últimos anos, a visão computacional teve uma explosão de aplicações em vários setores. Na área da saúde, a visão computacional tem sido usada para análise de imagens médicas,

diagnóstico de doenças e recomendações personalizadas de tratamento. Na indústria automotiva, a visão computacional é usada para carros autônomos e sistemas de assistência ao motorista. No varejo, a visão computacional é usada para gerenciamento de estoque, análise de comportamento do cliente e sistemas de pagamento baseados em reconhecimento facial.

Uma das aplicações mais recentes da visão computacional está no campo da realidade aumentada (AR) e realidade virtual (VR). As tecnologias AR e VR usam algoritmos de visão computacional para sobrepor informações digitais ao mundo real, criando experiências imersivas. Por exemplo, AR pode ser usado para sobrepor informações digitais em produtos em uma loja, fornecendo aos clientes informações adicionais sobre o produto e aprimorando a experiência de compra. À medida que a visão computacional continua avançando, espera-se que suas aplicações cresçam ainda mais, tornando-se um campo empolgante a ser observado nos próximos anos.

2.2 Expressões faciais

As expressões faciais configuram-se como uma das importantes formas de comunicação não verbal e possibilita transmitir emoções e sentimentos de uma maneira rápida e eficaz. Através da observação da expressão das pessoas com quais interagimos, esse recurso natural permite que detectemos informações sobre seu estado emocional, intenções e, inclusive, suas reações aos eventos apresentados em nosso ambiente (WHALEN et al., 2013).

As expressões indicam o estado emocional de pessoas e também animais (Figura 2.2). Há cerca de 150 anos atrás, essa foi a linha central na justificativa da teoria da evolução pela seleção natural proposta por Charles Darwin. Naquele momento, Darwin defendia a hipótese de que a expressão facial representa uma das formas de linguagem e comunicação universal, algo tão presente quanto importante em vida social desde o surgimento da humanidade (DARWIN, 1872). Apesar de seus argumentos terem sido considerados controversos pela comunidade científica, centenas de valiosos estudos empíricos e debates quanto à essa universalidade vêm sendo produzidos até os dias atuais.

Um dos mais renomados psicólogos na área de expressões faciais, Paul Ekman (1971) ressalta as expressões faciais como "o canal mais rápido e universal de comunicação emocional". Em sua pesquisa *Universals and cultural differences in facial expressions of emotion*, se propôs a encontrar informações acerca de significados, funções, origens, categorizações, atribuindo às expressões faciais a função de transmitir emoções de forma universal. Para isto, observou em diferentes culturas diversos sinais não verbais de comunicação, tais como; expressões faciais, gestos e postura corporal. Ekman (1971) concluiu que as expressões faciais humanas de alegria, medo, nojo, raiva e surpresa eram as mesmas independente da cultura.

De acordo com Costa-Vieira e Souza (2014), o reconhecimento de expressões faciais é uma área relevante na pesquisa em neurociências e psicologia. Em sua investigação, os autores investigaram a capacidade de jovens brasileiros em reconhecer expressões faciais e prosódia



Figura 1 – Exemplo de expressões faciais em pessoas e animais. Na parte Superior temos varias fotografias de uma mulher, enquanto na parte inferior vemos fotografias de cachorros e um gato

emocional. Os resultados deste estudo fornecem evidências de que existe uma habilidade de reconhecimento facial, no entanto, ainda há espaço para melhoria na identificação correta das expressões. Além disso, os resultados destacam a importância de treinar habilidades de reconhecimento emocional na população jovem, pois isso pode ter impactos positivos em suas vidas sociais e pessoais.

Recentemente, em um trabalho publicado na renomada revista *Nature*, os autores Cowen et al. (2021) descobriram que 16 tipos de expressões faciais coexistem sistematicamente com contextos sociais específicos em 144 países. Em diversas regiões geográficas, há 70% de sobreposição nas associações entre contexto e expressão facial—evidência de universalidade substancial em nosso mundo ciber conectado.

2.3 Reconhecimento de faces

O reconhecimento de faces é baseado em princípios matemáticos e tecnológicos que permitem identificar indivíduos a partir de imagens capturadas. Segundo Turaga et al. (2011), esses princípios incluem extração de características, comparação de características e classificação. A extração de características consiste em identificar pontos únicos ou padrões na face, como o formato dos olhos, nariz e boca. A comparação de características implica a comparação desses pontos ou padrões com os de uma imagem previamente armazenada, enquanto a classificação envolve determinar a semelhança entre as imagens comparadas.

De acordo com Phillips et al. (2000), outro conceito importante é a normalização, que envolve ajustar as imagens de forma a remover variações como posição, orientação, iluminação e escala, tornando a comparação mais precisa. Além disso, a utilização de bancos de dados

de referência é fundamental para o desempenho do reconhecimento de faces, pois permite o treinamento dos algoritmos e a comparação das imagens capturadas com as imagens armazenadas.

As pessoas possuem a habilidade de reconhecer milhares de rostos aprendidos ao longo da vida mesmo após anos de separação. O reconhecimento é robusto mesmo diante de fatores como diferentes expressões faciais, envelhecimento, óculos, barbas ou mudanças no estilo de cabelo. Porém desenvolver um modelo computacional de reconhecimento facial é bastante difícil, pois os rostos são estímulos visuais complexos e multidimensionais. Isso torna o reconhecimento facial uma tarefa de visão computacional de alto nível, na qual muitas técnicas são envolvidas(AGARWAL et al., 2010).

Um sistema de reconhecimento facial deve possuir a habilidade de detectar a face de uma pessoa a partir de uma imagem e compará-la com imagens anteriormente inseridas no banco de dados. As fotografias fazem parte do nosso cotidiano e são encontradas facilmente em documentos oficiais e nas redes sociais, tornando a implementação do reconhecimento acessível. Os problemas que afetam a precisão do reconhecimento facial são a variação de iluminação, ruído, escala, nitidez, dentre outros. As técnicas de pré-processamento são usadas no tratamento destes problemas que reduzem o desempenho do reconhecimento facial.

Atualmente o aprendizado profundo como as redes neurais convolucionais (CNNs) tornaram-se populares na área de processamento de imagens e reconhecimento de faces. Porém estas técnicas necessitam de uma grande quantidade de dados para um melhor desempenho. Os métodos *Eigenfaces*(Turk e Pentland (1991)) e *Fisherfaces*(Belhumeur, Hespanha e Kriegman (1997)) apresentam vantagens sobre outras técnicas de reconhecimento facial por demonstrarem desempenho mais estável em pequenos conjuntos de dados, possuir fácil implementação e consequentemente exigir menos tempo de computação.

2.4 Pré-Processamento de Imagens

O pré-processamento de imagens refere-se às técnicas usadas para melhorar a qualidade das imagens digitais antes da análise ou processamento. Isso pode incluir tarefas como redução de ruído, aprimoramento de contraste e suavização de imagem. De acordo com Marengoni e Stringhini (2010), processamento de imagens é um processo no qual se obtém de uma imagem um conjunto de valores numéricos, que podem ou não compor uma outra imagem. Por outro ângulo, a visão computacional procura simular a visão humana, porém, a saída é uma interpretação parcial ou total da imagem. Resumidamente, processar uma imagem significa transformá-la gradativamente para facilitar a retirada de informações nela contidas.

O pré-processamento de imagens tem aplicações em uma ampla gama de setores, incluindo saúde, agricultura e visão computacional. Na área da saúde, o pré-processamento de imagens é usado para melhorar a precisão das imagens e diagnósticos médicos, enquanto na agricultura é usado para identificar tipos de culturas e detectar doenças. Na visão computacional, o pré-

processamento de imagem é usado para melhorar a detecção e o reconhecimento de objetos, bem como o reconhecimento e rastreamento facial.

Para Gonzalez e Woods (2000), os passos que vão do pré-processamento de imagens até a visão computacional podem ser divididos em três níveis: baixo, médio e alto. Os de baixo nível compreendem manipulações primitivas, como diminuição do ruído e aumento de contraste. O nível médio envolve o particionamento da imagem em regiões ou objetos, a descrição desses objetos, e classificação dessas partes individuais. Os processos de nível alto consistem em atribuir significado ao conjunto de objetos reconhecidos anteriormente, analisando a imagem e realizando funções cognitivas associadas à visão.

Ruídos são contaminações que podem ser geradas por diversos motivos durante a aquisição, compressão e transmissão da imagem causando à distorção e perda de informações da imagem. Durante a aquisição por exemplo, podemos ter problemas com o sensor utilizado, a iluminação do ambiente, as condições climáticas no instante de captura da imagem e a câmera. Com a presença de ruído, algumas tarefas de processamento de imagem são prejudicadas. Logo, o pré-processamento de imagens é fundamental para a redução dessas contaminações e consequentemente uma boa análise das informações (FAN et al., 2019; MARENGONI; STRINGHINI, 2010).

A detecção de faces já é um método de pré-processamento nível médio que consiste em isolar o rosto de pessoas em uma imagem. Sua função é encontrar em uma determinada imagem, uma ou mais faces, e também pode reconhecer elementos como os olhos, bocas, nariz, entre outros (SANTANA; ROCHA et al., 2015). A escolha do método de detecção de faces depende de vários fatores, como a necessidade de tempo de processamento, a precisão da detecção, o tamanho do conjunto de dados de treinamento, entre outros.

O pré-processamento de imagens é uma etapa crítica na análise de imagens, ajudando a melhorar a qualidade e a confiabilidade das imagens digitais. Embora existam desafios e limitações a serem considerados, o desenvolvimento de novas técnicas e tecnologias de pré-processamento continuará a desempenhar um papel importante no avanço da análise e interpretação de imagens no futuro.

2.5 Extração de características

A extração de características em aplicações de análise de dados e aprendizado de máquina é utilizada para reduzir a dimensionalidade de um conjunto, afim de se obter as informações mais relevantes. Ao executar as técnicas de reconhecimento de faces, nem todos os pixels da imagem são essenciais e assim os dados redundantes podem ser eliminados sem prejudicar o desempenho do método. Além disso, variações no objeto a ser reconhecido como translação, rotação e escala afetam a precisão dos classificadores. Por isso é necessária a utilização de algoritmos que reduzam a dimensionalidade de forma eficiente proporcionando uma representação dos padrões de uma

forma robusta a essas alterações(SILVA, 2015).

Nesse sentido, duas técnicas amplamente utilizadas são a Análise de Componentes Principais (PCA) e a Análise Discriminante Linear (LDA). O PCA visa transformar um conjunto de dados em um novo conjunto de variáveis, chamadas componentes principais, que não são correlacionadas e capturam a maior parte da variabilidade dos dados. Jolliffe (2002) destaca que o PCA "pode ser uma técnica muito útil para reduzir a dimensionalidade dos dados enquanto retém a maior parte das informações no conjunto de dados original". Já o LDA é uma técnica supervisionada que visa encontrar uma combinação linear de recursos que maximize a separação entre as classes nos dados e tem se mostrado eficaz em problemas de classificação, especialmente ao lidar com dados de alta dimensão. Deng et al. (2005) destacam que o LDA é uma técnica de extração de recursos eficaz para tarefas de classificação de alta dimensão.

Estes métodos multivariados otimizam a matriz de transformação com diferentes intenções. O PCA otimiza a matriz de transformação ao encontrar as maiores variações no espaço de recurso original. A LDA procura a maior proporção de variação entre classes e a menor variação dentro de classe ao projetar a característica original para um subespaço.

Turk e Pentland (1991) propuseram uma abordagem para extração de características baseada em PCA para identificar as características mais distintas de uma face. Essa abordagem hoje conhecida como *Eigenfaces* foi amplamente utilizada e melhorada ao longo dos anos, tornando-se uma das técnicas mais populares e amplamente utilizadas na área de reconhecimento facial. Belhumeur, Hespanha e Kriegman (1997) propuseram uma abordagem para extração de características baseada na combinação PCA-LDA para identificar as características mais distintas de uma face conhecida como *Fisherfaces*. A combinação de PCA e LDA na extração de características tem sido amplamente utilizada em várias aplicações, como reconhecimento facial e reconhecimento de objetos na visão computacional.

Ao executar as técnicas de reconhecimento de faces, nem todos os pixels da imagem são essenciais e assim os dados redundantes podem ser eliminados sem prejudicar o desempenho do método. Além disso, variações no objeto a ser reconhecido como translação, rotação e escala afetam a precisão dos classificadores. Por isso é necessária a utilização de algoritmos que reduzam a dimensionalidade de forma eficiente proporcionando uma representação dos padrões de uma forma robusta a essas alterações(SILVA, 2015).

2.5.1 Análise de componentes principais - PCA

A Análise de Componentes Principais, do inglês *Principal Component Analysis (PCA)* é uma técnica estatística multivariada proposta por Karl Pearson (1901). A ideia principal desse método é reduzir a dimensionalidade do conjunto de dados original com o mínimo de perda de informação possível. Para tanto, a *PCA* transforma o conjunto de variáveis originais intercorrelacionadas em um novo conjunto de variáveis não correlacionadas (ortogonais), denominadas

de componentes principais. Essas novas variáveis correspondem a uma combinação linear de todas as variáveis originais (SILVA, 2016).

Perceba na figura 2.5.1 que o método *PCA* é particularmente útil quando as variáveis dentro do conjunto de dados estão altamente correlacionadas, o que indica que existem variáveis que medem o mesmo evento. Devido a essa redundância, o *PCA* pode ser usado para reduzir as variáveis originais em um número menor de novas variáveis (componentes principais), desse modo, o máximo de informação em termos de variação total contida nos dados é retido em ordem de estimação dos componentes e medido pelo seu seu autovalor (KASSAMBARA, 2017; PEREIRA, 2017).

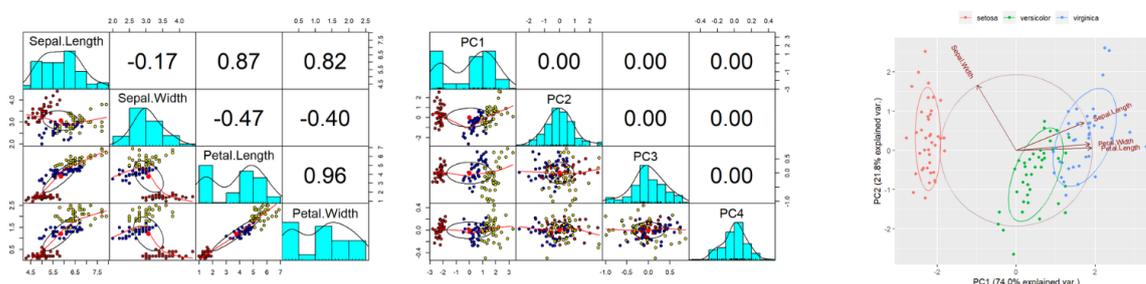


Figura 2 – Componentes principais obtidos do banco de dados Iris, muito usados em exemplos de aplicações de técnicas de redução de dimensionalidade.

Conforme mencionado anteriormente, os autovalores medem a quantidade de variação retida por cada componente principal. Os autovalores apresentam valores altos para os primeiros componentes e pequenos para os componentes subsequentes. Ou seja, os primeiros componentes correspondem às direções com a quantidade máxima de variação no conjunto de dados. A seleção dos componentes principais é um passo importante no *PCA*, uma vez que é necessário escolher um número adequado de componentes que seja capaz de capturar a maior parte da informação presente nos dados (FERREIRA, 2011; KASSAMBARA, 2017).

De acordo com Kaiser (1961), os autovalores podem ser usados para determinar o número de componentes principais a serem retidos após o *PCA*, isso dependerá do campo de aplicação e do conjunto de dados específico. Na prática, tendemos a olhar para os primeiros componentes principais para encontrar padrões interessantes nos dados. Nesse sentido, KASSAMBARA (2017) ressalta que :

Autovalores > 1. Indica que o componente é responsável por mais variância do que por uma das variáveis originais em dados padronizados. Isso é comumente usado como um ponto de corte para o qual os componentes são retidos. Isso só é verdade quando os dados são padronizados.

Ponto de corte. Pode-se limitar o número de componentes a esse número que representa uma certa fração da variância total.

Um método alternativo e muito utilizado para determinar o número de componentes principais é olhar para um *Scree Plot*, que é o gráfico de autovalores ordenados do maior para o menor. O número de componentes é determinado no ponto, além do qual os autovalores restantes são todos relativamente pequenos e de tamanho comparável (JOLLIFFE, 2002).

A partir do gráfico abaixo (Figura 3), podemos querer parar no quinto componente principal, por exemplo. Assim, 87% das informações (variâncias) contidas nos dados são retidas pelos cinco primeiros componentes principais.

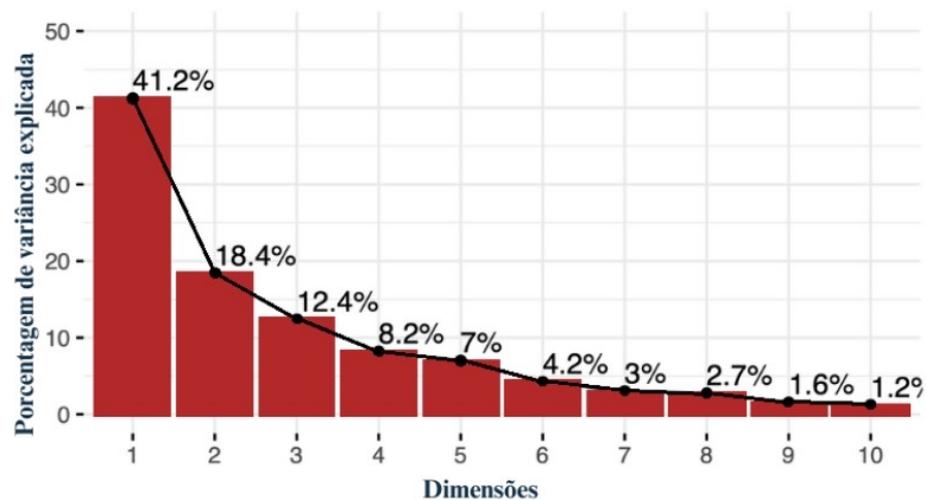


Figura 3 – Porcentagem de variância explicada pelos componentes principais em ordem decrescente.

A interpretação dos resultados do *PCA* envolve a compreensão da importância relativa dos componentes principais, a identificação de padrões nos dados e a seleção dos componentes principais para a redução de dimensionalidade. A aplicação adequada do *PCA* pode levar a conclusões importantes sobre os dados e contribuir para a solução de problemas em diversas áreas.

Diante do exposto, pode-se concluir que a análise de componentes principais agrupa os indivíduos de acordo com sua variação, isto é, os indivíduos são agrupados segundo suas variâncias, ou melhor, segundo seu comportamento dentro da população (Figura 2.5.1), representado pela variação do conjunto de características que define o indivíduo. Pode-se dizer que a técnica agrupa os indivíduos de uma população segundo a variação de suas características.

Por sua versatilidade em solucionar problemas intrínsecos à análise de dados, a *PCA* é amplamente utilizada em diversas áreas científicas, incluindo biologia, engenharia, finanças e

visão computacional. As principais soluções possíveis que essa técnica disponibiliza incluem:

Classificação. *PCA* pode ser utilizada para classificar dados em diferentes grupos, identificando padrões e características distintas em cada grupo (WOLD; RUHE; WOLD, 1987);

Extração de características. *PCA* pode ser usada para extrair características importantes dos dados, reduzindo a dimensão do conjunto de dados e fornecendo uma representação mais simples e compreensível. Para um contexto de modelagem, favorece a redução da complexidade do modelo podendo melhorar seu ajuste (JOLLIFFE, 2002);

Visualização de dados multidimensionais. A *PCA* pode ser usada para representar dados em dimensões menores, permitindo uma melhor visualização e compreensão dos padrões subjacentes nos dados, podendo ser útil em tarefas como segmentação, identificando *clusters* de clientes ou detectando anomalias (KARAMIZADEH et al., 2013);

Redução de dimensionalidade. *PCA* é frequentemente utilizada para reduzir a dimensionalidade de conjuntos de dados, permitindo uma melhor visualização e análise dos dados. Isso é especialmente útil quando se lida com grandes quantidades de dados e é necessário reduzir a complexidade do problema (ABDI; WILLIAMS, 2010);

Reconhecimento facial e de movimento. Sistemas de monitoramento de vídeo podem ser aprimorados com o uso dessa técnica, permitindo a identificação rápida e precisa de pessoas suspeitas ou em situações de perigo. Além disso, sistemas de reconhecimento facial podem ser utilizados para controle de acesso em locais restritos, como empresas, instituições financeiras e governamentais. Esses sistemas também têm potencial para auxiliar na localização de pessoas desaparecidas ou procuradas pelas autoridades.

Dentre as aplicabilidades do *PCA* mencionadas anteriormente, aqui serão ressaltadas brevemente outras aplicações passivas de serem realizadas com o *PCA*, em especial, o reconhecimento facial, por se tratar do objeto de estudo deste trabalho. Existe uma gama de aplicações possíveis para o *PCA* no contexto de reconhecimento facial, por exemplo: Na aplicação da lei, pode ser utilizado na identificação de suspeitos e rastreamento de criminosos. Também é usado em sistemas de segurança, como segurança de aeroportos, para detectar ameaças e prevenir o terrorismo. No marketing, é usado para personalizar anúncios e melhorar a experiência do cliente. Na área da saúde, é usado para diagnosticar certas condições médicas, como a síndrome de *Down* e distúrbios genéticos raros.

As principais vantagens do *PCA* são evidenciadas por Karamizadeh et al. (2013) e listadas abaixo;

Baixa sensibilidade ao ruído, a diminuição dos requisitos de capacidade e memória e maior eficiência devido aos processos que ocorrem em dimensões menores; as vantagens completas do *PCA* estão listadas abaixo:

1. Eliminação da redundância de dados a partir dos componentes ortogonais (DÖRSCH et al., 2022);
2. Redução da complexidade no agrupamento de imagens com o uso de PCA (ASADI; RAO; SAIKRISHNA, 2010);
3. Representação menor do banco de dados, pois apenas as imagens de treinamento são armazenadas na forma de suas projeções de modo reduzida (DÖRSCH et al., 2022);
4. Redução do ruído uma vez que é escolhida a base de variação máxima e assim as pequenas variações no fundo são automaticamente ignoradas.

Duas desvantagens principais estão associadas ao *PCA*:

1. A matriz de covariância é difícil de ser avaliada precisamente;
2. Mesmo a invariância mais simples não pode ser capturada pelo *PCA*, a menos que os dados de treinamento forneçam explicitamente essa informação.

Outra vantagem da *PCA* é que ela é uma técnica não supervisionada, o que significa que não requer informações prévias sobre as variáveis ou os grupos nos dados. Isso a torna uma ferramenta valiosa em aplicações em que as informações prévias sobre os dados são limitadas ou incertas.

Além das desvantagens mencionadas anteriormente, a *PCA* é sensível a *outliers* e também não é recomendada quando se tem mais variáveis do que unidades amostrais. Ao reduzir o número de variáveis, há perda da informação de variabilidade das variáveis originais. *PCA* nem sempre funciona; às vezes mesmo com a redução o número de componentes ainda continua grande. É o caso de variáveis originais pouco correlacionadas (HONGYU, 2015).

A análise de componentes principais (*PCA*), também conhecida como expansão de *Karhunen-Loeve*, foi apresentada por Hotelling (1933) em sua forma algébrica. A seguir são apresentadas as etapas envolvidas nessa técnica;

1. Dada a matriz de dados \mathbf{X} , composta por m atributos ou características observadas em n indivíduos de uma população. As características observadas são representadas pelas variáveis $X_1, X_2, X_3, \dots, X_m$. A matriz de dados é de dimensão $n \times m$

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix},$$

2. Vetor coluna com o valor médio para cada atributo.

$$\bar{\mathbf{X}} = \frac{1}{n} \sum_{i=1}^n x_i = \begin{bmatrix} \frac{1}{n} \sum_{i=1}^n x_{i1} \\ \frac{1}{n} \sum_{i=1}^n x_{i2} \\ \vdots \\ \frac{1}{n} \sum_{i=1}^n x_{im} \end{bmatrix}$$

3. A variância é uma medida de dispersão que indica o quanto os valores dos atributos se afastam da média.

A variância da i - ésima componente do vetor \mathbf{X} é calculada por:

$$\sigma_i^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.1)$$

4. Covariância mede a intensidade da relação entre duas características. Se eles são linearmente independentes, sua covariância é nula.

$$\text{Cov}(X_i, X_j) = \mathbf{S}_{ij}^2 = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{n-1} \quad (2.2)$$

5. Calculo da matriz de covariância: Matriz simétrica quadrada (m, m) das covariâncias

$$\text{Cov}(X) = \begin{bmatrix} \text{cov}(m_1, m_1) & \text{cov}(m_2, m_1) & \dots & \text{cov}(m, m_1) \\ \text{cov}(m_1, m_2) & \text{cov}(m_2, m_2) & \dots & \text{cov}(m, m_2) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(m_1, m_m) & \text{cov}(m_2, m_m) & \dots & \text{cov}(m_m, m_m) \end{bmatrix}$$

Seja \mathbf{p} um vetor-coluna ($m, 1$), a projeção de \mathbf{X} na direção \mathbf{p} é dada por $\hat{X} = \mathbf{p}^T X$. O objetivo do *PCA* é encontrar o vetor \mathbf{p} que maximiza a variância de \hat{X} .

$$\mathbf{p}_1 = \text{argmax}_{p=1} [\text{var}(\mathbf{p}^T \mathbf{X}_1)]$$

Diz-se que \mathbf{p}_1 é um componente principal de \mathbf{X} . Os demais podem ser calculados de forma semelhante.

$$\mathbf{p}_1 = \text{argmax}_{p=1} [\text{var}(\mathbf{p}^T \mathbf{X}_1)]$$

⋮

$$\mathbf{p}_1 = \operatorname{argmax}_{p=1} [\operatorname{var}(\mathbf{p}^T \mathbf{X}_1)]$$

\mathbf{p}_1 é um componente principal de \mathbf{X} . Os demais podem ser calculados de forma semelhante.

6. Obtenção dos autovalores e autovetores: A matriz de covariância pode ser diagonalizada através da decomposição em autovalores e autovetores. Os autovetores representam as direções dos novos eixos que maximizam a variância dos dados, enquanto os autovalores correspondem à variância explicada por cada novo eixo. Essa diagonalização pode ser realizada através da resolução da equação:

$$\Sigma \vec{v} = \lambda \vec{v} \quad (2.3)$$

onde \vec{v} é o autovetor, λ é o autovalor e Σ é a matriz de covariância. A matriz de covariância é geralmente usada em aplicações de redução de dimensionalidade, enquanto a matriz de correlação é usada em aplicações de extração de recursos.

7. Seleção de componentes principais: Os componentes principais são então selecionados com base nos autovalores, sendo que aqueles com autovalores maiores são considerados mais importantes na representação dos dados. É possível selecionar um número específico de componentes principais para a redução de dimensionalidade ordenando os autovalores em ordem decrescente e selecionando os autovetores correspondentes aos maiores autovalores ou então utilizar todos os componentes principais para a extração de recursos.
8. Projeção dos dados no novo espaço: A matriz de projeção pode ser obtida a partir dos autovetores selecionados. Os dados originais podem ser projetados nos novos eixos através da multiplicação da matriz de dados pela matriz de projeção. A matriz de projeção P é composta pelos k autovetores correspondentes aos k maiores autovalores da matriz de covariância, ou seja:

$$P = [v_1 \ v_2 \ \cdots \ v_k] \quad (2.4)$$

onde v_1, v_2, \dots, v_k são os autovetores correspondentes aos k maiores autovalores da matriz de covariância. Para projetar os dados originais X nos novos eixos, basta multiplicar X pela matriz de projeção P :

$$Y = X \cdot P \quad (2.5)$$

onde Y é a matriz de dados projetados nos novos eixos.

2.5.1.1 Eigenfaces

O *Eigenfaces*, também conhecido como autofaces, é um método baseado em uma variação do PCA introduzido por Turk e Pentland (1991) e aplicado ao reconhecimento de faces. Considerara-se cada pixel de uma imagem como um atributo, representado por uma dimensão no espaço de dados.

O método consiste em compor um espaço de dados formado pelas imagens de face e, por meio da PCA, gerar um espaço de características que reduz a dimensionalidade das imagens, chamado de espaço de faces. Os vetores de características originários desse processo são utilizados para realizar o reconhecimento de faces. Os vetores (autovetores) usados no reconhecimento facial são chamados de *Eigenfaces* porque, na prática, eles são vetores que representam características faciais típicas, como olhos, nariz e boca(SILVA, 2016). Nas Figuras 4 e 5 vemos com mais detalhes o funcionamento do método.

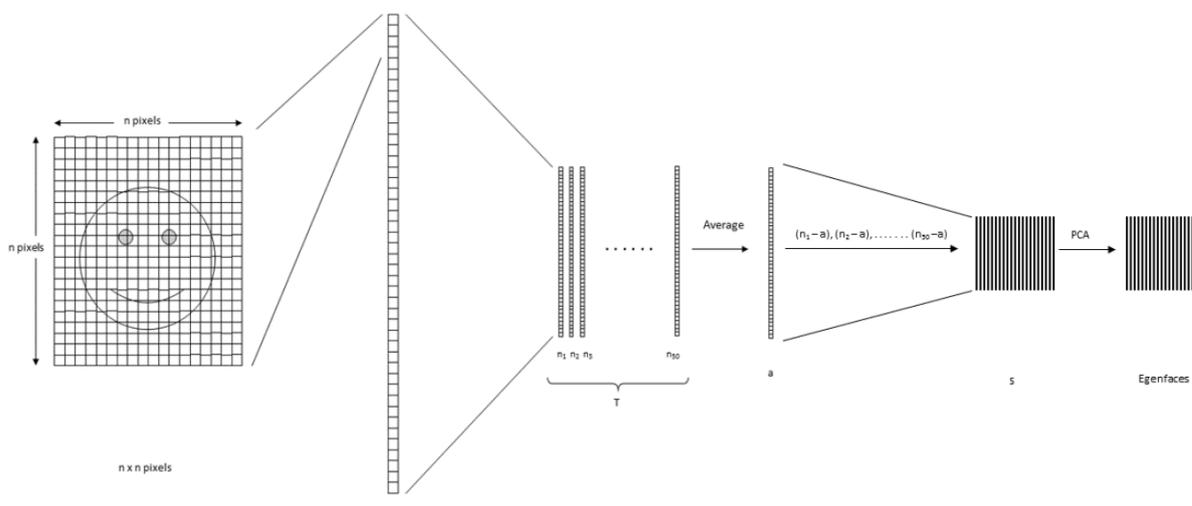


Figura 4 – Redução de características em uma imagem após aplicação do *Eigenfaces*. Imagem adaptada de Dinalankara (2017).

Silva (2016) descreveu o método *Eigenfaces* em 7 etapas:

1. Cálculo da face média: A face média é calculada como a média de todas as faces registradas na base de dados, ou seja, é a imagem que representa o padrão geral das faces na base de dados. A fórmula para o cálculo da face média é dada por:

$$\mu = \frac{1}{M} \sum_{i=1}^M x_i \quad (2.6)$$

onde μ é a face média, M é o número de faces na base de dados e x_i é a i -ésima face.

2. Subtração pela face média: Nesta etapa, cada face na base de dados é subtraída pela face média. A fórmula para o cálculo da subtração é dada por:

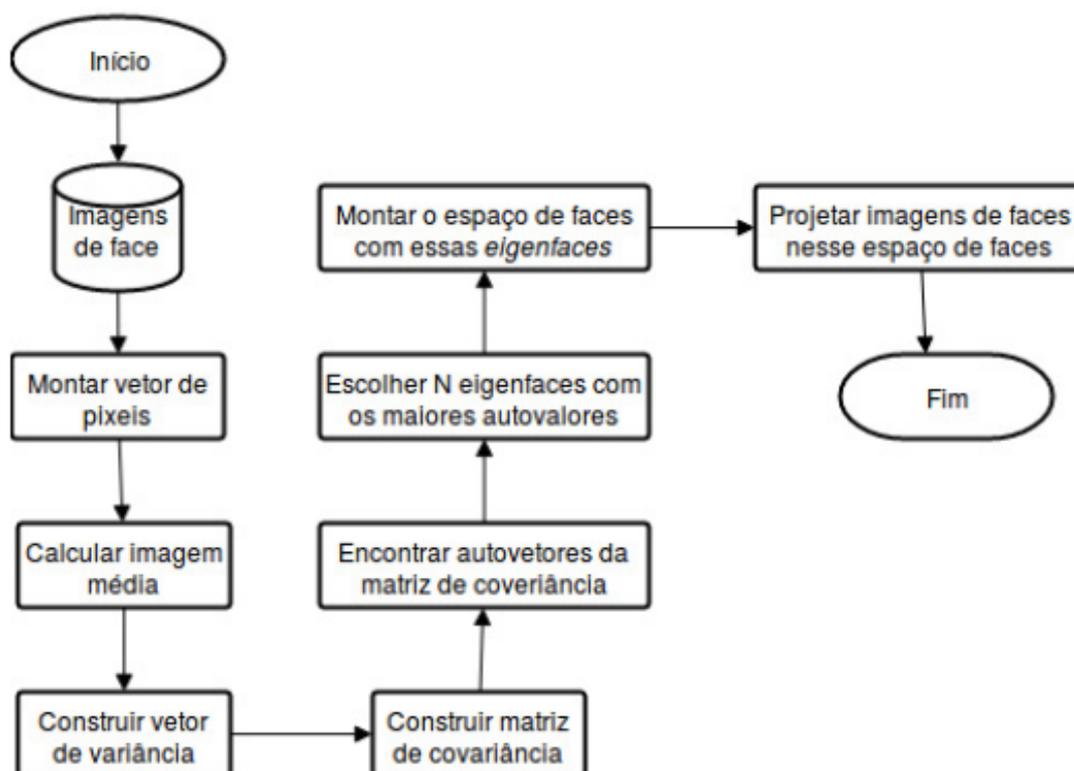


Figura 5 – Diagrama detalhando os passos do algoritmo *Eigenfaces*. Imagem adaptada de Lima e Neto (2017).

$$x'_i = x_i - \mu \quad (2.7)$$

onde x'_i é a i -ésima imagem resultante da subtração de uma imagem de treinamento pela face média. e x_i é a i -ésima face do conjunto de treinamento.

3. Transformação das imagens em vetores:

A transformação de uma imagem em um vetor consiste em concatenar as linhas da imagem, unindo o último pixel de cada linha com o primeiro pixel da linha seguinte, formando um vetor de pixels, como ilustrado na Figura 6.

4. Construção da matriz de covariância: A matriz de covariância é construída com base nas imagens subtraídas da face média transformadas em vetores. A fórmula para o cálculo da matriz de covariância é dada por:

$$S = \frac{1}{M} \sum_{i=1}^M x'_i (x'_i)^T \quad (2.8)$$

onde S é a matriz de covariância, M é o número de faces na base de dados, x'_i é a i -ésima imagem subtraída da face média e $(x'_i)^T$ é o transposto desta i -ésima imagem.

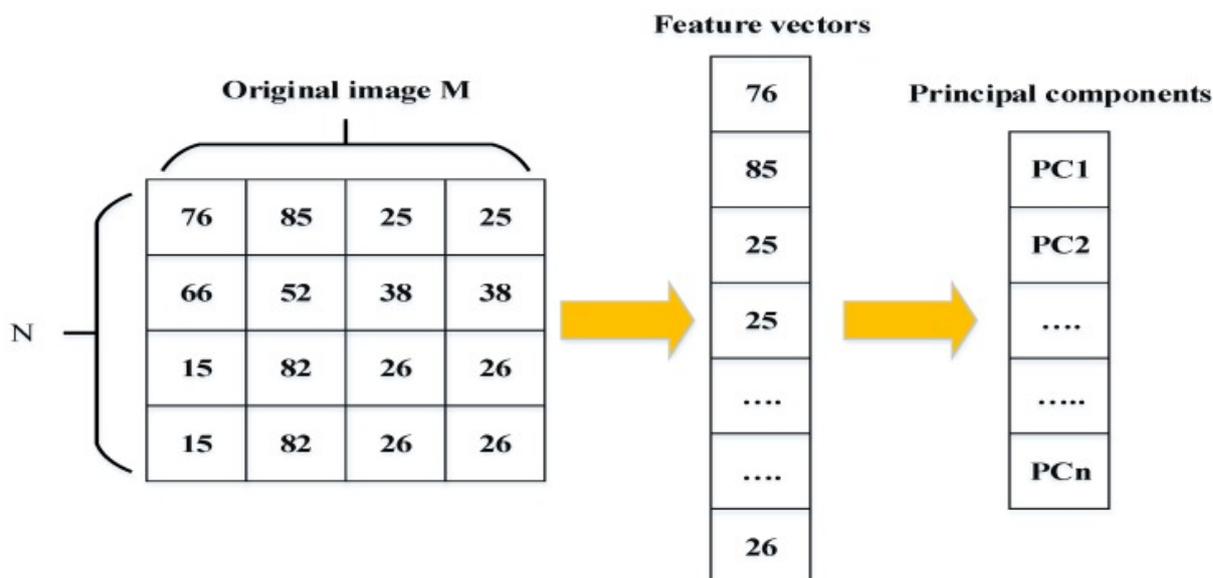


Figura 6 – Transformação de imagem em vetor, o que possibilita o cálculo das componentes principais. Imagem adaptada de Kortli et al. (2020).

5. Cálculo das *Eigenfaces*: A partir da matriz de covariância obtida na etapa 4, é possível calcular os autovetores e autovalores. Esses autovetores são chamados de *Eigenfaces* e representam as direções principais de variação na base de dados. A equação usada para calcular as *Eigenfaces* é:

$$S\vec{u} = \lambda\vec{u} \tag{2.9}$$

onde S é a matriz de covariância, \vec{u} é o autovetor e λ é o autovalor. O autovetor associado ao maior autovalor é o *Eigenfaces* que captura a direção de variação mais importante. Os autovetores são ordenados em ordem decrescente de seus respectivos autovalores, para que os mais significativos venham primeiro. Na Figura 5 vemos alguns exemplos dessas *Eigenfaces*.

6. Cálculo dos Vetores de Características: Os vetores de características são obtidos pela projeção dos vetores de imagem subtraídos pela face média nas *Eigenfaces*. A equação usada para calcular os vetores de características é:

$$\vec{w} = \vec{u}^T(x'_i - \mu) \tag{2.10}$$

onde \vec{w} é o vetor de características, \vec{u} é o autovetor (*Eigenface*), x'_i é o vetor de imagem subtraído pela face média e μ é a face média.

Ao projetar cada imagem do conjunto nas *Eigenfaces*, podemos obter os coeficientes que indicam a contribuição de cada *Eigenfaces* para a imagem. Por exemplo, considere uma



Figura 7 – Conjunto de exemplos de *Eigenfaces* redimensionadas para o tamanho das imagens do conjunto de treinamento. Imagem adaptada de Silva (2016).

imagem de um rosto que possui uma grande contribuição da primeira *Eigenface*, uma pequena contribuição da segunda *Eigenface* e nenhuma contribuição das demais *Eigenfaces*. Isso pode ser representado por um vetor de coeficientes, que pode ser usado para reconstruir a imagem original.

Esses coeficientes são usados para identificar as características mais relevantes de cada imagem. Por exemplo, uma imagem com coeficientes altos nas primeiras *Eigenfaces* pode ter características como olhos, sobrancelhas e nariz bem definidos, enquanto uma imagem com coeficientes altos nas últimas *Eigenfaces* pode ter características mais específicas, como rugas ou sardas.

O processo de projeção das imagens nos *Eigenfaces* é o que permite a redução da dimensionalidade das imagens, tornando possível representá-las de forma mais compacta.

7. Cálculo de Similaridade: Finalmente, o cálculo de similaridade é feito comparando os vetores de características das faces desconhecidas com os vetores de características das faces conhecidas. A similaridade pode ser calculada com uma medida de distância como a distância Euclidiana. A equação usada para calcular a distância Euclidiana é:

$$d = \sqrt{\sum_{i=1}^n (w_i - w'_i)^2} \quad (2.11)$$

onde d é a distância, w_i é o vetor de características da face desconhecida, w'_i é o vetor de características da face conhecida e n é o número de características.

A escolha dos autovetores com maior autovalor permite selecionar as direções de variação mais importantes, o que resulta em uma representação compacta e eficiente das imagens de faces.

2.5.2 Análise Discriminante Linear - LDA

O LDA (Linear Discriminant Analysis), ou Análise Discriminante Linear, é um método estatístico que visa classificar objetos a partir de variáveis descritivas. Esse método se baseia na ideia de que as variáveis descritivas de uma população de objetos possuem distribuições normais, e busca maximizar a separação entre as classes de objetos a partir da projeção dos dados em um subespaço dimensional reduzido. LDA gera um novo conjunto de dados de menor dimensionalidade que representa as classes dos dados originais (SILVA, 2016).

Ele foi proposto inicialmente por Fisher (1936) como uma técnica para análise de variância em dados com múltiplas classes, e posteriormente foi desenvolvido como uma técnica de redução de dimensionalidade para problemas de classificação. Desde então, o LDA tem sido amplamente utilizado em diversas áreas, sendo um dos métodos mais antigos e conhecidos nessa área. Segundo Belhumeur, Hespanha e Kriegman (1997), a LDA é uma técnica de classificação linear que tem se mostrado eficaz para o reconhecimento de face, pois busca maximizar a separação entre as classes de objetos a partir da projeção dos dados em um subespaço dimensional reduzido.

O primeiro passo do algoritmo LDA é a computação da matriz de dispersão entre classes (ou interclasses) e a matriz de dispersão dentro de classes (ou intraclasse). A matriz de dispersão interclasses mede a variação entre as médias das classes, enquanto a matriz de dispersão intraclasse mede a variação dentro de cada classe. A partir dessas matrizes, é possível calcular a matriz de dispersão total, que é a soma da matriz de dispersão interclasses e intraclasse.

Em seguida, o algoritmo LDA computa os vetores próprios da matriz de dispersão total, que correspondem às direções de projeção que maximizam a separação entre as classes. Esses vetores são ordenados de acordo com seus respectivos valores próprios, que indicam a quantidade de variação explicada por cada direção de projeção.

Finalmente, o algoritmo LDA seleciona as primeiras k direções de projeção com os maiores valores próprios e projeta os dados de entrada nesse novo espaço de características de menor dimensão. Essas projeções podem então ser usadas como entrada para um classificador, como por exemplo uma máquina de vetores de suporte (SVM).

Antes de aplicar LDA aos dados, geralmente são necessárias técnicas de pré-processamento para melhorar a qualidade e a eficácia do modelo. Uma técnica de pré-processamento comum usada na LDA é a normalização dos dados. A normalização pode ser feita por diferentes métodos, como normalização por z-score, normalização por min-max, normalização por escala e outros. A normalização ajuda a remover a influência de valores discrepantes, ajustar a escala dos dados e facilitar a comparação entre diferentes recursos.

Outra técnica de pré-processamento usada na LDA é a redução de dimensionalidade. Isso pode ser feito usando técnicas de seleção de recursos ou técnicas de extração de recursos. As técnicas de seleção de recursos selecionam um subconjunto de recursos relevantes, enquanto as técnicas de extração de recursos transformam os recursos originais em um novo conjunto de recursos de menor dimensão. As técnicas comuns de redução de dimensionalidade incluem Análise de Componentes Principais (PCA), Análise Fatorial (FA) e Análise Canônica de Correlação (CCA).

É importante notar que, embora a normalização e a redução de dimensionalidade possam melhorar a qualidade dos dados, essas técnicas também podem introduzir alguma perda de informação. Portanto, é necessário avaliar cuidadosamente os efeitos dessas técnicas em relação à eficácia da LDA.

Além disso, a Análise Discriminante Linear é computacionalmente mais simples e mais rápida do que outros métodos de classificação, o que a torna uma alternativa viável para o reconhecimento de face em larga escala. Contudo, o desempenho da LDA é fortemente afetado pela quantidade de variáveis descritivas utilizadas, e pode sofrer com a presença de ruído ou variações nas imagens.

Uma particularidade da técnica de análise discriminante linear é a projeção dos dados com o intuito de maximizar a separação de classes, considerando as diferenças entre elas. Para o seu procedimento, essa técnica leva em consideração a informação *a priori* sobre as classes presentes no conjunto de dados¹ (HASTIE et al., 2009).

Em contrapartida, a separação das classes é obtida por PCA através da maximização da variação total dos dados, sem levar em consideração nenhuma informação sobre as classes dos dados².

A técnica é amplamente utilizada em diversas áreas, incluindo reconhecimento de padrões e análise de dados. Segundo Hastie et al. (2009), a LDA é uma técnica de classificação supervisionada que busca encontrar uma combinação linear de variáveis que maximize a separação entre duas ou mais classes de dados. Isso a torna uma ferramenta valiosa em aplicações como reconhecimento de faces, análise de dados biológicos e análise de sentimentos em textos.

2.5.2.1 *Fisherfaces*

Fisherfaces é uma técnica de reconhecimento de face desenvolvida por Belhumeur, Hespanha e Kriegman (1997) que utiliza uma combinação de Análise de Componentes Principais (PCA) e Análise Discriminante Linear (LDA). O objetivo deste método é encontrar uma representação dos rostos na forma de uma base de características, que maximiza a separabilidade discriminante entre classes.

Silva (2016) descreveu o método *Fisherfaces* em 8 etapas:

¹ Procedimento supervisionado

² Procedimento não supervisionado

1. Cálculo da face média por classe: para cada classe, é necessário calcular a média de todas as imagens de rosto nessa classe. Isso pode ser feito pela divisão da soma de todas as imagens na classe pela quantidade de imagens na classe. Matematicamente, seja Z_{ij} uma imagem de face como sendo uma matriz de tamanho N (da mesma forma que (2.6)) projetada no subespaço PCA, em que a j -ésima imagem de uma classe de imagens C_i .

A face média (μ_i) da classe C_i , é obtida por:

$$\mu_i = \frac{1}{N_i} \sum_{j=1}^{N_i} Z_{ij}. \quad (2.12)$$

2. Cálculo da face média geral: é necessário calcular a média de todas as imagens de rosto independentemente da classe. Isso pode ser feito pela divisão da soma de todas as imagens pela quantidade total de imagens.

A média total (μ) dessas faces é dada por:

$$\mu = \frac{1}{\sum_{i=1}^r N_i} \sum_{i=1}^r \sum_{j=1}^{N_i} Z_{ij}, \quad (2.13)$$

onde N_i é o número de faces pertencentes a cada uma das r classes pré determinadas.

3. Transformação das imagens em vetores: cada imagem de rosto precisa ser transformada em um vetor para que possa ser manipulado matematicamente. Isso pode ser feito através da conversão da imagem em uma representação numérica.
4. Construção da matriz de dispersão intra-classe: a matriz de dispersão intra-classe é construída a partir da diferença entre cada imagem na classe e a média da classe. Isso pode ser feito pela subtração de cada vetor de imagem da face média da classe.
5. Construção da matriz de dispersão intraclasses: a matriz de dispersão intraclasses (dentro de cada classe) S_W é definida por:

$$S_W = \sum_{i=1}^r \sum_{j=1}^{C_i} (Z_{ij} - \mu_i) (Z_{ij} - \mu_i)^T = \sum_{i=1}^r S_{W_i}, \quad (2.14)$$

onde S_{W_i} é uma matriz de covariância da classe C_i e S_W representa a dispersão das imagens de faces Z_{ij} em relação à face média de sua classe.

6. Construção da matriz de dispersão interclasses: a matriz de dispersão interclasses é construída a partir da diferença entre a média geral de todas as imagens de rosto e a média de cada classe. Isso pode ser feito pela subtração da face média geral de cada face média de classe.

A matriz de dispersão interclasses (entre classes) S_B é definida por:

$$S_B = \sum_{i=1}^r N_i (\mu_i - \mu) (\mu_i - \mu)^T, \quad (2.15)$$

ou seja, S_B representa a dispersão das faces médias de cada classe em relação à média total.

7. Cálculo das *Fisherfaces*: as *Fisherfaces* são obtidas a partir da solução das matrizes de dispersão intra-classe e inter-classe. Essas faces são os *eigenvectors* das matrizes de dispersão que maximizam a separação entre as classes.
8. Cálculo dos vetores de características: os vetores de características são obtidos a partir da projeção de cada imagem de rosto nas *Fisherfaces*. Isso pode ser feito pela multiplicação da imagem pela matriz de *Fisherfaces*.
9. Cálculo da similaridade: a similaridade é calculada a partir da distância entre os vetores de características de duas imagens de rosto. A distância mais curta entre os vetores indica a maior similaridade entre as imagens.

3 Material e Métodos

A princípio serão detalhadas as informações quanto ao banco de dados, software e bibliotecas utilizadas para atingir os objetivos propostos e testar as hipóteses experimentais formuladas. As etapas a serem seguidas para a implementação do sistema de reconhecimento facial estão descritas no fluxograma da Figura 8.

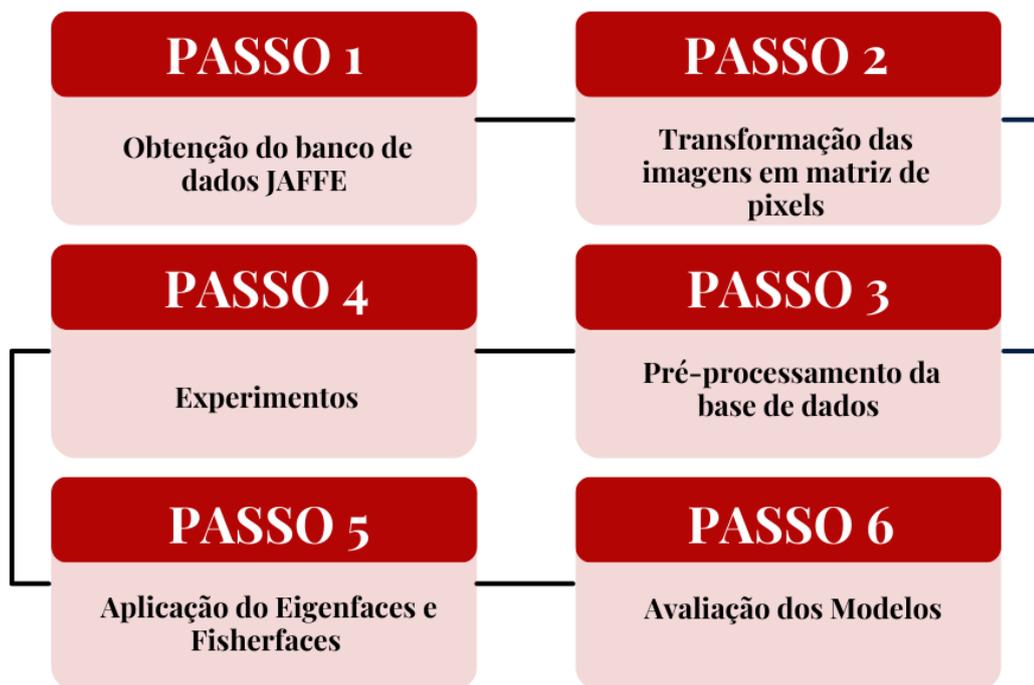


Figura 8 – Fluxograma do sistema de reconhecimento facial proposto.

3.1 Banco de dados

O banco de dados *JAFFE* (Japanese Female Facial Expression) foi criado em 1998 por Miyuki Kamachi, Michael Lyons e Jiro Gyoba da Universidade de Kyushu, no Japão (LYONS; KAMACHI; GYOBA, 2020; LYONS, 2021). A construção desse banco teve como objetivo disponibilizar imagens para os diversos estudos que envolvem expressões faciais de mulheres japonesas. A base *JAFFE* é amplamente utilizada para treinamento e validação de algoritmos de

reconhecimento facial e de expressões por pesquisadores da área de visão computacional, AM, IA e estatística.

O conjunto de dados possui 213 imagens em sete expressões faciais básicas (alegria, tristeza, raiva, surpresa, medo, aversão (nojo) e neutra) de dez mulheres japonesas entre as idades de 19 e 30 anos. As imagens foram adquiridas sob condições controladas de iluminação e fundo neutro. Cada imagem está em tons de cinza, com 256 níveis de intensidade, e tem uma resolução de 256×256 pixels de 8 bits. Todas as imagens estão disponibilizadas para *download* em formato *.tiff* (Figura 9).

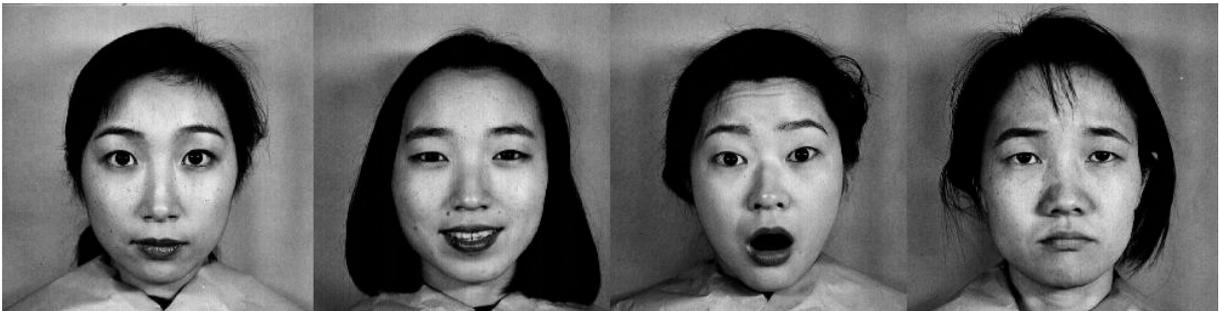


Figura 9 – Imagens de indivíduos do banco *JAFFE* em diferentes expressões faciais.

A figura 10 mostra a identificação de cada imagem no conjunto de dados. Cada descrição guarda as seguintes informações;

INICIAIS DO NOME. Inicialmente, duas letras maiúsculas representam as iniciais do nome de cada indivíduo;

IDENTIFICAÇÃO DA EXPRESSÃO. Duas letras maiúsculas representam a expressão de cada indivíduo na imagem.

REPETIÇÃO. Algarismo informando a quantidade de imagens fornecidas do mesmo indivíduo.

ID. Algarismo informando a identificação de cada indivíduo no banco.

Para obter-se a base de dados *JAFFE* é necessário realizar uma solicitação aos pesquisadores que a criaram. Antes de baixá-la, é importante conferir as restrições de uso e licenças.¹ As imagens podem ser utilizadas apenas para pesquisas científicas não comerciais e para qualquer outro uso requer aprovação prévia.

As imagens disponíveis no banco *JAFFE* encontram-se em escala de cinza. A opção de conversão em escala de cinza é muito útil para imagens capturadas que não precisam corresponder a detalhes coloridos, além de simplificar os algoritmos e eliminar as complexidades relacionadas

¹ Para informações detalhadas quanto à base de dados, restrições e licenças acesse: <<https://zenodo.org/record/3451524#.ZAYqcXbMLIU>>



Figura 10 – Exemplo da imagem de um indivíduo da base *JAFFE* com sua respectiva identificação.

Tabela 1 – Distribuição das imagens do banco de imagens *JAFFE*. Abreviaturas: NE - Neutro, HA - feliz, SA - Tristeza, SU- Surpresa, AN - Raiva, DI - Nojo.

Iniciais dos Nomes	Expressões Faciais							Total
	AN (Raiva)	DI (Nojo)	FE (Medo)	HA (Feliz)	NE (Neutra)	SA (Triste)	SU (Surpresa)	
KA	3	3	4	4	3	3	3	23
KL	3	4	3	3	3	3	3	22
KM	3	2	3	4	3	4	3	22
KR	3	3	3	2	3	3	3	20
MK	3	3	3	3	3	3	3	21
NA	3	3	3	3	3	3	3	21
NM	3	2	3	3	3	3	3	20
TM	3	3	3	3	3	3	3	21
UY	3	3	3	3	3	3	3	21
YM	3	3	4	3	3	3	3	22
Total	30	29	32	31	30	31	30	213

aos requisitos computacionais. Isso ocorre porque a escala de cinza compacta uma imagem em seu pixel mínimo. Diante disso, todas as imagens utilizadas são compostas de tons de cinza com valores médios de *pixel* que variam entre 0 (Preto) e 255 (Branco).

3.2 Python

Toda a implementação necessária para o desenvolvimento do **sistema** de reconhecimento de pessoas mediante um conjunto de imagens desses indivíduos exercendo expressões básicas foram escritas em linguagem Python Core Team (2023), do mesmo modo que a análise dos resultados. *Python* é uma linguagem amplamente utilizada em diversos campos, incluindo reconhecimento de imagens. Possui uma vasta comunidade de desenvolvedores e bibliotecas de código

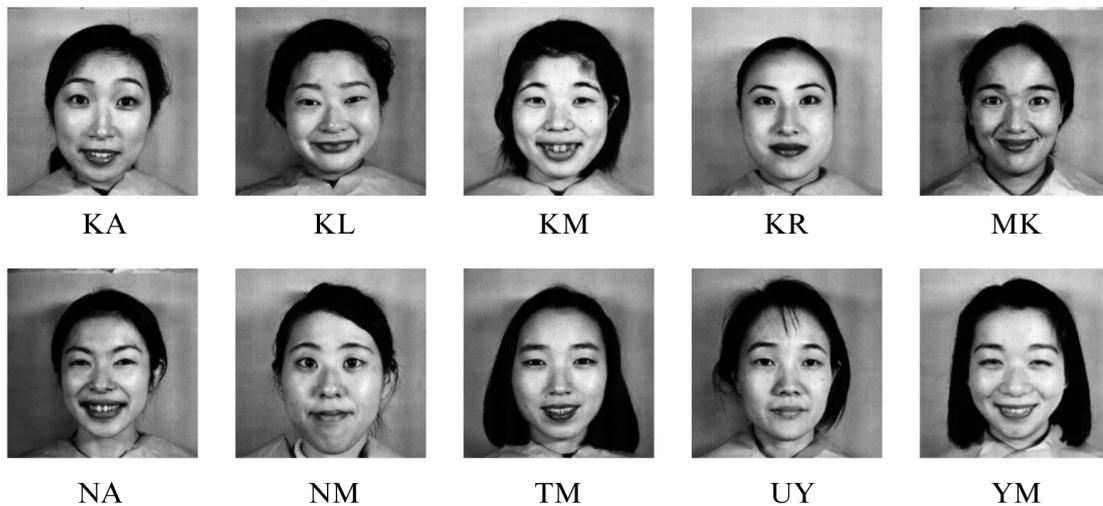


Figura 11 – Imagens com a expressão de felicidade de todas as mulheres da base *JAFFE* e suas respectivas iniciais do nome.

aberto, tornando-se uma escolha popular para aplicações de visão computacional, inteligência artificial, aprendizado de máquinas e estatística. Além disso, a linguagem Python é conhecida por sua acessibilidade, simplicidade e rápida progressão de desempenho ou performance da aprendizagem pelo usuário.

Para compilação dos códigos utilizou-se o *Jupyter Notebook*. Refere-se a um ambiente de desenvolvimento interativo (IDE)² proposto por Kluyver et al. (2016). O *Jupyter Notebook* corresponde a um produto totalmente em código aberto que proporciona aos usuários diversas funcionalidades de forma gratuita, além de permitir executar código em um navegador da web. O projeto foi nomeado pela primeira vez *IPython* e mais tarde renomeado para *Jupyter*, mais precisamente em 2014. Ele suporta mais de 40 linguagens, incluindo Julia, Python, R e Scala.

3.2.1 Manipulação das figuras com Pyplot

Os gráficos e figuras presentes neste trabalho foram obtidos através de funções pertencentes ao módulo³ **pyplot** da biblioteca **Matplotlib**. Essa biblioteca é usada para criar visualizações estáticas, animadas e interativas em Python. O módulo **pyplot** detém uma coleção de funções que são responsáveis pela criação e manipulação de figuras e gráficos.

² Integrated Development Environment

³ Módulo é um arquivo Python com instruções e comandos que pode conter recursos e estruturas como funções e variáveis.

3.3 Transformação da imagem em matriz de pixels

A imagem de cada indivíduo é representada por uma matriz de *pixels* com dimensão 256×256 , como ilustrado na Figura (12). A conversão das imagens em matrizes de *pixels* foram obtida a partir da função `array()`, que está disponível na biblioteca `numpy`. A biblioteca `numpy` suporta o processamento de grandes arranjos e matrizes, juntamente com uma grande coleção de funções matemáticas de alto nível para operar sobre estas matrizes.

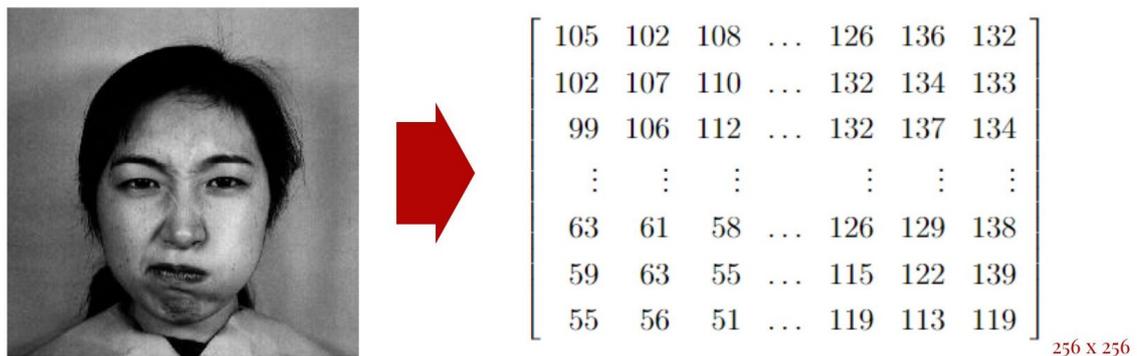


Figura 12 – Transformação da imagem em matriz de pixels. Cada *pixel* na imagem pode ser representado na escala de cinza com valores que variam entre 0 (Preto) e 255 (Branco).

3.4 Pré-processamento da base de dados

Realizou-se um pré-processamento no banco de imagens JAFFE a fim de separá-lo de acordo com as expressões faciais e ter o rótulo associando as imagens à pessoa a quem pertencem. Nesta etapa foram extraídas da descrição de cada imagem as informações referentes às iniciais da pessoa e expressão apresentada na imagem usando a função `listdir()` da biblioteca `os`.

Em seguida, com o auxílio da função `dataframe()` da biblioteca `Pandas` foi construído um *dataframe* contendo as imagens em formato de matriz de pixels, as iniciais do nome e o ID da expressão facial. Por fim a coluna referente ao ID foi renomeada para valores numéricos pois as funções *Eigenfaces* e *Fisherfaces* necessitam de um vetor numérico como rótulo no treinamento. A Tabela 2 mostra o como ficou a disposição dos dados.

3.5 OpenCV

A implementação dos métodos *Eigenfaces* e *Fisherfaces* que são objetos de estudo neste trabalho podem ser encontrados na biblioteca `OpenCv` (*Open Source Computer Vision Library*).

Tabela 2 – *Dataframe* contendo os dados usados para treinamento após processo de pré-processamento e organização do banco de dados JAFFE

ID	IMAGENS	INICIAIS DO NOME	ID DA EXPRESSÃO
1	[[105, 102, 108, 118, 114, 117, 113, 122, 113,...	AN	0
2	[[114, 115, 116, 118, 118, 114, 113, 118, 116,...	AN	0
3	[[96, 98, 114, 103, 102, 103, 110, 104, 103, 1...	AN	0
4	[[109, 118, 102, 108, 119, 109, 114, 110, 109,...	DI	0
5	[[114, 118, 115, 114, 116, 107, 114, 120, 111,...	DI	0
...
211	[[166, 165, 161, 166, 167, 174, 162, 178, 166,...	SU	9
212	[[143, 155, 161, 154, 151, 150, 155, 152, 153,...	SU	9
213	[[157, 151, 162, 167, 167, 164, 173, 162, 169,...	SU	9

Essa biblioteca foi desenvolvida pela **Intel**⁴ em 1999 com o intuito de criar uma ferramenta para ajudar os pesquisadores em visão computacional a desenvolver seus projetos de forma mais rápida e eficiente (BRADSKI, 2000). A **OpenCV** é escrita em linguagem *C++* e oferece suporte a várias linguagens de programação, incluindo *Python*. Além de reconhecimento de faces, a biblioteca provê diversas funções relacionadas à análise de imagens, detecção de objetos, processamento de vídeo e inteligência artificial.

A função **createEigenFaceRecognizer()** é parte da **OpenCV** e é usada para reconhecer rostos usando a técnica *Eigenfaces*. Ela examina todas as imagens de treinamento de todas as pessoas como um todo e tenta extrair os componentes que são relevantes e úteis e descarta o resto (BRADSKI, 2000). O *Eigenfaces* é treinado extraindo componentes principais e mantém um registro de quais componentes pertencem a qual pessoa. Assim, sempre que for introduzido uma nova imagem no algoritmo, ele repete o mesmo processo da seguinte forma (SuperDataScience Team, 2017):

1. Extrai os componentes principais da nova imagem.
2. Compara esses recursos com a lista de elementos armazenados durante o treinamento.
3. Encontra aqueles com a melhor correspondência.
4. Retorna o rótulo da 'pessoa' associado a esse componente de melhor correspondência.

Os principais parâmetros que podem ser fornecidos para esta função são: *num_components*(número de componentes) e *threshold*(limiar).

⁴ A Intel é uma multinacional americana líder mundial na fabricação de processadores para computadores pessoais. É conhecida também por seus avanços na pesquisa e desenvolvimento de tecnologias avançadas na área de inteligência artificial.

- `num_components`: Indica o número de componentes principais a serem mantidos. É ele quem especifica o número de componentes principais a serem mantidos na técnica *Eigenfaces*. O valor padrão é especificado como 0, o que significa que todos os componentes principais são mantidos. No entanto, em geral, é comum reduzir o número de componentes principais para reduzir o tempo de processamento e o espaço de armazenamento necessário para o modelo. O valor do parâmetro pode ser ajustado com base no tamanho do conjunto de dados de treinamento, na complexidade da tarefa de reconhecimento facial e nos recursos disponíveis para processamento.
- `threshold`: O *threshold* (limiar, em português) é um parâmetro que pode ser fornecido na função `createEigenFaceRecognizer()` para determinar se uma face é reconhecida ou não. Quando uma imagem de teste é comparada com as imagens de treinamento, o *threshold* determina se as diferenças entre as imagens são grandes o suficiente para serem consideradas como uma correspondência ou se devem ser consideradas como não correspondentes. Quanto maior o valor do *threshold*, mais restritivo é o reconhecimento facial e menor é a probabilidade de falsos positivos, mas também aumenta a probabilidade de falsos negativos.

A função *Fisherfaces* do **OpenCV** possui funcionamento e parâmetros semelhantes aos da função *Eigenfaces*. O número de componentes também pode ser ajustado usando o parâmetro `num_components` na função `createFisherFaceRecognizer()`. Porém para o *Fisherfaces*, o número máximo de autovetores que podem ser extraídos é $C - 1$, onde C é o número de classes presentes no conjunto de treinamento. Isso ocorre porque, para separar as classes, precisamos encontrar um subespaço de características que maximize a razão das variações entre classes e dentro das classes. Qualquer projeção adicional não traria informações adicionais discriminatórias. Além disso, teria um efeito negativo na eficiência computacional do algoritmo e na capacidade de generalização do modelo.

Existem algumas considerações que devem ser consideradas, sobre usar os métodos *Eigenfaces* e *Fisherfaces*: as imagens de treinamento e teste devem estar em escala de cinza e devem ter o mesmo tamanho (BRADSKI, 2000).

3.6 Criação das bases de Treinamento e Teste

Como visto na seção 2.3, diversas áreas podem ser beneficiadas pela técnica de reconhecimento facial. No entanto, nesse trabalho procuramos relacionar a aplicabilidade dos algoritmos a um problema mundial e de contexto real, dessa forma, esperamos que a escolha dos experimentos e entendimento dos resultados fiquem mais claros, revelando a contribuição desse trabalho.

Os experimentos foram definidos fixando uma hipótese de aplicação no contexto de segurança pública, especificamente, no reconhecimento ou identificação facial de pessoas desa-

parecidas ou perdidos. Com o intuito de contrastar a hipótese formulada em nosso objetivo, a definição do conjunto de treinamento e teste incluíram imagens com as seguintes emoções:

EXPERIMENTO 1: Treinamento - Neutra. A escolha de treinarmos os algoritmos com essa expressão se dá pelo fato de que, em um contexto real a chance de adquirir imagens de desaparecidos ou perdidos expressando essa emoção é alta, visto que tal expressão está presente na maioria dos documentos pessoais, quiçá em todos. Esse conjunto possui disponível 30 imagens correspondentes a 10 indivíduos;

EXPERIMENTO 2: Treinamento - Felicidade. A escolha de treinarmos os algoritmos com esse tipo de expressão se dá por dois motivos, primeiro, para um contexto real, na ausência de documentos pessoais ou imagens do indivíduo expressando emoção neutra, é comum ter em posse da família do desaparecido ou perdido a disponibilidade em seus arquivos pessoais registro(s) desse suposto indivíduo expressando tal emoção. Salvo que a chance de encontrar imagens com essa expressão alocadas em redes sociais também é altíssima. O segundo motivo está relacionado aos resultados obtidos em pesquisas sobre o processo emocional, psicologia e envelhecimento. Tais estudos afirmam que memórias para informações emocionais mostram-se melhor do que memórias para informações neutras, vide (SCHNITZSPAHN et al., 2014). Isso possibilita validar esse comportamento experimental através dos algoritmos escolhidos, reconhecendo a eficiências no reconhecimento a partir de cada expressão e comparando-os. Esse conjunto teve disponível 31 imagens correspondentes a 10 indivíduos;

EXPERIMENTO 3: Treinamento - Neutro + Felicidade. A escolha dessa combinação de expressões está relacionada a disposição desses registros, para um contexto real. Além de verificarmos se a combinação dessas expressões influenciam no poder de reconhecimento dos métodos. Esse conjunto teve disponíveis 61 imagens correspondentes a 10 indivíduos.

Os métodos foram aplicados aos 3 experimentos através das funções *Eigenfaces* e *Fisherfaces*, disponíveis na biblioteca **OpenCV**. Para treinar os algoritmos foi utilizada a função **train()**. Dentro dos parênteses foi preciso especificar o *array* onde estão os rostos ou imagens de treinamento, enquanto o segundo parâmetro corresponde aos rótulos. É possível salvar este modelo treinado e depois lê-lo em outro *script*, o que evita treiná-lo novamente. Para isso, usa-se a função **write()** e dentro dos parênteses devemos especificar o nome que queremos atribuir ao modelo, junto com a extensão *.XML*. Isso é uma vantagem, pois em alguns casos o treinamento leva tempo, ainda mais se o banco de dados for extenso.

Após aplicados à base de treino, os algoritmos estão prontos para inferir se uma imagem pertence ou não a um certo indivíduo. A previsão dos rótulos identificadores de cada indivíduo foram obtidos com a aplicação da função **predict()** à base de teste. Desse modo, aloca-se dentro dos parênteses o *array* correspondente as imagens de teste.

BASE DE DADOS - TESTE As expressões consideradas para o treinamento dos métodos foram ignoradas no conjunto de teste, o que refletiu uma base de dados com 152 imagens correspondentes a 10 indivíduos, distribuídas em 5 emoções básicas; tristeza, raiva, surpresa, medo e nojo. A tabela 3 mostra para cada emoção, as quantidades de imagens consideradas na base de teste. Vale salientar que as iniciais dos nomes (rótulos) dos indivíduos não são disponibilizadas.

Tabela 3 – Quantidade de imagens utilizadas na base de teste

AN (Raiva)	DI (Nojo)	FE (Medo)	SA (Triste)	SU (Surpresa)	Total
30	29	32	31	30	152

3.7 Avaliação dos modelos

A avaliação da eficiência dos algoritmos de reconhecimento de faces é uma tarefa importante para garantir a qualidade e confiabilidade dos resultados obtidos. Para tal, existem diversas ferramentas estatísticas que podem ser utilizadas. As métricas de avaliação de desempenho utilizadas neste trabalho foram:

- Matriz de Confusão: é uma tabela composta por quatro elementos principais: verdadeiros positivos (VP), falsos positivos (FP), verdadeiros negativos (VN) e falsos negativos (FN).

Tabela 4 – Matriz de Confusão

Valor Real	Valor Predito	
	Positivo	Negativo
Positivo	Verdadeiro Positivo(VP)	Falso Negativo(FN)
Negativo	Falso Positivo(FP)	Verdadeiro Negativo(VN)

VP indicam a quantidade de amostras classificadas corretamente como pertencentes à classe positiva, FP indicam incorretamente como pertencentes à classe positiva, VN indicam corretamente como pertencentes à classe negativa e FN indicam incorretamente como pertencentes à classe negativa. A partir desses elementos, podemos calcular diversas métricas de avaliação de desempenho, como precisão, sensibilidade, acurácia, entre outras.

- Acurácia: é a proporção entre o número de acertos e o total de amostras. É dada pela equação:

$$\text{Acurácia} = \frac{\text{VP} + \text{VN}}{\text{Total}} \quad (3.1)$$

Em que "VP" corresponde ao número de verdadeiros positivos, "VN" corresponde ao número de verdadeiros negativos e "Total" corresponde ao número total de amostras.

- Sensibilidade: também conhecida como *Recall*, é uma métrica de avaliação de desempenho comumente usada em sistemas de classificação. Ela representa a proporção de exemplos positivos corretamente identificados pelo classificador em relação ao número total de exemplos positivos.

A sensibilidade pode ser expressa como:

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (3.2)$$

No entanto, é importante destacar que uma sensibilidade elevada não necessariamente implica em uma boa precisão do classificador, pois ela pode gerar muitos falsos positivos (FP), ou seja, classificar instâncias como positivas mesmo que sejam negativas. Portanto, a sensibilidade deve ser avaliada em conjunto com outras métricas de avaliação de desempenho, como a precisão e F1-Score (SILVA, 2015).

- Especificidade: é uma medida que avalia a capacidade de um modelo de reconhecimento de faces de identificar corretamente os indivíduos que não são o alvo. É calculada como a proporção de verdadeiros negativos (VN) em relação ao total de negativos (VN + FP), ou seja:

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (3.3)$$

A especificidade é uma medida complementar à sensibilidade, que avalia a capacidade de identificar corretamente os indivíduos alvo. Juntas, essas medidas permitem avaliar a qualidade do modelo de reconhecimento de faces em relação aos erros de classificação.

Uma alta especificidade indica que o modelo tem baixa taxa de falsos positivos, ou seja, poucas pessoas são incorretamente identificadas como sendo o alvo. Por outro lado, uma baixa especificidade indica que o modelo tem alta taxa de falsos positivos, o que pode comprometer a segurança do sistema de reconhecimento de faces.

- Precisão: mede a proporção de acertos entre os positivos previstos. Ela é calculada por meio da equação:

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (3.4)$$

- F1-Score: é a média harmônica entre Precisão e Sensibilidade. É dada pela equação:

$$\text{F1-Score} = 2 * \frac{\text{Precisão} * \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}} \quad (3.5)$$

A vantagem do F1 Score é que ele combina precisão e sensibilidade em uma única métrica, o que torna a avaliação do desempenho do modelo mais abrangente.

A **Scikit-learn** é outra biblioteca *Python* de código aberto que suporta aprendizagem supervisionada e não supervisionada. Também fornece várias ferramentas para ajuste de modelo, pré-processamento de dados, seleção de modelo, avaliação de modelo, e muitos outros utilitários. O **sklearn.metrics** é um módulo da biblioteca **Scikit-learn** que inclui funções de pontuação, métricas de desempenho e cálculos de distância. Para calcular a matriz de confusão usou-se a função **confusion_matrix()** que recebe como parâmetros os valores reais e os valores preditos pelo classificador. Outra função usada foi **classification_report()**, que também recebe estes mesmos parâmetros e retorna um relatório de texto mostrando o resultado das principais métricas de classificação.

4 Resultados e Discussões

Neste capítulo são apresentados os resultados alcançados durante todo o trabalho.

4.1 *Eigenfaces*

Nesta seção serão apresentados os testes realizados para definir o número ótimo de dimensões dos vetores de características para o *Eigenfaces*. Os vetores de características utilizados são extraídos pela técnica PCA, que reduz a dimensionalidade das imagens.

É interessante avaliar o número ideal de componentes principais para que se consiga reduzir a dimensionalidade dos dados. O algoritmo *Eigenfaces* foi inicialmente treinado com as imagens neutras. As *Eigenfaces* geradas para as 30 imagens de treino são apresentadas na Figura 13.

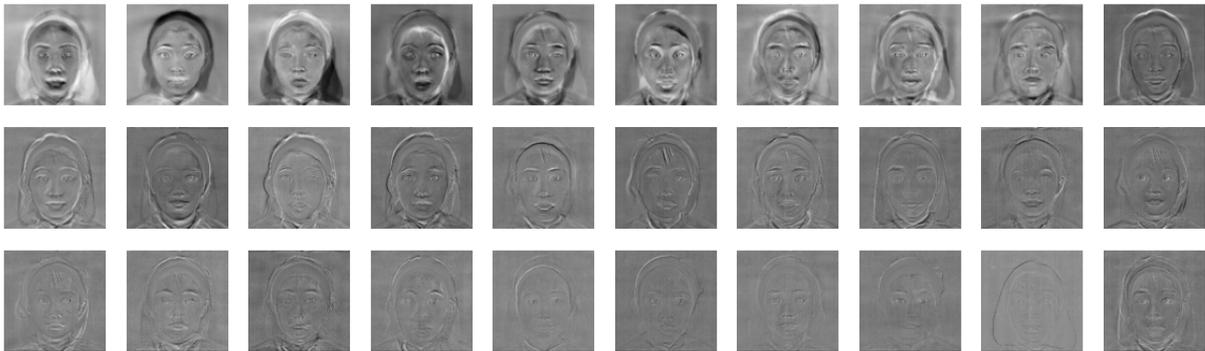


Figura 13 – Eigenfaces geradas a partir das imagens neutras do dataset JAFFE.

Para definir o número ideal de componentes no *Eigenfaces*, calculou-se a variância explicada para cada componente. No gráfico da Figura 14, está traçada a relação entre a variância acumulada das *Eigenfaces* e o número *Eigenfaces* geradas pela PCA. A variância acumulada, no eixo vertical, é uma medida que expressa o percentual da variação do banco de imagens presentes nas *Eigenfaces*. O número de componentes, no eixo horizontal, indica o número total de *Eigenfaces* geradas pela PCA para o banco de imagens neutras. A linha vertical indica que a partir do nono componente, o aumento da variância passa a não ser muito significativo.

A tabela 5 mostra a variância cumulativa explicada. Vemos que com 9 componentes o modelo ainda explica aproximadamente 92,38 da variância total. De acordo com o que foi apresentado no gráfico e na tabela, decidimos reduzir o número de componentes para 9, pois assim tivemos uma redução de dimensionalidade expressiva para o método *Eigenfaces*.

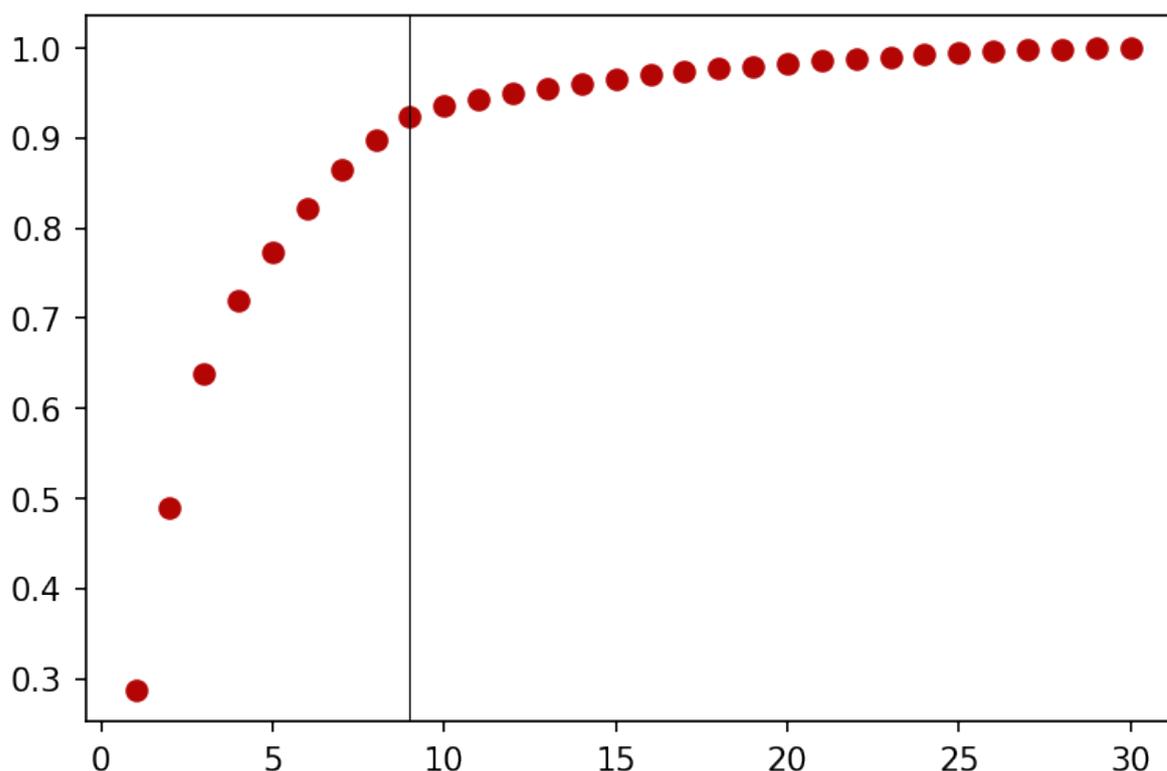


Figura 14 – Variância cumulativa explicada para o treinamento com as expressões neutras.

O modelo foi novamente treinado com os três bancos de treino utilizados neste trabalho (neutro, feliz, neutro+feliz) e classificado usando 9 componentes principais (*Eigenfaces*). Após isso, com o auxílio da função *confusion_matrix()*, foi calculada a matriz de confusão¹ para cada experimento. Através dessa matriz, foram obtidos a especificidade, sensibilidade, precisão e F-score para cada indivíduo.

4.1.1 EXPERIMENTO 1: Treinamento - Neutra

A Tabela 6 apresenta os resultados obtidos para o treinamento do modelo *Eigenfaces* utilizando as imagens neutras. As classificações das pessoas com as iniciais "KA", "KL", "KR", "NM", "TM" e "UY" apresentaram uma especificidade de 100%. Uma especificidade de 100% indica que o modelo não apresentou falsos positivos, ou seja, nenhuma das imagens das pessoas com as iniciais "KA", "KL", "KR", "NM", "TM" e "UY" foram incorretamente identificadas como sendo o alvo. As demais pessoas também apresentaram uma alta especificidade, com valores superiores a 94,16%, indicando que poucas imagens foram incorretamente identificadas como sendo o alvo.

A sensibilidade apresentou uma maior variabilidade. As classificações das imagens das pessoas com as iniciais "KL", "KM", "MK", "NM" e "UY" apresentaram uma sensibilidade de

¹ As matrizes de confusão geradas para todos os experimentos podem ser encontradas no anexo 1

Tabela 5 – Variância acumulada para o treinamento do *Eigenfaces* com as expressões neutras.

Número de componentes	Variância Acumulada
1	0,287501
2	0,490390
3	0,638495
4	0,718899
5	0,773104
6	0,821969
7	0,865675
8	0,897317
9	0,923750
10	0,935412
11	0,942618
12	0,949249
13	0,955202
14	0,960811
15	0,965698
16	0,969741
17	0,973511
18	0,976907
19	0,979893
20	0,982733
21	0,985463
22	0,988011
23	0,990342
24	0,992422
25	0,994318
26	0,996045
27	0,997724
28	0,999032
29	1.000.000
30	1.000.000

Tabela 6 – Resultado do modelo *Eigenfaces* treinado com as imagens neutras.

Iniciais do Nome	Especificidade	Sensibilidade	Precisão	F1-score	Support
KA	1	0.3125	1	0.4762	16
KL	1	1	1	1	16
KM	0.9927	1	0.9375	0.9677	15
KR	1	0.9333	1	0.9655	15
MK	0.9416	1	0.6522	0.7895	15
NA	0.9489	0.8	0.6316	0.7059	15
NM	1	1	1	1	14
TM	1	0.7333	1	0.8462	15
UY	1	1	1	1	15
YM	0.9706	0.9375	0.7895	0.8571	16

100%. Isso indica que o modelo não apresentou falsos negativos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das pessoas com as iniciais "KL", "KM", "MK", "NM" e "UY" foram incorretamente identificadas como pertencentes a outra pessoa. Por outro lado, a classificação das imagens pertencentes à pessoa com as iniciais "KA" apresentou uma sensibilidade de apenas 0,3125, ou seja, 68,75% das imagens foram identificadas como pertencentes a outra pessoa. As demais pessoas apresentaram uma alta sensibilidade, com valores entre 73,33% e 93,75%.

A precisão apresentou valores entre 63,16% e 100%. As classificações das imagens das pessoas com as iniciais "KA", "KL", "KR", "NM", "TM" e "UY" apresentaram uma precisão de 100%. Isso mostra que o modelo não apresentou falsos positivos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das demais pessoas com foram incorretamente identificadas como pertencentes as pessoas com as iniciais "KA", "KL", "KR", "NM", "TM" e "UY". O menor valor observado foi referente à pessoa com as iniciais "NA", que apresentou uma precisão de apenas 63,16%.

O F1-Score é a média harmônica entre Precisão e Sensibilidade. Para o método *Eigenfaces* treinado com as imagens neutras, as classificações das imagens das pessoas com as iniciais "KL", "NM" e "UY" apresentaram um F1-Score de 100%. A classificação das imagens pertencentes a pessoa com as iniciais "KA" havia apresentado uma baixa sensibilidade e por isso apresentou o menor F1-Score, de apenas 47,62%. As demais pessoas também apresentaram um F1-Score alto, variando entre 70,59% e 96,77%.

A coluna "Support" indica o número de imagens testadas de cada pessoa e esses valores foram usados para ponderar as medias dessas métricas e poder comparar os métodos *Eigenfaces* e *Fisherfaces* na seção 4.3.

4.1.2 EXPERIMENTO 2: Treinamento - Felicidade

Tabela 7 – Resultado do modelo *Eigenfaces* treinado com as imagens expressando felicidade.

Iniciais do Nome	Especificidade	Sensibilidade	Precisão	F1-score	Support
KA	0.9706	0.875	0.7778	0.8235	16
KL	1	1	1	1	16
KM	1	0.8667	1	0.9286	15
KR	1	0.9333	1	0.9655	15
MK	0.9781	0.9333	0.8235	0.875	15
NA	0.9781	0.8	0.8	0.8	15
NM	1	0.7857	1	0.88	14
TM	0.9781	1	0.8333	0.9091	15
UY	0.9854	1	0.8824	0.9375	15
YM	1	0.8125	1	0.8966	16

A Tabela 7 apresenta os resultados obtidos para o treinamento do modelo *Eigenfaces*

utilizando as imagens com expressão de felicidade. As classificações das pessoas com as iniciais "KL", "KM", "KR", "NM" e "YM" apresentaram uma especificidade de 100%. Isso indica que o modelo não apresentou falsos positivos, ou seja, nenhuma das imagens das pessoas com as iniciais "KL", "KM", "KR", "NM" e "YM" foram incorretamente identificadas como sendo o alvo. As demais pessoas também apresentaram uma alta especificidade, com valores superiores a 97,06%, indicando que poucas imagens foram incorretamente identificadas como sendo o alvo.

A sensibilidade apresentou melhores resultados e uma variabilidade menor que para o modelo treinado com as imagens neutras. As classificações das imagens das pessoas com as iniciais "KL", "TM" e "UY" apresentaram uma sensibilidade de 100%. Isso indica que o modelo não apresentou falsos negativos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das pessoas com as iniciais "KL", "TM" e "UY" foram incorretamente identificadas como pertencentes a outra pessoa. As demais pessoas também apresentaram uma alta sensibilidade, com valores entre 78,57% e 93,33%.

A precisão apresentou valores entre 77,78% e 100%. As classificações das imagens das pessoas com as iniciais "KL", "KM", "KR", "NM" e "YM" apresentaram uma precisão de 100%. Isso mostra que o modelo não apresentou falsos positivos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das demais pessoas foram incorretamente identificadas como pertencentes às pessoas com as iniciais "KL", "KM", "KR", "NM" e "YM". O menor valor observado foi referente à pessoa com as iniciais "KA", que apresentou uma precisão de 77,78%.

Para o método *Eigenfaces* treinado com as imagens felizes, apenas as classificações das imagens das pessoas com as iniciais "KL" apresentaram um F1-Score de 100%. As demais pessoas também apresentaram um F1-Score alto, variando entre 80% e 96,55%.

4.1.3 EXPERIMENTO 3: Treinamento - Neutra + Felicidade

Tabela 8 – Resultado do modelo *Eigenfaces* treinado com as imagens expressando emoções neutra e felicidade.

Iniciais do Nome	Especificidade	Sensibilidade	Precisão	F1-score	Support
KA	0.9779	0.75	0.8	0.7742	16
KL	1	1	1	1	16
KM	1	0.9333	1	0.9655	15
KR	1	0.9333	1	0.9655	15
MK	0.9635	0.9333	0.7368	0.8235	15
NA	0.9708	0.8	0.75	0.7742	15
NM	1	0.8571	1	0.9231	14
TM	0.9781	0.8	0.8	0.8	15
UY	1	1	1	1	15
YM	0.9779	0.8125	0.8125	0.8125	16

A Tabela 8 apresenta os resultados obtidos para o treinamento do modelo *Eigenfaces*

utilizando as imagens neutras e felizes juntas, totalizando 61 imagens de treino. As classificações das pessoas com as iniciais "KL", "KM", "KR", "NM" e "UY" apresentaram uma especificidade de 100%. Como já vimos, uma especificidade de 100% indica que o modelo não apresentou falsos positivos, ou seja, nenhuma das imagens das pessoas com as iniciais "KL", "KM", "KR", "NM" e "UY" foram incorretamente identificadas como sendo o alvo. As demais pessoas também apresentaram uma alta especificidade, com valores superiores a 96,35%, indicando que poucas imagens foram incorretamente identificadas como sendo o alvo.

As classificações das imagens das pessoas com as iniciais "KL" e "UY" apresentaram uma sensibilidade de 100%, ou seja, nenhuma das imagens das pessoas com as iniciais "KL" e "UY" foram incorretamente identificadas como pertencentes a outra pessoa. A classificação das imagens pertencentes a pessoa com as iniciais "KA" apresentou uma sensibilidade de 0,75, ou seja, 25% das imagens foram identificadas como pertencentes a outra pessoa. As demais pessoas também apresentaram uma alta sensibilidade, com valores entre 80% e 93,33%.

A precisão apresentou valores entre 73,68% e 100%. As classificações das imagens das pessoas com as iniciais "KL", "KM", "KR", "NM" e "UY", assim como na especificidade, apresentaram uma precisão de 100%. Isso mostra que o modelo não apresentou falsos positivos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das demais pessoas foram incorretamente identificadas como pertencentes as pessoas com as iniciais "KL", "KM", "KR", "NM" e "UY". O menor valor observado foi referente à pessoa com as iniciais "MK", que apresentou uma precisão de apenas 73,68%.

Para o método *Eigenfaces* treinado com as 61 imagens, as classificações das imagens das pessoas com as iniciais "KL" e "UY" apresentaram um F1-Score de 100%. As demais pessoas também apresentaram um F1-Score alto variando entre 77,42% e 96,55%.

4.2 *Fisherfaces*

O algoritmo *Fisherfaces* foi também treinado com as imagens neutras. As *Eigenfaces* geradas a partir 30 imagens de treino são apresentadas na Figura 15. Vemos que método retornou apenas 9 componentes principais e isso se deve ao fato de que na função *Fisherfaces*, o número máximo de autovetores que podem ser extraídos é $C - 1$, onde C é o número de classes presentes no conjunto de treinamento. Como temos 10 classes (indivíduos) diferentes, a função retornou 9 autovetores.



Figura 15 – *Fisherfaces* geradas a partir das imagens neutras do dataset JAFFE.

Como já houve uma redução de dimensionalidade gerada pelo próprio *Fisherfaces*, decidimos prosseguir a classificação com todos os 9 componentes, o que contribuiu para a comparação dos métodos, pois o número ideal encontrado no *Eigenfaces* também foi 9.

Semelhantemente à seção anterior, o modelo foi treinado com os três bancos de treino utilizados neste trabalho (neutro, feliz, neutro+feliz) e classificado usando 9 componentes principais (*Fisherfaces*).

4.2.1 EXPERIMENTO 1: Treinamento - Neutra

Tabela 9 – Resultado do modelo *Fisherfaces* treinado com as imagens neutras.

Iniciais do Nome	Especificidade	Sensibilidade	Precisão	F1-score	Support
KA	1	0.4375	1	0.6087	16
KL	1	1	1	1	16
KM	1	1	1	1	15
KR	1	0.9333	1	0.9655	15
MK	1	0.9333	1	0.9655	15
NA	0.927	1	0.6	0.75	15
NM	1	0.8571	1	0.9231	14
TM	1	0.8	1	0.8889	15
UY	0.9854	1	0.8824	0.9375	15
YM	0.9706	1	0.8	0.8889	16

A Tabela 9 apresenta os resultados obtidos para o treinamento do modelo *Fisherfaces* utilizando as imagens neutras. As classificações das pessoas com as iniciais "KA", "KL", "KM", "KR", "MK", "NM" e "TM" apresentaram uma especificidade de 100%. Uma especificidade de 100% indica que o modelo não apresentou falsos positivos, ou seja, nenhuma das imagens das pessoas com as iniciais "KA", "KL", "KM", "KR", "MK", "NM" e "TM" foram incorretamente identificadas como sendo o alvo. As três pessoas restantes também apresentaram uma alta especificidade, com valores superiores a 92,7%, indicando que poucas imagens foram incorretamente identificadas como sendo o alvo.

A sensibilidade, assim como no treinamento utilizando a técnica *Eigenfaces* apresentou uma maior variabilidade quando treinado com as imagens neutras. As classificações das imagens das pessoas com as iniciais "KL", "KM", "NA", "UY" e "YM" apresentaram uma sensibilidade de 100% para o método *Fisherfaces*. Isso indica que o modelo não apresentou falsos negativos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das pessoas com as iniciais "KL", "KM", "NA", "UY" e "YM" foram incorretamente identificadas como pertencentes a outra pessoa. Por outro lado, a classificação das imagens pertencentes à pessoa com as iniciais "KA" apresentou uma sensibilidade de apenas 43,75%, indicando que mais de 50% das imagens foram identificadas como pertencentes a outra pessoa. As demais pessoas também apresentaram uma alta sensibilidade, com valores entre 80% e 93,33%.

As classificações das imagens das pessoas com as iniciais "KA", "KL", "KM", "KR", "MK", "NM" e "TM" apresentaram uma precisão de 100%. Isso mostra que o modelo não apresentou falsos positivos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das demais pessoas com foram incorretamente identificadas como pertencentes as pessoas com as iniciais "KA", "KL", "KM", "KR", "MK", "NM" e "TM". O menor valor de precisão observado foi referente à pessoa com as iniciais "NA", que apresentou uma precisão de apenas 60%.

O F1-Score é a média harmônica entre Precisão e Sensibilidade. Para o método *Eigenfaces* treinado com as imagens neutras, as classificações das imagens das pessoas com as iniciais "KL" e "KM" apresentaram um F1-Score de 100%. A classificação das imagens pertencentes a pessoa com as iniciais "KA" havia apresentado uma baixa sensibilidade e por isso apresentou o menor F1-Score, de apenas 60,87%. As demais pessoas também apresentaram um F1-Score alto, variando entre 75% e 96,55%.

Lembrando que a coluna "Support" indica o número de imagens testadas de cada pessoa e esses valores foram usados para ponderar as medias dessas métricas e poder comparar os métodos *Eigenfaces* e *Fisherfaces* na próxima seção.

4.2.2 EXPERIMENTO 2: Treinamento - Felicidade

Tabela 10 – Resultado do modelo *Fisherfaces* treinado com imagens expressando felicidade.

Iniciais do Nome	Especificidade	Sensibilidade	Precisão	F1-score	Support
KA	0.9926	0.9375	0.9375	0.9375	16
KL	1	1	1	1	16
KM	1	0.9333	1	0.9655	15
KR	1	0.9333	1	0.9655	15
MK	0.9927	0.9333	0.9333	0.9333	15
NA	0.9927	1	0.9375	0.9677	15
NM	1	0.8571	1	0.9231	14
TM	0.9708	0.9333	0.7778	0.8485	15
UY	0.9854	1	0.8824	0.9375	15
YM	0.9926	0.8125	0.9286	0.8667	16

A Tabela 10 apresenta os resultados obtidos para o treinamento do modelo *Fisherfaces* utilizando as imagens com expressão de felicidade. As classificações das pessoas com as iniciais "KL", "KM", "KR" e "NM" apresentaram uma especificidade de 100%. Isso indica que o modelo não apresentou falsos positivos, ou seja, nenhuma das imagens das pessoas com as iniciais "KL", "KM", "KR" e "NM" foram incorretamente identificadas como sendo o alvo. As demais pessoas também apresentaram uma alta especificidade, com valores superiores a 97,08%, indicando que poucas imagens foram incorretamente identificadas como sendo o alvo.

A sensibilidade apresentou melhores resultados e uma variabilidade menor que para o modelo treinado com as imagens neutras. As classificações das imagens das pessoas com as iniciais "KL", "NA" e "UY" apresentaram uma sensibilidade de 100%. Isso indica que o modelo não apresentou falsos negativos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das pessoas com as iniciais "KL", "NA" e "UY" foram incorretamente identificadas como pertencentes a outra pessoa. As demais pessoas também apresentaram uma alta sensibilidade, com valores superiores a 81,25%.

A precisão apresentou valores entre 77,78% e 100%. As classificações das imagens das pessoas com as iniciais "KL", "KM", "KR" e "NM" apresentaram uma precisão de 100%. Isso mostra que o modelo não apresentou falsos positivos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das demais pessoas com foram incorretamente identificadas como pertencentes as pessoas com as iniciais "KL", "KM", "KR" e "NM". O menor valor observado foi referente à pessoa com as iniciais "TM", que apresentou uma precisão de 77,78%.

Para o método *Eigenfaces* treinado com as imagens felizes, apenas as classificações das imagens das pessoas com as iniciais "KL" apresentaram um F1-Score de 100%. As demais pessoas também apresentaram um F1-Score alto, variando entre 84,85% e 96,77%.

4.2.3 EXPERIMENTO 3: Treinamento - Neutra + Felicidade

Tabela 11 – Resultado do modelo *Fisherfaces* treinado com imagens expressando emoções neutra e felicidade.

Iniciais do Nome	Especificidade	Sensibilidade	Precisão	F1-score	Support
KA	1	1	1	1	16
KL	1	1	1	1	16
KM	1	1	1	1	15
KR	1	1	1	1	15
MK	1	0.9333	1	0.9655	15
NA	0.9927	1	0.9375	0.9677	15
NM	1	0.9286	1	0.963	14
TM	1	0.8	1	0.8889	15
UY	0.9927	1	0.9375	0.9677	15
YM	0.9779	1	0.8421	0.9143	16

A Tabela 11 apresenta os resultados obtidos para o treinamento do modelo *Fisherfaces* utilizando as imagens neutras e felizes juntas, totalizando 61 imagens de treino. As classificações das pessoas com as iniciais "KA", "KL", "KM", "KR", "NM" e "TM" apresentaram uma especificidade de 100%. Como já vimos, uma especificidade de 100% indica que o modelo não apresentou falsos positivos, ou seja, nenhuma das imagens das pessoas com as iniciais "KA", "KL", "KM", "KR", "NM" e "TM" foram incorretamente identificadas como sendo o alvo. As de-

mais pessoas também apresentaram uma especificidade próxima a 100%, com valores superiores a 97,79%, indicando que poucas imagens foram incorretamente identificadas como sendo o alvo.

As classificações das imagens de quase todas as pessoas apresentaram sensibilidades próximas ou iguais a 100%, ou seja, poucas imagens das foram incorretamente identificadas como pertencentes a outra pessoa. As classificações que apresentaram os menores valores de sensibilidade foram das imagens pertencentes a pessoa com as iniciais "TM", apresentando uma sensibilidade de 0,8, ou seja, 20% das imagens foram identificadas como pertencentes a outra pessoa.

A precisão apresentou valores entre 93,75% e 100%. As classificações das imagens das pessoas com as iniciais "KA", "KL", "KM", "KR", "NM" e "TM", assim como na especificidade, apresentaram uma precisão de 100%. Isso mostra que o modelo não apresentou falsos positivos para as imagens das pessoas com essas iniciais, ou seja, nenhuma das imagens das demais pessoas com foram incorretamente identificadas como pertencentes as pessoas com as iniciais "KA", "KL", "KM", "KR", "NM" e "TM". O menor valor observado foi referente à pessoa com as iniciais "YM", que apresentou uma precisão de apenas 84,21%.

Para o método *Fisherfaces* treinado com as 61 imagens, as classificações apresentaram um F1-Score alto variando entre 88,89% e 100%.

4.3 *Eigenfaces vs Fisherfaces*

Nesta seção é apresentada a comparação entre os resultados obtidos neste trabalho. Através dos resultados individuais apresentados nas últimas seções, foram obtidas as médias ponderadas para acurácia, especificidade, sensibilidade, precisão e F-score.

4.3.1 EXPERIMENTO 1: Treinamento - Neutra

Para melhor apresentar as diferenças entre os métodos, o eixo referente às proporções médias estimadas para cada medida avaliativa iniciou em 0,75. Vemos na Figura 16 que o *Eigenfaces* apresentou uma acurácia de aproximadamente 86,84%, valor que corresponde à proporção entre o número de acertos e o total de amostras. O método *Fisherfaces* apresentou uma acurácia de 89,47%.

Ambos os métodos apresentaram uma alta especificidade, com valores próximos a 99%, indicando que poucas imagens foram incorretamente identificadas como sendo o alvo. A sensibilidade de 86,64% do método *Eigenfaces* indica proporção de casos positivos corretamente identificados pelo classificador em relação ao número total de exemplos positivos. A sensibilidade do *Fisherfaces* foi superior e atingiu uma média de 89,47%. A precisão de 0,9010 indica que o modelo *Eigenfaces* acertou corretamente 90,1% entre os positivos previstos. O *Fisherfaces* apresentou uma precisão de 92,79%. O F1-Score do método *Eigenfaces* treinado com as imagens

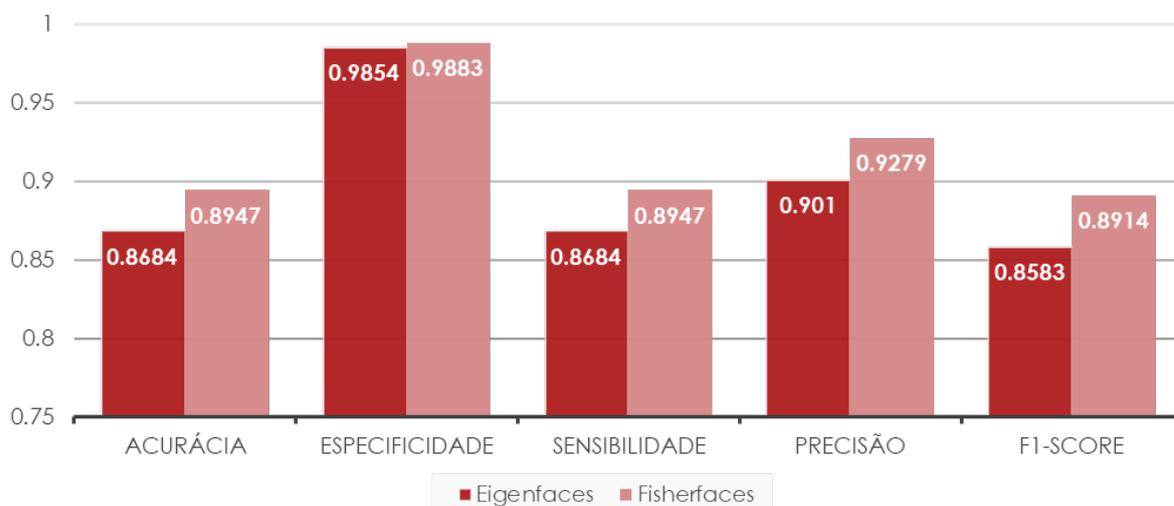


Figura 16 – Médias ponderadas dos resultados obtidos para o treinamento utilizando imagens neutras.

neutras mostrou que média harmônica entre Precisão e Sensibilidade foi de 85,83%. Para o método *Fisherfaces*, o F1-Score foi também superior, apresentando o valor de 89,14%.

4.3.2 EXPERIMENTO 2: Treinamento - Felicidade

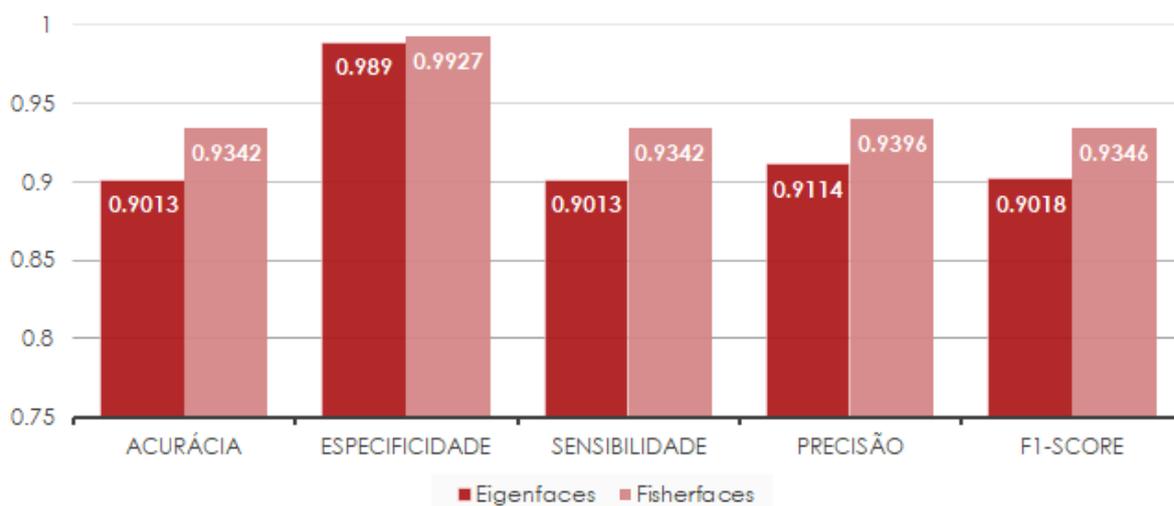


Figura 17 – Médias ponderadas dos resultados obtidos para o treinamento utilizando imagens expressando emoção felicidade.

As médias apresentadas na Figura 17 indicam que, quando treinados com as imagens felizes, o *Eigenfaces* apresentou uma acurácia de aproximadamente 90,13%, valor que corresponde à proporção entre o número de acertos e o total de amostras. O método *Fisherfaces* apresentou uma acurácia de 93,42%.

Ambos os métodos apresentaram uma alta especificidade, com valores próximos a 99%, indicando que poucas imagens foram incorretamente identificadas como sendo o alvo. A sensibilidade de 90,13% do método *Eigenfaces* indica proporção de casos positivos corretamente identificados pelo classificador em relação ao número total de exemplos positivos. A sensibilidade do *Fisherfaces* foi superior e atingiu média de 93,42%. A precisão de 0,9114 indica que o modelo *Eigenfaces* acertou corretamente 91,14% entre os positivos previstos. O *Fisherfaces* apresentou uma precisão de 93,96%. O F1-Score do método *Eigenfaces* treinado com as imagens neutras mostrou que média harmônica entre Precisão e Sensibilidade foi de 90,18%. Para o método *Fisherfaces*, o F1-Score foi também superior, apresentando o valor de 93,46%.

Os resultados obtidos com essa nova configuração indicam uma taxa média de acerto ligeiramente superior em relação ao treinamento com imagens de expressão facial neutra.

4.3.3 EXPERIMENTO 3: Treinamento - Neutra + Felicidade

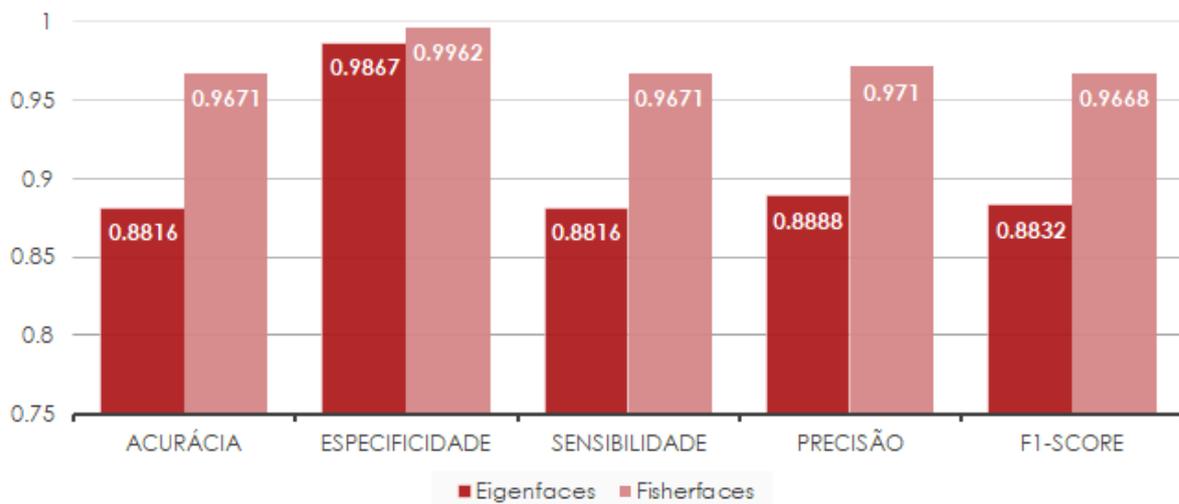


Figura 18 – Médias ponderadas dos resultados obtidos para o treinamento utilizando imagens expressando emoções neutra e felicidade.

As médias apresentadas na Figura 18 indicam que, quando treinados com as imagens neutras felizes, o *Eigenfaces* apresentou uma acurácia de aproximadamente 88,16%, valor que corresponde à proporção entre o número de acertos e o total de amostras. O método *Fisherfaces* apresentou uma acurácia de 96,71%.

Ambos os métodos apresentaram uma alta especificidade, com valores próximos a 100%, indicando que poucas imagens foram incorretamente identificadas como sendo o alvo. A sensibilidade de 88,16% do método *Eigenfaces* indica proporção de casos positivos corretamente identificados pelo classificador em relação ao número total de exemplos positivos. A sensibilidade do *Fisherfaces* foi superior e atingiu média de 96,71%. A precisão de 0,8888 indica que o modelo *Eigenfaces* acertou corretamente 88,88% entre os positivos previstos. O *Fisherfaces* apresentou

uma precisão de 96,71%. O F1-Score do método *Eigenfaces* treinado com as imagens neutras mostrou que média harmônica entre Precisão e Sensibilidade foi de 88,32%. Para o método *Fisherfaces*, o F1-Score foi também superior, apresentando o valor de 96,68%.

Neste último caso, o número de imagens de treino foi duas vezes maior que os casos anteriores. Observou-se que o método *Eigenfaces* apresentou resultados semelhante aos treinamentos anteriores, sendo que, o treinamento usando as 61 imagens foi ligeiramente melhor que o modelo treinado com as 30 imagens neutras e ligeiramente inferior ao treinado com as 31 imagens felizes.

O método *Fisherfaces* se comportou de forma semelhante ao *Eigenfaces* quando se comparou os dois primeiros treinamentos, apresentando um resultado ligeiramente melhor das expressões felizes em relação as expressões neutras. Porém houve um aumento significativo do *Fisherfaces* no último treinamento com o banco de dados maior. A taxa de acerto deste último experimento foi de 96,68%, valor maior que 89,14% e 93,46% observados respectivamente nos treinamentos com as imagens neutras e com as imagens felizes.

4.4 Discussão dos resultados

Em geral, os resultados obtidos sugerem que a expressão facial utilizada como imagem de treino pode influenciar de forma significativa o desempenho do modelo de reconhecimento facial. Além disso, as métricas de desempenho calculadas neste trabalho indicam que diferentes expressões faciais podem apresentar um grau de dificuldade variado na identificação de pessoas desaparecidas, o que pode ser relevante na escolha de imagens para serem utilizadas em sistemas de busca de pessoas desaparecidas.

O menor valor de precisão observado foi referente à pessoa com as iniciais "NA" para os dois modelos quando treinados com as imagens neutras. Grande parte das vezes os modelos classificaram as imagens pertencentes à pessoa com as iniciais "KA". É possível observar uma certa semelhança entre as mulheres. Apenas as classificações das imagens das pessoas com as iniciais "KL" apresentaram um F1-Score de 100% para os dois modelos (Ver Figura 11 do capítulo 3).

Todos os experimentos apresentaram uma especificidade próxima a 100% indicando que poucas imagens foram incorretamente identificadas como sendo o alvo. Isso é uma vantagem em sistemas de reconhecimento de desaparecidos pois evita-se criar falsas expectativas aos familiares desses desaparecidos. Por ter uma robustez maior, o método *Fisherfaces* apresentou um desempenho melhor em relação ao classificador *Eigenfaces*.

Ao analisar os resultados, podemos ver que no geral a expressão facial de felicidade foram as que apresentaram a maior taxa de acerto no reconhecimento facial. Isso pode ser atribuído ao fato de que essas expressões envolvem uma maior variação na aparência facial em comparação com a expressão neutra, principalmente nas partes da boca e olhos, o que torna o treinamento

com a expressão de felicidade mais abrangente.

Além disso, os resultados mostraram que o uso da técnica *Fisherfaces* foi mais eficiente em relação à técnica *Eigenfaces*, apresentando uma taxa de reconhecimento ligeiramente superior em comparação com a técnica *Eigenfaces*. Isso sugere que a técnica *Fisherfaces* pode ser mais adequada para lidar com variações em expressões faciais.

Em termos de discussão, é importante notar que a precisão do reconhecimento facial pode ser afetada por diversos fatores, como condições de iluminação, ângulo de visão, qualidade da imagem e variações nas expressões faciais. Esses fatores podem levar a erros no reconhecimento de indivíduos, especialmente em cenários reais de aplicação, como a busca por pessoas desaparecidas. Portanto, é necessário aprimorar as técnicas de reconhecimento facial para aumentar a precisão e robustez do sistema. Além disso, é importante coletar e utilizar conjuntos de dados maiores e mais variados para melhorar a generalização do modelo em diferentes cenários.

5 Considerações Finais

5.1 Conclusão

O presente estudo apresentou uma implementação de um sistema de reconhecimento de expressões faciais por meio de imagens usando Python e OpenCV. Para isso, foi realizado um pré-processamento no banco de imagens JAFFE. Os algoritmos de reconhecimento facial foram treinados usando as funções *Eigenfaces* e *Fisherfaces* da biblioteca OpenCV e foram testados com 152 imagens contendo expressões faciais diferentes das que foram utilizadas no treinamento. A avaliação da eficiência dos algoritmos de reconhecimento de faces foi realizada por meio de diversas ferramentas estatísticas, incluindo a matriz de confusão, acurácia, sensibilidade, especificidade, precisão, recall e F1 score.

Os resultados obtidos indicaram que o método *Fisherfaces* apresentou taxas ligeiramente superiores que o método *Eigenfaces*, mas no geral, ambos os algoritmos apresentaram resultados promissores no reconhecimento de expressões faciais. Este estudo pode ser utilizado como uma base para futuras pesquisas na área de reconhecimento facial e expressão facial, e pode ser estendido para outras aplicações, como reconhecimento de emoções em tempo real para análise de interações sociais.

Embora os reconhecedores faciais *Eigenfaces* e *Fisherfaces* sejam bons, existem algoritmos de reconhecimento facial mais avançados que usam uma combinação de redução de características e redes e redes neurais. No entanto, é importante ressaltar que o uso de tecnologias de reconhecimento facial pode levantar preocupações sobre privacidade e segurança de dados, especialmente em relação a possíveis usos indevidos ou discriminação. Portanto, é crucial que essas tecnologias sejam desenvolvidas e utilizadas de forma responsável, ética e transparente, com medidas de segurança e proteção de dados adequadas.

5.2 Trabalhos Futuros

Os resultados deste estudo podem ter implicações significativas para aplicações de reconhecimento facial, como em sistemas de reconhecimento de desaparecidos. No entanto, é importante ressaltar que este estudo apresenta algumas limitações, como o tamanho reduzido do conjunto de dados de treinamento e teste e a falta de variedade étnica das imagens, o que pode afetar a generalização dos resultados para outras populações. Portanto, recomenda-se que futuros estudos incluam conjuntos de dados maiores e mais diversificados para avaliar a eficácia dos algoritmos em diferentes populações.

Pode-se também usar o índice Kappa para comparar a concordância dos modelos ou avaliar

o desempenho do algoritmo de reconhecimento facial em relação a um avaliador humano, por exemplo. Isso pode fornecer informações valiosas sobre a precisão e confiabilidade do algoritmo, além de permitir a comparação com outras técnicas de reconhecimento facial disponíveis na literatura. Testes estatísticos como o teste t-Student ou o teste ANOVA, podem ser usados para comparar o desempenho de diferentes algoritmos e verificar se há diferenças significativas entre eles.

Outra sugestão é construir um sistema de reconhecimento facial baseado em um ambiente real, com a utilização dos algoritmos em conjunto, a fim de realizar os testes aplicados neste trabalho.

5.3 Materiais Elaborados

Os materiais técnicos elaborados durante a execução deste projeto estão disponíveis em um repositório público disponível em: <<https://github.com/PauloVitorOficial/ReconhecimentoDeFACES>>

Referências

- ABDI, H.; WILLIAMS, L. J. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, Wiley Online Library, v. 2, n. 4, p. 433–459, 2010. 10
- AGARWAL, M. et al. Face recognition using principle component analysis, eigenface and neural network. In: IEEE. *2010 International conference on signal acquisition and processing*. [S.l.], 2010. p. 310–314. 5
- ASADI, S.; RAO, C.; SAIKRISHNA, V. A comparative study of face recognition with principal component analysis and cross-correlation technique. *International Journal of Computer Applications*, Citeseer, v. 10, n. 8, p. 17–21, 2010. 11
- 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*, IEEE, v. 19, n. 7, p. 711–720, 1997. 2, 5, 7, 18, 19
- BRADSKI, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000. 27, 28
- COSTA-VIEIRA, H. A.; SOUZA, W. C. d. O reconhecimento de expressões faciais e prosódia emocional: Investigação preliminar em uma amostra brasileira jovem. *Estudos de Psicologia (Natal)*, SciELO Brasil, v. 19, p. 119–127, 2014. 3
- COWEN, A. S. et al. Sixteen facial expressions occur in similar contexts worldwide. *Nature*, Nature Publishing Group UK London, v. 589, n. 7841, p. 251–257, 2021. 4
- DARWIN, C. *The expression of the emotions in man and animals*. Albermale Street: John Murray, 1872. v. 1. 3
- DENG, H.-B. et al. A new facial expression recognition method based on local gabor filter bank and pca plus lda. *International Journal of Information Technology*, v. 11, n. 11, p. 86–96, 2005. 7
- DINALANKARA, L. Face detection & face recognition using open computer vision classifiers. *ResearchGate*, 2017. x, 14
- DÖRSCH, A. et al. Are average faces master faces? In: IEEE. *2022 International Workshop on Biometrics and Forensics (IWBF)*. [S.l.], 2022. p. 1–6. 11
- EKMAN, P. Universals and cultural differences in facial expressions of emotion. In: UNIVERSITY OF NEBRASKA PRESS. *Nebraska symposium on motivation*. [S.l.], 1971. 3
- FAN, L. et al. Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, Springer, v. 2, n. 1, p. 1–12, 2019. 6
- FERREIRA, D. F. Estatística multivariada. In: *Estatística Multivariada*. [S.l.: s.n.], 2011. p. 676–676. 8
- FISHER, R. A. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, Wiley Online Library, v. 7, n. 2, p. 179–188, 1936. 18
- GONZALEZ, R. C.; WOODS, R. E. *Processamento de imagens digitais*. [S.l.]: Editora Blucher, 2000. 6

- HASTIE, T. et al. *The elements of statistical learning: data mining, inference, and prediction*. [S.l.]: Springer, 2009. v. 2. 19
- HONGYU, K. *Comparação do GGE biplot-ponderado e AMMI-ponderado com outros modelos de interação genótipo × ambiente*. Tese (Doutorado) — Tese (Doutorado em Estatística e Experimentação Agronômica). Piracicaba: USP . . . , 2015. 11
- HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, Warwick & York, v. 24, n. 6, p. 417, 1933. 11
- JOLLIFFE, I. T. *Principal component analysis*. [S.l.]: Springer, 2002. 7, 9, 10
- KAISER, H. F. A note on guttman's lower bound for the number of common factors1. *British Journal of Statistical Psychology*, v. 14, p. 1–2, 1961. 8
- KARAMIZADEH, S. et al. An overview of principal component analysis. *Journal of Signal and Information Processing*, Scientific Research Publishing, v. 4, n. 3B, p. 173, 2013. 10
- KASSAMBARA, A. *Practical Guide To Principal Component Methods in R: PCA, M(CA), FAMD, MFA, HCPC, factoextra*. Sthda.com, 2017. (Multivariate Analysis). ISBN 9781975721138. Disponível em: <<https://books.google.com.br/books?id=eFEyDwAAQBAJ>>. 8
- KLUYVER, T. et al. Jupyter notebooks – a publishing format for reproducible computational workflows. In: LOIZIDES, F.; SCHMIDT, B. (Ed.). *Positioning and Power in Academic Publishing: Players, Agents and Agendas*. [S.l.], 2016. p. 87 – 90. 25
- KORTLI, Y. et al. Face recognition systems: A survey. *Sensors*, MDPI, v. 20, n. 2, p. 342, 2020. x, 16
- LIMA, J. R. de; NETO, R. F. O. Eigenface vs random forest: Um estudo comparativo em reconhecimento facial. In: SBC. *Anais do XVII Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais*. [S.l.], 2017. p. 428–441. x, 15
- LYONS, M. J. "excavating ai"re-excavated: Debunking a fallacious account of the jaffe dataset. *arXiv preprint arXiv:2107.13998*, 2021. 22
- LYONS, M. J.; KAMACHI, M.; GYOBA, J. Coding facial expressions with gabor wavelets (ivc special issue). *arXiv preprint arXiv:2009.05938*, 2020. 22
- MARENGONI, M.; STRINGHINI, D. Tutorial: Introdução à visão computacional usando opencv. *Revista de Informática Teórica e Aplicada*, v. 16, n. 1, p. 125–160, 2010. 2, 5, 6
- PEARSON, K. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, Taylor & Francis, v. 2, n. 11, p. 559–572, 1901. 7
- PEREIRA, T. Estatística multivariada i–uma abordagem aplicada utilizando o software r. *Ouro Preto: Departamento de Estatística, Universidade Federal de Ouro Preto*, 2017. 8
- PHILLIPS, P. J. et al. The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 22, n. 10, p. 1090–1104, 2000. 4
- Python Core Team. *Python: A dynamic, open source programming language*. [S.l.], 2023. Disponível em: <<https://www.python.org/>>. 24

- SANTANA, L. M. Q. de; ROCHA, F. G. et al. Processo de detecção facial utilizando viola; jones. *Interfaces Científicas-Exatas e Tecnológicas*, v. 1, n. 1, p. 35–40, 2015. 6
- SCHNITZSPAHN, K. et al. Mood impairs time-based prospective memory in young but not older adults: The mediating role of attentional control. *Psychology and Aging*, v. 29, 05 2014. 29
- SILVA, A. L. *Redução de características para classificação de imagens de faces*. [S.l.]: MOSSORÓ-RN, 2016. x, 8, 14, 17, 18, 19
- SILVA, J. I. S. d. *Reconhecimento facial em imagens de baixa resolução*. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2015. 7, 31
- SuperDataScience Team. *Face recognition using OpenCV and Python: A beginner's guide*. 2017. Disponível em: <<https://www.superdatascience.com/blogs/opencv-face-recognition>>. Acesso em: 01 de fevereiro 2023. 27
- TURAGA, P. et al. Statistical computations on grassmann and stiefel manifolds for image and video-based recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 33, n. 11, p. 2273–2286, 2011. 4
- TURK, M.; PENTLAND, A. Eigenfaces for recognition. *Journal of cognitive neuroscience*, MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , v. 3, n. 1, p. 71–86, 1991. 5, 7, 14
- WHALEN, P. J. et al. Neuroscience and facial expressions of emotion: The role of amygdala–prefrontal interactions. *Emotion Review*, SAGE Publications Sage UK: London, England, v. 5, n. 1, p. 78–83, 2013. 3
- WOLD, S.; RUHE, A.; WOLD, H. Principal component analysis. *Chemometrics and intelligent laboratory systems*, Elsevier, v. 2, n. 1-3, p. 37–52, 1987. 10

Apêndices

Tabela 14 – Matriz de confusão Eigenfaces Feliz

ID REAL	ID PREDITO									
	KA	KL	KM	KR	MK	NA	NM	TM	UY	YM
KA	14	0	0	0	1	1	0	0	0	0
KL	0	16	0	0	0	0	0	0	0	0
KM	2	0	13	0	0	0	0	0	0	0
KR	0	0	0	14	0	1	0	0	0	0
MK	0	0	0	0	14	1	0	0	0	0
NA	1	0	0	0	2	12	0	0	0	0
NM	1	0	0	0	0	0	11	0	2	0
TM	0	0	0	0	0	0	0	15	0	0
UY	0	0	0	0	0	0	0	0	15	0
YM	0	0	0	0	0	0	0	3	0	13

Tabela 15 – Matriz de confusão Fisherfaces Feliz

ID REAL	ID PREDITO									
	KA	KL	KM	KR	MK	NA	NM	TM	UY	YM
KA	15	0	0	0	1	0	0	0	0	0
KL	0	16	0	0	0	0	0	0	0	0
KM	1	0	14	0	0	0	0	0	0	0
KR	0	0	0	14	0	0	0	1	0	0
MK	0	0	0	0	14	1	0	0	0	0
NA	0	0	0	0	0	15	0	0	0	0
NM	0	0	0	0	0	0	12	0	2	0
TM	0	0	0	0	0	0	0	14	0	1
UY	0	0	0	0	0	0	0	0	15	0
YM	0	0	0	0	0	0	0	3	0	13

Tabela 16 – Matriz de confusão Eigenfaces Neutro e Feliz

ID REAL	ID PREDITO									
	KA	KL	KM	KR	MK	NA	NM	TM	UY	YM
KA	12	0	0	0	2	2	0	0	0	0
KL	0	16	0	0	0	0	0	0	0	0
KM	1	0	14	0	0	0	0	0	0	0
KR	0	0	0	14	0	1	0	0	0	0
MK	0	0	0	0	14	1	0	0	0	0
NA	1	0	0	0	2	12	0	0	0	0
NM	1	0	0	0	1	0	12	0	0	0
TM	0	0	0	0	0	0	0	12	0	3
UY	0	0	0	0	0	0	0	0	15	0
YM	0	0	0	0	0	0	0	3	0	13

APÊNDICE B – Médias ponderadas dos resultados

Tabela 18 – Médias ponderadas dos resultados obtidos para o treinamento utilizando as imagens neutras.

	Acurácia	Especificidade Ponderada	Sensibilidade Ponderada	Precisão Ponderada	F1-Score Ponderado
Eigenfaces	0.8684	0.9854	0.8684	0.9010	0.8583
Fisherfaces	0.8947	0.9883	0.8947	0.9279	0.8914

Tabela 19 – Médias ponderadas dos resultados obtidos para o treinamento utilizando as imagens Felizes

	Acurácia	Especificidade Ponderada	Sensibilidade Ponderada	Precisão Ponderada	F1-Score Ponderado
Eigenfaces	0.9013	0.9890	0.9013	0.9114	0.9018
Fisherfaces	0.9342	0.9927	0.9342	0.9396	0.9346

Tabela 20 – Médias ponderadas dos resultados obtidos para o treinamento utilizando as imagens neutras e Feliz juntas.

	Acurácia	Especificidade Ponderada	Sensibilidade Ponderada	Precisão Ponderada	F1-Score Ponderado
Eigenfaces	0.8816	0.9867	0.8816	0.8888	0.8832
Fisherfaces	0.9671	0.9962	0.9671	0.9710	0.9668