



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA
MESTRADO ACADÊMICO EM ENGENHARIA DE TELEINFORMÁTICA

PATRÍCIA JAMILE DE OLIVEIRA MARTINS

**SEGMENTAÇÃO DE REGIÕES DE INTERESSE EM IMAGENS 3D DE FACES PARA
RECONHECIMENTO DE INDIVÍDUOS: UMA ABORDAGEM BASEADA EM
UNIDADES DE AÇÃO MUSCULAR E ATRIBUTOS GEOMÉTRICOS DE
COVARIÂNCIA**

FORTALEZA

2022

PATRÍCIA JAMILE DE OLIVEIRA MARTINS

SEGMENTAÇÃO DE REGIÕES DE INTERESSE EM IMAGENS 3D DE FACES PARA
RECONHECIMENTO DE INDIVÍDUOS: UMA ABORDAGEM BASEADA EM UNIDADES
DE AÇÃO MUSCULAR E ATRIBUTOS GEOMÉTRICOS DE COVARIÂNCIA

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Engenharia de Teleinformática

Orientador: Prof. Dr. George André Pereira Thé

Coorientador: Prof. Dr. José Marques Soares

FORTALEZA

2022

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

M345s Martins, Patrícia Jamile de Oliveira.

Segmentação de Regiões de Interesse em Imagens 3D de Faces para Reconhecimento de Indivíduos : uma Abordagem Baseada em Unidades de Ação Muscular e Atributos Geométricos de Covariância / Patrícia Jamile de Oliveira Martins. – 2022.

70 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2022.

Orientação: Prof. Dr. George André Pereira Thé.

Coorientação: Prof. Dr. José Marques Soares.

1. Nuvem de Pontos. 2. Reconhecimento Facial. 3. Fatiamento de Nuvens. 4. Extratores de Características. I. Título.

CDD 621.38

PATRÍCIA JAMILE DE OLIVEIRA MARTINS

SEGMENTAÇÃO DE REGIÕES DE INTERESSE EM IMAGENS 3D DE FACES PARA
RECONHECIMENTO DE INDIVÍDUOS: UMA ABORDAGEM BASEADA EM UNIDADES
DE AÇÃO MUSCULAR E ATRIBUTOS GEOMÉTRICOS DE COVARIÂNCIA

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Engenharia de Teleinformática

Aprovada em: 04 de Março de 2022

BANCA EXAMINADORA

Prof. Dr. George André Pereira Thé (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. José Marques Soares (Coorientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Eduardo Furtado de Simas Filho
Universidade Federal da Bahia (UFBA)

Prof^a. Dra. Michela Mulas
Universidade Federal do Ceará (UFC)

À minha família, por todo o apoio e incentivo durante essa jornada.

AGRADECIMENTOS

À minha mãe, Maria das Graças, e ao meu pai, Francisco Genésio, por todos os esforços, amor e incentivos que dedicaram a mim. Aos meus familiares por todo o suporte. Ao meu marido, Jefferson Figueiredo, pelo companheirismo, compreensão e incentivo nas horas boas e ruins.

Ao Prof. Dr. José Marques Soares e ao Prof. Dr. George A. P. Thé por me orientarem em minha dissertação, por todo o apoio e compreensão durante todo o percurso. Foram muitos os momentos de dificuldade, durante a trajetória deste trabalho, e sempre pude contar com o olhar humano, o carinho e a compreensão de ambos.

Aos amigos do Grupo Interfaces, Gilderlane, Artur, Polycarpo, Izaías, Luís, Prof^o Natália, Marcus, Stella, João Pedro, Allyson e Robson, por todas as conversas, discussões de ideias e momentos compartilhados.

Aos amigos e colegas de trabalho do Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), *campus* Acopiara e *campus* Canindé por todo apoio e suporte que recebi.

"E ainda sou bem moço pra tanta tristeza,
Deixemos de coisas, cuidemos da vida,
Senão chega a morte ou coisa parecida
E nos arrasta moço sem ter visto a vida"

(Na hora do almoço - Belchior)

RESUMO

Neste trabalho é proposta uma investigação sobre diferentes formas de segmentar imagens de faces humanas representadas por nuvens de pontos 3D. Objetiva-se encontrar subnuvens, denominadas fatias, que contenham maior significância para extração de características visando aplicações de reconhecimento de indivíduos. Tendo por base a análise das regiões de maior variabilidade da face e soluções propostas por trabalhos anteriores neste mesmo domínio, é feita uma comparação entre os resultados de classificação com e sem fatiamento das faces, utilizando características geométricas no processo de extração de informação. Com o objetivo de investigar as regiões específicas da face 3D para aprimoramento dos resultados da classificação, dois tipos de segmentação foram avaliados: o fatiamento triaxial e o fatiamento em Pizza com superposição, sendo este último uma contribuição deste projeto. Os resultados mostram-se promissores, indicando que um refinamento da técnica pode gerar uma classificação de indivíduos robusta às naturais deformações da superfície da face.

Palavras-chave: Nuvem de pontos. Reconhecimento facial. Fatiamento de nuvens. Extratores de características.

ABSTRACT

In this work is proposed an investigation into different ways on slicing process applied on 3D Point Cloud data, representing individuals faces, on individual face recognition problem. The objective is to find subclouds, called slices, that contain greater significance for the extraction of features aiming at individual recognition applications. Based on the analysis of regions of greater variability of the face and solutions proposed by previous works in this same domain, a comparison is made between the classification results with and without slicing the faces, using geometric characteristics in the information extraction process. The purpose of the comparisons is to investigate if selecting specific regions of the 3D face can improve classification results, two types of segmentation were evaluated: triaxial slicing and pizza slicing with superposition, the latter being a contribution of this project. The results are promising, indicating that a refinement of the technique can generate a classification of individuals robust to the natural deformations of the surface of the face.

Keywords: Point Cloud. Facial Recognition. Point Cloud Slicing. Feature Extraction.

LISTA DE FIGURAS

Figura 1 – Exemplos de representações de imagens 3D: nuvem de pontos (esquerda), cuboides (meio) e malhas triangulares(direita). Imagem retirada de (HOANG <i>et al.</i> , 2019).	17
Figura 2 – Exemplos de Representações de Imagens 3D: Mapa de Profundidade (esquerda), Nuvem de Pontos (centro) e Malha de Triângulos (direita).	21
Figura 3 – Exemplos de Casos de Oclusão.	22
Figura 4 – Exemplos de Expressões Faciais.	22
Figura 5 – Músculos da Face Superior.	23
Figura 6 – Ações Musculares da Face Superior.	24
Figura 7 – Contornos detectados (em vermelho) de uma nuvem de pontos (em preto) que representa uma construção em um ambiente externo	28
Figura 8 – <i>Landmarks</i> fornecidas na base Bosphorus.	36
Figura 9 – Etapas do processo de classificação de uma imagem 3D	36
Figura 10 – Descrição dos cenários de experimentação utilizados.	38
Figura 11 – Nuvem de pontos em diferentes etapas de pré-processamento.	39
Figura 12 – Fatiamento triaxial, seções frontais, sagitais e transversais.	41
Figura 13 – Detalhamento do fatiamento de Pizza com sobreposição. Em vermelho estão as <i>landmarks</i> fornecidas pela base Bosphorus.	43
Figura 14 – Fatiamento em pizza com sobreposição, seções central, Superior e Inferior, e Direita e Esquerda	44
Figura 15 – Fatiamento em Pizza com sobreposição, todas as fatias (<i>UP, DN, RT, LT, CT</i>), visão lateral.	44
Figura 16 – Tempo de Execução <i>versus</i> Classificadores, todas as configurações	53
Figura 17 – Taxa de acerto <i>versus</i> Classificador, Configuração Pizza C1	54
Figura 18 – Taxa de acerto <i>versus</i> Classificador, Configuração Triaxial	55
Figura 19 – Taxa de acerto <i>versus</i> Classificador, Configuração <i>FULL</i>	56
Figura 20 – Taxa de acerto <i>versus</i> Tempo de Execução, Cenário Rank-1a (A), Configuração Pizza C1	57
Figura 21 – Taxa de acerto <i>versus</i> Tempo de Execução, Cenário Rank-1b (B), Configuração Pizza C1	58

Figura 22 – Taxa de acerto <i>versus</i> Tempo de Execução, Cenário ROC-I (C), Configuração Pizza C1	58
Figura 23 – Taxa de acerto <i>versus</i> Tempo de Execução, Cenário Rank-1a, Configuração Triaxial, Fatias ST, classificador KNN-1	59
Figura 24 – Aplicação Web para análise de imagens 3D	62
Figura 25 – Figuras obtidas utilizando filtros nas características descritas	63

LISTA DE TABELAS

Tabela 1 – Distribuição por subconjuntos dos exemplares da base de dados Bosphorus .	37
Tabela 2 – Maiores taxas de acerto dos classificadores obtidas de todos os fatiamentos.	47
Tabela 3 – Configurações do fatiamento em pizza com sobreposição utilizados.	48
Tabela 4 – Melhores taxas de acerto obtidas com fatiamento em pizza com sobreposição nas diferentes configurações escolhidas.	48
Tabela 5 – Resultados obtidos com fatiamento em pizza com sobreposição na Configuração 1 e utilizando cada segmento individualmente e combinações dos mesmos	49
Tabela 6 – Maiores taxas de acerto dos classificadores obtidas com fatiamento triaxial e outros conjuntos de fatia.	51
Tabela 7 – Resultados obtidos com combinações do fatiamento triaxial, fatiamento de pizza com sobreposição e face inteira.	52

LISTA DE ABREVIATURAS E SIGLAS

<i>AUC</i>	<i>Area Under the Curve</i>
<i>EGI</i>	<i>Extended Gaussian Image</i>
<i>FACS</i>	<i>Facial Action Coding System</i> (Sistema de Codificação de Ações Faciais)
<i>FNTFGP</i>	<i>Fast Nonnegative Tensor Factorization Based On Graph-Preserving</i>
<i>KNN</i>	<i>K-Nearest Neighbors</i>
<i>LDA</i>	<i>Linear Discriminant Analysis</i>
<i>LiDAR</i>	<i>Light Detection and Ranging</i>
<i>PCA</i>	<i>Principal Component Analysis</i>
<i>RBF</i>	<i>Radial Basis Function</i>
<i>ROC</i>	<i>Receiver Operating Characteristic</i>
<i>SVM</i>	<i>Support Vector Machine</i>
ADs	Descritores de Ação (Action Descriptors)
AUs	Unidades de Ação Muscular (Action Units)
IFCE	Instituto Federal de Educação, Ciência e Tecnologia do Ceará
Ms	Movimentos (Movements)

LISTA DE SÍMBOLOS

C_i	Matriz de Covariância de um conjunto de pontos
G_1	Somatório
G_2	Omnivariância
G_3	Autoentropia
G_4	Anisotropia
G_5	Planaridade
G_6	Linearidade
G_7	Variação de superfície
G_8	Esfericidade
G_9	Verticalidade
A	Cenário de experimentação Rank-1a
B	Cenário de experimentação Rank-1b
C	Cenário de experimentação ROC-I
F_i	Segmentos frontais do Fatiamento Triaxial
S_i	Segmentos sagitais do Fatiamento Triaxial
T_i	Segmentos transversais do Fatiamento Triaxial
<i>FULL</i>	Nuvem de pontos da Face inteira
<i>CT</i>	Segmento Central do Fatiamento em Pizza com Sobreposição
<i>UP</i>	Segmento Superior do Fatiamento em Pizza com Sobreposição
<i>DN</i>	Segmento Inferior do Fatiamento em Pizza com Sobreposição
<i>RT</i>	Segmento Direito do Fatiamento em Pizza com Sobreposição
<i>LT</i>	Segmento Esquerdo do Fatiamento em Pizza com Sobreposição

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Objetivos	18
1.2	Publicação	18
1.3	Organização da Dissertação	19
2	FUNDAMENTAÇÃO TEÓRICA	20
2.1	Nuvens de Pontos	20
2.2	Faces 3D	21
2.3	Sistema de Codificação de Ação Facial (FACS)	22
2.4	Reconhecimento de Faces 3D	23
2.5	Extratores de Características	25
2.5.1	<i>Extratores Geométricos Baseados em Covariância</i>	26
2.6	Classificadores	28
2.6.1	<i>Classificador K-Nearest Neighbors (KNN)</i>	28
2.6.2	<i>Classificador Gaussiano Naïve-Bayes</i>	29
2.6.3	<i>Classificador Support Vector Machines (SVM)</i>	30
2.6.4	<i>Classificador Árvore de Decisão (Decision Tree)</i>	32
2.6.5	<i>Classificador Floresta Aleatória (Random Forest)</i>	32
2.6.6	<i>Classificador Linear Discriminant Analysis (LDA)</i>	33
2.7	Considerações Finais	33
3	METODOLOGIA	35
3.1	Etapas do Processo de Classificação	36
3.1.1	<i>Cenários de experimentação</i>	37
3.2	Pré-processamento	38
3.3	Fatiamento	40
3.3.1	<i>Fatiamento Triaxial</i>	41
3.3.2	<i>Fatiamento em Pizza com sobreposição</i>	42
3.4	Classificação	43
3.5	Desempenho das experimentações	46
3.6	Considerações Finais sobre Metodologia	46
4	RESULTADOS	47

4.1	Resultados do Fatiamento Pizza com Sobreposição	48
4.2	Resultados do Fatiamento Triaxial	50
4.3	Resultados da Combinação de Fatiamento Triaxial, Fatiamento em Pizza com Sobreposição e Face Inteira	51
4.4	Desempenho dos Algoritmos de Classificação	52
4.5	Considerações Finais sobre Resultados	57
5	CONCLUSÕES E TRABALHOS FUTUROS	60
5.1	Trabalhos Futuros	60
	REFERÊNCIAS	64

1 INTRODUÇÃO

O reconhecimento de faces em imagens tridimensionais (3D) tem ganhado popularidade, em parte, pelos recentes avanços da tecnologia de sensoriamento e sua difusão comercial. Porém, um dos desafios neste domínio é a grande deformação da superfície da face causada por expressões, como um sorriso, ou manifestações de tristeza, medo ou nojo, entre outras. Outras dificuldades se devem por algum tipo de oclusão, como cabelos, pelos faciais, óculos, dedos ou outros objetos cobrindo parte da face. Assim, os sistemas reconhecedores de faces devem ser suficientemente robustos a deformações dessa natureza para automatizar a classificação.

As regiões ao redor dos olhos, sobrancelhas e boca apresentam grande variação em diferentes expressões e também são aquelas que contêm mais informações sobre a face, sendo áreas bastante utilizadas na seleção de *landmarks* em reconhecimento de faces 2D, como discutido em (ZHAO *et al.*, 2003) e (SHEPLEY, 2019). Estima-se que haja a possibilidade de uso dessas informações contidas nesses elementos para elaboração de algoritmos de reconhecimentos de indivíduos mais robustos a expressões faciais diferentes. Entretanto, a variabilidade de algumas regiões pode, ainda, representar um desafio para a extração de características invariantes à expressividade humana.

Estudos catalogaram regiões de alta variabilidade na face do indivíduo durante a manifestação de expressões. Tais estudos podem ser usados para dar apoio a pesquisas que buscam contornar o problema da deformação da superfície da face no reconhecimento automático de indivíduos. As regiões de alta variabilidade da face são denominadas Unidades de Ação Muscular (Action Units) (AUs) no *Facial Action Coding System* (Sistema de Codificação de Ações Faciais) (FACS) (EKMAN; FRIESEN, 1978). Além de catalogar as regiões de deformação, esse estudo discute as relações típicas entre as expressões emocionais e as contrações musculares destas unidades combinadas. As AUs são frequentemente consideradas em sistemas de reconhecimento de expressões faciais, como exemplo, (LI; KOBATAKE, 1995; GUAN, 2021; YOUNIS *et al.*, 2021).

As imagens 3D oferecem informações de características geométricas da face capazes de melhorar o desempenho de sistemas de reconhecimento em comparação àqueles baseados em imagens 2D (HARIRI *et al.*, 2016) (SOLTANPOUR *et al.*, 2017). Há diferentes formas de representação da imagem 3D. Dentre elas, temos as nuvens de pontos (que são usadas neste trabalho), as malhas triangulares e os cubóides (ou *voxels*). A Figura 1 apresenta exemplos das três formas de representação de imagem 3D mencionadas.

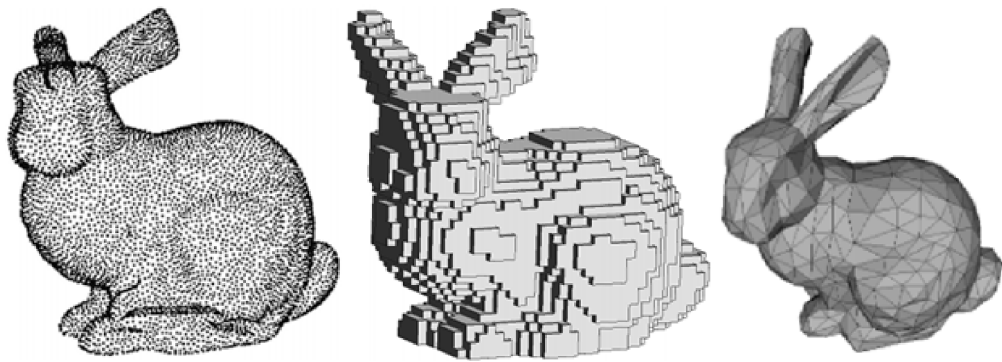


Figura 1 – Exemplos de representações de imagens 3D: nuvem de pontos (esquerda), cuboides (meio) e malhas triangulares(direita). Imagem retirada de (HOANG *et al.*, 2019).

Para o pré-processamento de dados de imagens 3D da face, pode ser necessária a realização da segmentação, dividindo a representação da face em regiões. Existem diferentes formas de realizar a segmentação da face, como as técnicas apresentadas em *Extended Gaussian Image (EGI)*, (LEE; MILIOS, 1991), (GORDON, 1991), (NAGAMINE *et al.*, 1992), (Tanaka *et al.*, 1998), (HESHER *et al.*, 2003), (JAFRI; ARABNIA, 2009) e nos *surveys* (BOWYER *et al.*, 2004), (KAVITA, 2016) e (Reddy *et al.*, 2018). Muitos desses trabalhos utilizam apenas faces neutras (sem expressões faciais), pois é o tipo de face mais comum em bases de dados tradicionais (BOWYER *et al.*, 2004). Neste trabalho, adotam-se os termos segmentação e fatiamento como a separação de um subconjunto da nuvem de pontos obtido por meio de um critério de seleção.

No que se refere aos métodos de extração de características, três principais categorias são apresentadas na literatura: (i) com base em atributos holísticos ou globais; (ii) com base em atributos locais e; (iii) métodos híbridos. Os trabalhos que se baseiam em atributos locais focam na extração dos atributos da face ou de regiões relevantes desta (região dos olhos e nariz ou *keypoints*, por exemplo). Por não requererem modelos completos da face, métodos baseados em atributos locais lidam melhor com situações de oclusão, auto-occlusão e expressões, sendo mais apropriados para propósitos de identificação e verificação (SOLTANPOUR *et al.*, 2017). Além disso, o fato de não requererem modelos completos da face faz com que as nuvens de pontos usadas em tais trabalhos sejam menores, pois usam apenas partes da face como dado para geração de atributos. Isso faz com que o processo de extração de características seja consideravelmente mais rápido.

Como a face não é um corpo rígido, as regiões de alta variabilidade prejudicam a classificação. Assim, visto que as taxas de acerto são maiores quando são utilizadas faces neutras (SOLTANPOUR *et al.*, 2017), o desafio de identificar indivíduos com diferentes expressões é

encontrar a melhor forma de extrair informações da face mesmo quando ela apresenta variações em sua superfície. Nesse sentido, selecionar segmentos da face e analisar as AUs pode ser útil, o que justifica a investigação realizada no presente trabalho.

1.1 Objetivos

Neste trabalho, tem-se como objetivo principal a proposição de uma forma de segmentação de faces representadas por nuvens de pontos para extração de características e classificação de indivíduos. Para isso, diferentes formas de segmentação são analisadas, buscando-se segmentos específicos, mais adequados para o problema do reconhecimento de faces. Extratores geométricos, utilizados em outros domínios de aplicação com o uso de nuvens 3D, também são investigados para o contexto de reconhecimento de faces. Serão utilizados exemplares da base de dados Bosphorus com expressões neutras e não-neutras, excluindo exemplares com oclusão e rotação.

Dentre os objetivos específicos a serem atingidos nesta dissertação, estão:

- diminuir o custo computacional de classificação indivíduos em imagens 3D;
- avaliar o efeito das AUs no processo de classificação de faces 3D;
- comparar diferentes formas de fatiamento da face 3D;
- utilizar diferentes tipos de classificadores como forma de avaliação, incluindo *K-Nearest Neighbors (KNN)*, *Naïve-Bayes*, *Linear Discriminant Analysis (LDA)*, árvore de decisão, floresta aleatória e diferentes combinações de *Support Vector Machine (SVM)*, *kernels* e hiperparâmetros;
- otimizar os mecanismos de classificação que utilizam fatiamento.

1.2 Publicação

Parte do trabalho descrito nesta dissertação foi publicado e apresentado em Outubro de 2019, (MARTINS *et al.*, 2019) no XV Encontro Nacional de Inteligência Artificial e Computacional (ENIAC), incluído na programação do BRACIS 2019, *Brazilian Conference on Intelligent Systems*, Salvador-BA, publicado nos Anais do Evento com o título "*Face slicing on facial recognition: An action units based investigation of the possible improvements by selecting the face regions in a point cloud 3D data*" (MARTINS *et al.*, 2019).

Neste evento, o artigo recebeu o prêmio *Outstanding Paper Award*.

1.3 Organização da Dissertação

A estrutura desta dissertação está organizada da seguinte forma:

- **Capítulo 2:** são discutidos os aspectos que fundamentam e dão suporte ao reconhecimento de faces em 3D;
- **Capítulo 3:** é descrita a metodologia proposta para realizar o fatiamento da face e os testes de hipóteses da investigação proposta;
- **Capítulo 4:** são apresentados e discutidos os resultados dos testes realizados.
- **Capítulo 5:** apresenta as principais conclusões alcançadas bem como suas limitações e possibilidades de melhorias.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os conceitos acerca de nuvens de pontos 3D, o sistema de codificação *FACS*, o problema de classificação de faces 3D, os extratores de características geométricas comumente utilizados para nuvens 3D e, por fim, diferentes tipos de classificadores.

2.1 Nuvens de Pontos

Uma nuvem de pontos é um conjunto de pontos no espaço de 3 dimensões, comumente representados por coordenadas cartesianas. Tais nuvens são produzidas por escâneres 3D diretamente ou reconstruídas por um conjunto de câmeras 2D, sendo pelo menos duas, em posições distintas (YIN *et al.*, 2006; Erdogmus; Marcel, 2013; SAVRAN *et al.*, 2008a).

As técnicas de obtenção de imagens 3D a partir de imagens 2D estão tornando-se comuns, pois alguns novos modelos de *smartphones* já vem com mais de uma câmera na parte de trás do aparelho. Essas 2 câmeras tanto ajudam a tirar fotos com múltiplos pontos focais, quanto possibilitam a reconstrução do ambiente 3D que está sendo capturado, funcionando como uma câmera estereoscópica.

Os sensores capazes de capturar imagens 3D, como os de infravermelho ou os de tempo-de-vôo (*Time-of-Flight*), têm se popularizado em ritmo acelerado graças ao barateamento de sua produção, à miniaturização e à redução de consumo de energia, facilitando a aquisição e análise desse tipo de informação (KIM *et al.*, 2020; 3DSOURCED.COM, 2021). Estes fatores possibilitam a integração desses sensores em equipamentos como *smartphones* e drones. Esses sensores são utilizados para gerar efeitos nas fotos, reconhecimento de gestos, desbloqueio do dispositivo através do padrão de veias da palma da mão e para auxiliar no desbloqueio do dispositivo através de reconhecimento facial (CHIBA, 2019; MARQUES, 2020).

A nuvem de pontos pode ser densa como a fornecida por equipamentos tais como tomógrafos que analisam as partes internas de órgãos humanos, como pode ser visto em (Salama *et al.*, 2018). Nuvens de pontos 3D de tomografias são largamente utilizadas em estudos para identificação e reconhecimento de doenças como câncer (DIEDERICH *et al.*, 2001; SANTOS *et al.*, 2014).

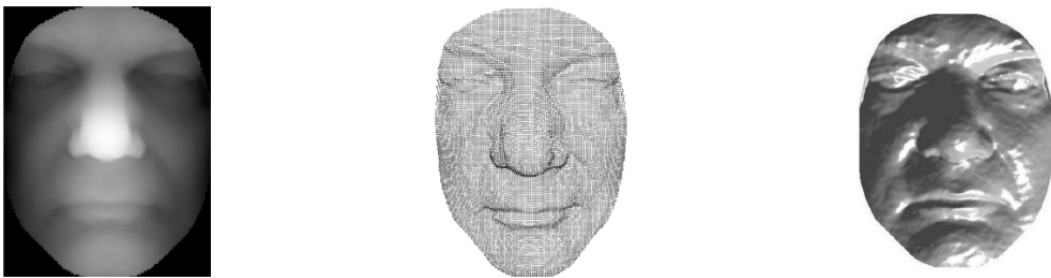
O uso de imagens 3D também está presente na indústria na área de metrologia e inspeção industrial. Nesse caso, objetos são escaneados e comparados, computacionalmente, a

modelos preexistentes de forma a identificar tanto as medidas quanto alguma alteração ocorrida que represente um risco ao bom funcionamento de algum equipamento (SCHUMACHER *et al.*, 2018; SKARŻYŃSKI; SUCHORZEWSKI, 2018).

A técnica de detecção de luz estruturada, *Light Detection and Ranging (LiDAR)*, é usada para muitas aplicações como realidade virtual, realidade aumentada, mapeamento terrestre, aplicações automotivas, navegação espacial, entre outras (VILLA *et al.*, 2021).

Além dessas aplicações, as nuvens de pontos também podem ser usadas para representar faces 3D. Porém, não contamos apenas com essa forma de representação de detalhes da face. A Figura 2 mostra uma face 3D representada pela nuvem de pontos e duas outras formas possíveis.

Figura 2 – Exemplos de Representações de Imagens 3D: Mapa de Profundidade (esquerda), Nuvem de Pontos (centro) e Malha de Triângulos (direita).



Fonte: adaptado de (BENNAMOUN *et al.*, 2015)

A seguir, discutiremos sobre as faces 3D. Após isso, será discutido um sistema de codificação para as formas como a superfície da face é deformada pelos músculos dela. Logo após, é explanado o contexto de reconhecimento facial partindo de dados de face 3D.

2.2 Faces 3D

As faces podem ser representadas de diferentes formas, dependendo dos objetivos para os quais são utilizadas. Além da 3D, algumas representações são baseadas em imagens 2D. Para algumas aplicações, as duas formas de representação podem ser combinadas. Soltanpour *et al.* (2017) mostram diferentes bancos de dados utilizados com aplicações 2D, 3D e combinadas.

A aquisição dos dados de face, assim como a aquisição de diferentes tipos de objetos, está sujeita à computação de erros associados, gerando variações entre uma amostra e outra, ainda que pertencentes ao mesmo indivíduo. Porém, existem outras formas de variação da imagem das faces, como oclusões e expressões faciais.

As oclusões, nesse contexto, são obstruções parciais ou totais da região de interesse da face. Podem ocorrer devido a objetos em frente à face (como mãos, parte do cabelo, óculos, anteparos) ou a más condições de iluminação no momento da aquisição. Alguns exemplos de oclusão são apresentados na Figura 3. As oclusões não são objeto de estudo deste trabalho e, portanto, não são utilizados dados que contenham esse tipo de variação.

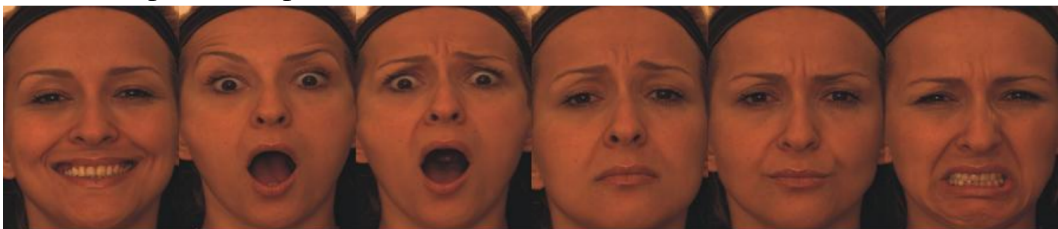
Figura 3 – Exemplos de Casos de Oclusão.



Fonte: adaptado de (SAVRAN *et al.*, 2008b)

Já as expressões são variações na superfície da face ocasionadas pelos músculos faciais. A Figura 4 apresenta exemplos de expressões faciais capturadas para um mesmo indivíduo. As variações da superfície da face ocasionadas por movimentos musculares desta região são catalogadas por sistemas resultantes de estudos realizados especificamente para este domínio, como é o caso do *Facial Action Coding System* (Sistema de Codificação de Ações Faciais) (FACS), discutido na seção a seguir.

Figura 4 – Exemplos de Expressões Faciais.



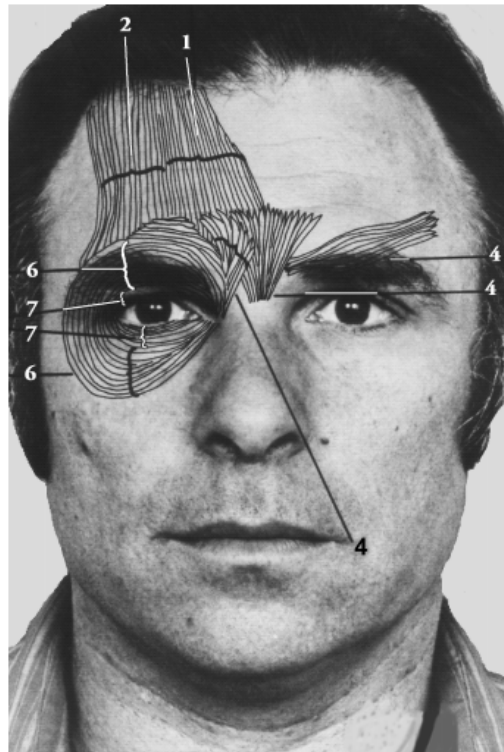
Fonte: adaptado de (SAVRAN *et al.*, 2008b)

2.3 Sistema de Codificação de Ação Facial (FACS)

Dentro do contexto de alterações da superfície da face através da realização de expressões emocionais, o *Facial Action Coding System* (Sistema de Codificação de Ações Faciais) (FACS) (EKMAN; FRIESEN, 1978) possibilita o estudo de emoções básicas, entre outros estados afetivos; leitura de incongruências emocionais; identificação e reconhecimento de microexpressões, macroexpressões, entre outros.

O *FACS* é um sistema para descrição visual de movimentos faciais discerníveis baseados em anatomia. Ele divide as expressões faciais em componentes individuais dos movimentos dos músculos, chamados Unidades de Ação Muscular (Action Units) (AUs), Descritores de Ação (Action Descriptors) (ADs) e Movimentos (Movements) (Ms). Ele é o sistema de codificação de ação facial mais conhecido e frequentemente utilizado para descrever a atividade da face. A Figura 5 traz as marcações dos músculos da face superior do próprio pesquisador Paul Ekman, em que diferentes músculos agem em conjunto para exibir alguma alteração na face. Já na Figura 6, são exibidas as marcações de AUs dos músculos da face superior.

Figura 5 – Músculos da Face Superior.

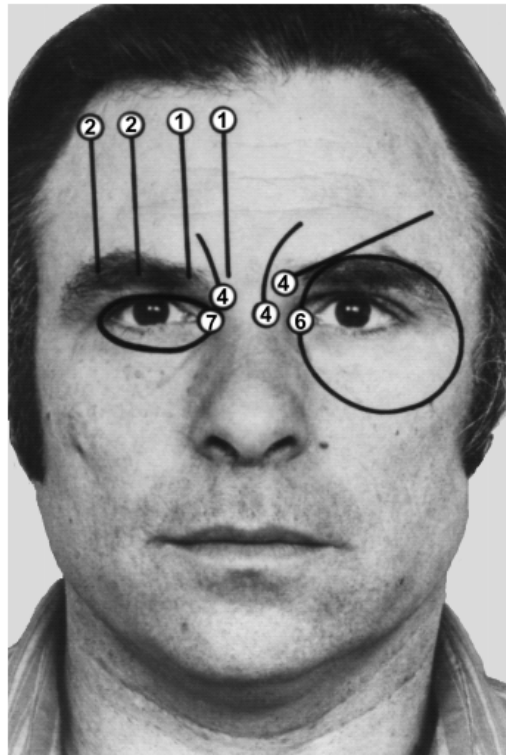


Fonte: adaptado de (EKMAN *et al.*, 2002)

2.4 Reconhecimento de Faces 3D

O reconhecimento de faces 3D traz como vantagens o funcionamento em condições de baixa luminosidade, além de oferecer maior precisão e evitar fraudes com o uso de fotos no lugar de uma pessoa de verdade. Soltanpour *et al.* (2017) classificam o reconhecimento facial em três grupos: holísticos, baseados em atributos e os híbridos. O primeiro se baseia em similaridades globais e utiliza descritores representando a face inteira. Como exemplo, têm-se técnicas como a *Principal Component Analysis (PCA)*, que é aplicada considerando todos os

Figura 6 – Ações Musculares da Face Superior.



Fonte: adaptado de (EKMAN *et al.*, 2002)

pontos da nuvem (RUSS *et al.*, 2006); o delta faces, que é aplicado diretamente na malha 3D (LI *et al.*, 2013) e; o *Fast Nonnegative Tensor Factorization Based On Graph-Preserving (FNFTFGP)*, desenvolvido por (FU *et al.*, 2016), que usa um tensor 4D para classificar expressões. Os métodos baseados em atributos são técnicas que buscam encontrar características locais na face ou focam em regiões específicas como nariz, boca, olhos, etc. Nessa categoria, existem métodos como o desenvolvido por (LI; DA, 2012) e (LEI *et al.*, 2012), que fazem combinações de regiões na face. Já (GUO *et al.*, 2016) e (DENG *et al.*, 2020) utilizam pontos chave como forma de descrever as regiões dinamicamente. Os algoritmos híbridos correspondem a uma combinação entre as duas categorias ou, ainda, à utilização conjunta de dados 2D e 3D (SOLTANPOUR *et al.*, 2017).

Alexandre *et al.* (2020) abordam aspectos relacionados às regiões de interesse na face, as quais podem ser divididas em três grupos: pontos chave, curvas e superfícies locais; ou combinações destas abordagens. Técnicas baseadas em pontos-chave têm o intuito de analisar as vizinhanças desses pontos a fim de descrever as características e atributos dessas adjacências. Zhong (2009) e Gilani *et al.* (2018) utilizam os autovalores da matriz de covariância extraída de uma vizinhança local como forma de selecionar pontos de alta relevância. Darom e Keller (2012) e Li *et al.* (2015) selecionam os pontos utilizando uma adaptação de um algoritmo de detecção projetado inicialmente para imagens 2D.

A abordagem de aprendizagem profunda se tornou um campo de pesquisa popular, com diversas aplicações em reconhecimento de faces 2D e reconhecimento de voz (LUO *et al.*, 2019). Alguns métodos propõem uma aprendizagem profunda por meio de redes neurais convolucionais. O método de Gilani e Mian (2018) alcança 100% de acurácia em um determinado cenário em 3D, entretanto necessita de uma quantidade grande de amostras para treinar a rede neural, incluindo amostras com oclusões geradas de maneira artificial. Zhang *et al.* (2019) alcançam resultados em torno de 98% de acurácia. Já os métodos convencionais, em geral, utilizam uma amostra por indivíduo para formar o conjunto de treino e, além disso, trabalham com uma quantidade menor de atributos. Xu *et al.* (2019) mostram um ganho significativo nas taxas de acurácia ao triplicar a quantidade de amostras no seu conjunto de treino usando redes neurais. Entretanto os resultados usando a abordagem de atributos locais mostram taxas próximas de 100% utilizando apenas uma amostra de cada indivíduo para formar a base de treino. Outro desafio relacionado ao uso de aprendizado profundo é a necessidade de um aparato computacional mais potente do que o de métodos convencionais. Essa abordagem requer, portanto, maior poder de processamento, como processadores mais rápidos e placas gráficas dedicadas, mais espaço de armazenamento e mais tempo de simulação. Como exemplo, Yu *et al.* (2021) chegam a realizar experimentos que levam 50 horas de treinamento em uma configuração de hardware de CPU i7-8700K e duas placas de vídeo GTX 1080TI.

2.5 Extratores de Características

Uma das etapas chave para viabilizar o processo de classificação dentro de um conjunto de dados é a extração de características. Cada dado a ser classificado pode possuir um número de informações variável, que podem ser mais ou menos significativas para um classificador. Na área de visão computacional, a extração de características é vista recorrentemente nas pesquisas (Hackel *et al.*, 2016). Isso ocorre porque extratores de características são técnicas desenvolvidas exatamente para extrair as (possivelmente) mais importantes informações contidas no dado bruto, fazendo com que haja uma redução significativa de sua representação matemática e direcionando o classificador para informações (ou características) que destacam melhor os dados. O resultado disso para o classificador é um aumento de desempenho, seja na velocidade da classificação, seja na taxa de sucesso do classificador.

Para o processo de classificação com imagens 3D, existem diversas técnicas específicas de extração de características, como extratores geométricos, momentos e agrupamento de

histogramas adaptativo (*Adaptive Histogram Binning*), como as apresentado em (Hackel *et al.*, 2016), (HACKEL *et al.*, 2016), (BLOMLEY *et al.*, 2014). Neste trabalho, serão utilizados os extratores geométricos baseados em covariância, discutidos em mais detalhes a seguir.

2.5.1 Extratores Geométricos Baseados em Covariância

Extratores Geométricos são extratores de características que buscam informações baseadas em elementos geométricos da imagem, como pontos, linhas e curvas. Dentre os extratores de características geométricas, existem os baseados em matriz de covariância de uma vizinhança de pontos. Como discutido em (Hackel *et al.*, 2016), diversos autores propuseram o uso desses tipos de extratores de características para o problema de classificação usando nuvens de pontos. Já em (WEINMANN *et al.*, 2017), é discutida a relevância dos extratores geométricos (inclusive os baseados em Covariância) para o problema de classificação com nuvens de pontos 3D.

A técnica de extração de característica baseada em covariância consiste em transformar uma selecionada vizinhança de pontos, com informações potencialmente correlacionadas entre si, em um conjunto de dados descorrelacionados (LIN *et al.*, 2014). Isso permite que informações importantes para a caracterização da imagem a ser classificada se destaquem, além de reduzir a probabilidade de dependência linear nas matrizes resultantes.

A presença de dependência linear em matrizes é indesejável, pois diversos processos de classificação envolvem transformações lineares dessas matrizes, podendo incluir, direta ou indiretamente, inversão de matrizes. Uma matriz que contém dependências lineares não possui matriz inversa e está, portanto, mal condicionada para o processamento da mesma. Isso ocorre porque a solução de qualquer sistema linear baseado nesse tipo de matriz gera infinitas soluções, tornando o cálculo da resposta (no caso, a classificação), altamente sensível a variações na entrada, prejudicando completamente o desempenho do classificador em questão (CHENEY; KINCAID, 2020). Assim, quanto mais próxima uma matriz está de uma matriz com dependências lineares, maior a probabilidade de seu condicionamento afetar o desempenho do processo.

Para realizar esse processo de descorrelação entre os dados, ou seja, a extração de características baseada em covariância, é feito o seguinte, para cada conjunto de pontos em uma vizinhança observada:

1. Calcula-se a matriz de covariância do conjunto;
2. Calcula-se os autovalores e autovetores associados dessa matriz de covariância;

3. Aplicam-se diferentes operações sobre os autovalores (e/ou autovetores). Os resultados dessas operações são justamente as características geométricas baseadas em covariância.

A matriz de covariância \mathbf{C}_i sobre a variância do i -ésimo conjunto de pontos é calculada da seguinte forma:

$$\mathbf{C}_i = \frac{1}{N} \sum_{n \in \mathbf{P}_i^N} (\mathbf{p}_n - \bar{\mathbf{p}})(\mathbf{p}_n - \bar{\mathbf{p}})^T \quad (2.1)$$

Nessa equação, \mathbf{P}_i^N representa o conjunto dos N vizinhos mais próximos de \mathbf{p}_i e $\bar{\mathbf{p}}$ representa o valor médio dos valores do ponto com sua vizinhança.

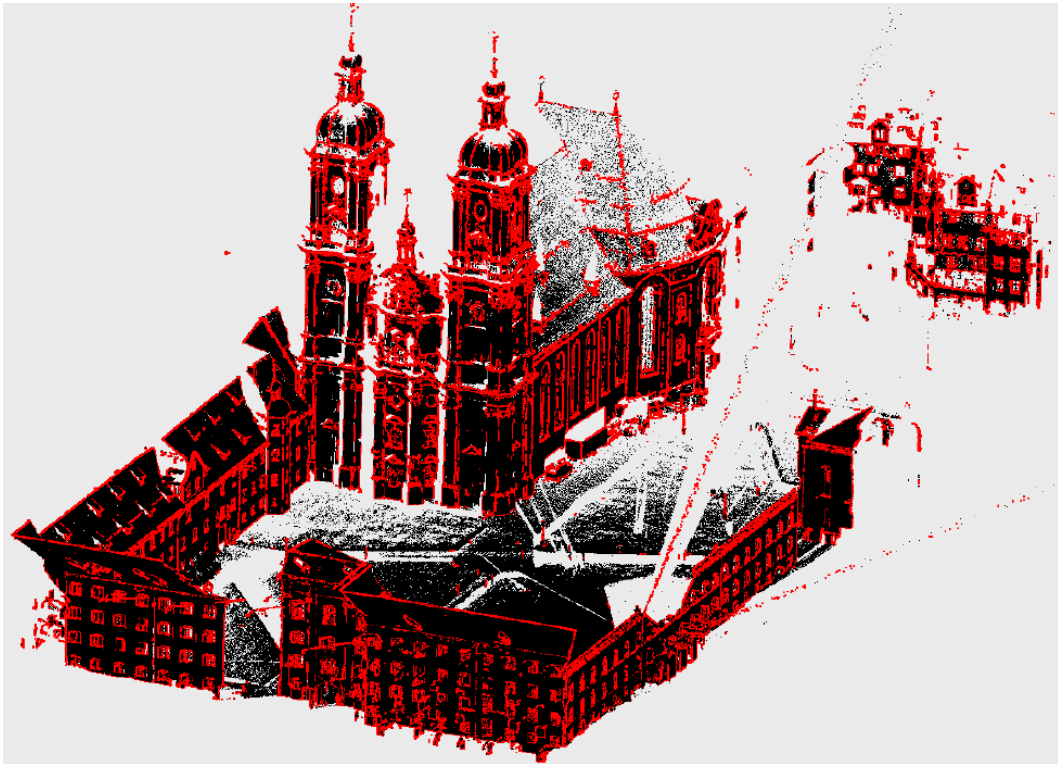
Os autovalores obtidos a partir dessa matriz são definidos por $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$, com seus respectivos autovetores denominados \mathbf{e}_1 , \mathbf{e}_2 e \mathbf{e}_3 . As características são, basicamente, combinações aritméticas desses elementos. Este trabalho usa as seguintes características, inspirados em (Hackel *et al.*, 2016):

- G_1 Somatório: $\lambda_1 + \lambda_2 + \lambda_3$
- G_2 Omnivariância: $(\lambda_1 \cdot \lambda_2 \cdot \lambda_3)^{1/3}$
- G_3 Autoentropia: $\sum_{i=1}^3 \lambda_i \cdot \ln(\lambda_i)$
- G_4 Anisotropia: $(\lambda_1 - \lambda_3)/\lambda_1$
- G_5 Planaridade: $(\lambda_2 - \lambda_3)/\lambda_1$
- G_6 Linearidade: $(\lambda_1 - \lambda_2)/\lambda_1$
- G_7 Variação de superfície: $\lambda_3/(\lambda_1 + \lambda_2 + \lambda_3)$
- G_8 Esfericidade: λ_3/λ_1
- G_9 Verticalidade: $1 - |\langle [0 \ 0 \ 1], \mathbf{e}_3 \rangle|$

Elas são utilizadas como detectores de borda em cenários de detecção de contornos em nuvem de pontos de larga escala de ambientes externos, no intuito de destacar construções em seu meio (Hackel *et al.*, 2016).

Esse processo de detecção de contornos faz com que as características extraídas por esse método forneçam informações geométricas locais em relação à nuvem de pontos. Com isso, esses extratores de características são utilizados com o objetivo de gerar informações importantes sobre o contorno do rosto e mudanças nas AUs quando há a expressão de uma emoção. A Figura 7 mostra um exemplo em que contornos de ambientes externos foram destacados em uma nuvem de pontos 3D utilizando os extratores de características geométricas discutidos neste documento.

Figura 7 – Contornos detectados (em vermelho) de uma nuvem de pontos (em preto) que representa uma construção em um ambiente externo



Fonte: Imagem extraída de (Hackel *et al.*, 2016)

2.6 Classificadores

No processo de classificação são utilizados classificadores de diferentes naturezas, desde algoritmos de baixa complexidade, como o "vizinho mais próximo", até classificadores estatísticos de diferentes abordagens. Nesta seção alguns tipos de classificadores são abordados.

2.6.1 Classificador *K-Nearest Neighbors (KNN)*

Trata-se de um algoritmo de baixa complexidade baseado em comparações de distância. Para cada classe, seleciona-se um número de amostras que as representem. Sempre que um dado novo é fornecido ao classificador, calcula-se a distância desse dado para cada amostra usada para representar as classes do problema. Assim, busca-se quais amostras estão mais próximas do novo dado a ser classificado. Assim, a classe selecionada será baseada nas K amostras mais próximas desse novo dado (vizinhos mais próximos, como o nome do algoritmo) (PEDREGOSA *et al.*, 2011a).

O valor K é definido como o número de vizinhos que será levado em conta nessa avaliação. Se $K = 1$, basta selecionar a amostra mais próxima ao dado fornecido e verificar a

classe dessa amostra mais próxima. A classe do novo dado será o mesmo desse vizinho. Se K for maior do que 1, teremos então uma votação. É esperado desse algoritmo que o dado fornecido esteja próximo a amostras que representam mais de uma classe. Assim, seleciona como classe não a classe pertencente ao vizinho mais próximo, mas a classe mais presente entre os K vizinhos mais próximos (maioria simples). Por exemplo: se $K = 5$, o algoritmo irá selecionar os 5 vizinhos mais próximos do dado novo e vai verificar, desses 5, qual a classe que contém mais membros selecionados. Em caso de empate, costuma-se selecionar aleatoriamente como critério de desempate.

O conceito de distância nesse problema pode ser determinado de várias formas, mas a mais comum costuma ser a distância euclidiana (obtida, subtraindo os dois vetores de atributos e calculando o módulo desse vetor resultante).

2.6.2 *Classificador Gaussiano Naïve-Bayes*

O método *Naïve-Bayes* consiste em realizar uma análise da distribuição de probabilidade de cada um dos atributos do problema afim de determinar qual é a classe mais provável (*likelihood*) para um novo dado a ser classificado. No caso dessa versão do método, assume-se que a distribuição dos atributos é sempre gaussiana para cada classe (PEDREGOSA *et al.*, 2011a). Assim, é calculada uma média e um desvio-padrão para cada atributo, considerando apenas os dados de treinamento de uma determinada classe. Esse processo é feito para cada classe.

Assim, quando um dado é fornecido ao classificador, ele calculará o resultado da função gaussiana para cada atributo de uma classe (considerando a média e o desvio-padrão específicos daquele atributo da classe). Isso vai gerar valores de probabilidade condicional para cada atributo dentro daquela classe. Com isso, calcula-se o produto dessas probabilidades. O resultado é multiplicado pela probabilidade de um dado pertencer àquela classe. Esse valor, chamado probabilidade a *priori*, é determinado pela quantidade de amostras presentes daquela classe, dividido pelo total de amostras de todo o conjunto de treinamento. No final, é gerado uma probabilidade a *posteriori* de aquele dado de entrada pertencer a essa classe.

O mesmo processo é repetido para cada uma das classes do problema, de modo que teremos um conjunto de probabilidades a *posteriori*, um para cada classe. A classe determinada pelo classificador será a com o valor de probabilidade mais alto.

O termo "*naïve*", que significa "ingênuo", se refere ao fato de que esse algoritmo assume que os atributos estão completamente descorrelacionados. Porém, a presença de algum

nível de correlação entre os atributos de um problema do mundo real é bastante comum. Assim, assume-se "ingenuamente" que os atributos sejam, na pior das hipóteses, pouco correlacionados, permitindo assim a boa performance do classificador. A vantagem desse classificador está na simplicidade de implementação e no fato de não necessitar de um grande volume de dados de treinamento para estimar as distribuições gaussianas.

2.6.3 *Classificador Support Vector Machines (SVM)*

As máquinas de vetores de suporte são sistemas de discriminantes binários entre conjuntos de dados. Assim, o treinamento de uma *SVM* baseia-se em achar uma margem otimizada entre dois conjuntos de dados que as separem da melhor forma possível (PEDREGOSA *et al.*, 2011a). Essa separação é avaliada através da distância geral entre as amostras das classes à margem, bem como o nível de permissividade da margem. O nível de permissividade da margem é definido com base numa associação entre cada conjunto de dados, gerando uma análise das melhores posições de margem de modo a aceitar algum erro de classificação enquanto maximiza o desempenho do classificador quanto ao seu nível de generalização do problema. Esse aspecto existe para tornar o sistema adaptável para tolerar *outliers* dentre as amostras de treino.

Sendo esse um classificador essencialmente binário, é importante se utilizar de algumas estratégias de comitê de classificadores. Assim, gera-se diversas margens entre as classes de modo a determinar a que classe um novo dado apresentado deve pertencer. Dentre as estratégias de comitê, temos *one vs one* (um contra um) e *one vs rest* (um contra o resto), detalhados a seguir:

- Na primeira estratégia, temos uma margem que distingue cada par de classes e cada um classificador desse sistema vai definir um voto entre uma de duas classes. A classificação final é determinada pela classe mais votada pelo comitê.
- Na segunda estratégia, cada classificador irá conferir uma "nota" para a sua classe, sendo quanto maior a nota, maior o grau de similaridade do dado apresentado com a classe que aquele classificador aprendeu. A classe decidida nesse caso será a do classificador que determinou a maior "nota" para o dado.

As margens que separam cada classe, geradas pelo *SVM*, são determinadas através de um processo de superdimensionamento do problema, transformando os dados em um espaço com mais dimensões do que originalmente. Isso ocorre porque, frequentemente, há superposição entre as classes, gerando dificuldades para um classificador aprender os limites de cada classe. Com

mais dimensões, se torna viável para o *SVM* traçar superfícies únicas (denominadas hiperplanos) que separam com mais facilidade as classes.

Para essa tarefa de adição de dimensionalidade do problema, são utilizadas diversas funções matemáticas de formato e ordem específicas (chamadas de *Kernel*), que geram diferentes modificações nos dados. Algumas vão facilitar mais a separação, outras menos. Não há uma regra exata para determinar que *kernel* melhor se adequa ao problema, sendo então importante utilizar *kernels* diferentes para determinar qual obtém melhor resultado. Dentre os possíveis *kernels*, temos o linear, o polinomial, a função de base radial (*Radial Basis Function (RBF)*) e a sigmoidal.

O *kernel* polinomial transforma a relação de dois dados *A* e *B* numa relação polinomial de ordem definida *d*. O caso do particular do *kernel* linear quando $d = 1$ é o *kernel* linear. O *kernel* do tipo *RBF* é definido por uma função exponencial (aplicando o módulo da diferença entre *A* e *B*, multiplicado por uma constante negativa γ), determinada pela série infinita de Taylor. O resultado prático desse *kernel* é que ele faz uma análise comparativa entre os dados utilizando virtualmente infinitas dimensões. O *kernel* sigmoidal é similar ao *RBF*, consistindo na aplicação do módulo da diferença entre *A* e *B* a uma tangente hiperbólica (que se trata de uma associação de exponenciais). O princípio é similar mas, enquanto a função exponencial tem uma característica curvatura em um sentido, a sigmoidal muda o sentido da curvatura, gerando um formato similar a uma letra "S", gerando assim uma diferenciação na forma que os dados serão organizados quando superdimensionados.

Quando o *SVM* é treinado, é importante observar dois parâmetros que determinam o comportamento geral do processo. Os parâmetros são *C* e γ . O parâmetro *C* é comum a todos os *kernels*, enquanto o parâmetro γ diz respeito apenas aos *kernels* de origem exponencial (*RBF* e sigmoidal). O parâmetro *C* ajusta o compromisso entre erro de classificação dentro dos dados de treinamento (geralmente por *outliers*) e o nível de simplicidade da margem definida. Um baixo valor de *C* faz com que a superfície de decisão seja mais suave (maior aceitabilidade de erro de classificação) enquanto que um alto valor faz o oposto, tornando o problema de treinamento mais complexo, conseqüentemente, uma margem definida mais detalhada. Já o parâmetro γ determina o nível de influência da distância entre os dois dados selecionados para análise do *kernel*. Isso significa que quanto maior o valor de γ , mais fraca fica a relação entre dois dados distantes entre si, destacando assim a importância na relação de dados mais próximos. A escolha desses parâmetros é empírica.

2.6.4 Classificador Árvore de Decisão (Decision Tree)

Trata-se de um método de decisão que também pode ser usado em aprendizado de máquinas. Consiste em um método que, ao treinar, determina regras de decisão simples, com base nos atributos dos dados e seus respectivos rótulos (PEDREGOSA *et al.*, 2011a). Quando todas as regras de decisão são formuladas, o algoritmo tem assim diversos caminhos em formato de árvore. Assim, cada nó da árvore é uma decisão sobre parte dos atributos de um dado de entrada. Essas decisões se propagam através de um caminho ramificado até que, ao chegar num nó sem ramificação (folha). Cada folha representa uma classe definida. Assim, quando um dado de entrada é fornecido à árvore, uma sequência de decisões é feita até que se chegue a uma das folhas, que indica a que classe esse dado provavelmente pertence.

2.6.5 Classificador Floresta Aleatória (Random Forest)

Esse método se trata, basicamente, de um conjunto de Árvores de Decisão que trabalham em paralelo. Dentre as desvantagens da Árvore de Decisão, temos que há uma possibilidade de o treinamento gerado causar uma alta sensibilidade aos dados de entrada, o que prejudica a capacidade de predição do modelo (PEDREGOSA *et al.*, 2011a).

Tendo isso em mente, a Floresta Aleatória busca desenvolver diversas Árvores de Decisão que operem em subconjuntos de dados de treinamento, com dimensionalidade reduzida (subconjunto dos atributos). Isso faz com que o método gere diversas Árvores de Decisão distintas, pois, caso todas fossem alimentadas com os mesmos dados, com todos os atributos, diversas Árvores seriam iguais ou muito similares. Assim, o processo de classificação se dá por um processo de votação entre cada árvore, sendo o dado de entrada classificado conforme o que a maioria das árvores decide.

A floresta é denominada aleatória porque o processo de geração das árvores de decisão se dá com dois passos aleatórios distintos:

- Seleção aleatória de dados dentro do conjunto de treinamento para cada árvore. Toda árvore recebe a mesma quantidade de dados, mas é permitido ao algoritmo repetir dados em diferentes subconjuntos.
- Seleção aleatória de atributos dentro dos atributos existentes nos dados de treinamento. Cada árvore treinará considerando apenas parte dos atributos possíveis. Isso fará com que cada árvore de decisão treine diferentes aspectos dos dados de treinamento, forçando as

árvores a serem distintas e complementares. Geralmente, a quantidade de atributos a ser selecionada é igual pra cada árvore, podendo haver atributos repetidos.

2.6.6 Classificador Linear Discriminant Analysis (LDA)

Este classificador busca separar os dados de treinamento utilizando um processo de redução de dimensionalidade do problema de modo que aumente a variabilidade entre cada classe, mas reduzindo a variabilidade interna dos dados das classes. Esse processo consiste em selecionar um modelo linear no qual os dados serão projetados, reduzindo a dimensão dos mesmos (PEDREGOSA *et al.*, 2011a).

Para isso, o algoritmo calcula a amostra média de cada classe (contendo os atributos médios das amostras da respectiva classe) e o nível de espalhamento de cada classe independentemente. Em seguida, é calculado um ponto central entre essas classes. Esse ponto central funciona como a referência para a posição da amostra média de cada classe. O objetivo é reduzir a dimensão do problema de modo a maximizar as distâncias das médias ao ponto central ao passo que é minimizado o espalhamento de cada classe. É usado, portanto, como critério de escolha do modelo, a razão entre a soma das distâncias de cada amostra média e a soma de cada nível de espalhamento. Assim, quanto maiores forem as distâncias e menores forem os espalhamentos, mais alto será o valor desse critério e melhor será o modelo.

A classificação consiste em projetar um dado de entrada nesse modelo linear determinado a fim de calcular a que classe o dado se assemelha.

É importante ressaltar que, como o método depende de uma razão entre médias e variâncias, não é possível treinar esse classificador em problemas contendo apenas uma amostra por classe. Isso ocorre porque a variância calculada será zero e o critério de treinamento do modelo depende da razão da soma das médias sobre as variâncias, gerando um valor indeterminado.

2.7 Considerações Finais

Neste capítulo foram apresentados os fundamentos teóricos que embasam este trabalho. Foi explicado como os dados de nuvens de pontos 3D são constituídos e como sua aquisição é feita. Além disso, vimos exemplos de aplicações desse formato de dados. Discutimos também sobre o que é o sistema *FACS*, que cataloga visualmente as múltiplas formas que a face se modifica conforme a pessoa esboça expressões faciais. Na sequência,

vimos o processo de extração de características geométricas, incluindo o que são, além de uma explicação sobre os extratores geométricos baseados em covariância e como são obtidos. Por fim, foram abordados diferentes tipos de classificadores que serão utilizados como ferramenta de avaliação das contribuições propostas neste trabalho. No próximo capítulo utilizaremos o sistema *FACS* e os extratores geométricos baseados em covariância para explorar diferentes formas de segmentação de nuvens de pontos da face.

3 METODOLOGIA

Um aspecto fundamental deste trabalho é a identificação de regiões da face, capturada como nuvem de pontos em 3D, que mais influenciam no processo de identificação de indivíduos e de emoções que suas faces esboçam. As AUs no rosto podem formar emoções que prejudicam a identificação de um indivíduo. Estima-se, portanto, que a redução ou a anulação da influência dessas regiões possam gerar um resultado positivo no processo de classificação, tornando-o mais robusto a expressões faciais.

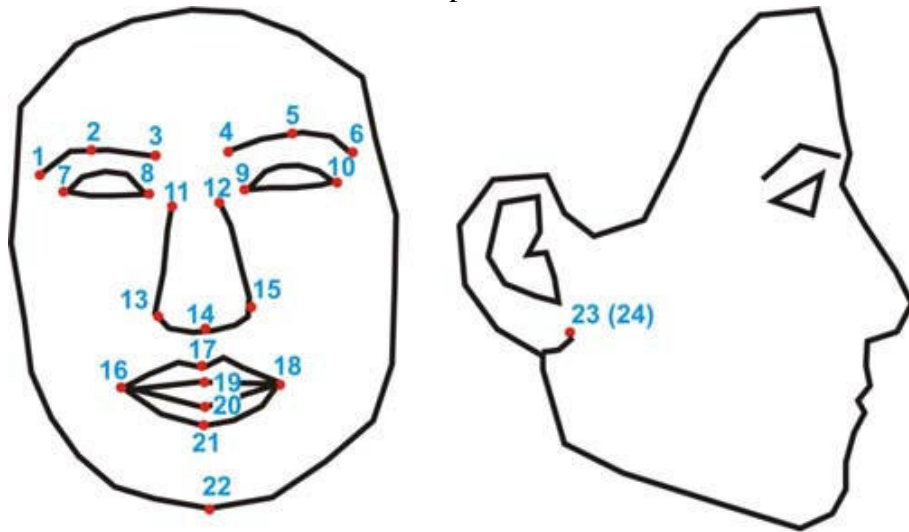
Para isso, neste trabalho, busca-se definir regiões (subnuvens) durante o processo de segmentação da face da nuvem de pontos que representam a face. Considera-se a ponta do nariz como referência para definição do eixo central e recortam-se partes da nuvem de pontos ao redor dele a fim de destacar informações de AUs da face.

No intuito de verificar a influência dos extratores de características geométricas, são utilizados diferentes tipos de extratores em conjunto com o novo tipo de corte. O objetivo é verificar se há um resultado de classificação que indica melhorias em relação a casos sem segmentação ou com outro tipo de segmentação. Caso isso ocorra, fica identificada como promissora essa composição de fatiamento em pizza.

Para realizar esta investigação, é necessária a utilização de um conjunto de dados de faces 3D. Muitas bases de dados têm sido utilizadas com o propósito de realizar o reconhecimento facial de imagens 3D, como: BU-3DFE de (YIN *et al.*, 2006), FRGC v2.0 de (Phillips *et al.*, 2005), Bosphorus de (SAVRAN *et al.*, 2008a), 3DMAD de (Erdogmus; Marcel, 2013), GavabDB de (MORENO; SÁNCHEZ, 2004), entre outros. A Bosphorus certamente é uma das mais desafiadoras, pois traz uma diversidade de expressões que não é vista em outra base de dados com a mesma finalidade. Isso exposto, optou-se por realizar o estudo proposto neste trabalho com a base de dados Bosphorus.

A Bosphorus possui amostras 3D com as respectivas coordenadas dos pontos-chave localizadas nas Unidades de Ação (Action Units - AU), definidas por (EKMAN; FRIESEN, 1976). Entre esses pontos chave, está a ponta do nariz, que é utilizada na abordagem proposta neste trabalho durante o pré-processamento e a extração de características. A Figura 8 traz os pontos fornecidos pela base. A base consiste em 4.666 exemplares faciais de 105 indivíduos em diversas condições distintas, capturadas usando o *scanner Inspeck* 3D. Além disso, não há uma distribuição equitativa nem de exemplares neutros nem de exemplares de expressões dentro da base, o que a torna mais desafiadora.

Figura 8 – Landmarks fornecidas na base Bosphorus.



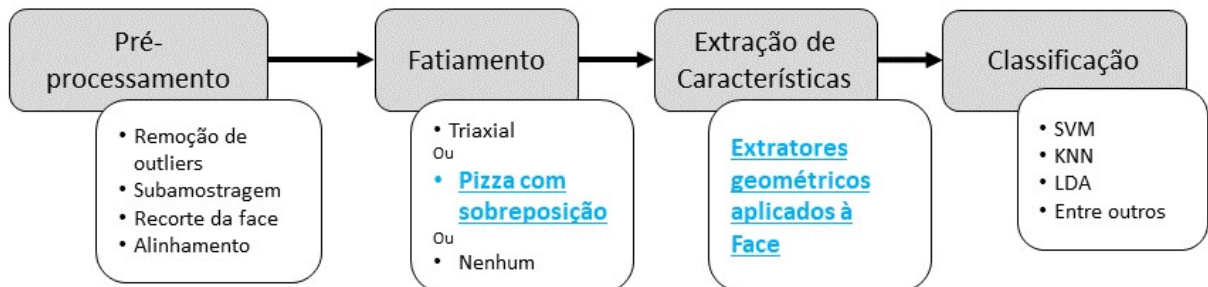
Fonte: Imagem extraída de (SAVRAN *et al.*, 2008a)

O conjunto de dados escolhido é composto pelas amostras neutras, todas as emoções (raiva, tristeza, alegria, medo, nojo e surpresa) e também unidades de ação da parte baixa do rosto (*Lower Face Action Units* - LFAU), unidades de ação da parte alta do rosto (*Upper Face Action Units* - UFAU) e unidades de ação combinadas (*Combined Action Unit* - CAU). Das 4.666 amostras disponíveis na base, portanto, são utilizadas 2.902. As demais amostras contêm elementos distintos do propósito desse estudo, como oclusão e diferentes graus de rotação. Na Tabela 1, está a descrição completa da distribuição de subconjuntos e quantidade de exemplares.

3.1 Etapas do Processo de Classificação

As etapas do processo de classificação realizadas na base de dados são exibidas na Figura 9 e são compostas por pré-processamento, fatiamento, extração de características e classificação. Em destaque, as etapas que são propostas neste trabalho.

Figura 9 – Etapas do processo de classificação de uma imagem 3D



Fonte: a autora.

Tabela 1 – Distribuição por subconjuntos dos exemplares da base de dados Bosphorus

Conjunto / Subconjunto	Rótulo	Exemplares (#)
Expressões	E	2902
Neutras	N	299
Raiva	ANGER	71
Nojo	DISGUST	69
Medo	FEAR	70
Felicidade	HAPPY	106
Tristeza	SAD	66
Surpresa	SURPRISE	71
<i>Lower Face Action Unit</i>	LFAU	1549
<i>Upper Face Action Unit</i>	UFAU	432
<i>Combined Action Unit</i>	CAU	169
Rotações		1365
<i>Yaw Rotation</i>	YR	735
<i>Pitch Rotation</i>	PR	419
<i>Cross Rotation</i>	CR	211
Oclusões	O	381
<i>Eye by Hand</i>	EYE	105
<i>Mouth by hands</i>	MOUTH	105
<i>Eye by Glasses</i>	GLASSES	104
<i>Face by Hair</i>	HAIR	67
Sem rótulo	IGN	18

3.1.1 Cenários de experimentação

Com o intuito de manter formas de comparação com outros trabalhos, utilizamos os cenários mais comuns descritos em (SOLTANPOUR *et al.*, 2017). A Figura 10 traz os 3 experimentos detalhados.

O primeiro experimento é o **Rank-1**, o mais completo e complexo, uma vez que utiliza como base de dados de treinamento apenas um exemplar neutro por indivíduo. É considerado o mais complexo porque o método deve ser de tal forma robusto que, com apenas um exemplar de um indivíduo, seja possível reconhecer outros exemplares do mesmo indivíduo com ou sem expressões. Dessa forma, os atributos devem ser praticamente invariantes a grandes deformações da face. Utiliza-se o primeiro exemplar neutro de cada indivíduo para formar o conjunto de treinamento. Neste caso, há dois tipos de sub-experimentos: **Rank-1a (A)**, Neutro vs Neutro, que utiliza 1 exemplar neutro de treinamento (105 exemplares) e usa os demais neutros como exemplares de teste (194 exemplares), e **Rank-1b (B)**, Neutro vs Não-Neutro, que utiliza 1 exemplar neutro de treinamento (105 exemplares) e usa todos os exemplares Não-Neutros como teste (2603 exemplares).

O terceiro experimento é o **ROC-I (C)**, o mais amplo e que faz o confronto entre todas os exemplares neutros como treinamento (299 exemplares) e todos os Não-Neutros como exemplares de teste (2603 exemplares).

Figura 10 – Descrição dos cenários de experimentação utilizados.

Cenário	Exemplares por indivíduo	Total de exemplares de treinamento	Exemplares de Teste	Total de Exemplares de teste
(A) Rank 1a	UM Neutro	105 exemplares	Neutros	194 exemplares
(B) Rank 1b	UM Neutro	105 exemplares	Não-Neutros	2603 exemplares
(C) ROC-I	TODOS os Neutros	299 exemplares	Não-Neutros	2603 exemplares

Fonte: a autora.

3.2 Pré-processamento

A base passou por algumas etapas de pré-processamento, utilizando funções da biblioteca PCL (*Point Cloud Library*)(RUSU; COUSINS, 2011), construída na linguagem C++. Na Figura 11 são mostradas algumas dessas etapas. A Figura 11a mostra a imagem no estado original, como recebida da base Bosphorus. A seguir, a descrição de cada etapa:

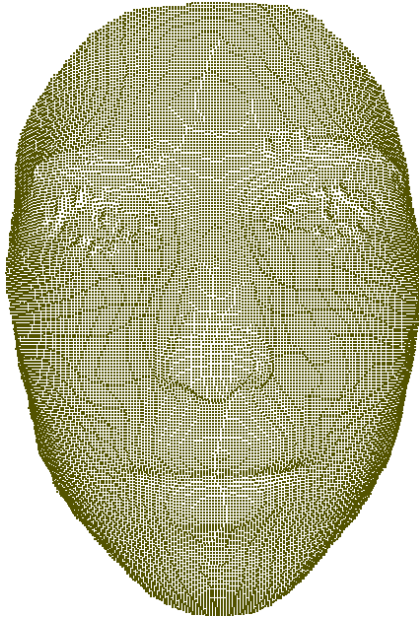
1. **Remoção de outliers** utilizando a função de remoção radial¹ de outliers com raio de busca igual a 5 mm e mínimo de 10 vizinhos dentro do raio. Esta etapa é realizada com o intuito de remover pontos que se diferenciam drasticamente da nuvem de pontos e podem piorar a qualidade da nuvem. A Figura 11b, exibe esse processamento em um exemplar.
2. **Subamostragem** utilizando malhas de cubóides², com cada *voxel* igual a um cubo de aresta 2 mm. Esta etapa é realizada com o intuito de reduzir a quantidade de pontos

¹ Função `pcl::RadiusOutlierRemover`

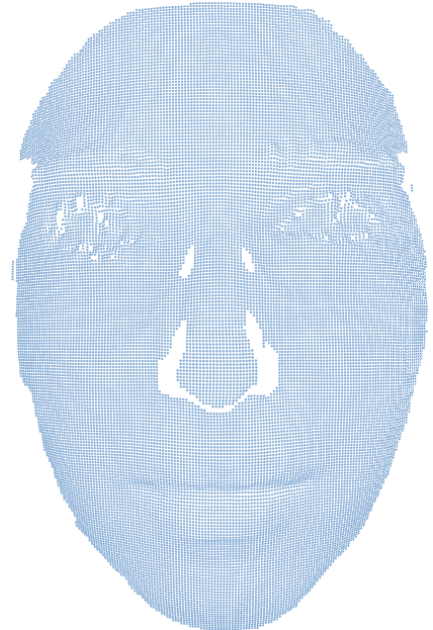
² Função `pcl::VoxelGrid`

Figura 11 – Nuvem de pontos em diferentes etapas de pré-processamento.

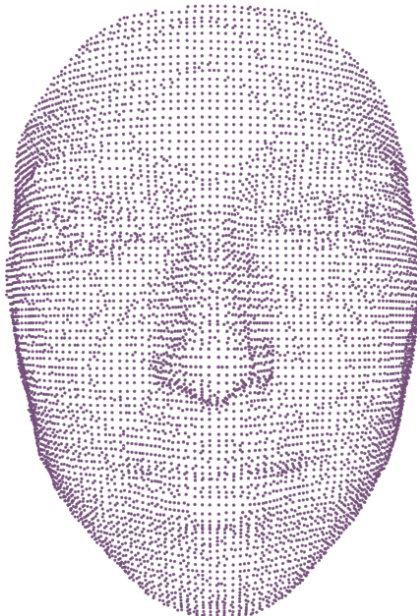
(a) Nuvem de pontos sem processamento. Total de 28452 pontos.



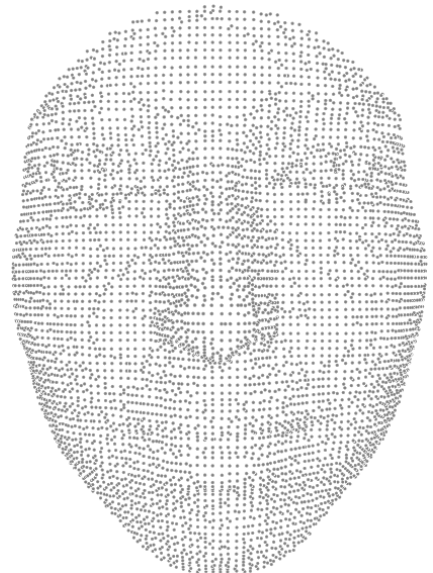
(b) Nuvem de pontos usando remoção de *outliers* radial. Total de 25369 pontos.



(c) Nuvem de pontos usando remoção de *outliers* radial e subamostragem de 2 mm. Total de 7621 pontos.



(d) Nuvem de pontos usando remoção de *outliers* radial, subamostragem de 2 mm e recorte de 80 mm. Total de 5753 pontos.



Fonte: a autora.

processados e pode levar a um menor tempo de processamento nas próximas fases da classificação. A Figura 11c, exibe a combinação dos dois processamentos em um exemplar.

3. **Recorte da face** para remover *background*, partindo da ponta do nariz e com raio 80 mm, no formato de esfera. Esta etapa é realizada com o intuito de remover o plano de fundo

e demais informações irrelevantes para a identificação da face. A Figura 11d, exibe a combinação dos três processamentos em um exemplar.

4. **Alinhamento dos exemplares em duas etapas**, na tentativa de deixar o resultado mais preciso. Em ambas as etapas é realizada uma segmentação esférica com 35 *mm* de raio, deslocada em 1,5 *cm* acima da ponta do nariz (+y) e utiliza-se o algoritmo *Iterative Closest Point* (ICP) (Besl; McKay, 1992).

A primeira etapa alinha todos os primeiros exemplares de faces neutras na coordenada (0,0,0), utilizando os seguintes parâmetros: número máximo de iterações igual a 400, $\epsilon = 0,0001$.

A segunda etapa do alinhamento é feita com todos os demais exemplares, mas desta vez utilizando os seguintes parâmetros: número máximo de iterações igual a 400, $\epsilon = 0,0001$. Esta etapa precisa ser realizada porque as características extraídas da face não são invariantes à rotação ou translação, sendo necessário que todos os exemplares comparados tenham o mesmo referencial³.

Nas etapas de recorte e de alinhamento da face, é necessário utilizar a coordenada na ponta do nariz, fornecidas na própria base de dados Bosphorus. Quando tal referência não é fornecida, é necessário utilizar uma etapa extra de pré-processamento para encontrar a sua localização e que não é o foco deste trabalho. Embora existam pesquisas nesta direção realizadas do grupo no qual este trabalho se insere (JÚNIOR *et al.*, 2019)

3.3 Fatiamento

Após a etapa de alinhamento, é realizado o fatiamento da face em diferentes segmentos. Esse procedimento ajuda no processamento mais rápido das imagens, pois, ao invés de trabalhar com toda a nuvem de pontos, trabalha-se com seções individuais ou um conjunto delas, com uma menor quantidade de pontos.

Foram avaliados dois tipos diferentes de fatiamento: o fatiamento triaxial, proposto em (Siqueira *et al.*, 2018), o fatiamento em pizza com sobreposição, que consiste na proposta apresentada, nesta dissertação, e combinações destas duas formas de fatiamento. Ambos os fatiamentos são descritos nas subseções a seguir.

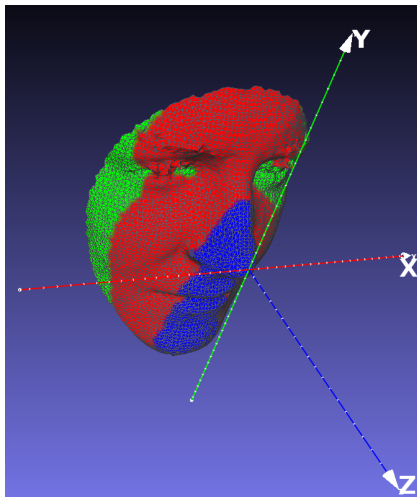
³ Função `pcl::IterativeClosestPoint`

3.3.1 Fatiamento Triaxial

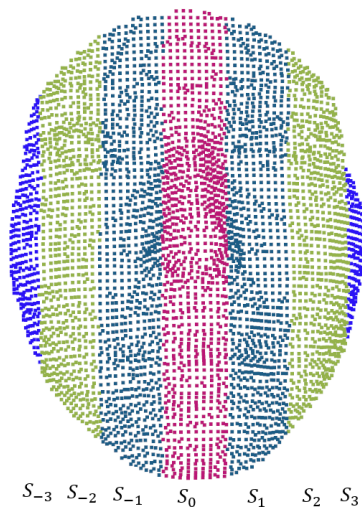
O fatiamento triaxial é composto de fatias frontais, ao longo do eixo z , fatias sagitais, ao longo do eixo x , e fatias transversais, ao longo do eixo y . A ideia por trás do método está relacionada à natureza das deformações na superfície da face sob variações de expressão. Diferentes regiões sofrem diferentes perfis de deformação ao longo de direções preferíveis, enquanto a expressão facial muda. Nos experimentos realizados neste trabalho, utilizam-se os mesmos parâmetros encontrados em (Siqueira *et al.*, 2018). Cada fatia tem uma largura de 20 mm e o fatiamento triaxial é apresentado na Figura 12.

Figura 12 – Fatiamento triaxial, seções frontais, sagitais e transversais.

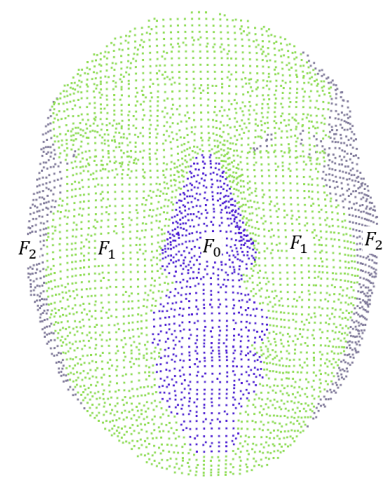
(a) Orientação da face em relação aos eixos x, y, z



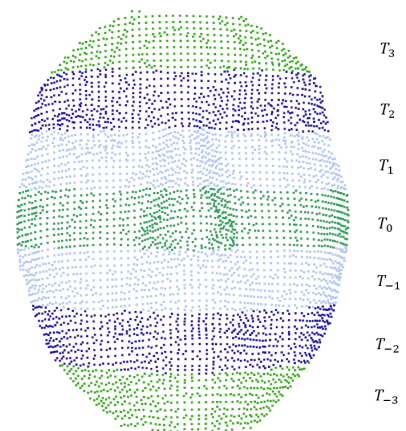
(c) Corte Sagital



(b) Corte Frontal



(d) Corte Transversal



Fonte: adaptado de (Siqueira *et al.*, 2018).

3.3.2 *Fatiamento em Pizza com sobreposição*

Proposto neste trabalho, este fatiamento é inspirado nas remoções das regiões de maior deformação na face durante a expressão de emoções e favorece a seleção de parâmetros independentes para cada eixo da face, podendo, assim, permitir a sobreposição de fatias. Essa sobreposição é inspirada nos movimentos da face, pois uma AU participa de mais de uma emoção (EKMAN; FRIESEN, 1978). Dessa maneira, um mesmo conjunto de músculos pode ser acionado na expressão de raiva e de medo, por exemplo. Na Figura 13, é apresentado o passo a passo da definição do modelo proposto. Nela, os pontos vermelhos são as *landmarks* fornecidas na base de dados.

A Figura 13a mostra a fatia *CT*, que engloba os pontos chave referentes ao nariz, ponta do nariz, cantos externos inferiores e superiores. A Figura 13b mostra as fatias *UP* e *DN*, elas contêm os principais pontos chaves da parte superior e inferior da face, respectivamente. A Figura 13c mostra as fatias *RT* e *LT*, elas contêm os pontos chave das laterais da face, indo da região externa das sobrancelhas até as laterais da boca.

A Figura 13d mostra a sobreposição das fatias *UP* e *RT*, que cobre a região do olho e sobrancelha. Na Figura 13e, as sobreposições das fatias *UP*, *RT* e *CT* são exibidas, concentram a região externa e interna dos olhos, além da área do nariz. Por fim, a Figura 13f mostra a visão frontal da combinação das fatias *UP*, *DN*, *RT*, *LT* e *CT*, que englobam todos os pontos chaves da face.

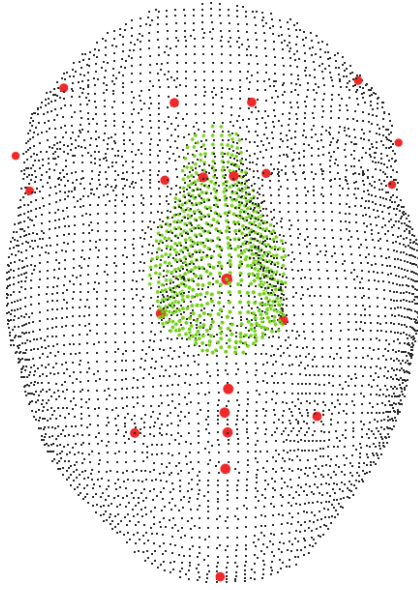
Após a análise descrita, os parâmetros escolhidos e apresentados na lista a seguir foram escolhidos por critérios empíricos, considerando-se a remoção das regiões em que frequentemente se encontram as AUs que promovem maior deformações na superfície da face para se obter um caso inicial (Figura 14). Com estes diferentes fatiamentos, objetiva-se investigar os cenários de segmentação que forneçam os melhores resultados.

- Fatia superior (*UP*) possui raio de corte externo de 80 *mm* e abertura de 150°;
- Fatia inferior (*DN*) possui raio de corte externo de 80 *mm* e abertura de 90°;
- Fatia esquerda (*LT*) possui raio corte interno de 40 *mm* e abertura de 120°;
- Fatia direita (*RT*) possui raio corte interno de 40 *mm* e abertura de 120°;
- Fatia central (*CT*) é uma esfera centrada na ponta do nariz com raio de 35 *mm* e deslocamento para cima (em +*y*) de 15 *mm*.

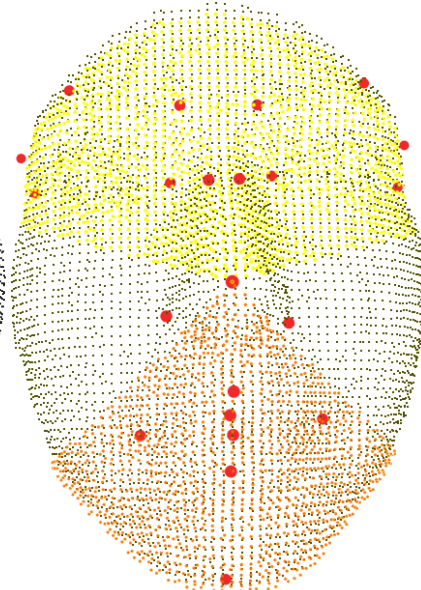
Na Figura 15 vemos as combinações de fatias em pizza sobrepostas através de uma visão parcialmente lateral. Nela os eixos *x*, *y* e *z* são apresentados. É possível perceber como a

Figura 13 – Detalhamento do fatiamento de Pizza com sobreposição. Em vermelho estão as *landmarks* fornecidas pela base Bosphorus.

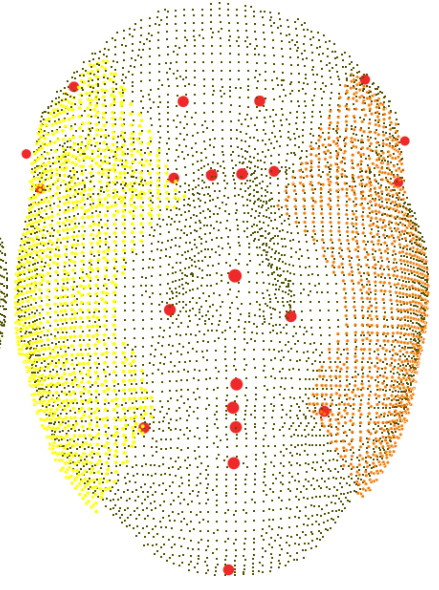
(a) Fatia *CT*



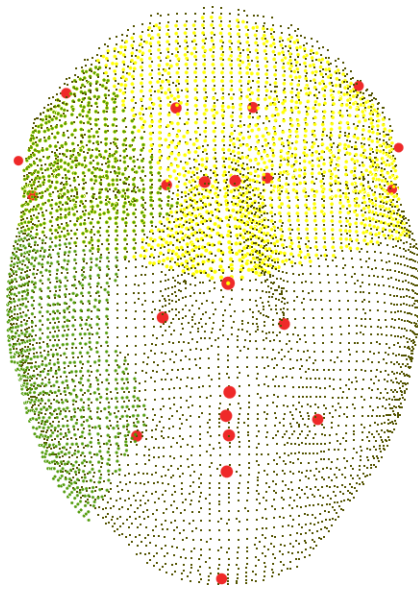
(b) Fatias *UP* e *DN*



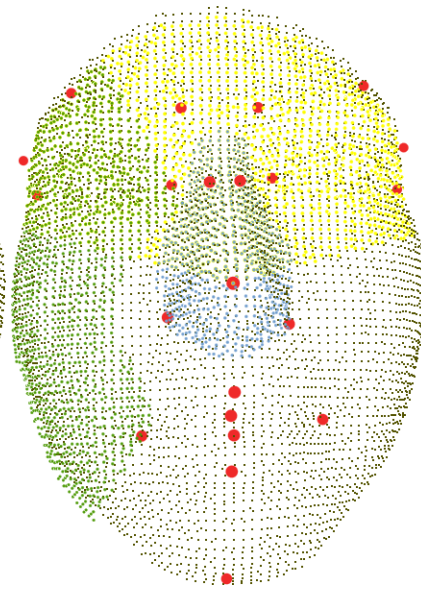
(c) Fatias *RT* e *LT*



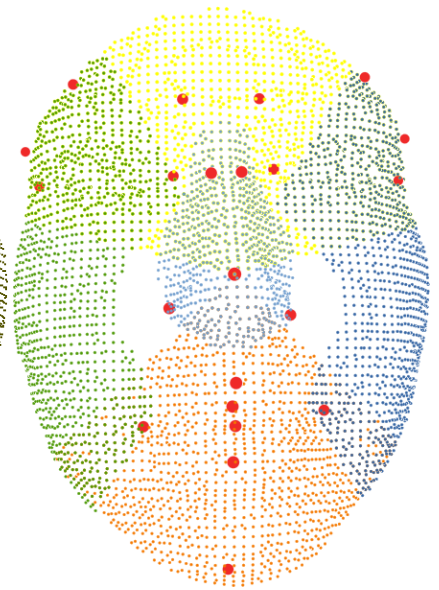
(d) Fatias *UP* e *RT*



(e) Fatias *UP*, *RT* e *CT*



(f) Fatias *UP*, *DN*, *RT*, *LT* e *CT*



Fonte: a autora.

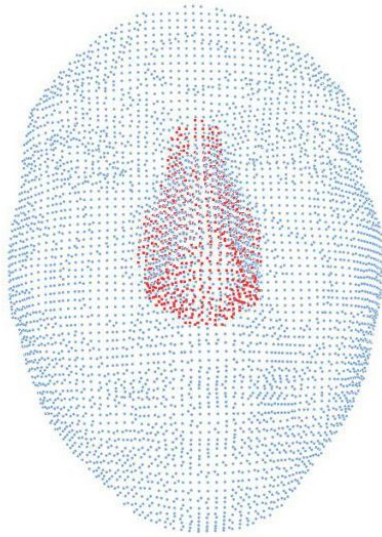
fatia *CT* engloba os pontos referentes ao nariz e também as sobreposições na área dos olhos e da lateral da boca.

3.4 Classificação

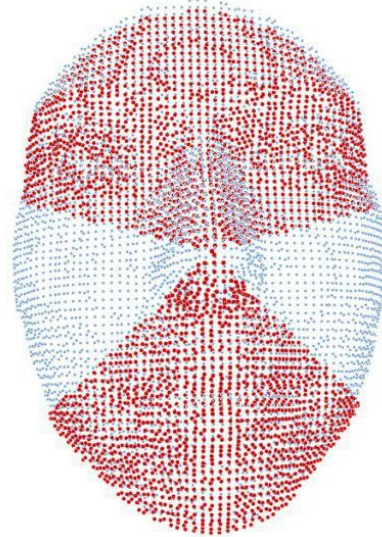
Seguindo os cenários descritos 3.1.1, o processo de classificação realizado consiste na fase de treinamento com parte da base de dados seguida de teste com a classificação dos dados

Figura 14 – Fatiamento em pizza com sobreposição, seções central, Superior e Inferior, e Direita e Esquerda

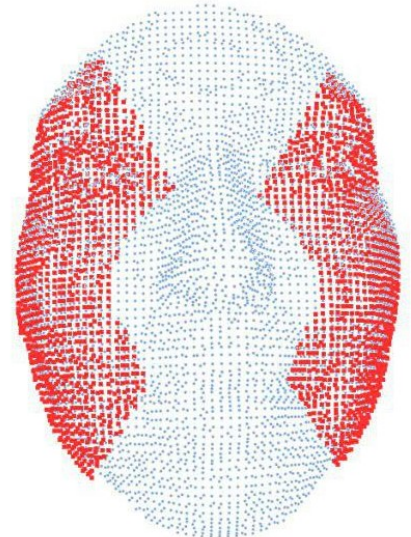
(a) Central



(b) Superior e Inferior

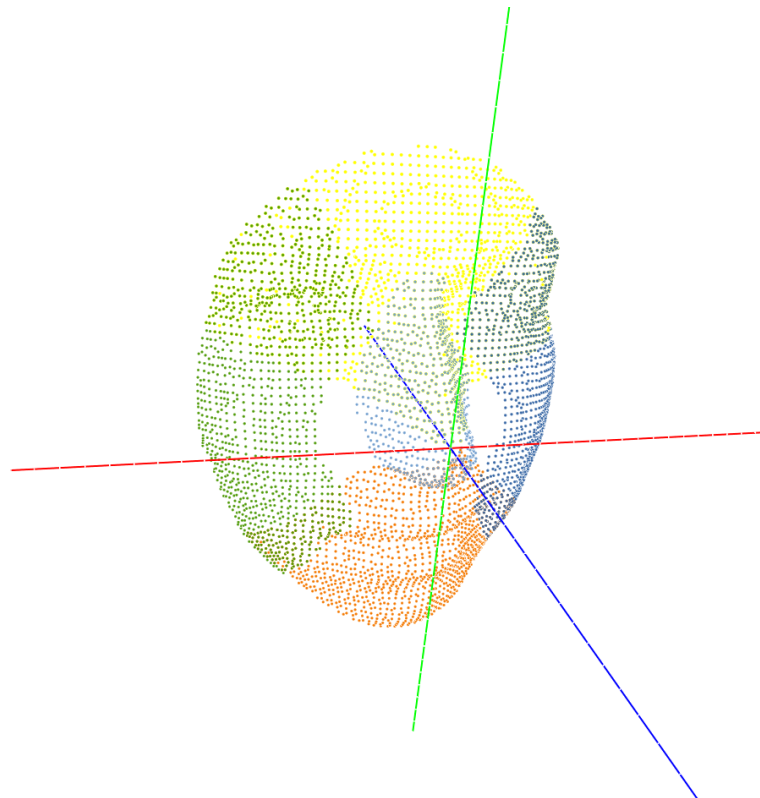


(c) Direita e esquerda



Fonte: a autora.

Figura 15 – Fatiamento em Pizza com sobreposição, todas as fatias (*UP, DN, RT, LT, CT*), visão lateral.



Fonte: a autora.

não usados em treinamento.

Desse teste são computados os acertos e erros dos classificadores a fim de termos

uma avaliação quantitativa dos resultados. Foram utilizados múltiplos classificadores por dois motivos: primeiro, para tornar o problema independente de classificador, permitindo assim uma avaliação mais confiável dos resultados quanto ao fatiamento proposto; segundo, para estimar o nível de complexidade de classificador necessário para solucionar o problema. Assim, foram selecionados alguns classificadores comumente utilizados, tanto entre os mais elementares como entre os mais complexos, descritos na Seção 2.6.

Para a realização de um processo exaustivo treinamento, testes e geração de resultados, foi utilizada a biblioteca *scikit-learn* (PEDREGOSA *et al.*, 2011b), implementada em linguagem *Python*. Uma boa parcela desses classificadores são variantes de *SVM*. Os classificadores são:

- *SVM*, com *kernel RBF*, parâmetros $\gamma = 0.125$ e $C = 8$, (aqui chamado de SVM_RBF);
- *SVM*, com *kernel* polinomial, parâmetros $\gamma = 0.125$ e $C = 8$, (aqui chamado de SVM_poly);
- *SVM*, com *kernel* sigmoidal, parâmetros $\gamma = 0.125$ e $C = 8$, (aqui chamado de SVM_sig);
- *SVM*, com *kernel* linear, parâmetros $\gamma = 0.125$ e $C = 8$, (aqui chamado de SVM_lin);
- *SVM*, com *kernel* polinomial, com configuração padrão (aqui chamado de SVM_pd);
- *SVM*, com *kernel* linear, com configuração padrão (aqui chamado de SVM_ld);
- *LDA*, parâmetros solução de mínimo quadrados e contração automática⁴, (aqui chamado de LDA);
- KNN, com $K = 1$ (aqui chamado de KNN1);
- KNN, com $K = 3$ (aqui chamado de KNN3);
- Um comitê de classificadores *SVM* com *kernel RBF* na configuração "Um contra todos",⁵ (aqui chamado de OneVsRest);
- Gaussiano do tipo *Naïve Bayes* (aqui chamado de GausNB);
- Árvore de Decisão (aqui chamado de DTree);
- Floresta Aleatória (aqui chamado de RandomF).

O classificador *LDA* é utilizado somente no caso ROC-I (C), discutido na Seção 3.1.1. O motivo é que esse classificador necessita de mais de um exemplar por classe para o treinamento.

⁴ Parâmetros *solver = lsqr* e *shrinkage = auto*

⁵ Parâmetros *kernel = rbf*, *degree = 1*, *probability = True*, *random - state = 0*

3.5 Desempenho das experimentações

Este trabalho tem como foco a investigação de diferentes formas de segmentação da face e o uso de características geométricas na identificação de indivíduos sendo a classificação uma das etapas necessárias, utilizada como ferramenta de avaliação. Em cada execução, o uso de diferentes combinações de extratores de características geométricas também é avaliado, visando encontrar as melhores combinações de segmento e características geométricas. Inicialmente, é utilizado o mesmo conjunto de atributos para todo conjunto de segmentos, por exemplo, conjunto de segmentos selecionados: *Superior*, *Central* e *Inferior*; conjunto de atributos selecionados: *G2* (Omnivariância), *G5* (Planaridade), *G8* (Esfericidade); total de combinações avaliadas: *Superior_G2*, *Superior_G5*, *Superior_G8*, *Central_G1*, *Central_G5*, *Central_G8*, *Inferior_G2*, *Inferior_G5* e *Inferior_G8*. Para cada conjunto de fatias selecionadas há 511 combinações possíveis de extratores que podem ser utilizadas. Embora não seja uma execução exaustiva de todas as possibilidades, apresenta-se um grande conjunto de combinações para avaliação.

Ensaio prévios mostram que o tempo de execução pode ser um fator limitante na execução dos experimentos propostos. Por isso, também será realizada uma avaliação da acurácia e tempo de execução dos classificadores pré-selecionados a fim de reduzir o conjunto de classificadores. Após essa avaliação, outras medidas de desempenho como análises de matriz de confusão, curva *Receiver Operating Characteristic (ROC)* e curva *Area Under the Curve (AUC)*, serão realizadas no conjunto de classificadores escolhidos.

3.6 Considerações Finais sobre Metodologia

Apresentamos neste capítulo as propostas de contribuições deste trabalho: uma nova forma de segmentação para nuvens de pontos da face e a avaliação deste método em contraposição e em combinação com métodos já existentes. Além disso, delimitamos os cenários de experimentação e critérios de refinamento para a seleção de ferramentas de avaliação utilizadas. A discussão dos impactos da segmentação proposta é feita no Capítulo 4, onde se avalia a capacidade de representação dos segmentos e seu uso em um cenário de classificação de teste.

4 RESULTADOS

No primeiro experimento realizado, foram considerados: a face inteira; o conjunto das fatias definidas pelo fatiamento triaxial (frontais, sagitais e transversais) e; o conjunto definido para o fatiamento em pizza com sobreposição (lado esquerdo, lado direito, lado superior, lado inferior e fatia central). Todos os extratores definidos na Seção 2.5 foram utilizados e os resultados são exibidos na Tabela 2, sendo **A** o cenário Rank-1a, **B** o cenário Rank-1b e **C** o cenário ROC-I. Os melhores resultados estão em negrito e vários classificadores alcançam o mesmo resultado máximo através de diferentes algoritmos e diferentes custos computacionais. Para cada um dos classificadores, os dois resultados com fatiamento têm taxas de acerto maiores que os resultados da Face Inteira. O classificador LDA só pode ser utilizado no cenário ROC-I pois requer mais de um exemplar de treinamento por indivíduo.

Tabela 2 – Maiores taxas de acerto dos classificadores obtidas de todos os fatiamentos.

classificador	Face Inteira			Pizza S.			Triaxial		
	A (%)	B (%)	C (%)	A (%)	B (%)	C (%)	A (%)	B (%)	C (%)
SVM-rbf	65,46	25,09	23,4	80,93	43,49	42,76	97,94	66,81	46,83
SVM-poly	59,79	23,36	21,17	72,16	37,92	41,30	86,08	48,29	52,25
SVM-sig	54,64	20,98	18,40	53,61	21,90	25,32	79,38	37,46	30,08
SVM-lin	65,46	25,09	25,32	80,93	43,49	46,95	97,94	66,81	68,23
SVM_pd	59,79	23,36	19,13	72,16	37,92	29,54	86,08	48,29	37,88
SVM_ld	65,46	25,09	22,17	80,93	43,49	46,60	97,94	66,81	66,23
KNN-1	65,46	25,09	24,36	80,93	43,49	44,83	97,94	66,81	66,88
KNN-3	14,95	15,75	24,32	19,59	24,55	41,07	16,49	31,39	57,28
OneVsRest	53,09	21,78	19,98	78,35	38,92	37,38	97,24	66,81	24,16
GaussNB	65,46	25,09	14,25	80,93	43,49	18,86	90,21	44,49	45,95
RandomF	54,64	20,48	21,28	46,39	22,82	30,16	32,47	14,68	21,32
DTree	21,13	9,30	18,21	31,96	14,64	23,28	31,96	14,71	21,51
LDA	-	-	25,85	-	-	53,78	-	-	72,11

Na Seção 4.1 são especificadas novas configurações para o Fatiamento de Pizza com Sobreposição. A Configuração 1 corresponde aos parâmetros detalhados na Seção 3.3 e as Configurações 2, 3, 4 foram propostas com a variação de alguns parâmetros isoladamente. Na Seção 4.2 são apresentadas diferentes combinações de fatias triaxiais. Na Seção 4.3 são apresentadas diferentes combinações de fatias triaxiais, fatias de pizza com sobreposição e da face inteira. Na Seção 4.4 é apresentada uma avaliação sobre a acurácia e o tempo de execução dos algoritmos selecionados.

4.1 Resultados do Fatiamento Pizza com Sobreposição

Para avaliar o desempenho do método de fatiamento proposto, foi realizado o experimento com diferentes configurações de ângulos e raios internos e externos, descritos na Tabela 3. Na Tabela 4, são apresentados os melhores resultados de taxa de acerto dos classificadores para cada configuração testada. Nos cenários **A** e **B**, no fatiamento em pizza com sobreposição, cinco classificadores obtiveram o mesma taxa de acerto máxima (SVM-rbf, SVM-lin, SVM-ld, KNN1 e GaussNB), então apenas a taxa de acerto é mencionada (o mesmo padrão é adotado nas demais tabelas). No cenário **C**, o classificador *LDA* obteve a maior taxa de acerto. A Configuração 1 apresentou os melhores resultados e foi escolhida para uma avaliação individual da taxa de acerto de cada fatia. Esses resultados são apresentados na Tabela 5.

Tabela 3 – Configurações do fatiamento em pizza com sobreposição utilizados.

Fatia	Parâmetro	Config.1	Config.2	Config.3	Config.4
Direita (<i>RT</i>)	ângulo (°)	120	90	90	120
	corte interno (<i>mm</i>)	40	40	30	30
Esquerda (<i>LT</i>)	ângulo (°)	120	90	90	120
	corte interno (<i>mm</i>)	40	40	30	30
Superior (<i>UP</i>)	ângulo (°)	150	120	120	150
	corte externo (<i>mm</i>)	80	80	100	100
Inferior (<i>DN</i>)	ângulo (°)	90	120	120	90
	corte externo (<i>mm</i>)	80	80	100	100
Centro (<i>CT</i>)	corte externo (<i>mm</i>)	35	35	40	45
	Deslocamento em +y (<i>mm</i>)	15	15	15	15

Tabela 4 – Melhores taxas de acerto obtidas com fatiamento em pizza com sobreposição nas diferentes configurações escolhidas.

	Configuração 1	Configuração 2	Configuração 3	Configuração 4
Rank-1a (A)	80,93%	77,84%	77,84%	77,84%
Rank-1b (B)	43,49%	42,76%	36,46%	42,76%
ROC-I (C)	53,78%	52,32%	47,29%	52,32%

A fatia que apresentou melhor taxa de acerto no cenário **A** foi a fatia Superior (*UP*) com 60,31% de taxa de acerto. Dentre as combinações testadas *UP* + *DN* teve a melhor taxa de acerto no cenário **A** e a combinação *CT* + *LT* os melhores percentuais nos cenários **B** e **C**.

É possível observar na Tabela 5, as diferenças sutis entre os resultados para o lado direito e para o lado esquerdo, sendo o lado direito mais significativo na identificação do indivíduo que o lado esquerdo. Essa assimetria entre os lados pode ser ressaltada e ajudar tanto

Tabela 5 – Resultados obtidos com fatiamento em pizza com sobreposição na Configuração 1 e utilizando cada segmento individualmente e combinações dos mesmos

Fatia	Cenário	Taxa de Acerto
Superior (<i>UP</i>)	A	60,31%
	B	20,98%
	C	22,13%
Inferior (<i>DN</i>)	A	50,52%
	B	19,52%
	C	21,13%
Lado Direito (<i>RT</i>)	A	42,27%
	B	20,36%
	C	22,86%
Lado Esquerdo (<i>LT</i>)	A	38,14%
	B	21,13%
	C	22,9%
Central (<i>CT</i>)	A	37,63%
	B	21,59%
	C	25,12%
<i>UP + LT</i>	A	68,56%
	B	29,97%
	C	37,01%
<i>UP + RT</i>	A	67,53%
	B	31,73%
	C	37,42%
<i>UP + DN</i>	A	71,13%
	B	29,62%
	C	35,27%
<i>CT + LT</i>	A	65,98%
	B	36,11%
	C	44,06%
<i>CT + RT</i>	A	62,89%
	B	35,88%
	C	42,95%
<i>UP + DN + LT</i>	A	70,10%
	B	31,69%
	C	39,84%
<i>UP + DN + RT</i>	A	70,10%
	B	31,73%
	C	40,41%

na identificação de indivíduos, objetivo deste trabalho, como na identificação de emoções. Tal assimetria é objeto de atenção desde muito tempo. Trabalhos como (SACKEIM *et al.*, 1978), investigam essa assimetria, propondo o uso de imagens compostas por apenas um dos lados da face concatenando a esse mesmo lado duplicado e espelhado. Os resultados mostram que são percebidas expressões mais significativas nas imagens compostas com o lado esquerdo da face.

Ekman (1980) publicou um artigo questionando metodologia e as conclusões de Sackeim *et al.* (1978). Nesse artigo, Ekman alerta para outras possibilidades que justificariam essa percepção sobre as faces com apenas lados esquerdos, como diferenças anatômicas e nervos periféricos. Ele também ressalta as diferenças entre expressões emocionais espontâneas e movimentos faciais voluntários, como emoções "posadas". Elas utilizam caminhos neurológicos diferentes e Ekman sugere que a assimetria de expressões faciais ocorre apenas em emoções deliberadas e não-espontâneas, tal como o sorriso ou expressão de felicidade. Entretanto, ele afirma que, para que se possa generalizar para as outras emoções, é necessária a realização de estudos de movimentos musculares envolvendo emoções negativas. Isso significa que, segundo Ekman, os resultados de Sackeim *et al.* (1978) não são conclusivos e requerem uma investigação mais detalhada a fim de determinar se a face humana realmente possui uma assimetria nas expressões no sentido de termos o lado esquerdo como mais informativo de estados de emoções.

Apesar das diferentes opiniões sobre a origem dessa assimetria, ela foi detectada estatisticamente em mais de um experimento (HAGER; EKMAN, 2005). A assimetria das expressões emocionais da face, com maior expressividade do lado esquerdo, também foi verificada quantitativamente em (Makioka *et al.*, 2016). Nesse estudo, Makioka *et al.* (2016) propuseram medidas objetivas para avaliação da assimetria facial e realizaram dois experimentos utilizando a base de dados BU-3DFE e as expressões neutra, raiva, felicidade, tristeza e surpresa. O primeiro experimento utilizou 15 participantes, que avaliaram qualitativamente imagens espelhadas de cada lado da face. O segundo experimento utilizou redes neurais multicamadas para gerar máscaras de emoções e essas máscaras aumentaram a taxa de reconhecimento em 4%.

4.2 Resultados do Fatiamento Triaxial

Nos testes realizados nesta investigação, os melhores resultados foram obtidos com o fatiamento Triaxial combinando as fatias sagitais (S_i) e transversais (T_i), confirmando a melhor configuração encontrada por Siqueira *et al.* (2018), autores desta abordagem. Na Tabela 6 são apresentadas essa e outras configurações, utilizando extratores de características diferentes das do autor. Além do conjunto ST (listado a seguir), em que S_{-i} representa fatias i posições à esquerda da fatia sagital central (S_0) e T_{-i} as fatias que estão i posições abaixo em relação à transversal central (T_0), foram testadas outras seleções:

- $ST(S_0, S_{-1}, S_1, S_{-2}, S_2, S_{-3}, S_3, T_0, T_{-1}, T_1, T_{-2}, T_2, T_{-3}, T_3)$,
- Direita (S_1, S_2, S_3),

- Esquerda (S_{-1}, S_{-2}, S_{-3}),
- Superior (T_1, T_2),
- Inferior (T_{-2}, T_{-3}),
- Central (T_0, S_0),
- Central Largo ($T_{-1}, T_0, T_1, S_{-1}, S_0, S_1$).

Tabela 6 – Maiores taxas de acerto dos classificadores obtidas com fatiamento triaxial e outros conjuntos de fatia.

Fatia	A (%)	B (%)	C (%)
<i>ST</i>	97,94	66,81	72,11
F_0, F_1, F_2, F_3	78,87	44,33	52,52
F_0, F_1, F_2	87,63	48,67	52,71
F_1	62,37	27,01	29,04
Direita	79,38	35,19	37,76
Esquerda	82,99	39,61	44,22
Superior	65,46	45,79	50,98
Inferior	60,31	16,02	18,56
Central	56,70	28,97	31,77
Central Largo	89,69	48,25	54,86

Na Tabela 6, de fatiamento triaxial, cinco classificadores obtiveram a mesma taxa máxima de acerto nos cenários **A** e **B**: SVM-rbf, SVM-lin, SVM-ld, KNN1 e OneVsRest. No cenário **C**, a maior taxa de acerto foi do classificador *LDA*. Em todos os casos, apenas o valor da taxa de acerto é exibido na tabela.

4.3 Resultados da Combinação de Fatiamento Triaxial, Fatiamento em Pizza com Sobreposição e Face Inteira

Na Tabela 7, são apresentadas diferentes combinações de seções de fatiamento triaxial, fatiamento de pizza com sobreposição e face inteira. Essas combinações apresentam taxas de acerto maiores que as obtidas utilizando apenas o fatiamento de pizza com sobreposição. Das combinações testadas, o conjunto *UP*, *CT* e *FULL*, em que *FULL* representa o uso da face inteira, obteve as maiores taxas de acertos nos cenários **A**, **B** e **C**, superando os valores individuais obtidos pelo fatiamento em pizza com sobreposição e a face inteira.

Tabela 7 – Resultados obtidos com combinações do fatiamento triaxial, fatiamento de pizza com sobreposição e face inteira.

Fatias	A (%)	B (%)	C (%)
<i>UP, S₀, S₋₁, S₁</i>	78,87	43,07	49,48
<i>UP, F₁</i>	82,99	42,37	46,60
<i>UP, F₀</i>	76,29	33,96	38,26
<i>UP, DN, F₀</i>	71,13	34,08	44,53
<i>CT, S₀, S₋₁, S₁</i>	81,96	45,45	50,63
<i>UP, FULL</i>	86,6	40,88	43,33
<i>UP, CT, FULL</i>	88,14	50,94	55,97

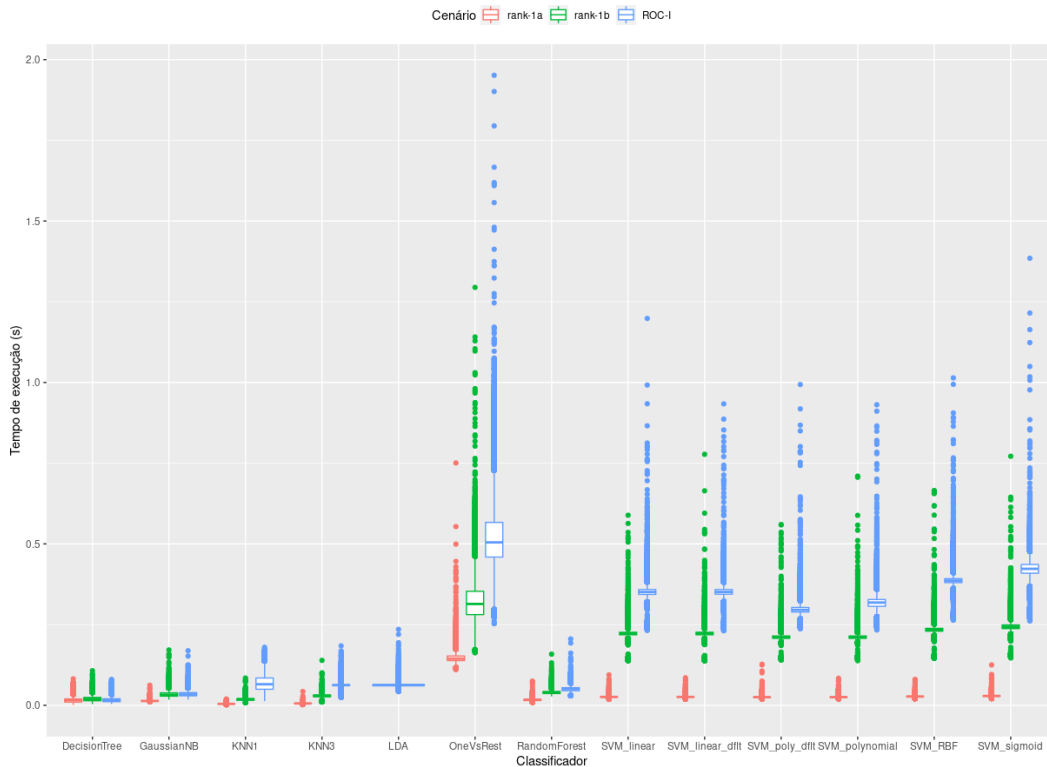
4.4 Desempenho dos Algoritmos de Classificação

Um dos fatores limitantes na execução de diferentes experimentos foi o tempo de execução dos mesmos. Vale lembrar que, para cada seleção de fatias avaliada, são realizados os testes utilizando todas as combinações de extratores de características. Por exemplo, a simulação de todas as combinações de fatias de uma configuração em pizza, 5 tipos de fatias e 9 tipos de extratores, limitando o mesmo conjunto de extratores para todas as fatias selecionadas, leva aproximadamente 24h de execução em um computador com processador Intel Core-I7 6ª geração, 16GB de RAM, placa de vídeo NVidia GeForce 940M 4GB, e SSD. A mesma execução para o fatiamento triaxial, 20 tipos de fatias e 9 tipos de extratores, não pode ser realizada completamente, sendo selecionadas combinações específicas de fatias para analisar, visto foram necessários 6 dias para executar 1% desta combinação. Nesta Seção faremos a análise de desempenho dos classificadores para encontrar uma seleção mais otimizada com o objetivo de possibilitar a execução de experimentos mais diversificados.

Na Figura 16, temos o diagrama de caixa (*box plot*) do tempo de execução de todos os cenários e em todas as configurações de fatiamento. O classificador que apresenta maior tempo de execução é o *OnevsRest*. Também é apresentada uma significativa diferença de tempos de execução entre os cenários Rank-1a (A), Rank-1b (B) e ROC-I (C), sendo ROC-I o maior deles. Isso é justificado pela diferença na quantidade de exemplares de treinamento e teste, sendo ROC-I o cenário mais desafiador. Vale ressaltar que o classificador *LDA* só pode ser executado no cenário ROC-I (C), pois necessita de mais de um exemplar de treinamento por classe. As execuções mais demoradas levam em torno de 2 segundos e as mais rápidas apenas alguns milissegundos.

Na Figura 17, temos o diagrama de caixa para o fatiamento em pizza com sobreposição, configuração 1. Na primeira seção do gráfico, na cor vermelha, estão os diagramas de

Figura 16 – Tempo de Execução *versus* Classificadores, todas as configurações



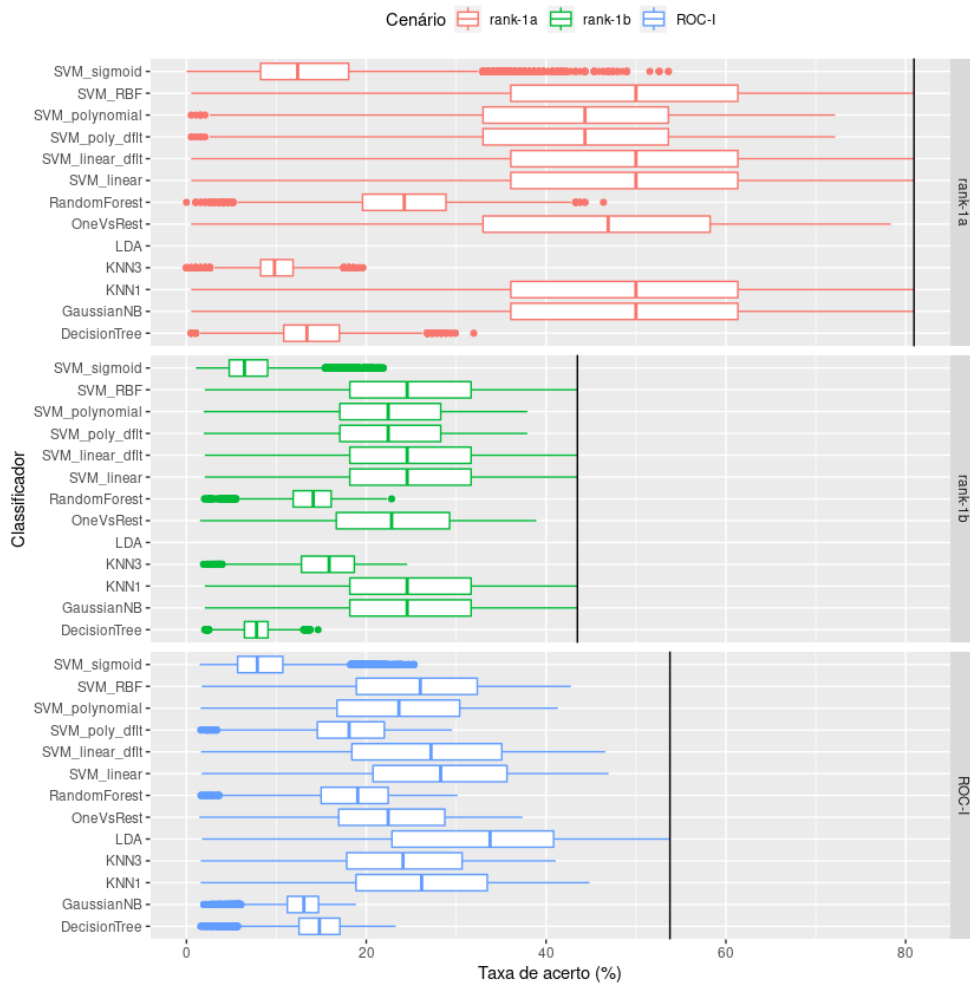
Fonte: a autora.

caixa de cada classificador para o cenário Rank-1a (**A**), a linha vertical representa a maior taxa de acerto alcançada pelos classificadores e está posicionada em 80,93%. Reparem que cinco classificadores, citados anteriormente, atingem esta taxa máxima. Cada dado representado em cada diagrama de caixa simboliza uma combinação de fatias e de extratores de características. Na segunda seção do gráfico, na cor verde, está o cenário Rank-1b (**B**), a linha vertical, representando a taxa máxima alcançada, está posicionada em 43,49%. Na terceira seção do gráfico, na cor azul, está o cenário ROC-I (**C**), com a taxa máxima alcançada posicionada em 53,78%.

Na Figura 18, temos os diagramas de caixa para o fatiamento triaxial. As maiores taxas de acerto na primeira seção, cenário Rank-1a (**A**), na cor vermelha, estão em 97,94%. Na segunda seção, cenário Rank-1b (**B**), na cor verde, a maior taxa de acerto está em 66,81%. Na terceira seção, cenário ROC-I (**C**), na cor azul, a maior taxa de acerto está em 72,11%. Cada coluna do diagrama representa a distribuição da taxa de acerto das combinações apresentadas na seção 4.2 e dos 9 extratores.

Na Figura 19, temos o diagrama de caixa para a face inteira (*FULL*). As maiores taxas de acerto na primeira seção, cenário Rank-1a (**A**), na cor vermelha, estão em 65,46%. Na segunda seção, cenário Rank-1b (**B**), na cor verde, a maior taxa de acerto está em 25,09%. Na terceira seção, cenário ROC-I (**C**), na cor azul, a maior taxa de acerto está em 25,85%. Cada

Figura 17 – Taxa de acerto *versus* Classificador, Configuração Pizza C1



Fonte: a autora.

coluna do diagrama representa a distribuição da taxa de acerto das combinações dos 9 extratores.

Na Figura 20, temos a taxa de acerto *versus* o tempo de execução para o cenário Rank-1a (A). Nesse gráfico temos a taxa de acerto representada no eixo y e o tempo de execução simbolizado através da escala de cor, que vai de um tom de verde até a cor preta, que representa os maiores tempos de execução, além da separação de dados no eixo x para cada classificador. A maior taxa de acerto encontrada foi sinalizada com uma linha horizontal em 80,93%. Neste contexto, o melhor classificador seria o de maior taxa de acerto, maior valor no eixo y e menor tempo de execução, cor mais clara. O classificador de maior custo computacional é o *OnevsRest*, distanciando-se bastante dos demais classificadores. Este é o cenário mais simples, onde o treinamento é feito com um exemplar neutro por indivíduo e o teste com os demais exemplares neutros.

Na Figura 21, temos a taxa de acerto *versus* o tempo de execução para o cenário Rank-1b. Neste caso, o classificador de maior custo computacional é o *OnevsRest* e, depois dele,

Figura 18 – Taxa de acerto *versus* Classificador, Configuração Triaxial



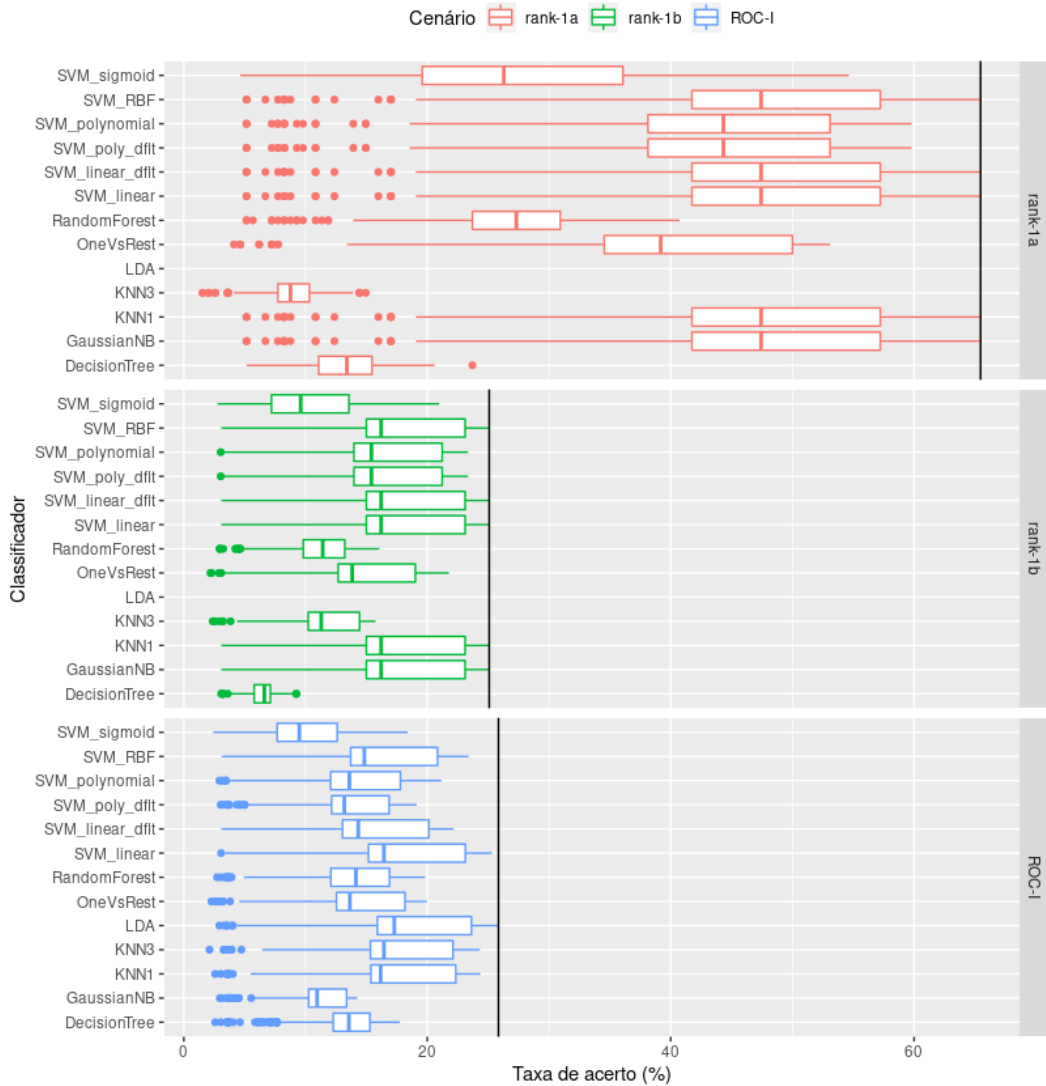
Fonte: a autora.

os classificadores do tipo *SVM*. Este é o cenário intermediário, onde o treinamento é feito com um exemplar neutro por indivíduo e o teste com exemplares não-neutros.

Na Figura 22, temos a taxa de acerto *versus* o tempo de execução para o cenário ROC-I. O classificador de maior custo computacional, neste cenário, é o *OnevsRest* com uma grande dispersão e, em seguida, os classificadores do tipo *SVM*. Este é o cenário mais complexo, onde o treinamento é feito com todos os exemplares neutros e o teste com exemplares não-neutros.

O primeiro classificador removido da investigação foi o *OnevsRest*, devido a sua baixa eficiência. Com isso, o tempo de execução utilizando a configuração de fatiamento em pizza com sobreposição caiu de 24h para 18h. O próximo passo é o refinamento do conjunto de classificadores *SVM* utilizados. Após esta etapa, poderemos realizar análises mais refinadas utilizando matriz de confusão para analisar as classes com menor acurácia para testar outras

Figura 19 – Taxa de acerto *versus* Classificador, Configuração *FULL*

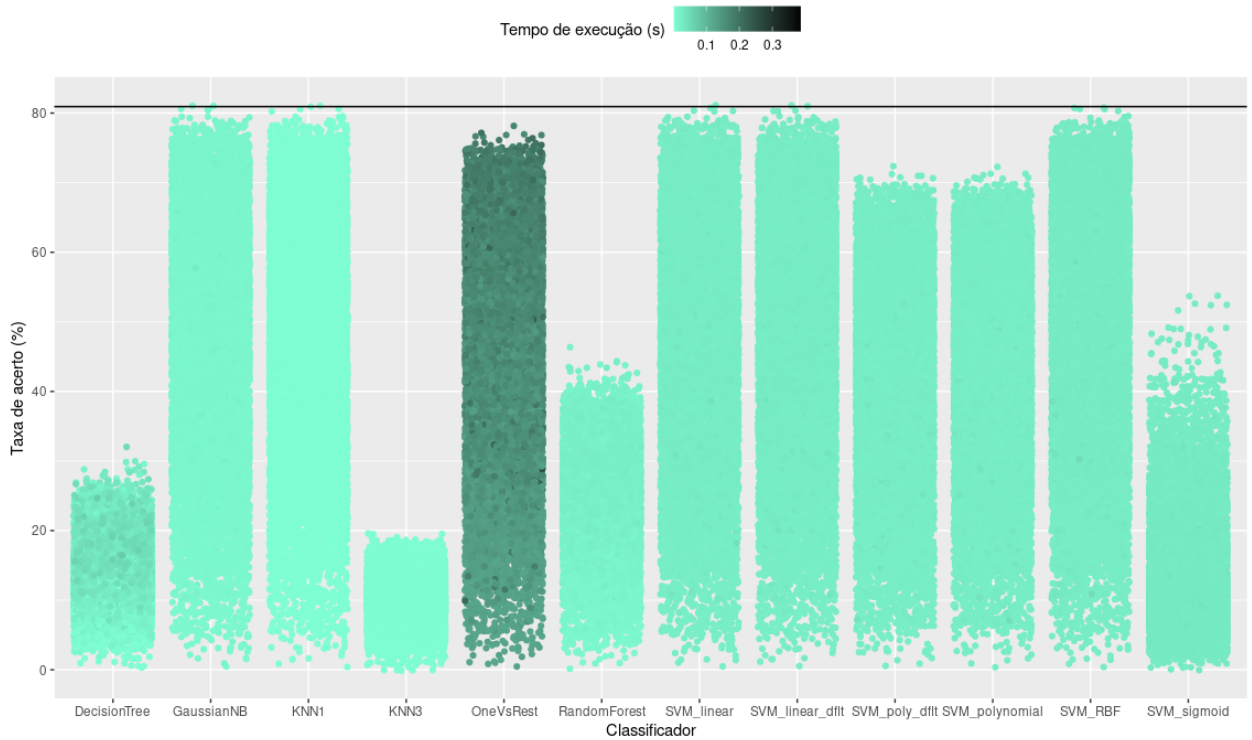


Fonte: a autora.

formas de pré-processamento e alinhamento.

Analisando os classificadores com tempos de execução mais baixos e com as maiores taxas de acerto, podemos destacar o KNN-1. Na Figura 23 são mostradas as taxas de acerto do KNN-1, no cenário Rank-1a (A), na configuração Triaxial *ST*, com cada extrator de características individualmente. Com um único extrator de característica (G_2 - omnivariância), é possível alcançar 92,78% de taxa de acerto. A segunda maior taxa de acerto é obtida com G_8 , esfericidade, com 92,27%, seguida por G_7 , variação de superfície, com 91,75%. O extrator com menor tempo foi G_4 , anisotropia, com 90,72%. Neste contexto, o maior valor alcançado foi 97,94% com cada uma das seguintes combinações de características: $(G_1, G_5, G_7, G_8, G_9)$, $(G_3, G_5, G_7, G_8, G_9)$, $(G_1, G_2, G_5, G_7, G_8, G_9)$ e $(G_2, G_3, G_5, G_7, G_8, G_9)$.

Figura 20 – Taxa de acerto *versus* Tempo de Execução, Cenário Rank-1a (A), Configuração Pizza C1

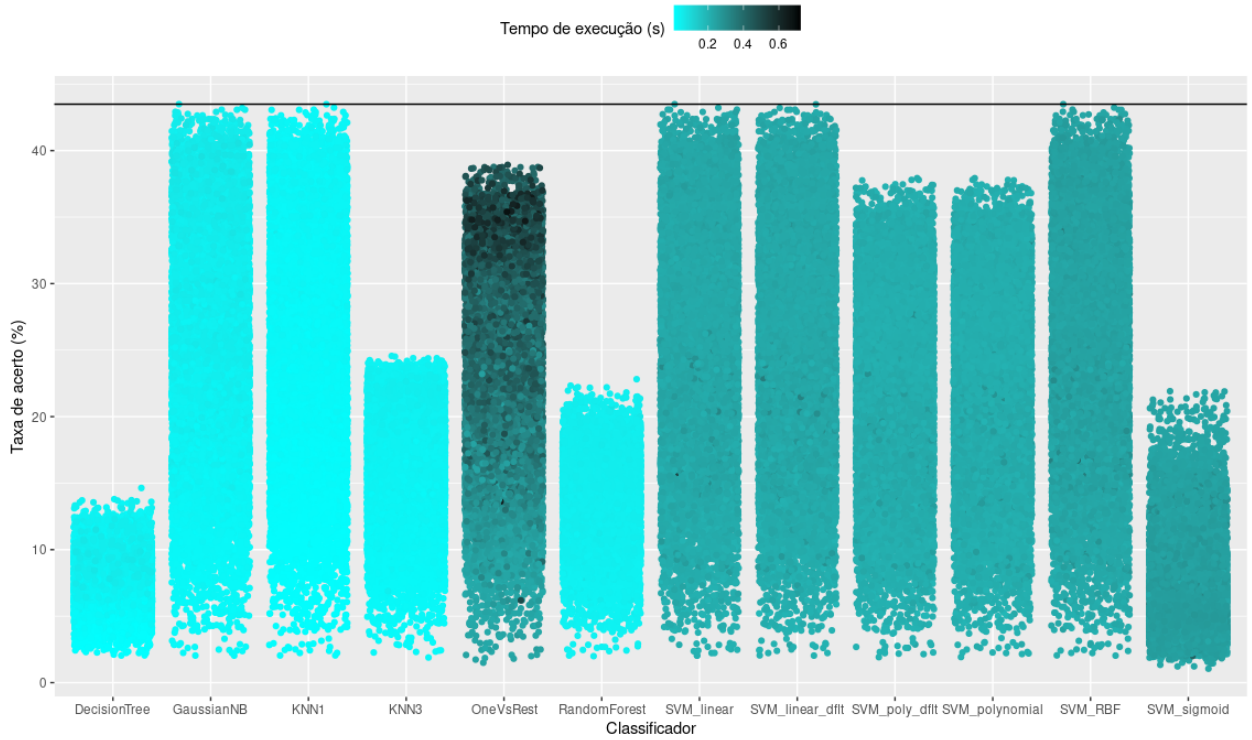


Fonte: a autora.

4.5 Considerações Finais sobre Resultados

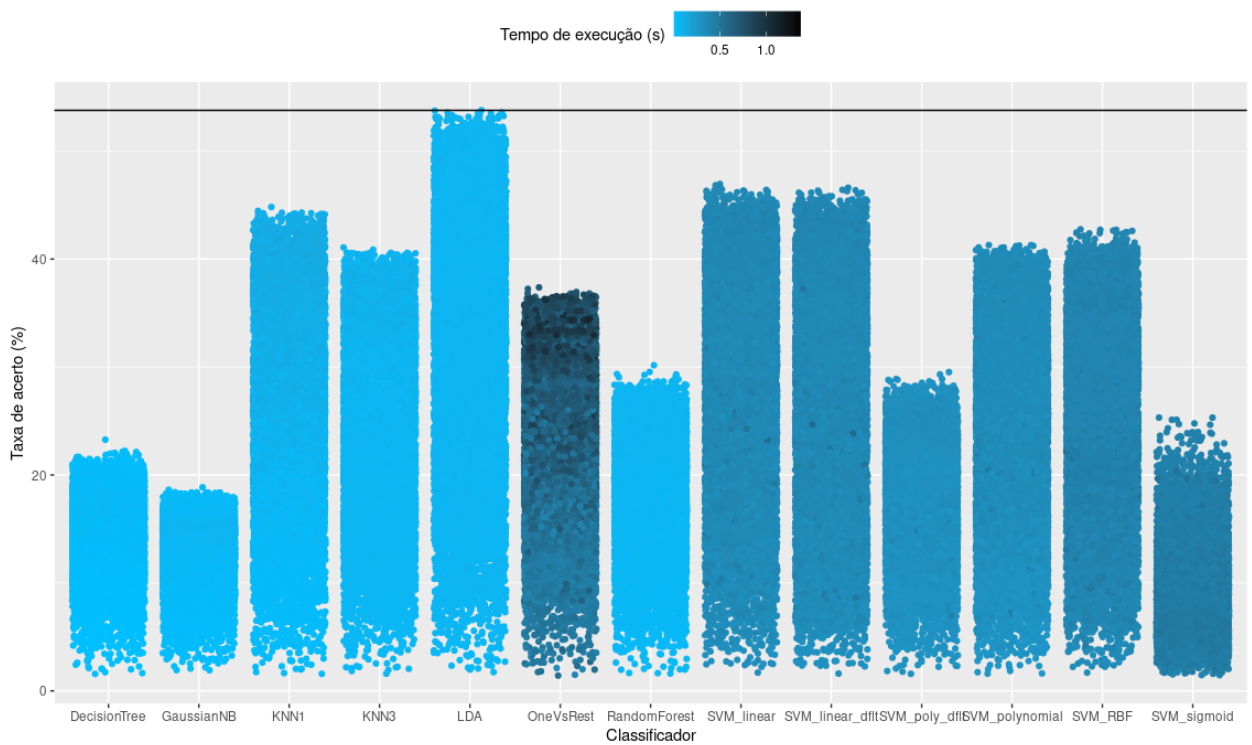
Apresentamos, neste capítulo, resultados de experimentos utilizando extratores de características baseado em covariância, usando diferentes formas e combinações de fatiamento de faces 3D no contexto de identificação de indivíduos. Os resultados de classificação na Base Bosphorus evidenciam a capacidade de descrição dos extratores e abrem caminhos de investigação para a continuação deste trabalho.

Figura 21 – Taxa de acerto *versus* Tempo de Execução, Cenário Rank-1b (B), Configuração Pizza C1



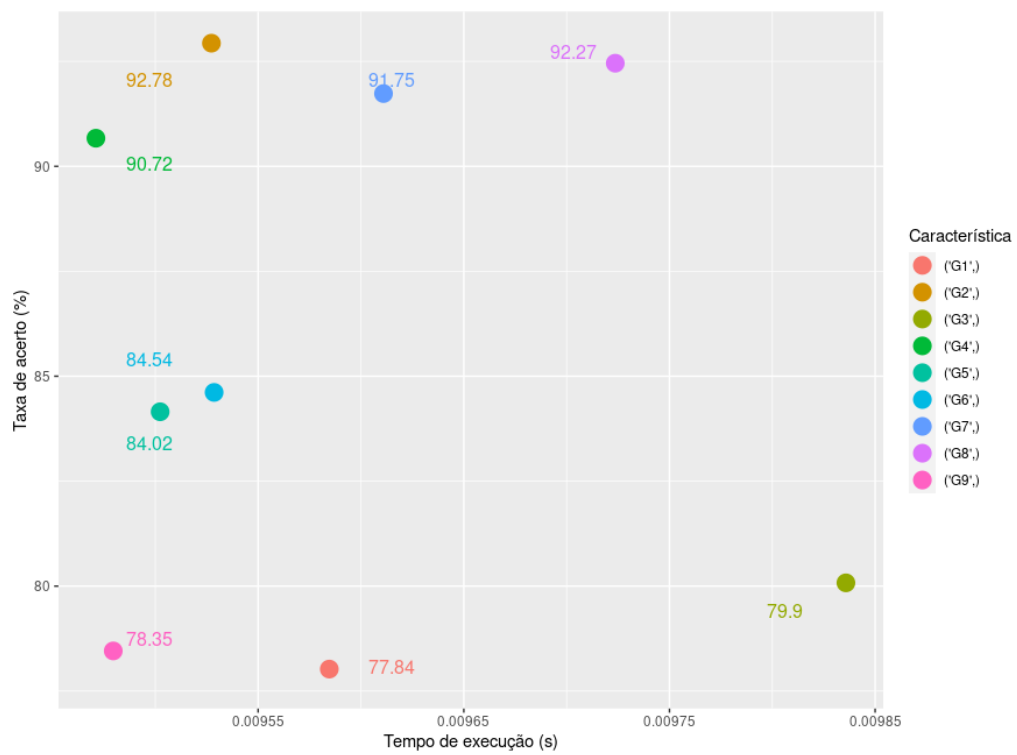
Fonte: a autora.

Figura 22 – Taxa de acerto *versus* Tempo de Execução, Cenário ROC-I (C), Configuração Pizza C1



Fonte: a autora.

Figura 23 – Taxa de acerto *versus* Tempo de Execução, Cenário Rank-1a, Configuração Triaxial, Fatias *ST*, classificador KNN-1



Fonte: a autora.

5 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho é contextualizado no reconhecimento de faces 3D por meio de métodos baseados em superfície local, utilizando suas características geométricas. Foi investigada a influência de diferentes técnicas de segmentação (fatiamento) para extração de características, tendo em perspectiva a aplicação na identificação de indivíduos via algoritmos de classificação. Durante as análises, verificou-se que a assimetria da face condiz com estudos realizados por outros pesquisadores, mostrando que as informações geométricas extraídas do lado esquerdo caracterizam melhor o indivíduo para os classificadores.

Os extratores geométricos são rápidos e permitem a geração de uma larga base de dados tratados para o problema de classificação em pouca quantidade de tempo, permitindo o processo de comparação proposto no artigo.

Foi possível verificar que o fatiamento da nuvem de pontos 3D de uma face pode aprimorar consideravelmente os resultados de classificação em comparação aos resultados obtidos sem o fatiamento.

Além disso, o método proposto de fatiamento em Pizza com Sobreposição se mostrou promissor, apesar de não superar os resultados do método de fatiamento Triaxial. Esse método também apresentou resultados melhores quando combinadas algumas fatias com a representação da face inteira. Observou-se que, ao seccionar a face considerando as AUs, é possível melhorar o resultado de classificação em relação ao não fatiamento. Isso sugere que o aprofundamento das investigações, em especial no sentido de identificar as melhores formas de segmentação considerando as diversas AUs, permitirá um refinamento da técnica, podendo-se chegar a uma classificação robusta a variações da superfície da face devidas às manifestações de emoção, uma vez que será possível considerar, com maior precisão, as regiões de maior variabilidade na face.

5.1 Trabalhos Futuros

O trabalho apresentado não esgota as possibilidades de investigação sobre o assunto. Várias hipóteses promissoras foram percebidas durante as experimentações e deverão ser exploradas em trabalhos futuros.

Os extratores de características geométricos baseados em covariância utilizados foram incorporados à Ferramenta de Análise de nuvens 3D proposta por Feitosa (2021). Esta ferramenta é exibida na Figura 24 e, na Figura 25, são apresentados filtros com os extratores, capazes

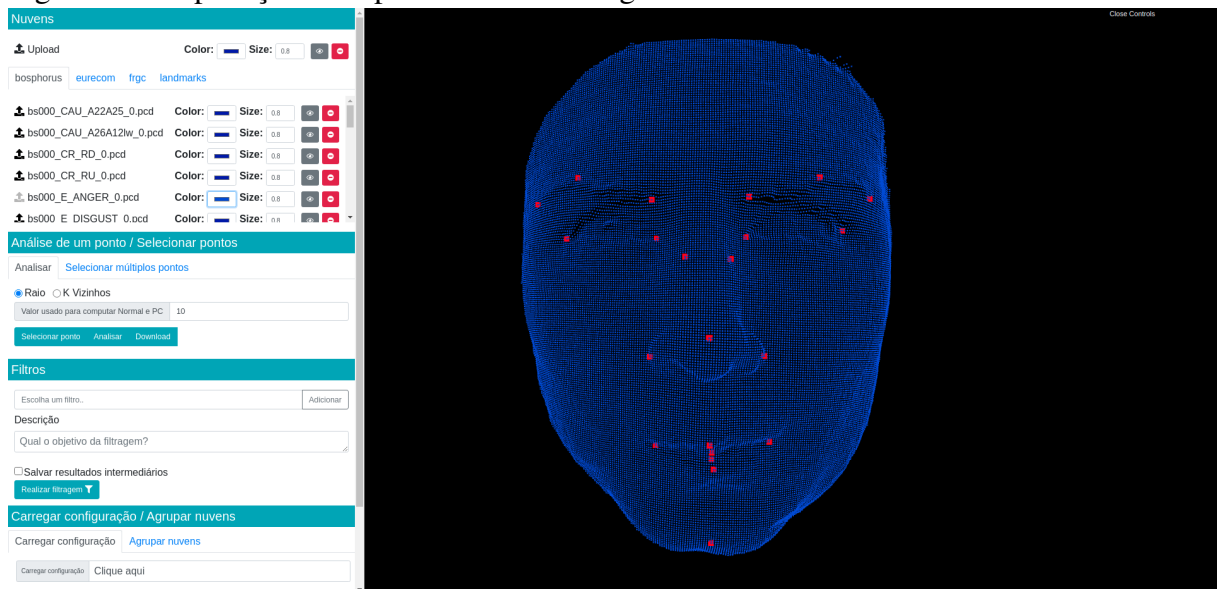
de destacar regiões de interesse para a identificação de indivíduos. Tais filtros podem ser incorporados às formas de fatiamento e combinações propostas.

A seguir são enumeradas as perspectivas de continuidade para este trabalho:

- Definir um novo modelo de segmentação baseado em regiões de intersecção entre o fatiamento triaxial e o fatiamento em pizza proposto nesta investigação. Prospecta-se, com isso, a remoção das regiões de grande variabilidade dentro dos segmentos definidos pelo fatiamento triaxial, o que, por hipótese, tende a melhorar a estabilidade das informações extraídas de cada segmento em situações de expressões emocionais;
- Aplicar *PCA* aos extratores de características geométricas;
- Reduzir a quantidade de classificadores;
- Refinar o conjunto restante de classificadores utilizados, com técnicas como *Grid Search* para otimizar os parâmetros de configuração;
- Utilizar tecnologias de computação paralela para melhorar a eficiência dos experimentos;
- Avaliar os métodos de fatiamento em outras bases de dados de faces 3D com expressão;
- Incorporar algum algoritmo de detecção da ponta do nariz, de forma a facilitar a integração com outras bases de dados, algumas necessitam de marcação manual;
- Analisar o efeito de uma nova etapa de registro nas fatias das faces com expressão, tendo em vista a deformação natural e utilizando apenas as fatias poderemos ter um custo de processamento pequeno;
- Utilizar e melhorar a Ferramenta de análise desenvolvida por (FEITOSA, 2021) para identificar problemas em potencial¹ e novas estratégias de pré-processamento;

¹ Foi possível detectar erros na marcação de *landmarks* fornecidas com a base de dados Bosphorus

Figura 24 – Aplicação Web para análise de imagens 3D



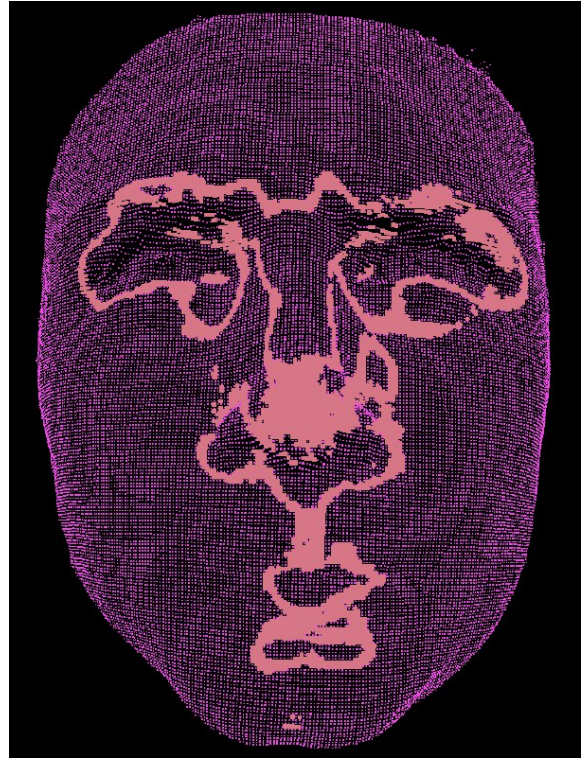
Fonte: (FEITOSA, 2021)

Figura 25 – Figuras obtidas utilizando filtros nas características descritas

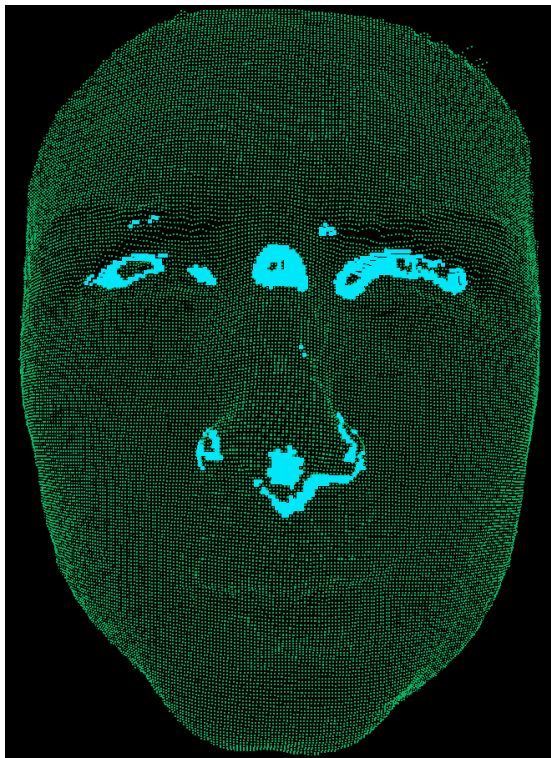
(a) G1 - Somatório



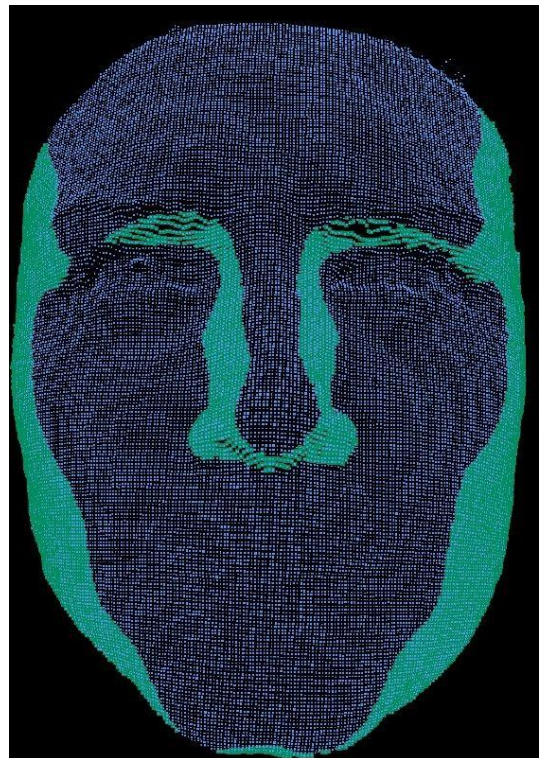
(b) G2 - Omnivariância



(c) G4 - Anisotropia



(d) G9 - Verticalidade



REFERÊNCIAS

- 3DSOURCED.COM. *The 10 Best 3D Cameras 2021 (For All Uses!) | 3DSourced*. 2021. Disponível em: <<https://www.3dsourced.com/rankings/best-3d-camera/>>.
- ALEXANDRE, G. R.; SOARES, J. M.; Pereira Thé, G. A. Systematic review of 3D facial expression recognition methods. **Pattern Recognition**, Elsevier Ltd, v. 100, p. 107108, 2020. ISSN 00313203. Disponível em: <<https://doi.org/10.1016/j.patcog.2019.107108>>.
- BENAMOUN, M.; GUO, Y.; SOHEL, F. Feature selection for 2d and 3d face recognition. **Wiley Encyclopedia of Electrical and Electronics Engineering**, p. 1–28, 2015.
- Besl, P. J.; McKay, N. D. A method for registration of 3-d shapes. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 14, n. 2, p. 239–256, Feb 1992. ISSN 0162-8828.
- BLOMLEY, R.; WEINMANN, M.; LEITLOFF, J.; JUTZI, B. Shape distribution features for point cloud analysis – a geometric histogram approach on multiple scales. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, II-3, p. 9–16, 2014.
- BOWYER, K. W.; CHANG, K.; FLYNN, P. A survey of approaches to three-dimensional face recognition. In: IEEE. null. [S.l.], 2004. p. 358–361.
- CHENEY, E. W.; KINCAID, D. **Numerical mathematics and computing**. [S.l.]: Brooks/Cole, 2020.
- CHIBA, M. F. **Câmeras com 'visão 3D' equipam smartphones**. 2019. Disponível em: <<https://www.folhadelondrina.com.br/mercado-digital/cameras-com-visao-3d-equipam-smartphones-2964351e.html>>.
- DAROM, T.; KELLER, Y. Scale-invariant features for 3-D mesh models. **IEEE Transactions on Image Processing**, IEEE, v. 21, n. 5, p. 2758–2769, 2012. ISSN 10577149.
- DENG, X.; DA, F.; SHAO, H.; JIANG, Y. A multi-scale three-dimensional face recognition approach with sparse representation-based classifier and fusion of local covariance descriptors R. **Computers and Electrical Engineering**, v. 85, p. 106700, 2020. Disponível em: <<https://doi.org/10.1016/j.compeleceng.2020.106700>>.
- DIEDERICH, S.; LENTSCHIG, M.; OVERBECK, T.; WORMANN, D.; HEINDEL, W. Detection of pulmonary nodules at spiral ct: comparison of maximum intensity projection sliding slabs and single-image reporting. **European radiology**, Springer, v. 11, n. 8, p. 1345–1350, 2001.
- EKMAN, P. Asymmetry in facial expression. **Science**, v. 209, p. 833–834, Aug 1980.
- EKMAN, P.; FRIESEN, W.; HAGER, J. Facial action coding system: The manual. **A human face**, 2002.
- EKMAN, P.; FRIESEN, W. V. Measuring facial movement. **Environmental psychology and nonverbal behavior**, v. 1, n. 1, p. 56–75, Sep 1976. ISSN 1573-3653. Disponível em: <<https://doi.org/10.1007/BF01115465>>.
- EKMAN, P.; FRIESEN, W. V. **Facial Action Coding System: A Technique for the Measurement of Facial Moviment**. Palo Alto: Consulting Psychologists Press, 1978.

Erdogmus, N.; Marcel, S. Spoofing in 2d face recognition with 3d masks and anti-spoofing with kinect. In: **2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)**. [S.l.: s.n.], 2013. p. 1–6.

FEITOSA, T. F. d. S. **Análise dos descritores de imagens 3D por meio de aplicação de filtros sucessivos: Processamento remoto e visualização por aplicação web**. 49 f. Monografia (Trabalho de Conclusão de Curso (Graduação)) — Universidade Federal do Ceará, Centro de Tecnologia, Curso de Engenharia de Computação, Fortaleza, 2021.

FU, Y.; RUAN, Q.; AN, G.; JIN, Y. Fast nonnegative tensor factorization based on graph-preserving for 3d facial expression recognition. In: **2016 IEEE 13th International Conference on Signal Processing (ICSP)**. [S.l.: s.n.], 2016. p. 292–297.

GILANI, S. Z.; MIAN, A. Learning from millions of 3d scans for large-scale 3d face recognition. In: **2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2018. p. 1896–1905.

GILANI, S. Z.; MIAN, A.; SHAFAIT, F.; REID, I. Dense 3d face correspondence. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 40, n. 7, p. 1584–1598, 2018.

GORDON, G. G. Face recognition based on depth maps and surface curvature. **SPIE Proc. Geometric Methods Comput. Vision**, v. 1570, p. 234–247, 09 1991.

GUAN, Y. Application of logistic regression algorithm in the diagnosis of expression disorder in parkinson’s disease. In: **2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)**. [S.l.: s.n.], 2021. v. 2, p. 1117–1120.

GUO, Y.; LEI, Y.; LIU, L.; WANG, Y.; BENNAMOUN, M.; SOHEL, F. EI3D: Expression-invariant 3D face recognition based on feature and shape matching. **Pattern Recognition Letters**, v. 83, p. 403–412, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.patrec.2016.04.003>>.

Hackel, T.; Wegner, J. D.; Schindler, K. Contour detection in unstructured 3d point clouds. In: **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2016. p. 1610–1618. ISSN 1063-6919.

HACKEL, T.; WEGNER, J. D.; SCHINDLER, K. Fast semantic segmentation of 3d point clouds with strongly varying density. **ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences**, III-3, p. 177–184, 2016.

HAGER, J. C.; EKMAN, P. The asymmetry of facial actions is inconsistent with models of hemispheric specialization. In: **What the Face Reveals Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)**. Oxford University Press, 2005. p. 40–58. Disponível em: <<https://doi.org/10.1093/acprof:oso/9780195179644.003.0003>>.

HARIRI, W.; TABIA, H.; FARAH, N.; BENOURETH, A.; DECLERCQ, D. 3d face recognition using covariance based descriptors. **Pattern Recognition Letters**, Elsevier BV, v. 78, p. 1–7, jul. 2016. Disponível em: <<https://doi.org/10.1016/j.patrec.2016.03.028>>.

HESHER, C.; SRIVASTAVA, A.; ERLEBACHER, G. A novel technique for face recognition using range images. In: **Seventh International Symposium on Signal Processing and**

Its Applications, 2003. Proceedings. [S.l.: s.n.], 2003. v. 2, p. 201 – 204 vol.2. ISBN 0-7803-7946-2.

HOANG, L.; LEE, S.-H.; KWON, O.-H.; KWON, K.-R. A deep learning method for 3d object classification using the wave kernel signature and a center point of the 3d-triangle mesh.

Electronics, v. 8, n. 10, p. 1196, 2019.

JAFRI, R.; ARABNIA, H. A survey of face recognition techniques. **JIPS**, v. 5, p. 41–68, 06 2009.

JÚNIOR, I. E. M.; NETO, P. S.; ANDRADE, M. V. L. M.; THÉ, G. A. P.; SOARES, J. M. Segmentação da oclusão em faces 3d com suporte de técnicas de registro locais. In: **XXXIX Congresso Nacional de Matemática Aplicada e Computacional (CNMAC)**. [S.l.: s.n.], 2019.

KAVITA, M. M. K. A survey of face recognition techniques. **International Journal of Scientific and Research Publications (IJSRP)**, Volume 6, July 2016.

KIM, H. M.; KIM, M. S.; LEE, G. J.; JANG, H. J.; SONG, Y. M. Miniaturized 3D depth sensing-based smartphone light field camera. **Sensors (Switzerland)**, v. 20, n. 7, 2020. ISSN 14248220.

LEE, J.; MILIOS, E. Matching range images of human faces. In: **[1990] Proceedings Third International Conference on Computer Vision**. [S.l.: s.n.], 1991. p. 722 – 726. ISBN 0-8186-2057-9.

LEI, Y.; BENNAMOUN, M.; EL-SALLAM, A. A. An efficient 3D face recognition approach based on the fusion of novel local low-level features. 2012. Disponível em: <<http://dx.doi.org/10.1016/j.patcog.2012.06.023>>.

LI, H.; HUANG, D.; MORVAN, J. M.; WANG, Y.; CHEN, L. Towards 3D Face Recognition in the Real: A Registration-Free Approach Using Fine-Grained Matching of 3D Keypoint Descriptors. **International Journal of Computer Vision**, Kluwer Academic Publishers, v. 113, n. 2, p. 128–142, jun 2015. ISSN 15731405.

LI, X.; DA, F. Efficient 3d face recognition handling facial expression and hair occlusion. **Image and Vision Computing**, v. 30, n. 9, p. 668–679, 2012. ISSN 0262-8856. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0262885612001163>>.

LI, X.; RUAN, Q.; AN, G. 3D facial expression recognition using delta faces. In: **IET Conference Publications**. Institution of Engineering and Technology, 2013. v. 2013, n. 641 CP, p. 234–239. ISBN 9781849197267. Disponível em: <<https://digital-library.theiet.org/content/conferences/10.1049/cp.2013.2415>>.

LI, Y.; KOBATAKE, H. Extraction of facial sketch images and expression transformation based on faces. In: **Proceedings., International Conference on Image Processing**. [S.l.: s.n.], 1995. v. 3, p. 520–523 vol.3.

LIN, C.-H.; CHEN, J.-Y.; SU, P.-L.; CHEN, C.-H. Eigen-feature analysis of weighted covariance matrices for lidar point cloud classification. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 94, p. 70 – 79, 2014. ISSN 0924-2716. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0924271614001117>>.

LUO, J.; HU, F.; WANG, R. 3d face recognition based on deep learning. In: **2019 IEEE International Conference on Mechatronics and Automation (ICMA)**. [S.l.: s.n.], 2019. p. 1576–1581.

Makioka, T.; Kuriyaki, Y.; Uchimura, K.; Satonaka, T. Quantitative study of facial expression asymmetry using objective measure based on neural networks. In: **2016 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)**. [S.l.: s.n.], 2016. p. 1–4.

MARQUES, A. **iPhone 12 Pro mede altura das pessoas usando sensor Li-DAR | Celular | Tecnoblog**. 2020. Disponível em: <<https://tecnoblog.net/378387/iphone-12-pro-mede-altura-das-pessoas-usando-sensor-lidar/>>.

MARTINS, P. J.; SOARES, J.; THÉ, G. Face slicing on facial recognition: An action units based investigation of the possible improvements by selecting the face regions in a point cloud 3d data. In: **XV Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2019)**. [s.n.], 2019. Disponível em: <http://www.bracis2019.ufba.br/Camera_Ready/199184_1.pdf>.

MORENO, A.; SÁNCHEZ, A. **GavabDB: a 3D face database**. 2004.

NAGAMINE, T.; UEMURA, T.; MASUDA, I. 3d facial image analysis for human identification. In: **[1992] Proceedings. 11th IAPR International Conference on Pattern Recognition**. [S.l.: s.n.], 1992. p. 324 – 327. ISBN 0-8186-2910-X.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

Phillips, P. J.; Flynn, P. J.; Scruggs, T.; Bowyer, K. W.; Jin Chang; Hoffman, K.; Marques, J.; Jaesik Min; Worek, W. Overview of the face recognition grand challenge. In: **2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)**. [S.l.: s.n.], 2005. v. 1, p. 947–954 vol. 1. ISSN 1063-6919.

Reddy, A. M.; Kishore, M. R.; Sreenivasulu, P.; Jyothi, V. A survey of face recognition system. **International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)**, Volume 5, April 2018.

RUSS, T.; BOEHNEN, C.; PETERS, T. 3d face recognition using 3d alignment for pca. In: **2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)**. [S.l.: s.n.], 2006. v. 2, p. 1391–1398.

RUSU, R. B.; COUSINS, S. 3d is here: Point cloud library (pcl). In: **In Robotics and Automation (ICRA), 2011 IEEE International Conference on**. [S.l.: IEEE, 2011. p. 1–4.

SACKEIM, H.; GUR, R.; SAUCY, M. Emotions are expressed more intensely on the left side of the face. **Science**, American Association for the Advancement of Science, v. 202, n. 4366, p. 434–436, 1978. ISSN 0036-8075. Disponível em: <<https://science.sciencemag.org/content/202/4366/434>>.

Salama, A.; Malekmohammadi, A.; Mohanna, S. Single-ended electrical impedance tomography. **The Journal of Engineering**, v. 2018, n. 4, p. 173–185, 2018. ISSN 2051-3305.

SANTOS, A. M.; FILHO, A. O. de C.; SILVA, A. C.; PAIVA, A. C. de; NUNES, R. A.; GATTASS, M. Automatic detection of small lung nodules in 3d ct data using gaussian mixture models, tsallis entropy and svm. **Engineering applications of artificial intelligence**, Elsevier, v. 36, p. 27–39, 2014.

SAVRAN, A.; ALYÜZ, N.; DIBEKLIOĞLU, H.; ÇELIKTUTAN, O.; GÖKBERK, B.; SANKUR, B.; AKARUN, L. Bosphorus database for 3d face analysis. In: SCHOUTEN, B.; JUUL, N. C.; DRYGAJLO, A.; TISTARELLI, M. (Ed.). **Biometrics and Identity Management**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 47–56. ISBN 978-3-540-89991-4.

SAVRAN, A.; ALYÜZ, N.; DIBEKLIOĞLU, H.; ÇELIKTUTAN, O.; GÖKBERK, B.; SANKUR, B.; AKARUN, L. Bosphorus database for 3d face analysis. In: SCHOUTEN, B.; JUUL, N. C.; DRYGAJLO, A.; TISTARELLI, M. (Ed.). **Biometrics and Identity Management**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 47–56. ISBN 978-3-540-89991-4.

SCHUMACHER, D.; SHARMA, R.; GRAGER, J.-C.; SCHRAPP, M. Scatter and beam hardening reduction in industrial computed tomography using photon counting detectors. **Measurement Science and Technology**, IOP Publishing, v. 29, n. 7, p. 075101, 2018.

SHEPLEY, A. J. **Face Recognition in Unconstrained Conditions: A Systematic Review**. 2019.

Siqueira, R.; Alexandre, G. R.; Soares, J. M.; Thé, G. A. P. Triaxial slicing for 3-d face recognition from adapted rotational invariants spatial moments and minimal keypoints dependence. **IEEE Robotics and Automation Letters**, v. 3, n. 4, p. 3513–3520, Oct 2018.

SKARŻYŃSKI, Ł.; SUCHORZEWSKI, J. Mechanical and fracture properties of concrete reinforced with recycled and industrial steel fibers using digital image correlation technique and x-ray micro computed tomography. **Construction and Building Materials**, Elsevier, v. 183, p. 283–299, 2018.

SOLTANPOUR, S.; BOUFAMA, B.; WU, Q. M. J. A survey of local feature methods for 3d face recognition. **Pattern Recognition**, v. 72, 08 2017.

Tanaka, H. T.; Ikeda, M.; Chiaki, H. Curvature-based face surface recognition using spherical correlation. principal directions for curved object recognition. In: **Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition**. [S.l.: s.n.], 1998. p. 372–377.

VILLA, F.; SEVERINI, F.; MADONINI, F.; ZAPPA, F. Spads and sipms arrays for long-range high-speed light detection and ranging (lidar). **Sensors**, v. 21, n. 11, 2021. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/21/11/3839>>.

- WEINMANN, M.; JUTZI, B.; MALLET, C. Geometric features and their relevance for 3d point cloud classification. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, Copernicus GmbH, v. 4, p. 157, 2017.
- XU, K.; WANG, X.; HU, Z.; ZHANG, Z. 3D Face Recognition Based on Twin Neural Network Combining Deep Map and Texture. **International Conference on Communication Technology Proceedings, ICCT**, IEEE, p. 1665–1668, 2019.
- YIN, L.; WEI, X.; SUN, Y.; WANG, J.; ROSATO, M. J. A 3d facial expression database for facial behavior research. In: **Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition**. Washington, DC, USA: IEEE Computer Society, 2006. (FGR '06), p. 211–216. ISBN 0-7695-2503-2. Disponível em: <<https://doi.org/10.1109/FGR.2006.6>>.
- YOUNIS, M. H.; SALIK, M.; GUL, S. T.; ALEEM, A. Automatic human facial affect classification using computational intelligence techniques. In: **2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC)**. [S.l.: s.n.], 2021. p. 1–7.
- YU, Y.; DA, F.; ZHANG, Z. **Few-Data Guided Learning Upon End-to-End Point Cloud Network for 3D Face Recognition**. [S.l.], 2021.
- ZHANG, Z.; DA, F.; YU, Y. Data-Free Point Cloud Network for 3D Face Recognition. 2019. Disponível em: <<http://arxiv.org/abs/1911.04731>>.
- ZHAO, W.; CHELLAPPA, R.; PHILLIPS, P. J.; ROSENFELD, A. Face recognition: A literature survey. **ACM Comput. Surv.**, ACM, New York, NY, USA, v. 35, n. 4, p. 399–458, dez. 2003. ISSN 0360-0300. Disponível em: <<http://doi.acm.org/10.1145/954339.954342>>.
- ZHONG, Y. Intrinsic shape signatures: A shape descriptor for 3d object recognition. In: **2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops**. [S.l.: s.n.], 2009. p. 689–696.