DISSERTAÇÃO DE MESTRADO PROFISSIONAL

ANÁLISE TENSORIAL PARA PREVENÇÃO DE FALSIFICAÇÕES EM SISTEMAS DE RECONHECIMENTO FACIAL: UMA PROPOSTA BASEADA EM CLUSTERIZAÇÃO

Caio César Rodrigues Garcez

Brasília, Março de 2025

UNIVERSIDADE DE BRASÍLIA

FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

UNIVERSIDADE DE BRASÍLIA
Faculdade de Tecnologia

DISSERTAÇÃO DE MESTRADO PROFISSIONAL

ANÁLISE TENSORIAL PARA PREVENÇÃO DE FALSIFICAÇÕES EM SISTEMAS DE RECONHECIMENTO FACIAL: UMA PROPOSTA BASEADA EM CLUSTERIZAÇÃO

Caio César Rodrigues Garcez

Dissertação de Mestrado Profissional submetida ao Departamento de Engenharia Elétrica como requisito parcial para obtenção do grau de Mestre em Engenharia Elétrica

Banca Examinadora

Prof Dr. Fábio Lúcio Lopes de Mendonça, PPEE/ENE/UnB _____
Presidente Orientador

Prof Dr. Daniel Alves da Silva, PPEE/UnB _____
Examinador Interno

Prof Dr. Jayme Milanezi Junior, UnDF _____
Examinador Externo

Prof Dr. Georges Daniel Amvame Nze, PPEE/ENE/UnB _____
Suplente

FICHA CATALOGRÁFICA

GARCEZ, CAIO CÉSAR RODRIGUES
ANÁLISE TENSORIAL PARA PREVENÇÃO DE FALSIFICAÇÕESEM SISTEMAS DE RECONHECIMENTO FACIAL:UMA PROPOSTA BASEADA EM CLUSTERIZAÇÃO [Distrito Federal] 2025.
xvi, 53 p., 210 x 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2025).
Dissertação de Mestrado Profissional - Universidade de Brasília, Faculdade de Tecnologia.
Departamento de Engenharia Elétrica

1. Inteligência Artificial	2. Visão Computacional
3. Métodos de Classificação	4. Reonhecimento Facial
I. ENE/FT/UnB	II. Título (série)

PUBLICAÇÃO: PPEE.MP.097

REFERÊNCIA BIBLIOGRÁFICA

GARCEZ, C.C.R (2025). *ANÁLISE TENSORIAL PARA PREVENÇÃO DE FALSIFICAÇÕESEM SISTEMAS DE RECONHECIMENTO FACIAL:UMA PROPOSTA BASEADA EM CLUSTERIZAÇÃO*. Dissertação de Mestrado Profissional, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 53 p.

CESSÃO DE DIREITOS

AUTOR:
TÍTULO: ANÁLISE TENSORIAL PARA PREVENÇÃO DE FALSIFICAÇÕESEM SISTEMAS DE RECONHECIMENTO FACIAL:UMA PROPOSTA BASEADA EM CLUSTERIZAÇÃO.
GRAU: Mestre em Engenharia Elétrica ANO: 2025
PUBLICAÇÃO: PPEE.MP.097

Depto. de Engenharia Elétrica (ENE) - FT
Universidade de Brasília (UnB)
Campus Darcy Ribeiro
CEP 70919-970 - Brasília - DF - Brasil

AGRADECIMENTOS

Agradeço, em especial, ao meu orientador, Professor Dr. Fábio Lúcio Lopes de Mendonça, pela orientação profissional e atenciosa nos momentos mais desafiadores deste trabalho, sempre demonstrando paciência diante das dúvidas e dificuldades surgidas ao longo da elaboração desta dissertação.

Estendo minha gratidão aos professores do Programa de Pós-Graduação em Engenharia Elétrica da Universidade de Brasília (PPEE/UnB), Rafael Timóteo de Sousa Júnior, Georges D. Amvame Nze, Edna Dias Canedo, Robson de Oliveira Albuquerque, William Ferreira Giozza, Geraldo Pereira Rocha Filho e Daniel Alves da Silva, pelas contribuições, apoio e incentivo indispensáveis ao desenvolvimento deste trabalho. Agradeço ainda aos membros da banca pelas valiosas observações e sugestões.

Agradeço o apoio técnico e computacional do Laboratório de Tecnologias para Tomada de Decisão - LATITUDE, da Universidade de Brasília, que conta com apoio do CNPq - Conselho Nacional de Pesquisa (Outorgas 312180/2019-5 PQ-2 e 465741/2014-2 INCT em Cibersegurança), da Advocacia Geral da União (Outorga AGU 697.935/2019), da Procuradoria Geral da Fazenda Nacional (Outorga PGFN 23106.148934/2019-67), da Polícia Federal (Outorga PF 03/2020), do Mestrado Profissional em Engenharia Elétrica, na área de concentração: Segurança Cibernética – 1ª Turma para Profissionais do Setor de Inteligência (Outorga ABIN 01/2019) ao Decanatos de Pesquisa e Inovação e de Pós-Graduação da Universidade de Brasília (Outorga 7129 FUB/EMENDA/DPI/COPEI/AMORIS) e do Projeto SISTER City –Sistemas Inteligentes Seguros e em Tempo Efetivo Real para Cidades Inteligentes (Outorga 625/2022) e a Fundação de Apoio a Pesquisa do Distrito Federal - FAP/DF.

RESUMO

Embora os Sistemas de Reconhecimento Facial (SRF) apresentem um desempenho cada vez mais satisfatório para aplicações como controle de acesso, a sua confiabilidade é continuamente ameaçada pelo aumento de ataques com biometria falsa (spoofing), tornando a detecção de vivacidade (liveness detection) um aspecto crucial para a segurança. Para enfrentar este desafio, este trabalho propõe um novo mecanismo de prevenção baseado na integração entre modelagem tensorial e aprendizado não supervisionado. A metodologia parte do pré-processamento rigoroso de imagens do dataset CelebA-Spoof, onde as faces são detectadas, recortadas e normalizadas, aplicando-se a equalização de histograma adaptativa (CLAHE) para mitigar variações de iluminação. As imagens processadas são então estruturadas em um tensor tridimensional que modela de forma abrangente as correlações complexas entre os pixels, a condição da amostra (autêntica ou falsa) e o conjunto de dados. Para extrair características intrínsecas e discriminativas, empregamos a Decomposição CANDECOMP/PARAFAC (CPD) por meio do algoritmo Alternating Least Squares (ALS). As componentes de posto unitário resultantes da fatoração são submetidas à redução de dimensionalidade não linear com a técnica t-distributed Stochastic Neighbor Embedding (t-SNE), que preserva a estrutura local e global dos dados em um espaço de baixa dimensão. Por fim, o algoritmo k-means é utilizado para clusterizar as representações em dois grupos distintos, permitindo a classificação de novas amostras e a identificação precisa de biometrias genuínas e falsas. Os resultados indicam que essa abordagem integrada demonstra ser um mecanismo poderoso para aumentar a segurança contra tentativas de falsificação.

Palavras-chave: Sistemas de Reconhecimento Facial, detecção de autenticidade, técnicas de clusterização.

ABSTRACT

Although Facial Recognition Systems (FRS) exhibit increasingly satisfactory performance for applications like access control, their reliability is continually threatened by the rise of presentation attacks (spoofing), making liveness detection a crucial aspect of security. To address this challenge, this paper proposes a novel prevention mechanism based on the integration of tensor modeling and unsupervised learning. The methodology begins with the rigorous preprocessing of images from the CelebA-Spoof dataset, where faces are detected, cropped, and normalized, followed by the application of Contrast Limited Adaptive Histogram Equalization (CLAHE) to mitigate illumination variations. The processed images are then structured into a three-dimensional tensor that comprehensively models the complex correlations among pixels, the presentation condition (real or spoof), and the samples themselves. To extract intrinsic and discriminative features, we employ the CANDECOMP/PARAFAC (CPD) decomposition, implemented through the Alternating Least Squares (ALS) algorithm. The resulting rank-one components from this

factorization are then subjected to nonlinear dimensionality reduction using the t-distributed Stochastic Neighbor Embedding (t-SNE) technique, which preserves the local and global data structure in a low-dimensional space. Finally, the k-means algorithm is used to cluster these representations into two distinct groups, allowing for the precise identification of genuine and spoofed biometrics. The results indicate that this integrated approach provides to be a powerful mechanism for enhancing security against spoofing attempts.

Keywords: Facial Recognition Systems, liveness detection, clustering techniques.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

1 INTRODUÇÃO

A visão computacional constitui um subcampo da inteligência artificial voltado ao desenvolvimento de métodos para aquisição, processamento e análise automatizada de imagens e vídeos. Esse domínio envolve as seguintes etapas sistemáticas: captação de dados visuais por sensores ópticos (câmeras RGB, térmicas e 3D), pré-processamento das imagens, extração de características relevantes e interpretação semântica, com o objetivo de converter insumos brutos em representações adequadas para a tomada de decisão automática [1]. A qualidade das imagens obtidas exerce influência determinante sobre o desempenho de todas as fases subsequentes, afetando a precisão na detecção, a eficiência na classificação e a fidelidade da compreensão de cenas complexas.

Neste contexto, o reconhecimento facial consiste em um método de identificação biométrica fundamentado na análise das características faciais de um indivíduo. Ele integra técnicas de visão computacional, aprendizado de máquina e processamento de imagem para capturar, extrair e comparar representações faciais. Esse procedimento se estrutura em etapas sequenciais, cada uma essencial para garantir a acurácia e a eficiência do sistema [2]. Primeiro, realiza-se a aquisição das imagens por câmeras RGB, infravermelhas ou 3D, em tempo real ou a partir de arquivos armazenados, sendo que variáveis como iluminação, ângulo de captura e resolução afetam diretamente a qualidade dos dados. Em seguida, aplica-se a detecção automática da face por meio de algoritmos como Haar Cascade [3], Histogram of Oriented Gradients [4] com SVM [5] ou redes neurais convolucionais [6], visando localizar e delimitar a região facial mesmo na presença de múltiplos sujeitos.

A adoção de sistemas de reconhecimento facial (SRF) tem se destacado em relação a outros métodos biométricos, como reconhecimento de voz e identificação por impressões digitais, principalmente devido à ausência de contato físico e ao caráter não intrusivo. Essa tecnologia evoluiu significativamente nos últimos anos e é amplamente empregada como etapa complementar em fluxos de autenticação voltados à segurança e ao controle de acesso [7]. Inicialmente, as abordagens baseavam-se em PCA [8] (Análise de Componentes Principais) e em LBP [9] (Padrões Binários Locais). Atualmente, arquiteturas de redes neurais profundas, como FaceNet [10], VGG-Face [11] e ArcFace [12], mapeiam imagens faciais em vetores de características (embeddings) num espaço multidimensional. A correspondência entre faces é então avaliada por meio de um limiar de similaridade configurável, permitindo ajustar o nível de segurança e gerar autenticações positivas ou negativas.

A biometria facial apoiada nos sistemas de reconhecimento facial consolidou-se como tema central em pesquisa e desenvolvimento. Neste contexto, as tentativas de falsificação, conhecidas como *spoofing*, em que um agente se faz passar por outra identidade para obter acesso não autorizado a um sistema, representam um desafio crítico à segurança e à confiabilidade dos processos de autenticação. Neste contexto, este trabalho propõe uma solução robusta para a prevenção de tentativas de falsificação em sistemas de reconhecimento facial, integrando técnicas avançadas de modelos tensoriais e algoritmos de clusterização. A metodologia inicia com o pré-processamento dos dados: utilizando como base o conjunto de dados *CelebASpoof* [13]. Essa preocesso envolve a detecção e o recorte preciso das faces, seguido da normalização das imagens e a equalização da iluminação por meio da técnica CLAHE [14], assegurando a consistência

dos dados.

Concluída a etada de pré-processamento, as imagens são postariomente organizadas em um tensor tridimensional que captura simultaneamente as relações intrínsecas entre os *pixels* da imagem, as condições de apresentação da amostra (face autêntica ou falsificação) e número total de amostras. Essa estrutura multidimensional é empregada para explorar de maneira abrangente as correlações complexas existentes nos dados faciais.

Para extração de características intrínsecas, o tensor previamente construído é fatorado por meio da Decomposição *CANDECOMP/PARAFAC* (CPD) [15], resultando em componentes de posto unitário que revelam vetores associados a *pixels*, condições e amostras. Essa fatoração é realizada iterativamente pelo algoritmo *Alternating Least Squares* (ALS) [16], garantindo a estabilidade numérica no processo. As componentes obtidas são então submetidas a uma redução de dimensionalidade não linear utilizando a técnica *t-distributed Stochastic Neighbor Embedding* (t-SNE) [17]. Este método projeta dados de alta dimensão em um espaço bidimensional, preservando a estrutura e permitindo a identificação de padrões.

Ao final, as representações dimensionais reduzidas passam por uma etapa de clusterização utilizando o algoritmo *k-means* [18]. Este método agrupa os dados em um número pré-definido de clusters ($k = 2$), buscando minimizar a variância dentro de cada grupo. A classificação de novas amostras ocorre por meio de sua projeção no espaço reduzido, seguida pela atribuição ao cluster cujo centroide esteja mais próximo, o que possibilita distinguir entre faces autênticas e tentativas de falsificação.

1.0.1 Objetivo Geral

O objetivo geral deste trabalho é propor e validar um mecanismo de prevenção de tentativas de falsificação em sistemas de reconhecimento facial, baseado na integração de modelagem tensorial e algoritmos de aprendizado não supervisionado, capaz de distinguir de forma robusta e eficiente entre amostras autênticas e falsificadas.

1.0.2 Objetivos Específicos

Para atingir o objetivo geral, este trabalho define os seguintes objetivos específicos:

- Realizar a curadoria e o pré-processamento do dataset *CelebA-Spoof* [13], incluindo detecção precisa de faces, recorte, normalização de resolução e equalização de iluminação via CLAHE [14].
- Construir a partir das imagens pré-processadas o tensor tridimensional:

$$\mathcal{X} \in \mathbb{R}^{P \times C \times N}$$

  em que:

  - P é o número de pixels de cada face, após vetorização (dimensão espacial);
  - $C = 2$ corresponde às duas condições de apresentação — autêntica (*live*) e falsificada (*spoof*) (dimensão semântica);

  - N indica o número total de indivíduos (identidades únicas) no conjunto de dados (dimensão identitária).

- Aplicar a Decomposição CANDECOMP/PARAFAC (CPD) [15] ao tensor $\mathcal{X} \in \mathbb{R}^{P \times C \times N}$, extraindo componentes de posto unitário que representem de modo interpretável as variabilidades relacionadas a pixels, condições de apresentação e identidades.

- Empregar o método t-SNE [17] às matrizes fatoradas pela CPD, projetando as representações latentes em um espaço bidimensional que preserve tanto as estruturas de vizinhança local quanto as relações de distância global entre amostras.

- Executar o algoritmo k-means [19] com $(k = 2)$ sobre as projeções t-SNE para separar amostras autênticas e falsificadas e avaliar a qualidade dos agrupamentos por meio da análise estatística da dispersão das distâncias intra-cluster ao centróide dos agrupamentos resultantes.

1.1 METODOLOGIA

Este trabalho propõe um sistema de prevenção contra tentativas de falsificação em sistemas de reconhecimento facial, utilizando decomposição tensorial e clusterização hierárquica para distinguir rostos autênticos de artefatos forjados. Para isso, é empregada a base de dados *CelebASpoof* [13].

O pré-processamento das imagens inclui detecção facial, recorte, redimensionamento, conversão para escala de cinza, normalização estatística (subtração da média e divisão pelo desvio padrão) e equalização de histograma local por meio do algoritmo CLAHE [14]. As imagens resultantes são organizadas em um tensor tridimensional que representa as relações entre pixels, condição de apresentação (autêntico ou tentativa de falsificação) e diferentes amostras.

Para a extração de padrões latentes, é aplicada a decomposição CANDECOMP/PARAFAC [15]. Em seguida, os fatores associados às amostras são projetados em um espaço bidimensional por meio da técnica t-SNE [17]. A clusterização dos pontos projetados é realizada com o algoritmo k-means, e a avaliação do desempenho do modelo é conduzida com base em métricas de acurácia e análise da matriz de confusão.

Essa abordagem visa ao desenvolvimento de um sistema robusto para detecção de tentativas de falsificação facial, com aplicação voltada a ambientes que demandam altos níveis de segurança.

1.2 CONTRIBUIÇÕES TÉCNICAS E CIENTÍFICAS

Este trabalho apresenta contribuições significativas tanto para o avanço científico quanto para a aplicação prática na detecção de tentativas de falsificação em sistemas de reconhecimento facial:

- Arquitetura modular e reprodutível que integra etapas de pré-processamento, modelagem tensorial, redução de dimensionalidade e clusterização, permitindo fácil adaptação a diferentes bases de dados e fluxos de processamento.

- Formulação de um tensor tridimensional $\mathcal{X} \in \mathbb{R}^{P \times C \times N}$ que explicita as dimensões espacial (pixels vetorizados), semântica (tipo de amostra) e identitária (identidade do sujeito), aumentando a interpretabilidade dos resultados.
- Otimização da decomposição CANDECOMP/PARAFAC por meio da fatoração Khatri-Rao [20] elevando a capacidade de discriminação entre amostras autênticas e falsas.

1.3 TRABALHOS PUBLICADOS

Durante o período de desenvolvimento desta dissertação, foi publicado o seguinte artigo, aceito na CIAWI 2023 – 21ª Conferência Ibero-Americana sobre a World Wide Web e Internet:

- [21] GARCEZ, C. C. R.; MARQUES, G. S.; CANEDO, E. D.; PRACIANO, B. G.; CALDAS FILHO, F. L.; MENDONÇA, F. L. L. *Prevenção de falsificação em sistemas de reconhecimento facial: uma proposta baseada em clusterização*. In: MIRANDA, P.; SANTORO, F. M.; COSTA, C. (Eds.). *Anais da 21ª Conferência Ibero-Americana WWW/Internet (CIAWI)*, 2023, p. 35–42. ISBN 978-989-8704-54-2.

1.4 ESTRUTURA DA DISSERTAÇÃO

Os demais capítulos deste trabalho estão organizados da seguinte forma:

- **Capítulo 2 – Fundamentação Teórica.** Este capítulo apresenta os conceitos teóricos que fundamentam o desenvolvimento da solução proposta, oferecendo a base necessária para a compreensão do trabalho.
- **Capítulo 3 – Trabalhos Relacionados.** São discutidos métodos propostos na literatura relacionados ao contexto de tentativas de falsificação em sistemas de reconhecimento facial, bem como sua comparação com a abordagem adotada nesta dissertação, em termos de objetivos, escopo e estratégias adotadas.
- **Capítulo 4 – Solução Proposta.** Neste capítulo é apresentada a proposta desenvolvida para mitigar tentativas de falsificação em sistemas de reconhecimento facial, estruturada a partir de técnicas de clusterização. A arquitetura da proposta, juntamente com seus principais componentes e os métodos utilizados.
- **Capítulo 5 – Avaliação de Desempenho.** São detalhados os experimentos realizados para avaliar o desempenho da solução proposta, incluindo a descrição dos conjuntos de dados utilizados, as métricas de avaliação adotadas e a análise dos resultados obtidos.
- **Capítulo 6 – Conclusão e Trabalhos Futuros.** Apresenta-se uma análise dos principais resultados do trabalho, juntamente com propostas de aprimoramento e perspectivas para investigações futuras.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta a fundamentação teórica que norteia o trabalho, iniciando com os princípios gerais da biometria e a sua estrutura funcional. A discussão será então direcionada às particularidades dos Sistemas de Reconhecimento Facial (SRF), examinando os desafios correntes que ameaçam sua segurança, com destaque para as tentativas de falsificação (conhecidas como ataques de apresentação ou, do inglês, *spoofing*). Como resposta a essas ameaças, será analisado o conjunto de técnicas de mitigação propostas na literatura. Por fim, a abordagem culminará na introdução de métodos tensoriais e a motivação para sua aplicação em algoritmos de agrupamento, do inglês *clustering*, delineando o conjunto de ferramentas conceituais e técnicas que sustentam os métodos a serem desenvolvidos nos capítulos seguintes.

2.1 SISTEMAS BIOMÉTRICOS E AUTENTICAÇÃO FACIAL

2.1.1 Conceitos Fundamentais da Biometria

A palavra biometria tem origem na junção dos termos gregos $\beta\iota o\varsigma$ (vida) e $\mu\varepsilon\tau\rho o\nu$ (medida), remetendo, em seu sentido mais amplo, à ciência dedicada à medição de fenômenos biológicos [22]. No entanto, no contexto da literatura científica e tecnológica atual, o termo assumiu um significado mais específico: a automação do reconhecimento de indivíduos com base em características fisiológicas e comportamentais singulares [23].

2.1.1.1 Características Biométricas

Uma característica biométrica é definida como um traço mensurável, de origem fisiológica ou comportamental, que apresenta distinção suficiente e estabilidade ao longo do tempo para ser utilizado na identificação ou verificação de um indivíduo [22]. Essas características são geralmente agrupadas em duas categorias principais:

- **Características fisiológicas**

  Referem-se a traços físicos e estruturais do corpo humano, geralmente determinados por fatores genéticos. Essas características tendem a apresentar alta estabilidade ao longo do tempo, o que as torna especialmente adequadas para aplicações biométricas.

  Exemplos típicos incluem:

  – Impressão digital.

  – Geometria da mão.

  – Padrão vascular da retina.

  – Estrutura da íris.

  - **Geometria facial**.

- **Características comportamentais**

  Correspondem a padrões de comportamento adquiridos e desenvolvidos ao longo do tempo, refletindo a atividade neuromuscular do indivíduo. Diferentemente dos traços fisiológicos, essas características são mais suscetíveis à variabilidade, tanto entre indivíduos quanto em diferentes contextos de captura. No entanto, apresentam vantagens importantes, especialmente em aplicações que exigem coleta de dados de forma não intrusiva.

  Exemplos comuns incluem:

  - Assinatura manuscrita,
  - Dinâmica de digitação,
  - Padrão de marcha,
  - Modulação da voz.

A adoção de características biométricas em sistemas de autenticação pode ocorrer por meio de diferentes estratégias de integração:

- **Sistemas unimodais**

  Estes sistemas tilizam uma única fonte biométrica para autenticação. Apresentam arquitetura mais simples e menor custo de implementação, mas são mais suscetíveis a limitações como a ausência do traço biométrico em parte da população, além de maior sensibilidade a ruídos, falhas de leitura e tentativas de falsificação.

- **Sistemas multimodais**

  Combinam duas ou mais fontes biométricas, buscando compensar as limitações dos sistemas unimodais. Essa abordagem aumenta a robustez do sistema, melhora a acurácia da identificação e reforça a resistência contra fraudes e ataques. Além disso, proporciona maior tolerância a falhas em sensores individuais e a variações nas condições de captura [24].

2.1.1.2 Estrutura e Template Biométrico

Nos sistemas biométricos em geral, incluindo os Sistemas de Reconhecimento Facial (SRF), não se armazena diretamente a biometria bruta ou a imagem do rosto capturado. Em vez disso, o sistema extrai uma **estrutura biométrica**, ou seja, um conjunto de informações discriminativas que representam, de forma abstrata e compacta, as características faciais do indivíduo. Essa estrutura pode assumir diferentes formas, tais como vetores compostos por medições entre pontos de referência faciais (por exemplo, olhos, nariz e boca), coeficientes extraídos a partir de transformações de textura como a Análise de Componentes Principais (PCA, do inglês, Principal Component Analysis) ou Padrões Binários Locais (LBP, do inglês, de Local Binary Patterns), ou ainda ativações internas de camadas profundas em redes neurais convolucionais (CNNs, do inglês Convolutional Neural Networks) treinadas especificamente para tarefas de reconhecimento facial.

A partir dessa estrutura, é gerado um **template biométrico**, que consiste em uma representação matemática compacta, normalizada e otimizada tanto para o armazenamento quanto para a etapa de comparação. A geração do template é, essencialmente, um processo de redução de dimensionalidade, cujo objetivo é preservar a variabilidade entre indivíduos e, simultaneamente, reduzir a variabilidade entre diferentes amostras de um mesmo indivíduo [24]. A qualidade, a representatividade e a segurança do template biométrico são fatores críticos para o desempenho do sistema de reconhecimento facial. Um template bem projetado permite maior precisão na identificação, além de contribuir para a robustez, escalabilidade e resiliência do sistema frente a ruídos, variações na captura e tentativas de fraude.

2.1.2 Arquitetura Funcional de um Sistema Biométrico

A operação de um sistema biométrico pode ser descrita como uma sequência de módulos funcionais interconectados, que vão desde a aquisição inicial do sinal biológico até a tomada de decisão final [23]. Esse fluxo define o percurso tanto para o reconhecimento legítimo quanto para potenciais tentativas de ataque, sendo, portanto, o principal alvo de análise em termos de desempenho e segurança.

- **Amostragem e Pré-processamento** ($S \rightarrow P$): O processo tem início no módulo de *Amostragem* (ou sensor), responsável por capturar uma representação digital (S) do fenômeno biométrico — no caso do reconhecimento facial, uma imagem do rosto. A qualidade dessa amostra depende diretamente das condições de aquisição e das propriedades físicas do sensor. Em seguida, o módulo de *Pré-processamento* (P) aplica normalizações fotométricas e geométricas, com o objetivo de reduzir variações indesejadas de iluminação, pose e escala.
- **Extração de Características** ($P \rightarrow T$): Este módulo constitui o núcleo do sistema biométrico. Nele, algoritmos especializados transformam a amostra pré-processada em um *template biométrico* (T), uma representação matemática discriminativa e compacta das características do indivíduo.
- **Controle de Qualidade** ($T \rightarrow Q$): Uma etapa de avaliação é então aplicada para verificar a adequação do template gerado. Amostras comprometidas — como imagens desfocadas, com oclusões ou capturadas sob condições inadequadas — podem ser rejeitadas neste ponto, evitando que prejudiquem o desempenho global do sistema.
- **Reconhecimento e Decisão** ($Q \rightarrow C$): Por fim, o módulo de *Reconhecimento* (ou Comparação) calcula uma pontuação de similaridade entre o template de consulta e os templates de referência armazenados na base de dados (E). Com base nessa pontuação, o sistema realiza uma *decisão de classificação* (C), tal como aceitar ou rejeitar a identidade reivindicada.

A robustez dessa arquitetura funcional é essencial para a confiabilidade do sistema biométrico. Vulnerabilidades em qualquer módulo podem afetar tanto a precisão quanto a segurança. A Figura (2.1) ilustra esse fluxo operacional.

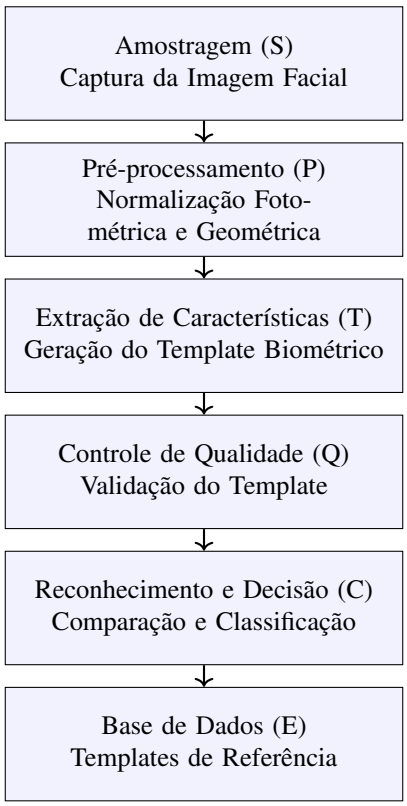

Figura 2.1: Arquitetura funcional de um sistema biométrico.

2.2 VISÃO COMPUTACIONAL

A Visão computacional é a área que estuda como obter, processar e interpretar de forma automatizada imagens e sequências de vídeo, convertendo sinais ópticos em representações numéricas que possam ser manipuladas por algoritmos em tarefas como detecção, segmentação, rastreamento e classificação de objetos [25]. Esse campo engloba quatro etapas fundamentais:

1. **Aquisição de dados visuais**: captura de imagens ou vídeo por meio de sensores, definindo resolução, taxa de quadros e espectro utilizado (RGB, infravermelho, profundidade).
2. **Pré-processamento**: aplicação de filtros e correções geométricas e fotométricas para reduzir ruído, corrigir distorções e uniformizar condições de iluminação e escala.
3. **Extração de características**: transformação dos pixels brutos em descritores (bordas, texturas, formas ou embeddings de redes neurais) que resumem aspectos relevantes da cena.
4. **Interpretação e decisão**: uso de modelos de aprendizado ou regras heurísticas para reconhecer padrões, identificar objetos ou inferir informações de alto nível.

O Reconhecimento Facial é uma aplicação especializada dentro da visão computacional, cuja finalidade é identificar e verificar indivíduos a partir de suas feições. A partir dos blocos básicos descritos acima, o pipeline de reconhecimento facial estrutura-se em:

- **Detecção facial**: identifica automaticamente as regiões da imagem que contêm rostos, sinalizando cada instância presente.
- **Alinhamento e recorte**: ajusta a orientação e extrai apenas a área da face, eliminando o fundo e outras regiões irrelevantes.
- **Normalização visual**: padroniza atributos como escala, orientação, contraste e iluminação para uniformizar todas as amostras.
- **Representação biométrica**: transforma a face pré-processada em um vetor de características numéricas, conforme o lfuxo detalhado na Figura (2.1).

Neste trabalho, a extração de descritores biométricos é aplicada ao conjunto de dados *CelebA-Spoof* [13]. Antes dessa etapa, todas as amostras são submetidas a um pré-processamento padronizado, garantindo homogeneidade e comparabilidade dos dados em cada fase do método proposto.

2.3 FUNDAMENTOS DO RECONHECIMENTO FACIAL

Nesta seção, será apresentada uma visão evolutiva dos fundamentos teóricos que sustentam os sistemas de reconhecimento facial, proporcionando uma compreensão clara e progressiva das abordagens desenvolvidas ao longo do tempo e destacando os principais paradigmas computacionais que moldaram a trajetória dessa tecnologia.

2.3.1 Histórico dos Paradigmas de Reconhecimento Facial

O desenvolvimento dos métodos de reconhecimento facial pode ser organizado em três fases distintas, cada uma caracterizada por transformações nas abordagens matemáticas utilizadas.

1. Métodos Estatísticos e Holísticos (1990–2010). Nesta fase, destacam-se os primeiros métodos computacionais que buscavam representar matematicamente a identidade facial por meio de projeções em subespaços de menor dimensionalidade. Destacam-se, nesse contexto, as técnicas baseadas em Análise de Componentes Principais (*Principal Component Analysis* — PCA), que permitiam capturar as variações mais relevantes entre diferentes rostos, reduzindo a dimensionalidade dos dados faciais e possibilitando uma representação mais discriminativa para tarefas de reconhecimento. Entre os principais métodos representativos dessa abordagem, destacam-se:

- **Eigenfaces** [26]. Proposta baseada em Análise de Componentes Principais (PCA), que projeta imagens faciais em um subespaço de menor dimensionalidade. Nesse espaço, os rostos são representados como combinações lineares de "autorretratos" (eigenfaces), extraídos de forma não supervisionada a partir da base de treinamento. Embora eficiente e computacionalmente simples, o método é sensível a variações de iluminação, pose e expressão facial.

- **Fisherfaces** [27]. Abordagem que combina PCA e Análise Discriminante Linear (LDA, do inglês *Linear Discriminant Analysis*), visando maximizar a separabilidade entre classes. Ao considerar informações de classe durante a projeção, o método torna-se mais robusto a mudanças de iluminação e expressões faciais.

- **LBP (Local Binary Patterns)** e **HOG (Histograms of Oriented Gradients)** [28]. Abordagens baseadas, respectivamente, em padrões de textura binária e gradientes de orientação, utilizadas para extrair características da imagem facial. Essas técnicas capturam informações de alta frequência e são frequentemente combinadas para aumentar a robustez do reconhecimento diante de variações sutis de pose, iluminação e oclusões parciais.

2. Transição para o Paradigma das Redes Neurais Convolucionais (CNNs, do inglês *Convolutional Neural Networks*) (2010–2014). O conjunto de dados LFW (*Labeled Faces in the Wild*, em tradução livre, "Rostos Rotulados em Ambientes Naturais") [29], lançado em 2007, desempenhou um papel importante na transição do paradigma holístico para métodos baseados em aprendizado profundo, ao expor as limitações dos métodos tradicionais em cenários não controlados, como variações de pose, iluminação, expressão facial, entre outros. Esse contexto impulsionou o surgimento de abordagens mais avançadas, capazes de lidar de forma mais eficaz com esses desafios. Entre as principais técnicas e avanços que caracterizam esse período de transição, destacam-se:

- **AlexNet** [30]. Arquitetura de rede neural convolucional profunda composta por oito camadas, das quais cinco são convolucionais. AlexNet venceu a competição ILSVRC-2012 (*ImageNet Large Scale Visual Recognition Challenge*) [31] com resultados expressivos, demonstrando o potencial das CNNs combinadas com grandes volumes de dados. Seu impacto foi significativo em várias aplicações de visão computacional, incluindo o reconhecimento facial.

3. Arquiteturas Avançadas de Redes Neurais Convolucionais e Aprendizado de Métrica (2014–presente). Com a consolidação das CNNs como abordagem dominante, este período foi marcado pelo desenvolvimento de arquiteturas especializadas e técnicas de aprendizado de métrica, voltadas à obtenção de representações faciais altamente discriminativas. Entre os principais avanços, destacam-se:

- **DeepFace** [32]. Rede neural composta por nove camadas, treinada em mais de 4 milhões de imagens rotuladas. Foi um dos primeiros sistemas a atingir desempenho comparável ao humano em tarefas de verificação facial, marcando um avanço significativo na precisão sob condições não controladas.

- **FaceNet** [10]. Introduziu o uso da função de perda *triplet loss* para aprendizado de representações embutidas em um espaço euclidiano, onde a distância entre vetores reflete diretamente a similaridade facial. Essa abordagem permitiu comparações mais precisas entre identidades, mesmo em bases de dados extensas.

- **ArcFace** [12], **CosFace** [33] e **SphereFace** [34]. Arquiteturas que incorporam funções de perda com margens angulares, com o objetivo de aumentar a separação entre classes distintas e reduzir

a variação intra-classe. Tais estratégias impulsionaram a generalização em ambientes reais e se tornaram padrão em competições e benchmarks.

Período	Abordagem	Métodos Chave	Destaques
1990–2010	Métodos estatísticos e holísticos	Eigenfaces (PCA) [26], Fisherfaces (LDA) [27], LBP, HOG [28]	**Foco:** Redução de dimensionalidade. **Limitação:** Sensível a variações de pose e iluminação.
2010–2014	Transição para CNNs	AlexNet [30]	**Foco:** Prova do potencial das CNNs em tarefas de larga escala. Melhor robustez a variações.
2014–Hoje	CNNs avançadas e aprendizado de métrica	DeepFace [32], FaceNet (Triplet Loss) [10], ArcFace [12], CosFace [33]	**Foco:** Aumentar a discriminação entre classes com funções de perda avançadas. Desempenho de nível humano.

Tabela 2.1: Evolução das abordagens em reconhecimento facial ao longo do tempo.

2.4 DESAFIOS CRÍTICOS AOS SISTEMAS DE RECONHECIMENTO FACIAL: A AMEAÇA DO SPOOFING

No decorrer da evolução dos sistemas de reconhecimento facial, intensificaram-se também os esforços para comprometer sua segurança por meio de técnicas de falsificação, com destaque para o *facial spoofing* (do inglês, falsificação facial) [35]. Essa técnica busca simular a identidade legítima de um indivíduo diante do sensor biométrico, utilizando artefatos como fotografias impressas, vídeos em reprodução ou máscaras tridimensionais que imitam o rosto da vítima, com o intuito de burlar o processo de autenticação [36]. As tentativas de falsificação facial têm se tornado cada vez mais frequentes em aplicações críticas, como sistemas bancários e mecanismos de controle de acesso, e permanecem um desafio significativo para os sistemas biométricos modernos. Nesta seção, são apresentados os principais tipos de ataques abordados na literatura.

2.4.1 Principais Categorias de Ataques

Os métodos de ataques de falsificação facial podem ser agrupados em duas categorias principais, de acordo com a natureza dos artefatos utilizados para enganar os sistemas de autenticação [35].

2.4.1.1 Ataques Bidimensionais (2D)

- **Fotografias Impressas.** Consistem na apresentação de imagens estáticas do rosto da vítima, geralmente impressas em papel de alta qualidade. São de fácil execução e têm ampla viabilidade prática, especialmente devido à disponibilidade de fotos pessoais em redes sociais.
- **Reprodução de Vídeo.** Utilizam vídeos previamente gravados da pessoa-alvo, exibidos em dispositivos como smartphones ou tablets. Por conterem movimentos faciais e expressões naturais, esses ataques são mais difíceis de detectar do que aqueles baseados em imagens estáticas.

2.4.1.2 Ataques Tridimensionais (3D)

- **Máscaras Faciais 3D.** Empregam máscaras tridimensionais com alto grau de realismo, confeccionadas com materiais como silicone, resina ou plástico. Essas máscaras reproduzem com precisão a geometria e textura do rosto da vítima, representando um dos desafios mais complexos para os mecanismos atuais de detecção de falsificação.

2.5 TÉCNICAS DE DETECÇÃO ANTI-SPOOFING

O avanço das técnicas de falsificação facial tem impulsionado o desenvolvimento de métodos especializados de detecção de vivacidade, do inglês *liveness detection*, e mecanismos anti-spoofing [37]. Essas abordagens visam distinguir entre uma apresentação autêntica e tentativas de burlar o sistema com representações artificiais. As estratégias adotadas podem ser agrupadas em três categorias principais:

2.5.1 Metodologias de Classificação

2.5.1.1 Técnicas Baseadas em Sensor

Essas abordagens utilizam componentes de hardware adicionais para capturar propriedades físicas da face humana. Sensores de profundidade, como câmeras TOF, do inglês *time-of-flight*, são eficazes contra ataques com imagens e vídeos 2D por mapearem a geometria tridimensional do rosto [38]. Sistemas com iluminação multiespectral analisam a resposta da pele em diferentes faixas de comprimento de onda, possibilitando a distinção entre tecidos biológicos e materiais sintéticos [9]. Sensores térmicos ou de textura capturam padrões de calor ou microtexturas da pele, sendo eficazes na detecção de rostos impressos ou mascarados [39].

2.5.1.2 Técnicas Baseadas em Características

São métodos que não requerem hardware adicional e operam sobre imagens convencionais. Métodos baseados em descritores como LBP (*Local Binary Patterns* — Padrões Binários Locais), HOG (*Histogram of Oriented Gradients* — Histograma de Gradientes Orientados) extraem microtexturas faciais difíceis de replicar em artefatos sintéticos [40]. Análises temporais detectam movimentos involuntários, como piscadas ou variações de expressão [41, 42]. O fluxo óptico, do inglês *optical flow*, contribui para identificar rigidez em rostos falsificados [43].

2.5.1.3 Técnicas Baseadas em Pontuação

Essas técnicas operam diretamente sobre as pontuações de similaridade produzidas pelo sistema biométrico, sem realizar uma análise explícita da imagem. Ataques de falsificação geralmente resultam em padrões de pontuação distintos dos observados em interações autênticas, o que possibilita sua identificação [44]. Estratégias de limiarização adaptativa ajustam automaticamente os critérios de aceitação com base no histórico e no contexto de uso [45]. Além disso, a aplicação de técnicas de fusão de classificadores, como SVM (Support Vector Machine), Random Forest e CNNs, tem se mostrado eficaz ao combinar múltiplos critérios decisórios sobre essas pontuações [46].

A Tabela (2.2) resume essa e outras abordagens utilizadas na detecção de ataques de falsificação facial.

Categoria	**Descrição**	**Métodos/Exemplos**	**Vantagens e Desvantagens**
Baseadas em Sensor	Utilizam sensores especializados para capturar sinais além do espectro visível.	Câmeras TOF [38], multiespectral [9], térmico [39]	**Vantagens:** Alta eficácia contra ataques 2D e materiais falsificados. **Desvantagens:** Alto custo e menor viabilidade em ambientes reais.
Baseadas em Características	Extraem padrões visuais e dinâmicos diretamente das imagens faciais.	LBP, HOG [40], piscadas [41], fluxo óptico [43]	**Vantagens:** Simples, econômicas e eficazes para ataques simples. **Desvantagens:** Menor robustez contra deepfakes ou máscaras realistas.
Baseadas em Pontuação	Analisam as respostas do sistema biométrico, sem examinar diretamente a imagem.	Limiarização adaptativa [45], fusão de classificadores [46]	**Vantagens:** Boa integração com sistemas existentes, detecta desvios estatísticos. **Desvantagens:** Dependência do comportamento interno do sistema.

Tabela 2.2: Metodologias de classificação para detecção de ataques de falsificação facial.

2.6 TÉCNICAS TENSORIAIS

Esta seção tem por objetivo apresentar os fundamentos dos métodos tensoriais, destacando sua relevância para a modelagem de dados biométricos multidimensionais. Serão abordados a motivação para seu emprego, os principais benefícios em relação a abordagens convencionais, em especial as redes neurais convolucionais, bem como os modelos de dados e as técnicas de decomposição multilinear mais representativos.

2.6.1 Motivação

Sistemas de Reconhecimento Facial (SRF) em ambientes de autenticação biométrica exigem representações que capturem simultaneamente variações espaciais, temporais e contextuais. Métodos tradicionais baseados em vetores ou matrizes, impõem restrições que comprometem relações particulares entre diferentes variáveis como identidade, expressão e iluminação [26, 27]. Redes neurais convolucionais (CNNs), embora dominantes em tarefas visuais, partem de suposições limitadas quanto à estrutura dos dados, como localidade e invariância espacial. Além disso, exigem grandes volumes de amostras rotuladas para alcançar o desempenho desejado [30, 47].

Representações tensoriais, embora já consolidadas na literatura, continuam sendo uma alternativa relevante por preservarem a estrutura de ordem superior dos dados [48]. Atualmente, seu emprego tem se expandido em cenários híbridos, nos quais representações extraídas por CNNs são posteriormente organizadas em estruturas tensoriais. Isso permite a aplicação de técnicas de decomposição e análise não supervisionada, como métodos de clusterização, promovendo maior robustez e interpretabilidade mesmo em condições com dados escassos ou rótulos limitados [15, 49].

2.6.2 Modelos de Dados Tensoriais

A generalização natural de vetores e matrizes leva ao conceito de *tensores*, que são estruturas de ordem superior capazes de representar dados multidimensionais [15]. Enquanto vetores modelam relações unidimensionais e matrizes capturam interações bidimensionais, tensores permitem modelar simultaneamente múltiplos relações de variação de dados, como espaço, tempo, identidade e condição de aquisição.

Uma imagem colorida pode ser representada como um tensor de terceira ordem com as seguintes dimensões: largura $\times$ altura $\times$ canais (RGB). Considerando o contexto de reconhecimento facial, conjuntos de imagens de múltiplos indivíduos sob diferentes condições podem ser modelados como tensores de ordem superior, sem necessidade de achatamento ou perda estrutural [48].

Os dados modelados em uma representação tensorial podem ser organizados como:

- **Tensores densos:** Apresentam valores definidos na maioria de suas entradas [15].
- **Tensores esparsos:** Contêm majoritariamente entradas nulas ou irrelevantes. São comuns em aplicações como análise de coocorrência ou eventos raros em bases de dados biométricos extensos [50].

No contexto deste trabalho, os dados faciais são organizados em um tensor tridimensional, o que permite capturar simultaneamente variações espaciais, contextuais e de identidade. A estrutura adotada é:

$$\mathcal{X} \in \mathbb{R}^{I \times J \times K} \tag{2.1}$$

em que:

- I representa a dimensão vetorizada da imagem facial, isto é, o número total de pixels ($I = h \cdot w$).
- J refere-se às condições sob as quais a imagem foi adquirida, distinguindo entre aquisições autênticas e tentativas de falsificação.
- K representa diferentes identidades presentes na base de dados.

A construção do tensor $\mathcal{X}$ definido na Equação (2.1) segue uma organização pré-definida dos dados faciais coletados. Suponha-se um conjunto de K indivíduos, cada um submetido a J condições distintas de aquisições. Para cada par (j, k), é capturada uma imagem bidimensional de tamanho $h \times w$, a qual é então transformada em um vetor coluna de dimensão:

$$\mathbf{x}_{jk} \in \mathbb{R}^{I}, \quad \text{com } I = h \cdot w \tag{2.2}$$

Esse vetor $\mathbf{x}_{jk}$ contém os valores de intensidade de pixel reorganizados, por exemplo, linha a linha. A associação entre vetor, condição e identidade é preservada, formando o elemento correspondente no tensor tridimensional:

$$\mathcal{X}(i, j, k) = \text{valor do } i\text{-ésimo pixel da imagem da identidade } k \text{ sob condição } j \tag{2.3}$$

A estrutura do tensor resultante pode ser visualizada da seguinte forma:

- $\mathcal{X}_{::k}$: fatia frontal, contém todas as amostras da identidade k sob diferentes condições.
- $\mathcal{X}_{:j:}$: fatia lateral, agrupa todas as imagens sob a condição j para todos os indivíduos.
- O modo I corresponde aos vetores de pixels das imagens.

Esse arranjo tensorial definido na Equação (2.3) preserva a estrutura dos dados permitindo posteriores aplicações de decomposições tensoriais em que são mantidas as correlações entre cada modo [15]. Essa abordagem é amplamente utilizada em tarefas de visão computacional e reconhecimento facial, dada sua capacidade de representar dados de ordem superior de forma interpretável [48]. Para maiores detalhes sobre notação e operações com tensores, recomenda-se a leitura do Apêndice (I.3).

Observação sobre a dimensão do modo de condição (J). No arranjo tensorial da Equação (2.1), a dimensão J controla a granularidade das *condições* (p. ex., *live* vs. *spoof* ou subtipos de ataque). Em

termos práticos, J pequeno (e.g., binário) tende a simplificar a separação de fatores ao longo dos demais modos (I e K), enquanto J elevado, com condições heterogêneas, pode exigir posto R maior na fatoração para manter poder de representação. Em decomposições multilineares, isso aparece como: (i) necessidade de núcleos maiores (Tucker) ou (ii) aumento do posto canônico (CPD), sob pena de *misturar* variação de condição em componentes espaciais (I) ou identitárias (K). Na prática, calibra-se R (ou R_1, R_2, R_3 no caso Tucker) de modo a preservar *variância explicada* por condição sem degradar a separabilidade em I e K.

2.6.3 Decomposições Multilineares

A análise de dados tensoriais, que envolvem informações organizadas em múltiplas dimensões (modos), é essencial em diversas áreas que lidam com dados complexos e estruturados. As decomposições multilineares constituem ferramentas poderosas para a extração de estruturas latentes, redução de dimensionalidade e compressão de dados de ordem superior. Elas generalizam conceitos clássicos da Álgebra Linear, como a Decomposição em Valores Singulares (SVD) e a Análise de Componentes Principais (PCA), discutidos no Apêndice I.2, oferecendo uma forma mais expressiva e interpretável de analisar dados multidimensionais.

2.6.3.1 Decomposição CANDECOMP/PARAFAC (CPD) e o Produto Khatri–Rao

A Decomposição CANDECOMP/PARAFAC (CPD) [51, 52], também conhecida como CP ou PARAFAC, é uma das técnicas de decomposição tensorial mais utilizadas na literatura em diversas aplicações. Essa abordagem representa um tensor como a soma de um número mínimo de componentes de posto um, sendo cada componente composta pelo produto externo de vetores, um para cada modo. Essa estrutura permite capturar diretamente as interações entre as diferentes dimensões dos dados, de maneira compacta, interpretável e adequada à análise de padrões latentes.

A CPD destaca-se por sua forma canônica, pela facilidade de interpretação dos fatores extraídos e por não exigir restrições adicionais para alcançar unicidade [15]. Além disso, sua aplicação tem se mostrado eficaz em cenários com dados de ordem superior, como ocorre no contexto biométrico. Por essas razões, essa decomposição será adotada como base para o modelo de dados desenvolvido neste trabalho.

A Equação (2.4) ilustra a forma geral da decomposição CPD para um tensor tridimensional.

$$\mathcal{X} \approx \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \tag{2.4}$$

em que:

- $\mathbf{a}_r \in \mathbb{R}^I$, $\mathbf{b}_r \in \mathbb{R}^J$, $\mathbf{c}_r \in \mathbb{R}^K$: São denomidados vetores de fatores e estão associados a modo do tensor $\mathcal{X}$ na r-ésima componente de posto um.
- $\circ$: Indica o produto externo.

- R: É o posto tensorial canônico (também chamado de posto PARAFAC), que é o menor número de componentes de posto um necessárias para representar o tensor.

A decomposição CPD também pode ser expressa em forma matricial por meio do conceito de desdobramento do tensor e da aplicação do produto de Khatri-Rao. Por exemplo, o desdobramento no modo-1 do tensor $\mathcal{X}$, denotado por $\mathbf{X}_{(1)}$, pode ser representado da seguinte forma:

$$\mathbf{X}_{(1)} = \mathbf{A}(\mathbf{C} \diamond \mathbf{B})^T \tag{2.5}$$

Nessa equação:

- $\mathbf{X}_{(1)} \in \mathbb{R}^{I \times (JK)}$: representa o desdobramento do modo-1 do tensor $\mathcal{X}$, obtido ao empilhar sequencialmente suas fatias ao longo desse modo.
- $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_R] \in \mathbb{R}^{I \times R}$: matriz cujas colunas são os vetores de fatores associados ao modo-1.
- $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_R] \in \mathbb{R}^{J \times R}$: matriz de fatores do modo-2.
- $\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R] \in \mathbb{R}^{K \times R}$: matriz de fatores do modo-3.
- $\diamond$: denota o produto de Khatri-Rao, uma operação fundamental na álgebra tensorial que corresponde ao produto de Kronecker coluna a coluna:

  $$\mathbf{A} \diamond \mathbf{B} = [\mathbf{a}_1 \otimes \mathbf{b}_1, \ldots, \mathbf{a}_R \otimes \mathbf{b}_R] \tag{2.6}$$

  onde $\mathbf{a}_r$ e $\mathbf{b}_r$ são as colunas correspondentes das matrizes $\mathbf{A}$ e $\mathbf{B}$, e $\otimes$ representa o produto de Kronecker.

A representação matricial é útil para a implementação prática da CPD e para a aplicação de métodos de otimização baseados em mínimos quadrados. A CPD permanece especialmente valorizada por sua unicidade sob condições leves [15] e pela interpretabilidade direta de seus fatores, o que a torna adequada para tarefas de análise de dados estruturados, como em aplicações biométricas. Para maiores informações sobre o desdobramento de tensores, produtos matriciais e demais operações associadas, recomenda-se a leitura do Apêndice (I.3).

2.6.3.2 Decomposição Tucker, Higher-Order SVD e Higher-Order Orthogonal Iteration

A decomposição de Tucker [53] é descrita como uma generalização direta da Decomposição em Valores Singulares (SVD, do inglês *Singular Value Decomposition*) para tensores de ordem superior. O objetivo é representar um tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ em termos de um tensor núcleo de menor dimensão e de matrizes fator que definem subespaços associados a cada modo. Formalmente, o método é definido como:

$$\mathcal{X} \approx \mathcal{G} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}, \tag{2.7}$$

em que $\mathcal{G} \in \mathbb{R}^{R_1 \times R_2 \times R_3}$ é o *tensor núcleo*, que preserva as interações entre os modos, e cada matriz fator $\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times R_n}$ contém bases ortogonais que definem os subespaços de dimensão reduzida associados ao modo n (com $I_1 = I$, $I_2 = J$, $I_3 = K$). Diferentemente da decomposição CPD (CANDECOMP/PARAFAC), a decomposição de Tucker permite especificar diferentes postos (R_1, R_2, R_3) para cada modo, oferecendo maior flexibilidade de modelagem. O tensor núcleo $\mathcal{G}$ atua como uma versão compactada do tensor original, mantendo as correlações entre os fatores.

A obtenção das matrizes fator é realizada por meio da *Higher-Order Singular Value Decomposition* (HOSVD). Este método pode ser interpretado como um procedimento específico para computar a decomposição de Tucker, em que aplica-se a SVD às matrizes de desdobramento $X_{(n)}$ de $\mathcal{X}$ em cada modo n e selecionam-se as $R_{(n)}$ primeiras componentes principais para compor $\mathbf{U}^{(n)}$. O tensor núcleo é então obtido por projeção:

$$\mathcal{G} = \mathcal{X} \times_1 \mathbf{U}^{(1)^\top} \times_2 \mathbf{U}^{(2)^\top} \times_3 \mathbf{U}^{(3)^\top} \tag{2.8}$$

A HOSVD fornece uma aproximação inicial, não necessariamente a que minimiza o erro de reconstrução para os postos escolhidos. Para obter a melhor aproximação no sentido da norma de Frobenius, emprega-se o algoritmo *Higher-Order Orthogonal Iteration* (HOOI), que refina iterativamente as matrizes fator. Em cada iteração, um modo é otimizado mantendo os demais fixos, até convergência, produzindo uma solução que minimiza o erro de reconstrução de forma global [54].

2.7 TÉCNICAS DE CLUSTERIZAÇÃO E VISUALIZAÇÃO

Esta seção apresenta os fundamentos de duas técnicas complementares amplamente utilizadas na análise de dados de alta dimensionalidade: o t-distributed Stochastic Neighbor Embedding (t-SNE) [17] e o algoritmo K-Means [19]. Enquanto o t-SNE é empregado para projetar dados complexos em um espaço de menor dimensionalidade, preservando suas relações locais, o K-Means atua sobre esse espaço reduzido para identificar agrupamentos com base em proximidade euclidiana. A combinação dessas duas técnicas permite não apenas visualizar padrões estruturais em representações biométricas, mas também organizar automaticamente os dados em grupos distintos de forma interpretável. A classificação de novas amostras ocorre por meio de sua projeção no espaço reduzido, seguida pela atribuição ao cluster mais próximo, o que possibilita distinguir entre faces autênticas e tentativas de falsificação.

2.7.1 Redução de Dimensionalidade com t-SNE

O *t-distributed Stochastic Neighbor Embedding* (t-SNE) [55] é uma técnica de redução de dimensionalidade projetada para preservar relações locais de vizinhança entre os dados. A ideia central do t-SNE é converter distâncias euclidianas no espaço de alta dimensão em distribuições de probabilidades que representam similaridades entre pares de pontos, tanto no espaço original quanto no espaço projetado.

Sejam $\mathbf{x}_i$ e $\mathbf{x}_j$ vetores no espaço de alta dimensionalidade, a similaridade entre eles é modelada por

uma distribuição gaussiana simétrica, conforme expressa pela equação:

$$p_{ij} = \frac{\exp\left(-|\mathbf{x}_i - \mathbf{x}j|^2/2\sigma_i^2\right)}{\sum k \neq l \exp\left(-|\mathbf{x}_k - \mathbf{x}_l|^2/2\sigma_k^2\right)} \tag{2.9}$$

em que:

- $\mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^D$: vetores de entrada no espaço de alta dimensão, onde D é a dimensionalidade original dos dados;
- $|\mathbf{x}_i - \mathbf{x}_j|^2$: distância euclidiana ao quadrado entre os vetores $\mathbf{x}_i$ e $\mathbf{x}_j$.
- σ_i: parâmetro de dispersão (desvio padrão) ajustado localmente para o ponto $\mathbf{x}i$, responsável por regular a largura da distribuição gaussiana.
- p_{ij}: probabilidade simétrica que representa o grau de similaridade entre os pontos $\mathbf{x}_i$ e $\mathbf{x}_j$ no espaço original, refletindo a chance de $\mathbf{x}_j$ ser considerado vizinho de $\mathbf{x}_i$.
- O denominador da expressão realiza a normalização global necessária para garantir que as probabilidades formem uma distribuição válida, ou seja, $\sum_{i \neq j} p_{ij} = 1$.

No espaço de baixa dimensão (por exemplo, $\mathbb{R}^2$), os vetores $\mathbf{y}_i$ e $\mathbf{y}_j$ representam as projeções dos dados originais, sendo modelados por uma distribuição de Student com um grau de liberdade:

$$q_{ij} = \frac{(1 + \|\mathbf{y}_i - \mathbf{y}_j\|^2)^{-1}}{\sum_{k \neq l}(1 + \|\mathbf{y}_k - \mathbf{y}_l\|^2)^{-1}} \tag{2.10}$$

O objetivo do t-SNE é minimizar a divergência de Kullback-Leibler entre as distribuições P e Q:

$$\mathrm{KL}(P \parallel Q) = \sum_{i \neq j} p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right) \tag{2.11}$$

Essa otimização resulta em uma projeção que preserva as vizinhanças locais dos dados, facilitando a visualização de agrupamentos, separabilidade entre classes e detecção de outliers em representações de alta dimensão. No contexto deste trabalho, o espaço reduzido obtido pelo t-SNE será utilizado como base para a etapa de agrupamento com o algoritmo K-Means a ser detalhado na próxima seção.

2.7.2 Agrupamento de dados com o método K-Means

O algoritmo *K-Means* [18] é um dos métodos mais empregados para clusterização em espaços vetoriais. Ele busca particionar um conjunto de dados em k grupos disjuntos, minimizando a variância intra-cluster com base na distância euclidiana entre os pontos e os centróides dos clusters.

Formalmente, dado um conjunto de n vetores no espaço reduzido gerado pelo t-SNE, denotado por $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}$ o algoritmo K-Means busca determinar k centróides $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k$ que minimizem a seguinte função:

$$J = \sum_{i=1}^{n} \sum_{j=1}^{k} r_{ij} |\mathbf{y}_i - \boldsymbol{\mu}_j|^2 \tag{2.12}$$

em que:

- $\mathbf{y}_i \in \mathbb{R}^d$: vetor que representa a i-ésima amostra no espaço de baixa dimensão.
- $\boldsymbol{\mu} j \in \mathbb{R}^d$: centróide do j-ésimo cluster.
- $r_{ij} \in 0, 1$: variável indicadora que assume valor 1 se $\mathbf{y}_i$ pertence ao cluster j e 0 caso contrário.
- J: função objetivo que representa a soma total das distâncias quadráticas entre os pontos e seus respectivos centróides, ou seja, a variância intra-cluster a ser minimizada.

O algoritmo é iterativo e alterna entre duas etapas:

- **Atribuição**: cada ponto é associado ao centróide mais próximo.
- **Atualização**: os centróides são recalculados como a média dos pontos atribuídos a cada cluster.

Como neste trabalho assume-se previamente que os dados correspondem a dois grupos distintos, compostos por amostras autênticas e tentativas de falsificação, o valor de k é fixado em 2.

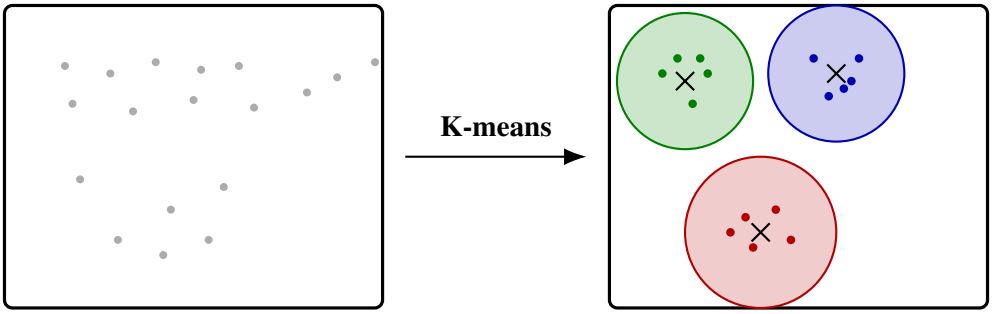

Figura 2.2: Exemplo ilustrativo do algoritmo K-Means: à esquerda, pontos sem rótulos; à direita, os mesmos pontos agrupados em três clusters com seus centróides.

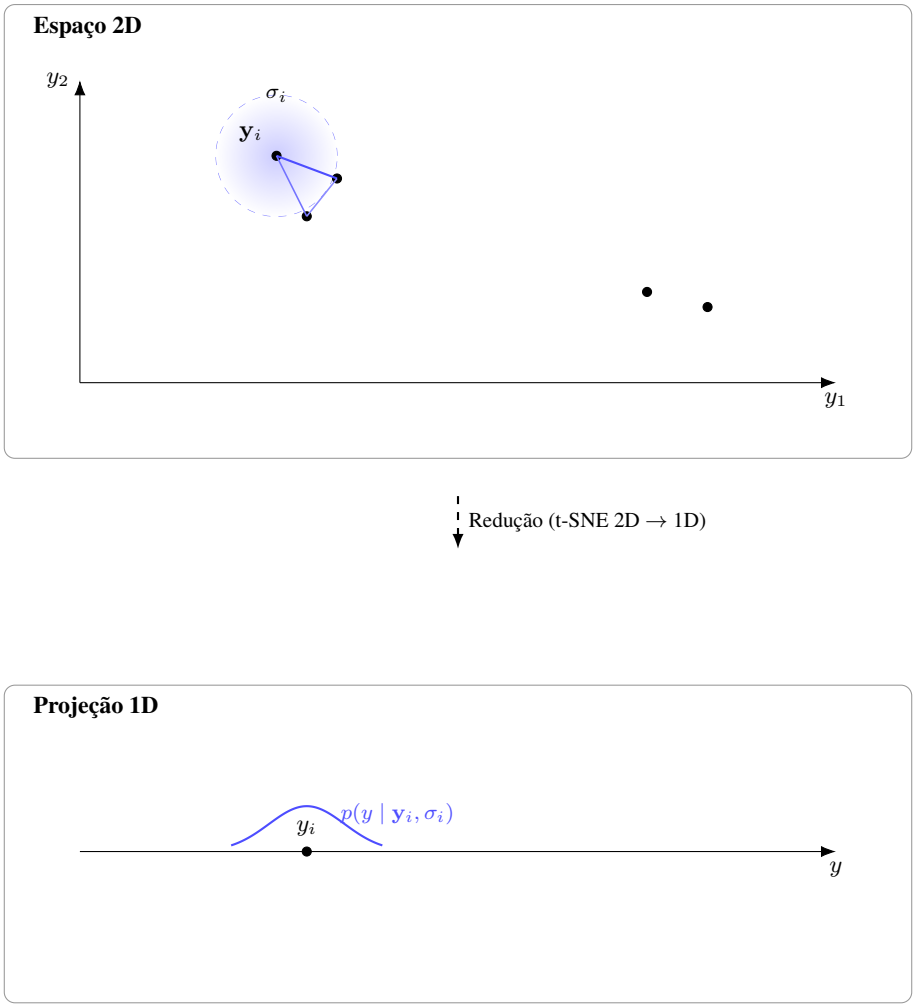

Figura 2.3: Visualização com a gaussiana: no espaço 2D, cada ponto $\mathbf{y}_i$ tem uma distribuição local $\mathcal{N}(\mathbf{y}_i, \sigma_i^2 I)$; na projeção 1D, essa distribuição aparece como uma curva gaussiana centrada em y_i.

3 TRABALHOS RELACIONADOS

Este capítulo apresenta uma revisão abrangente da literatura relacionada à detecção de falsificação em sistemas biométricos faciais, com ênfase em abordagens baseadas em decomposição tensorial e técnicas de aprendizado de máquina, do inglês *machine learning*. A compreensão das metodologias existentes, bem como de suas limitações, é fundamental para contextualizar a proposta deste trabalho, destacando sua originalidade e relevância na literatura.

3.1 DECOMPOSIÇÕES TENSORIAIS

As decomposições tensoriais têm se destacado por sua capacidade de modelar dados multidimensionais de forma estruturada, como sequências temporais de imagens, dados multiespectrais ou coleções de imagens com variações em iluminação, pose e identidade. Essas técnicas oferecem vantagens significativas na extração de representações latentes e na redução de dimensionalidade dos dados [15].

3.1.1 Aplicações no Reconhecimento e Clusterização Facial

A Decomposição Tucker [56] (ou MLSVD) é amplamente utilizada para modelagem de faces em bases de dados multidimensionais. Por exemplo, ela permite separar as variações de identidade, expressão e iluminação em conjuntos de dados de faces. A referência [48] introduziram o conceito de *TensorFaces*, demonstrando como a Decomposição Tucker pode capturar as relações multilineares entre esses modos, levando a representações mais robustas e compactas para o reconhecimento facial. Posteriormente, a referência [49] aprofundou essa ideia propondo o método *Multi-linear Principal Component Analysis (MPCA)*, que estende a PCA para tensores, permitindo extrair componentes principais para cada modo do tensor de faces de forma interligada, com o objetivo de melhorar o desempenho de tarefas de reconhecimento facial.

A Decomposição CANDECOMP/PARAFAC (CPD) [16, 52] também encontrou diversas aplicações no reconhecimento facial e na clusterização de identidades. A referência [8] explora a generalização da PCA para o uso de decomposições tensoriais em reconhecimento. Outro trabalho relevante nesse contexto é trabalho de [57], que propõe uma formulação para o reconhecimento facial como uma equação de produto de Kronecker. Foi demonstrado como um conjunto de imagens faciais pode ser modelado como um tensor e fatorado usando decomposições tensoriais, permitindo que a variação dentro de cada classe de identidade seja explicada por um produto de Kronecker de matrizes de menor dimensão. Essa abordagem foca na clusterização de identidades baseada nas estruturas de produto de Kronecker inferidas dos dados tensoriais. Outros estudos também empregam decomposições tensoriais para problemas de clusterização de dados multimodais, como em reconhecimento de gestos [58] e análise de vídeos de vigilância [59]. É importante destacar que os trabalhos revisados nesta seção têm como foco principal a identificação e o agrupamento de indivíduos no contexto do reconhecimento facial, sem abordar diretamente a distinção entre faces autênticas e falsificadas.

3.2 DETECÇÃO DE TENTATIVAS DE FALSIFICAÇÃO

No contexto da distinção entre uma apresentação facial genuína e uma tentativa de falsificação, os ataques podem assumir diversas formas, incluindo o uso de fotografias impressas, vídeos exibidos em dispositivos eletrônicos, máscaras tridimensionais, entre outros artifícios.

3.2.1 Abordagens Tradicionais

As metodologias propostas na literatura podem ser categorizadas em:

- **Baseadas em Características de Textura:** Métodos que analisam a textura da pele para identificar artefatos introduzidos por ataques. Exemplos incluem LBP (Local Binary Patterns) [60], Descritores de Padrões de Gradiente Orientado (HOG) [4] e Wavelets [9]. Esses métodos buscam irregularidades na textura, ruído ou falta de micro-estruturas da pele viva.
- **Baseadas em Análise de Sinais Fisiológicos:** Focam na detecção de sinais de vitalidade, como o fluxo sanguíneo (pulso) detectado por variações de cor na pele (PPG - Photoplethysmography) [61], ou piscar de olhos e movimentos naturais da cabeça [62].
- **Baseadas em Informação de Profundidade 3D:** Utilizam sensores de profundidade (RGB-D) para analisar a geometria do rosto. Ataques 2D (fotos, vídeos) não possuem informações de profundidade consistentes com um rosto humano real [63].

Embora eficazes em certos cenários, muitas dessas abordagens tradicionais sofrem com a falta de generalização para ataques não vistos e alta dependência de condições de aquisição dos dados.

3.2.2 Abordagens Baseadas em Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNNs) consolidaram-se como o estado da arte na detecção de tentativas de falsificação facial [64]. Sua principal vantagem reside na capacidade de aprender representações hierárquicas e discriminativas diretamente a partir dos dados brutos, dispensando a definição manual de características. Em geral, as abordagens baseadas em CNN operam sobre as imagens de entrada de forma direta ou com o auxílio de etapas preliminares de pré-processamento simples.

Diversas arquiteturas de CNNs têm sido propostas na literatura para a detecção de tentativas falsificação em sistemas de reconhecimento facial. Entre os principais métodos, detacam-se:

- **Redes convolucionais compactas:** Têm por objetivo aplicações em tempo real. Neste contexto, arquiteturas como MobileNet [65] e ShuffleNet [64]. A referência [66] propôs uma CNN sensível a detalhes de alta frequência para detectar tentativas de ataques por impressões e vídeos.
- **Redes Multi-Escala e Multi-Modais:** Estratégias que combinam múltiplas escalas de análise, como a Feature Pyramid Network (FPN) [67]. A integração de diferentes modalidades com CNNs paralelas ou de fusão, têm se mostrado eficazes para capturar artefatos visuais em diferentes níveis, desde detalhes sutis até padrões globais [68, 69].

- **Funções de Perda Avançadas e Supervisão Auxiliar:** Esses métodos empregam funções de perda específicas e tarefas auxiliares para enriquecer o aprendizado. Entre elas, destacam-se o uso de mapas de profundidade estimados [69] e mapas de reflexão de superfície [70], que induzem a rede a extrair representações discriminativas e mais generalizáveis. Estratégias como aprendizado contrastivo e busca por arquiteturas eficientes também têm sido exploradas com bons resultados [71].
- **Generalização e Adaptação de Domínio:** Um dos principais desafios em detecção de falsificação facial com CNNs é garantir a capacidade de generalização frente a ataques desconhecidos e variações entre bases de dados. Para enfrentar esse problema, técnicas como aprendizado adversarial de domínio, do inglês *Domain Adversarial Training* [72] e meta-aprendizagem [73] têm sido empregadas, visando aprimorar a robustez e a transferibilidade dos modelos para diferentes contextos.

É importante destacar que a maioria dos métodos baseados em CNNs não incorpora decomposições tensoriais complexas como etapa preliminar no processamento dos dados. Em geral, esses métodos exploram diretamente a capacidade das camadas convolucionais de extrair representações hierárquicas a partir das imagens de entrada, muitas vezes utilizando apenas pré-processamentos simples.

3.2.3 Comparativo das Abordagens

A tabela 3.1 resume as características e o foco principal das duas grandes categorias de soluções discutidas, destacando a lacuna que este trabalho busca preencher.

Tipo de Solução	**Foco Principal**	**Uso de Tensores no Pré-processamento de Dados**	**Aplicação Direta em Detecção de Tentativas de Falsificação**
Soluções Tensoriais Clássicas [48, 49]	Reconhecimento facial, identificação de identidades, modelagem de variação de pose/iluminação.	**Sim.** utilizam decomposições tensoriais como base fundamental para extrair características.	**Não**.
Soluções Baseadas em CNN [69, 66, 68, 67, 71, 64]	Detecção de ataques de apresentação (distinção *live* vs. *spoof*). Aprendizado de características discriminativas diretamente das imagens.	**Não**. Processam imagens diretamente ou com pré-processamentos simples.	**Sim**.
Método proposto neste trabalho	**Sim.**	**Sim.**	**Sim.**

Tabela 3.1: Comparativo entre Soluções Baseadas em Tensores e Soluções Baseadas em CNN para Problemas Biométricos Faciais.

A Tabela (3.1) evidencia que, embora métodos baseados em decomposição tensorial sejam eficazes na captura de estruturas complexas em dados faciais, seu uso na detecção de tentativas de falsificação ainda é limitado e pouco sistematizado. Por outro lado, abordagens baseadas em CNN têm sido amplamente empregadas para esse fim, mas geralmente sem explorar as potencialidades oferecidas pelas decomposições tensoriais no estágio de pré-processamento. A proposta deste trabalho busca justamente preencher essa lacuna, integrando a estruturação tensorial como uma etapa prévia à extração de características, de modo a potencializar a capacidade discriminativa na detecção de falsificações.

4 MÉTODO PROPOSTO

Este capítulo apresenta a solução desenvolvida para a prevenção de falsificações em sistemas de reconhecimento facial. A abordagem proposta combina técnicas de representação tensorial com algoritmos de clusterização, visando identificar tentativas de falsificação de maneira robusta e eficiente.

4.1 PREPARAÇÃO E PRÉ-PROCESSAMENTO DOS DADOS

4.1.1 O conjunto de dados CelebASpoof

Com o objetivo de validar a eficácia das técnicas propostas, a curadoria e a integridade dos dados utilizados são fundamentais. Seguindo critérios estabelecidos em diversas referências da literatura, este estudo adota um conjunto de dados amplamente reconhecido e homologado: o *CelebASpoof* [13]. Esse dataset é estruturado em diretórios distintos, correspondentes às duas classes que compõem as amostras: autênticas, do inglês *live*, e falsificadas, do inglês *spoof*. Cada imagem é acompanhada por um arquivo de texto contendo as coordenadas da caixa de detecção, do inglês *bounding box*, o que permite a segmentação precisa da região facial para análises subsequentes. Essa organização proporciona uma categorização objetiva das amostras, facilitando sua manipulação computacional e garantindo maior reprodutibilidade experimental. A estrutura hierárquica do diretório do conjunto de dados pode ser visualizada na Figura (4.1).

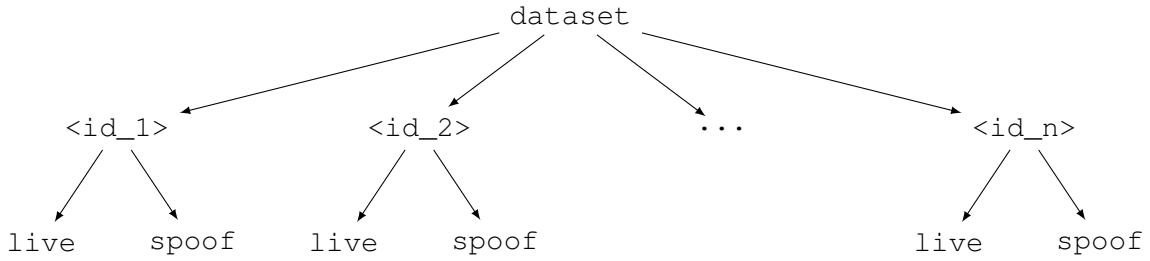

Figura 4.1: Estrutura hierárquica do diretório do conjunto de dados com IDs de indivíduos e categorias `live`/`spoof`.

4.1.2 Pré-processamento

O pré-processamento é uma etapa essencial para garantir a qualidade, padronização e robustez dos dados antes da aplicação dos métodos analíticos. Esta fase compreende três procedimentos principais:

- **Detecção e recorte facial (do inglês, *bounding box*)**
  As coordenadas fornecidas pelo conjunto de dados são utilizadas para extrair com precisão a região facial de cada imagem original $I \in \mathbb{R}^{H \times W}$, por meio do recorte definido por:

$$I_{\text{face}} = I\big(x_{\min} : x_{\max},\, y_{\min} : y_{\max}\big) \tag{4.1}$$

em que:

- $x_{\text{min}}, x_{\text{max}}$ definem os limites horizontais da *bounding box*.
- $y_{\text{min}}, y_{\text{max}}$ correspondem aos limites verticais da *bounding box*.
- H, W representam a altura e a largura da imagem original.
- I_{face} é a subimagem resultante contendo exclusivamente a região da face.

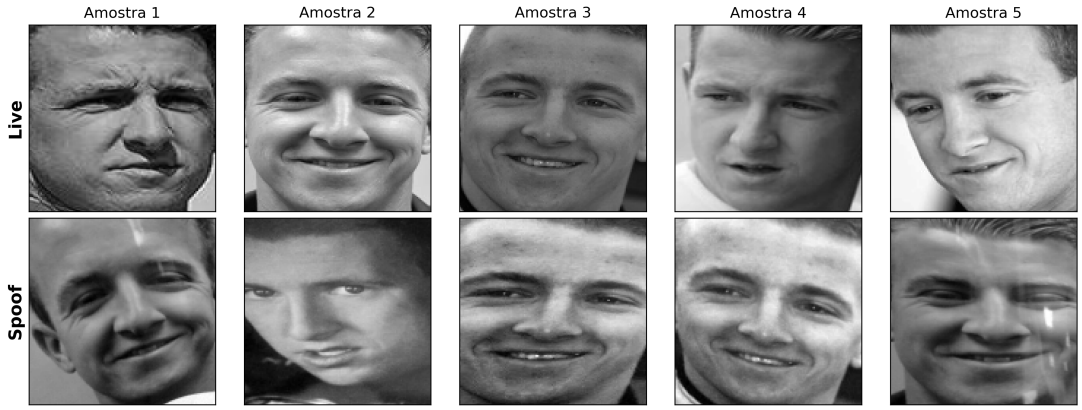

Figura 4.2: Exemplos de recortes faciais obtidos a partir das bounding boxes fornecidas. A linha superior mostra amostras *live*, enquanto a inferior apresenta amostras *spoof*.

- **Normalização das imagens para uma resolução uniforme**
  Cada face recortada I_{face} é redimensionada para uma resolução fixa (H_f, W_f), garantindo uniformidade nas entradas do pipeline de processamento. O redimensionamento é realizado por meio da seguinte transformação:

$$I_{\text{norm}}(x', y') = I_{\text{face}}\left(\frac{x'}{W_f}W,\ \frac{y'}{H_f}H\right) \tag{4.2}$$

  em que:

  - x', y' são as coordenadas na imagem normalizada, com $x' \in [0, W_f)$ e $y' \in [0, H_f)$.
  - H_f, W_f correspondem à altura e à largura desejadas após o redimensionamento.
  - H, W representam a altura e a largura da imagem da face original.
  - I_{norm} é a imagem resultante da face redimensionada.

- **Equalização e padronização da iluminação**
  Para reduzir variações de iluminação, aplica-se *Contrast Limited Adaptive Histogram Equalization* (CLAHE) [14]:

$$I_{\text{eq}}(x, y) = \text{CLAHE}\left(I_{\text{norm}}(x, y)\right), \tag{4.3}$$

  em que:

  - I_{eq} imagem resultante após equalização de histograma;

– $\mathrm{CLAHE}(\cdot)$ operador de equalização adaptativa de contraste.

A Figura (4.3) apresenta a média das imagens faciais para cada condição de apresentação (*live* e *spoof*), após a aplicação das etapas de pré-processamento descritas anteriormente.

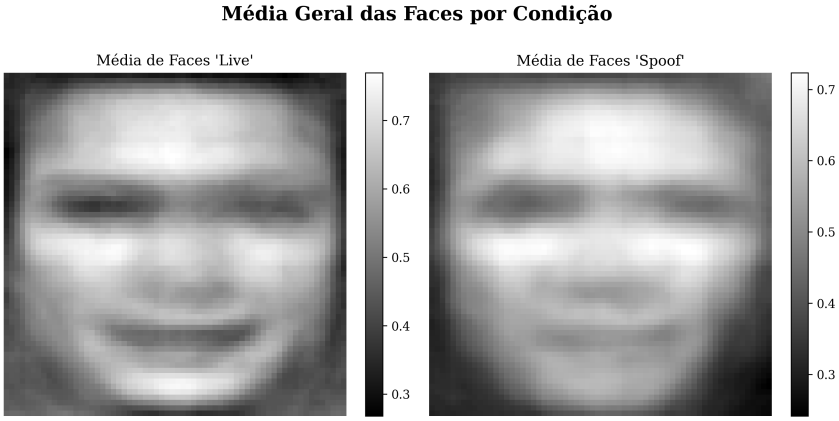

Figura 4.3: Imagens médias resultantes para as condições autêntica (*live*) e falsificada (*spoof*) após o pré-processamento

4.2 CONSTRUÇÃO DO TENSOR MULTIDIMENSIONAL

Após o pré-processamento, todas as imagens faciais são reorganizadas em uma estrutura de dados multidimensional que representa o conjunto completo de amostras. Essa estrutura é formalizada por meio da construção de um tensor tridimensional:

$$\mathcal{X} \in \mathbb{R}^{P \times C \times N}, \tag{4.4}$$

em que:

- P corresponde ao número de pixels por imagem após a vetorização, dado por $P = H_f \cdot W_f$, sendo H_f e W_f as dimensões normalizadas da imagem facial.
- C representa o número de condições de apresentação distintas, como autêntica (*live*) e falsificada (*spoof*).
- N denota o número total de indivíduos (identidades únicas) presentes no conjunto de dados.

A construção do tensor $\mathcal{X}$ inicia-se com a vetorização das imagens faciais pré-processadas. Cada imagem bidimensional $I_{\text{eq}} \in \mathbb{R}^{H_f \times W_f}$, conforme definido na Equação (4.3), é transformada em um vetor coluna $\mathbf{x} \in \mathbb{R}^P$, onde $P = H_f \cdot W_f$, por meio da concatenação dos pixels linha a linha. Em seguida, essas representações vetoriais são organizadas de acordo com duas dimensões adicionais: a condição de apresentação da amostra (autêntica ou falsificada) e a identidade do indivíduo a quem ela pertence. Essa estruturação resulta em um tensor tridimensional $\mathcal{X}$, cuja organização contempla as seguintes dimensões:

- **Espacial**: corresponde aos *pixels vetorizados* de cada imagem.
- **Semântica**: representa a *condição da amostra*, indicando se é autêntica (*live*) ou falsificada (*spoof*).
- **Identitária**: refere-se ao *indivíduo* associado à imagem.

Esse arranjo viabiliza a aplicação de técnicas de decomposição multilinear, como a decomposição CANDECOMP/PARAFAC (CPD), que permitem explorar as correlações latentes entre as diferentes dimensões do tensor construído. Ao preservar simultaneamente a variabilidade individual, associada às características faciais únicas de cada sujeito, e os aspectos contextuais relacionados às condições de apresentação, como iluminação, expressão ou tentativa de falsificação, essa representação fornece uma base matemática sólida e interpretável para tarefas de detecção de falsificação facial. Como resultado, o modelo obtido torna-se mais robusto na distinção entre amostras autênticas e falsificadas, contribuindo para uma análise mais precisa e confiável.

A Figura (4.4) ilustra a estrutura tridimensional do tensor construído a partir das imagens faciais vetorizadas, organizadas por condição de apresentação e identidade individual.

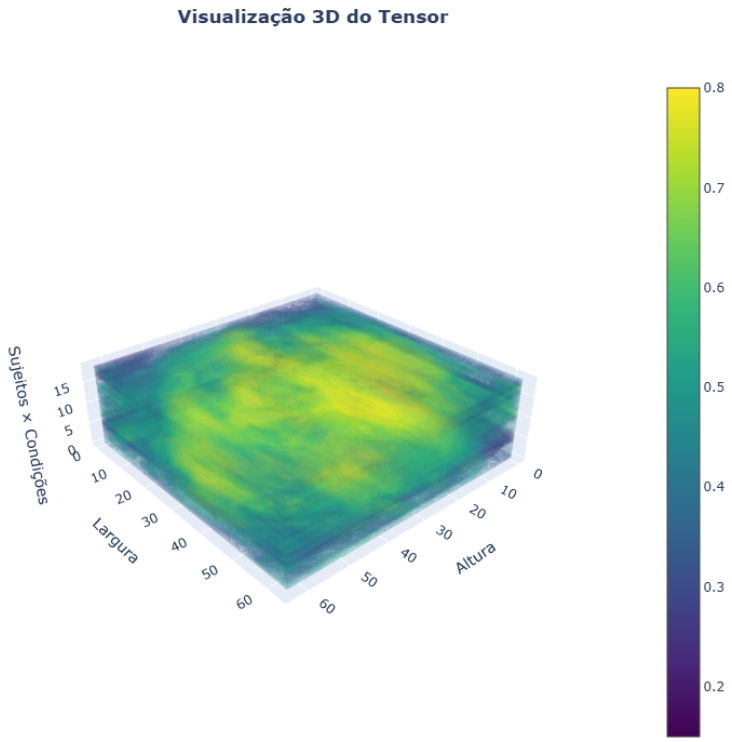

Figura 4.4: Representação esquemática do tensor formado a partir de dados de 10 indivíduos.

4.3 DECOMPOSIÇÃO TENSORIAL E CLUSTERIZAÇÃO PARA DETECÇÃO DE FRAUDE

Esta seção descreve o fluxo do processo proposto a partir da decomposição tensorial das imagens pré-processadas e culminando na clusterização voltada à detecção robusta de tentativas de falsificação facial. O objetivo é transformar os dados brutos de imagem em representações latentes compactas e semanticamente significativas, que possam ser agrupadas de forma a facilitar a distinção entre amostras autênticas e falsificadas.

4.3.1 Decomposição Tensorial

Para decompor o tensor tridimensional $\mathcal{X} \in \mathbb{R}^{P \times C \times N}$, emprega-se a técnica CANDECOMP/PARAFAC (CPD) detalhada previamente na Seção (2.6.3.1). Esta decomposição multilinear expressa o tensor como uma soma de componentes de posto unitário. A CPD permite fatorar $\mathcal{X}$ como:

$$\mathcal{X} \approx \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \tag{4.5}$$

em que:

- $\circ$ denota o produto externo.
- R é o posto da decomposição.
- Os vetores $\mathbf{a}_r \in \mathbb{R}^P$, $\mathbf{b}_r \in \mathbb{R}^C$ e $\mathbf{c}_r \in \mathbb{R}^N$ formam as componentes fatoradas associadas, respectivamente, às dimensões espacial, semântica e identitária do tensor.

A decomposição CPD fornece uma representação latente e interpretável das imagens faciais, separando variabilidades associadas a cada uma das dimensões. Além disso, a CPD é útil para capturar padrões consistentes que emergem entre condições e indivíduos, mesmo em presença de ruído ou variações sutis entre as amostras.

A decomposição CPD foi aplicada ao tensor representado na Figura (4.4), construído a partir de amostras faciais de 15 indivíduos. Com o intuito de capturar de maneira eficiente as variações latentes presentes nas três dimensões do tensor, foi adotado um posto de decomposição igual a $R = 15$. Essa escolha visa fornecer uma representação suficientemente expressiva para discriminar padrões relevantes, mantendo a interpretabilidade dos fatores extraídos.

As componentes resultantes da decomposição são apresentadas nas figuras a seguir. A Figura (4.5) mostra os fatores latentes obtidos para a dimensão semântica, que corresponde às condições de apresentação: autêntica ou falsificada. Cada ponto do gráfico representa uma condição, projetada sobre o espaço definido pelas duas primeiras componentes latentes extraídas da decomposição. Já a Figura (4.6) exibe a visualização das cinco primeiras componentes associadas à dimensão espacial (referente aos pixels), comumente referidas como *eigenfaces*. Tais componentes evidenciam padrões visuais recorrentes que contribuem para a modelagem das faces, destacando variações relevantes entre as amostras.

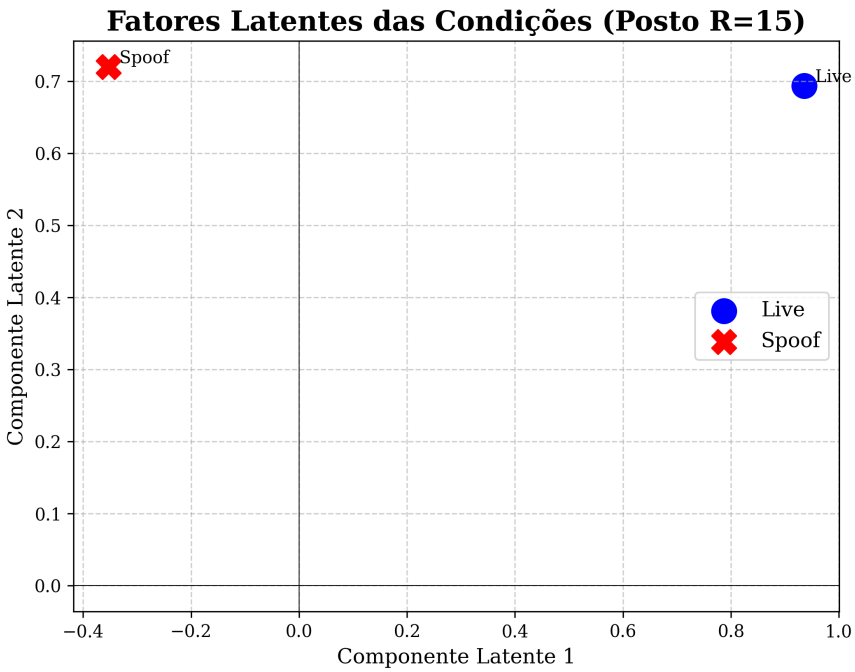

Figura 4.5: Fatores latentes obtidos para a dimensão das condições de apresentação (autêntica e falsificada), a partir da decomposição CPD com posto $R = 15$. Cada ponto representa uma condição projetada no espaço das duas primeiras componentes latentes.

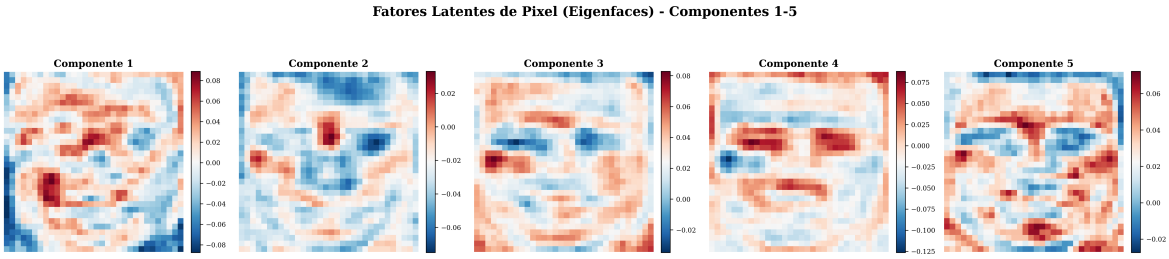

Figura 4.6: Visualização das cinco primeiras componentes latentes associadas à dimensão dos pixels (modo espacial), também conhecidas como *eigenfaces*. As variações observadas em cada componente refletem padrões relevantes extraídos da decomposição tensorial.

4.3.2 Redução de Dimensionalidade

Após a decomposição tensorial, obtêm-se representações latentes de alta dimensionalidade associadas aos indivíduos presentes no conjunto de dados. Para permitir a visualização dessas representações e viabilizar sua posterior agrupamento, aplica-se uma etapa de redução de dimensionalidade.

Neste trabalho, emprega-se o método *t-distributed Stochastic Neighbor Embedding* (t-SNE), descrito na Seção (2.7.1), para a etapa de redução de dimensionalidade. Essa técnica é eficaz na preservação das estruturas locais do espaço original dos dados, facilitando a visualização de agrupamentos latentes. Ao projetar as representações fatoradas em um espaço bidimensional, o t-SNE permite observar de forma clara a separabilidade entre amostras autênticas e falsificadas, destacando padrões discriminativos extraídos pela

decomposição tensorial.

4.3.3 Clusterização

Com as representações latentes projetadas para um espaço bidimensional pelo t-SNE, aplica-se o algoritmo de clusterização k-means detalhado na Seção (2.7.2), com o objetivo de agrupar as amostras de maneira não supervisionada.

O k-means é um algoritmo baseado em centroide que busca particionar os dados em k grupos, minimizando a variância intra-cluster. Neste contexto, considera-se $k = 2$, correspondente às duas possíveis condições de apresentação: autêntica (*live*) e falsificada (*spoof*). Essa etapa visa avaliar a capacidade das representações fatoradas e reduzidas em separar as classes, mesmo sem o uso explícito de rótulos. Quando bem-sucedido, o agrupamento das amostras em regiões distintas no espaço latente evidencia que as características extraídas são discriminativas e informativas para o problema de detecção de falsificação facial.

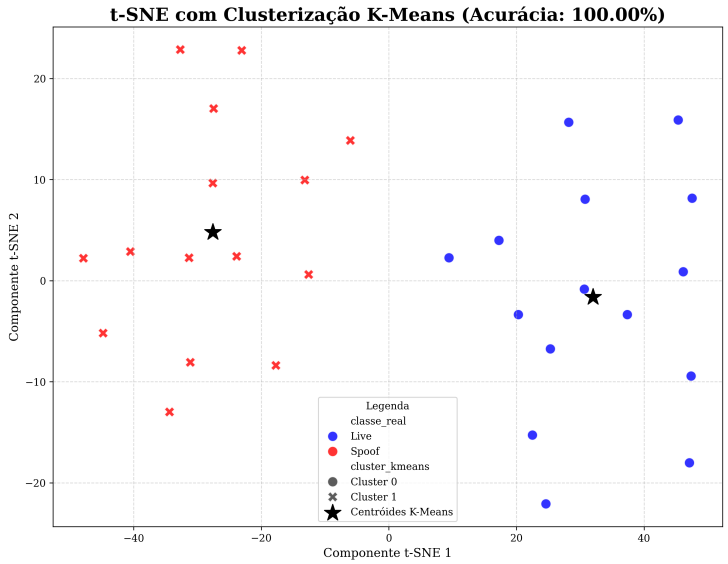

Figura 4.7: Visualização do agrupamento K-Means sobre os dados projetados via t-SNE. Os pontos azuis representam amostras autênticas (live) e os vermelhos correspondem a falsificações (spoof). As estrelas pretas indicam os centróides estimados pelo algoritmo K-Means. A separação clara entre os dois grupos demonstra a eficácia das representações fatoradas e reduzidas em discriminar entre as classes de forma não supervisionada.

Neste exemplo ilustrativo, observa-se uma acurácia de 100% na separação entre amostras autênticas e falsificadas. Tal resultado decorre do fato de que a decomposição foi realizada sobre o próprio conjunto de treinamento, com posto $R = 15$, igual ao número de indivíduos presentes no tensor. Isso configura uma situação de solução fechada, onde o modelo tem acesso total à variabilidade do conjunto. Esse experimento tem por finalidade apenas demonstrar o funcionamento do pipeline proposto, desde a decomposição até a clusterização. A avaliação formal de desempenho, com métricas adequadas e separação entre treino e teste, será conduzida na Seção (5).

5 AVALIAÇÃO E DESEMPENHO

Este capítulo apresenta uma análise detalhada da implementação experimental do método proposto. São descritos os parâmetros experimentais, as etapas de execução do pipeline, os resultados obtidos e as métricas de desempenho alcançadas.

5.1 CONFIGURAÇÃO EXPERIMENTAL

A configuração experimental foi projetada para avaliar a eficácia da decomposição tensorial CPD (CANDECOMP/PARAFAC) na detecção de tentativas de falsificação.

5.1.1 Parâmetros do Experimento

Os parâmetros do experimento foram definidos considerando o equilíbrio entre precisão dos resultados e viabilidade computacional. A Tabela (5.1) apresenta um resumo detalhado de todos os parâmetros utilizados.

Parâmetro	Valor	Justificativa
Tamanho das Imagens	(32×32) pixels	Resolução reduzida para otimizar processamento, mantendo características faciais essenciais e reduzindo significativamente o espaço dimensional.
Número Máximo de Sujeitos	250	Subconjunto representativo do dataset, balanceando diversidade com recursos computacionais.
Posto do Tensor	25	Dimensionalidade do espaço latente.
Tamanho do Lote de Análise	500 amostras	Processamento eficiente da análise de clustering, com otimização específica para algoritmos t-SNE e K-Means.

Tabela 5.1: Parâmetros Experimentais e suas Justificativas.

5.1.2 Configuração do Hardware e Software

O experimento foi executado em ambiente computacional otimizado para processamento tensorial e análise de imagens. A Tabela (5.2) detalha as especificações e configurações utilizadas.

Componente	Especificação	Função no Experimento
Processamento	GPU/CPU (CUDA quando disponível)	Aceleração de operações tensoriais e processamento de imagens através do framework PyTorch.
Detecção Facial	MTCNN (Multi-task CNN) [74]	Localização automática e extração de regiões faciais com alta precisão e robustez.
Framework Tensorial	TensorLy com backend NumPy	Implementação otimizada da decomposição CPD (PARAFAC) com suporte a operações de grande escala.
Gerenciamento de Memória	Garbage Collection + CUDA Cache	Otimização automática de recursos com limpeza periódica de variáveis temporárias e cache GPU.
Visualização	Matplotlib + Seaborn + Plotly	Geração de gráficos científicos de alta qualidade para análise dos resultados.

Tabela 5.2: Configuração de Hardware e Software do Ambiente Experimental.

5.2 METODOLOGIA EXPERIMENTAL

Esta seção detalha a modelagem dos dados em uma estrutura tensorial e as estratégias de otimização de memória que viabilizaram o processamento.

5.2.1 Estrutura do Tensor

- **Primeira Dimensão (1024):** Pixels das imagens $(32 \times 32 = 1024)$.
- **Segunda Dimensão (2):** Condições de imagem. Autêntica ou falsa.
- **Terceira Dimensão (250):** Total de sujeitos coletados.

5.2.2 Otimizações de Memória

Para gerenciar eficientemente a memória durante a construção do tensor, foram implementadas as seguintes estratégias:

- **Processamento em Lotes:** Divisão dos sujeitos em grupos menores (quatro indivúduos).
- **Garbage Collection:** Limpeza automática de variáveis temporárias.
- **Cache Intermediário:** Armazenamento temporário em disco quando necessário.
- **Monitoramento de Memória:** Acompanhmento contínuo do uso de RAM.

5.2.3 Decomposição Tensorial

A segunda etapa realiza a decomposição do tensor construído, conforme visto na Seção (4.3.1), extraindo os fatores latentes que representam as características discriminativas.

5.2.3.1 Configuração da Decomposição PARAFAC

A decomposição foi configurada com parâmetros otimizados para garantir convergência estável. A Tabela (5.3) apresenta os parâmetros utilizados na decomposição tensorial. Este processo produz três matrizes de fatores com características específicas e funções distintas no processo de detecção. A Tabela (5.4) detalha as propriedades de cada fator.

Parâmetro PARAFAC	**Valor**	**Impacto na Decomposição**
Número Máximo de Iterações	1000	Permite convergência adequada mesmo para tensores complexos, evitando parada prematura do algoritmo.
Tolerância de Convergência	1×10^{-6}	Critério rigoroso de parada que garante alta precisão na decomposição dos fatores latentes.
Inicialização	Aleatória	Evita mínimos locais e garante exploração adequada do espaço de soluções durante a otimização.
Normalização de Fatores	Ativada	Previne instabilidades numéricas e facilita a interpretação dos componentes resultantes.
Posto do Tensor	25	Dimensionalidade do espaço latente balanceada entre capacidade representativa e eficiência computacional.

Tabela 5.3: Configuração dos Parâmetros da Decomposição PARAFAC.

Fator	**Dimensões**	**Interpretação**	**Função na Detecção**
Fator A	$\mathbb{R}^{1024\times25}$	Padrões espaciais das imagens (eigenfaces) que capturam variações principais na estrutura facial.	Representa características visuais discriminativas entre diferentes tipos de apresentação facial.
Fator B	$\mathbb{R}^{2\times25}$	Características latentes das condições no espaço reduzido.	Base fundamental para separação entre faces reais e falsificadas no espaço latente.
Fator C	$\mathbb{R}^{250\times25}$	Características individuais dos sujeitos.	Permite análise de variabilidade inter-sujeito.

Tabela 5.4: Características e Funções dos Fatores Resultantes da Decomposição CPD.

5.2.3.2 Configuração da Decomposição HOSVD

A decomposição HOSVD foi configurada para projetar os dados em subespaços ortogonais, reduzindo redundâncias entre os modos do tensor. A Tabela (5.5) apresenta os parâmetros utilizados. O processo gera um tensor núcleo e matrizes fator ortogonais para cada modo, permitindo representar a variabilidade dos dados de forma compacta e estruturada.

Parâmetro HOSVD	Valor	Impacto na Decomposição
Posto de Truncamento por Modo	$[25, 25, 25]$	Define a dimensionalidade de cada modo, garantindo representação compacta sem perda excessiva de informação.

Tabela 5.5: Configuração dos Parâmetros da Decomposição HOSVD.

5.2.3.3 Configuração da Decomposição HOSVD + HOOI

Após a decomposição inicial via HOSVD, aplica-se o método HOOI para refinar iterativamente as matrizes fator e o tensor núcleo. O objetivo é minimizar o erro de reconstrução e melhorar a separação entre amostras autênticas e tentativas de falsificação. A Tabela (5.6) apresenta os parâmetros de configuração utilizados nesta etapa.

Parâmetro HOOI	Valor	Impacto na Decomposição
Número Máximo de Iterações	200	Permite convergência estável sem sobrecarga computacional excessiva.
Tolerância de Convergência	1×10^{-6}	Garante refinamento preciso das matrizes fator e do tensor núcleo.
Atualização Alternada de Modos	Ativada	Otimiza cada modo separadamente, acelerando a convergência para solução de menor erro.
Critério de Parada	Estabilização do erro de reconstrução	Evita iterações desnecessárias quando não há ganho significativo de desempenho.

Tabela 5.6: Configuração dos Parâmetros da Decomposição HOSVD + HOOI.

5.2.4 Clusterização

A etapa final avalia a capacidade discriminativa dos fatores através de técnicas de clusterização não supervisionada.

5.2.4.1 Redução Dimensional com t-SNE

O t-SNE foi aplicado com configurações otimizadas para preservar a estrutura local dos dados no espaço bidimensional. A Tabela (5.7) apresenta os parâmetros utilizados.

Parâmetro t-SNE	Valor	Justificativa
Perplexidade	25	Valor padrão escolhido por proporcionar um bom equilíbrio entre a fidelidade à estrutura local dos dados e a visualização da sua organização global.
Número de Iterações	1000	Garante convergência adequada do algoritmo de otimização t-SNE mesmo para datasets complexos.
Semente Aleatória	42	Reprodutibilidade dos resultados entre execuções independentes do experimento.

Tabela 5.7: Configuração dos Parâmetros do Algoritmo t-SNE.

5.3 ANÁLISE DE RESULTADOS

Para quantificar o desempenho da clusterização, foram utilizadas métricas de avaliação detalhadas na Tabela (5.8). A análise se concentrou na acurácia do agrupamento e na matriz de confusão.

Métrica	Fórmula	Interpretação
Acurácia	$\frac{TP+TN}{TP+TN+FP+FN}$	Proporção de amostras corretamente classificadas. Varia entre 0 e 1, sendo 1 a classificação perfeita.
Matriz de Confusão	$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$	Visualização detalhada dos erros de classificação, mostrando verdadeiros/falsos positivos e negativos.

Tabela 5.8: Métricas de Avaliação Utilizadas no Experimento.

Em que:

- **TP (*True Positive*):** número de amostras que pertencem a uma determinada classe e foram corretamente atribuídas a ela pelo modelo.
- **TN (*True Negative*):** número de amostras que não pertencem à classe em questão e foram corretamente excluídas pelo modelo.
- **FP (*False Positive*):** número de amostras que não pertencem à classe, mas foram incorretamente atribuídas a ela.
- **FN (*False Negative*):** número de amostras que pertencem à classe, mas foram incorretamente atribuídas a outra.

5.3.1 Agrupamento e Validação com K-Means

Após a redução de dimensionalidade com t-SNE, o algoritmo K-Means foi aplicado para agrupar os dados em dois clusters (k=2), com o objetivo de separar as amostras autênticas das falsificadas. Vale res-

saltar que este procedimento foi executado sobre o conjunto de testes, a partir das características geradas pela fatoração do tensor. O resultado desse processo é visualizado na Figura (5.1) em que é demonstrada uma clara separação entre os dois grupos. O Cluster 0 (pontos azuis) concentra a grande maioria das amostras da classe *Live* (círculos azuis), enquanto o Cluster 1 (cruzes vermelhas) agrupa predominantemente as amostras *Spoof* (círculos vermelhos). As estrelas pretas indicam os centróides calculados para cada um dos clusters. A alta sobreposição entre os clusters e as classes reais já sugere visualmente o bom desempenho do método.

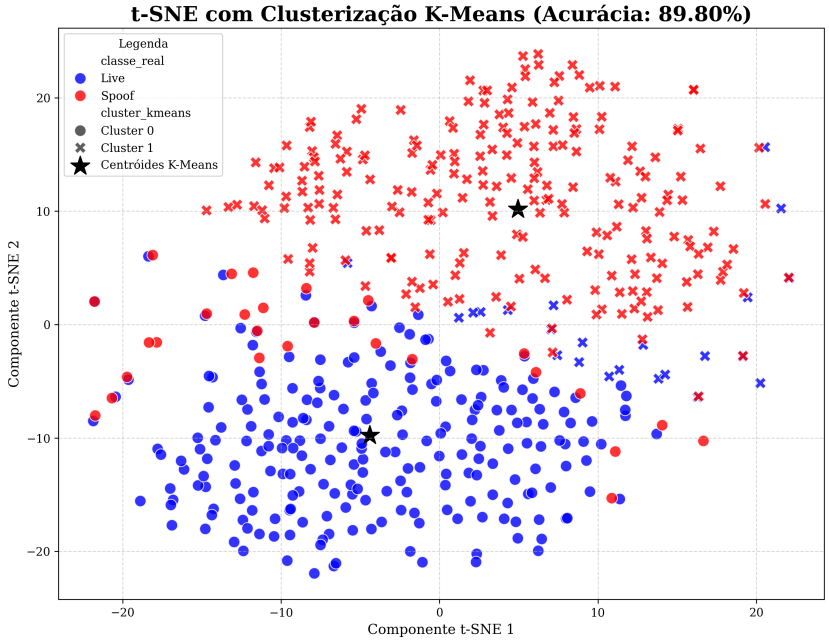

Figura 5.1: Resultado da visualização t-SNE com clustering K-Means: pontos azuis (live), pontos vermelhos (spoof), estrelas pretas (centróides).

5.3.2 Análise da Matriz de Confusão

A matriz de confusão, ilustrada na Figura (5.2), quantifica a correspondência entre os clusters gerados pelo algoritmo K-Means e as classes reais. Nesse contexto, o *Cluster 0* corresponde à classe *Previsto Live*, enquanto o *Cluster 1* representa a classe *Previsto Spoof*, permitindo a avaliação dos acertos e erros de agrupamento em relação às categorias originais.

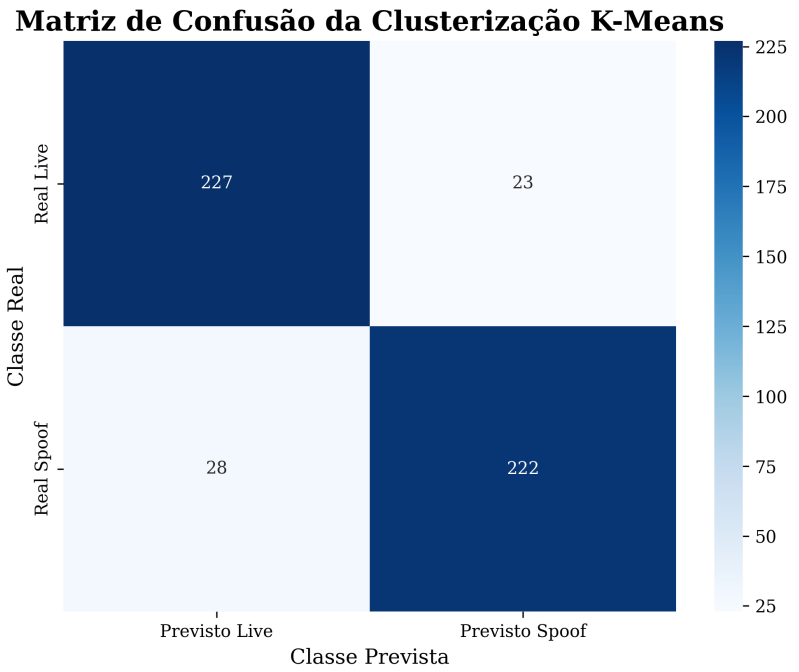

Figura 5.2: Matriz de confusão da clusterização K-Means, comparando a classe real das amostras com a classe prevista pelo agrupamento.

A partir destes valores, a acurácia global do modelo é calculada como $\frac{227+222}{227+222+23+28} = \frac{449}{500} = 89.80\%$, confirmando o valor exibido na Figura (5.1).

5.3.3 Comparativo de Desempenho entre Métodos

Esta seção tem como objetivo comparar o desempenho das diferentes técnicas de decomposição tensorial apresentadas na Seção (2.6.3). A Figura (5.3) apresenta a acurácia obtida por cada uma das abordagens. Observa-se que a decomposição HOSVD supera a CPD, refletindo sua maior capacidade de projetar os dados em subespaços ortogonais e, assim, reduzir redundâncias entre os modos. Além disso, a combinação HOSVD+HOOI atinge o melhor resultado entre os métodos comparados, evidenciando o ganho de desempenho proporcionado pela etapa de refinamento iterativo das matrizes fator e do tensor núcleo, que reduz o erro de reconstrução. Esse refinamento adicional confere maior robustez às variações de iluminação e pose, resultando em separação mais consistente entre amostras autênticas e tentativas de falsificação.

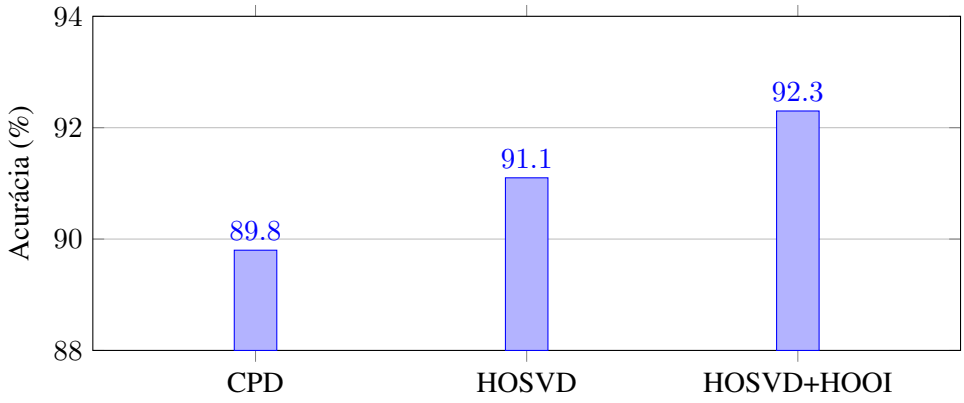

Figura 5.3: Acurácia obtida por cada técnica de decomposição tensorial.

5.3.4 Sensibilidade do t-SNE para Visualização e Agrupamento em 2D

A influência do parâmetro de perplexidade no desempenho do pipeline t-SNE seguido do método *k*-means é avaliada. A Figura (5.4) mostra a variação da acurácia em função da perplexidade, considerando como entrada os fatores obtidos pela decomposição HOSVD+HOOI. Nota-se que valores muito baixos de perplexidade tendem a superenfatizar vizinhanças locais, fragmentando os agrupamentos, enquanto valores excessivamente altos borram a separação entre classes. De forma geral, a faixa entre 25–35 proporciona a melhor separação visual entre entre amostras autênticas e tentativas de falsificação, resultando no melhor valor de acurácia observado em torno de 30.

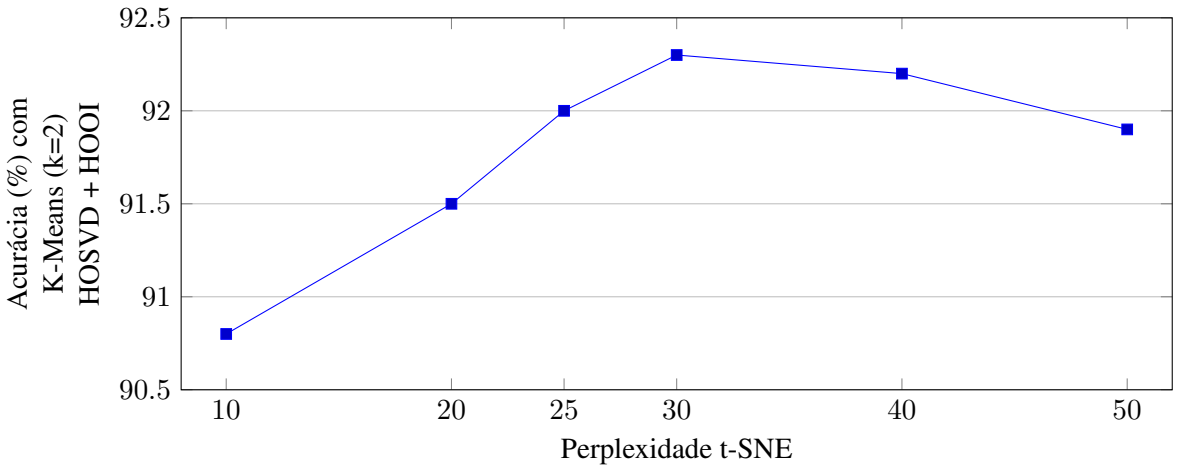

Figura 5.4: Variação da acurácia em função da perplexidade na projeção 2D (t-SNE) seguida de *k*-means. Pico observado em perplexidade ≈ 30.

5.4 CONCLUSÕES DA AVALIAÇÃO

A avaliação do pipeline demonstrou que a combinação entre extração de características por decomposição tensorial, redução de dimensionalidade com t-SNE e clusterização por meio do K-Means constitui uma abordagem eficaz para detecção de fraude facial. O sistema atingiu uma acurácia de 92,30%, evi-

denciando uma alta capacidade de discriminar entre amostras autênticas e falsificadas. A inspeção visual apresentada na Figura (5.1), juntamente com a análise quantitativa da Figura (5.2), revela uma forte correspondência entre os clusters formados no espaço de características reduzido e as classes reais. Os erros, embora minoritários, estão distribuídos de forma equilibrada entre falsos positivos e falsos negativos, o que sugere ausência de viés significativo do modelo em relação a uma das classes. Esses resultados reforçam a viabilidade da abordagem proposta como uma solução para o problema de detecção de falsificações faciais.

6 CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho propôs um método para a detecção de tentativas de falsificação em sistemas de reconhecimento facial. A abordagem emprega uma estratégia de clusterização em duas etapas, combinando técnicas avançadas de extração de representações via decomposição tensorial com redução de dimensionalidade não linear por t-SNE. O objetivo principal foi distinguir, de forma robusta, imagens faciais autênticas e falsificadas. Os resultados apresentados no Capítulo (5) evidenciam o potencial da abordagem, destacando sua capacidade de reforçar a segurança na autenticação facial e mitigar ataques com maior eficácia.

6.1 TRABALHOS FUTUROS

Apesar dos resultados promissores, o campo de detecção de falsificações faciais continua em rápida evolução. Com isso, destacam-se as seguintes direções para continuidade e aprimoramento da pesquisa:

- **Exploração de outras decomposições tensoriais:** Investigar variantes como *Tensor Train* (TT) [75] ou *Tensor Ring* (TR) [76], que podem oferecer diferentes propriedades de esparsidade e custo computacional. A comparação com a CPD pode revelar ganhos em desempenho ou interpretabilidade.
- **Integração com redes neurais profundas:** Explorar arquiteturas híbridas que combinem decomposição tensorial com redes profundas. O método [77] mostra que representações tensoriais podem ser integradas a modelos neurais para aprendizado mais eficiente, esparso e interpretável.
- **Explorar com diferentes métodos de clusterização:** Avaliar algoritmos como DBSCAN adaptativo [78], Gaussian Mixture Models (GMM) [79] ou métodos baseados em grafos [80], buscando melhorar a separabilidade entre amostras genuínas e falsificadas em diferentes cenários.
- **Incorporação de técnicas de rastreamento de subespaços:** Incluir métodos de rastreamento de subespaços, como projeções adaptativas [81], decomposição incremental [82], e técnicas robustas a ruído [83], com o objetivo atualizar representações tensoriais com menor custo computacional.

Essas direções buscam aprimorar o desempenho do método proposto e fomentar o desenvolvimento de soluções mais robustas, interpretáveis e com maior capacidade de generalização no contexto da segurança biométrica, por meio do uso de representações tensoriais.

REFERÊNCIAS BIBLIOGRÁFICAS

1 MILANO, D. de; HONORATO, L. B. Visao computacional. *UNICAMP Universidade Estadual de Campinas FT Faculdade de Tecnologia*, 2014.

2 COSTA, R. S.; KREMER, B. Inteligência artificial e discriminação: desafios e perspectivas para a proteção de grupos vulneráveis frente às tecnologias de reconhecimento facial. *Revista Brasileira de Direitos Fundamentais & Justiça*, v. 16, n. 1, 2022.

3 SOO, S. Object detection using haar-cascade classifier. *Institute of Computer Science, University of Tartu*, v. 2, n. 3, p. 1–12, 2014.

4 DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: IEEE. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. [S.l.], 2005. v. 1, p. 886–893.

5 CHANDRA, M. A.; BEDI, S. Survey on svm and their application in image classification. *International Journal of Information Technology*, Springer, v. 13, n. 5, p. 1–11, 2021.

6 KU, H.; DONG, W. Face recognition based on mtcnn and convolutional neural network. *Frontiers in Signal Processing*, v. 4, n. 1, p. 37–42, 2020.

7 ROGERS, S. J.; DAWSON, G.; VISMARA, L. A. Autismo: compreender e agir em família. *Lisboa: Lidel*, p. 213–235, 2015.

8 RASMUSSEN, C. Generalized principal component analysis. *Neural networks*, Elsevier, v. 4, n. 6, p. 709–714, 1991.

9 KOMULAINEN, J.; HADID, A.; PIETIKÄINEN, M. Context based face anti-spoofing. In: IEEE. *2013 IEEE sixth international conference on biometrics: theory, applications and systems (BTAS)*. [S.l.], 2013. p. 1–8.

10 SCHROFF, F.; KALENICHENKO, D.; PHILBIN, J. Facenet: A unified embedding for face recognition and clustering. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 815–823.

11 QAWAQNEH, Z.; MALLOUH, A. A.; BARKANA, B. D. Deep convolutional neural network for age estimation based on vgg-face model. *arXiv preprint arXiv:1709.01664*, 2017.

12 DENG, J.; GUO, J.; XUE, N.; ZAFEIRIOU, S. Arcface: Additive angular margin loss for deep face recognition. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2019. p. 4690–4699.

13 ZHANG, Y.; YIN, Z.; LI, Y.; YIN, G.; YAN, J.; SHAO, J.; LIU, Z. Celeba-spoof: Large-scale face anti-spoofing dataset with rich annotations. In: SPRINGER. *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16*. [S.l.], 2020. p. 70–85.

14 REZA, A. M. Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. *Journal of VLSI signal processing systems for signal, image and video technology*, Springer, v. 38, p. 35–44, 2004.

15 KOLDA, T. G.; BADER, B. W. Tensor decompositions and applications. *SIAM review*, SIAM, v. 51, n. 3, p. 455–500, 2009.

16 HARSHMAN, R. A. Foundations of the parafac procedure: Models and conditions for an "explanatory" multimodal factor analysis. *UCLA Working Papers in Phonetics*, v. 16, p. 1–84, 1970.

17 WATTENBERG, M.; VIÉGAS, F.; JOHNSON, I. How to use t-sne effectively. *Distill*, v. 1, n. 10, p. e2, 2016.

18 VATS, S.; SHARMA, V.; RAWAT, P.; RATRA, A. K-means clustering over distributed environment: A review. *Uncertainty in Computational Intelligence-Based Decision Making*, Elsevier, p. 173–185, 2025.

19 ZHANG, Z.; CHEN, X.; WANG, C.; WANG, R.; SONG, W.; NIE, F. Structured multi-view k-means clustering. *Pattern Recognition*, Elsevier, v. 160, p. 111113, 2025.

20 LIU, S.; TRENKLER, G. et al. Hadamard, khatri-rao, kronecker and other matrix products. *International Journal of Information and Systems Sciences*, v. 4, n. 1, p. 160–177, 2008.

21 GARCEZ, C. C. R.; MARQUES, G. S.; CANEDO, E. D.; PRACIANO, B. G.; FILHO, F. L. C.; MENDONçA, F. L. L. Prevenção de falsificação em sistemas de reconhecimento facial: uma proposta baseada em clusterização. In: MIRANDA, P.; SANTORO, F. M.; COSTA, C. (Ed.). *21ª Conferência Ibero Americana WWW/INTERNET (CIAWI)*. [S.l.]: IADIS, 2023. p. 35–42. ISBN 978-989-8704-54-2.

22 VELARDO, C.; DUGELAY, J.-L.; DANIEL, L.; DANTCHEVA, A.; ERDOGMUS, N.; KOSE, N.; MIN, R.; ZHAO, X. Introduction to biometry. In: *Multimedia Image and Video Processing*. [S.l.]: CRC Press, 2011. p. 397–418.

23 WAYMAN, J. L. Generalized biometric identification system model. *NATIONAL BIOMETRIC TEST CENTER COLLECTED WORKS*, p. 25, 2000.

24 BAČA, M.; SCHATTEN, M.; GOLENJA, B. Modeling biometrics systems in uml. In: *IIS2007 International Conference on Intelligent and Information Systems Proceedings*. [S.l.: s.n.], 2007. v. 18, p. 23–27.

25 VOULODIMOS, A.; DOULAMIS, N.; DOULAMIS, A.; PROTOPAPADAKIS, E. Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, Wiley Online Library, v. 2018, n. 1, p. 7068349, 2018.

26 TURK, M.; PENTLAND, A. Eigenfaces for recognition. *Journal of cognitive neuroscience*, MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , v. 3, n. 1, p. 71–86, 1991.

27 BELHUMEUR, P. N.; HESPANHA, J. P.; KRIEGMAN, D. J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 19, n. 7, p. 711–720, 1997.

28 GHORBANI, M.; TARGHI, A. T.; DEHSHIBI, M. M. Hog and lbp: Towards a robust face recognition system. In: IEEE. *2015 Tenth International Conference on Digital Information Management (ICDIM)*. [S.l.], 2015. p. 138–141.

29 HUANG, G. B.; MATTAR, M.; BERG, T.; LEARNED-MILLER, E. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. In: *Workshop on faces in'Real-Life'Images: detection, alignment, and recognition*. [S.l.: s.n.], 2008.

30 KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, v. 25, 2012.

31 RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATHY, A.; KHOSLA, A.; BERNSTEIN, M. et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, Springer, v. 115, p. 211–252, 2015.

32 HO, H.-T.; NGUYEN, L. V.; LE, T. H. T.; LEE, O.-J. Face detection using eigenfaces: A comprehensive review. *IEEE Access*, IEEE, 2024.

33 WANG, H.; WANG, Y.; ZHOU, Z.; JI, X.; GONG, D.; ZHOU, J.; LI, Z.; LIU, W. Cosface: Large margin cosine loss for deep face recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 5265–5274.

34 LIU, W.; WEN, Y.; YU, Z.; LI, M.; RAJ, B.; SONG, L. Sphereface: Deep hypersphere embedding for face recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 212–220.

35 HASSAN, R.; KAUR, M.; RAJPOOT, A. K. Automated face spoofing detection using machine learning: A review. In: IEEE. *2022 9th International Conference on Computing for Sustainable Global Development (INDIACom)*. [S.l.], 2022. p. 293–297.

36 BAGGA, M.; SINGH, B. Spoofing detection in face recognition: A review. In: *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*. [S.l.: s.n.], 2016. p. 2037–2042.

37 CHAKRABORTY, S.; DAS, D. An overview of face liveness detection. *arXiv preprint arXiv:1405.2227*, 2014.

38 YU, Z.; QIN, Y.; LI, X.; ZHAO, C.; LEI, Z.; ZHAO, G. Deep learning for face anti-spoofing: A survey. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 45, n. 5, p. 5609–5631, 2022.

39 ANJOS, A.; MARCEL, S. Counter-measures to photo attacks in face recognition: a public database and a baseline. In: IEEE. *2011 international joint conference on Biometrics (IJCB)*. [S.l.], 2011. p. 1–7.

40 MÄÄTTÄ, J.; HADID, A.; PIETIKÄINEN, M. Face spoofing detection from single images using micro-texture analysis. In: IEEE. *2011 international joint conference on Biometrics (IJCB)*. [S.l.], 2011. p. 1–7.

41 PAN, G.; SUN, L.; WU, Z.; LAO, S. Eyeblink-based anti-spoofing in face recognition from a generic webcamera. In: IEEE. *2007 IEEE 11th international conference on computer vision*. [S.l.], 2007. p. 1–8.

42 KOLLREIDER, K.; FRONTHALER, H.; FARAJ, M. I.; BIGUN, J. Real-time face detection and motion analysis with application in "liveness" assessment. *IEEE Transactions on Information Forensics and Security*, IEEE, v. 2, n. 3, p. 548–558, 2007.

43 YANG, J.; LEI, Z.; LI, S. Z. Learn convolutional neural network for face anti-spoofing. *arXiv preprint arXiv:1408.5601*, 2014.

44 SEBASTIEN, M.; NIXON, M.; LI, S. Handbook of biometric anti-spoofing: trusted biometrics under spoofing attacks. Springer, 2014.

45 MOHAMMADI, A.; BHATTACHARJEE, S.; MARCEL, S. Deeply vulnerable: a study of the robustness of face recognition to presentation attacks. *Iet Biometrics*, Wiley Online Library, v. 7, n. 1, p. 15–26, 2018.

46 GALBALLY, J.; MARCEL, S.; FIERREZ, J. Biometric antispoofing methods: A survey in face recognition. *Ieee Access*, IEEE, v. 2, p. 1530–1552, 2014.

47 ZHANG, C.; BENGIO, S.; HARDT, M.; RECHT, B.; VINYALS, O. Understanding deep learning requires rethinking generalization. *arXiv preprint arXiv:1611.03530*, 2016.

48 VASILESCU, M. A. O.; TERZOPOULOS, D. Multilinear analysis of image ensembles: Tensorfaces. In: SPRINGER. *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part I 7*. [S.l.], 2002. p. 447–460.

49 LU, H.; PLATANIOTIS, K. N.; VENETSANOPOULOS, A. N. Mpca: Multilinear principal component analysis of tensor objects. *IEEE transactions on Neural Networks*, IEEE, v. 19, n. 1, p. 18–39, 2008.

50 NICKEL, M.; TRESP, V. Tensor factorization for multi-relational learning. In: SPRINGER. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. [S.l.], 2013. p. 617–621.

51 HITCHCOCK, F. L. Representation of data by means of three-way arrays. *Journal of Mathematics and Physics*, MIT Press, v. 43, n. 1-4, p. 32–43, 1964.

52 CARROLL, J. D.; CHANG, J.-J. Analysis of three-way data by means of three-mode principal components analysis. *Psychometrika*, Springer, v. 35, n. 3, p. 283–319, 1970.

53 BHATT, V.; KUMAR, S.; SAINI, S. Tucker decomposition and applications. *Materials Today: Proceedings*, Elsevier, v. 46, p. 10787–10792, 2021.

54 LATHAUWER, L. D.; MOOR, B. D.; VANDEWALLE, J. On the best rank-1 and rank-(r 1, r 2,..., rn) approximation of higher-order tensors. *SIAM journal on Matrix Analysis and Applications*, SIAM, v. 21, n. 4, p. 1324–1342, 2000.

55 MAATEN, L. Van der; HINTON, G. Visualizing data using t-sne. *Journal of machine learning research*, v. 9, n. Nov, p. 2579–2605, 2008.

56 TUCKER, L. R. Some mathematical notes on three-mode factor analysis. *Psychometrika*, Springer, v. 31, n. 3, p. 279–311, 1966.

57 BOUSSÉ, M.; VERVLIET, N.; DEBALS, O.; LATHAUWER, L. D. Face recognition as a kronecker product equation. In: IEEE. *2017 IEEE 7th International Conference on Biometrics (ICB)*. [S.l.], 2017. p. 238–243.

58 ZHOU, Y.; CHEN, J.; CHEN, X.; SU, M.; GU, J. Tensor decomposition for action recognition in surveillance videos. *Journal of Electronic Science and Technology*, v. 11, n. 4, p. 357–362, 2013.

59 ZHOU, Y.; CHEN, J.; CHEN, X.; SU, M.; GU, J. Tensor decomposition for action recognition in surveillance videos. *Journal of Electronic Science and Technology*, v. 11, n. 4, p. 357–362, 2013.

60 MÄÄTTÄ, J.; HADID, A.; PIETIKÄINEN, M. Face anti-spoofing using directional local binary patterns. *2012 11th IEEE International Conference on Automatic Face Gesture Recognition (FG)*, IEEE, p. 577–582, 2012.

61 LI, J.; MA, X.; LIU, J.; LI, Y.; LI, X. Generalized face anti-spoofing by detecting pulse from face videos. *IEEE transactions on cybernetics*, IEEE, v. 44, n. 5, p. 592–602, 2014.

62 PAN, L.; ZHANG, J.; ZHANG, X. Biometric face anti-spoofing: A review. In: IEEE. *2019 Chinese Automation Congress (CAC)*. [S.l.], 2019. p. 4537–4542.

63 WANG, P.; CHEN, Z.; MA, L.; WANG, X.; ZHANG, J. Face anti-spoofing with rgb-d cameras. In: IEEE. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.], 2015. p. 1306–1310.

64 ZHANG, S.; WANG, Y.; WANG, P.; FU, M.; WANG, J.; WU, Z. Deep learning for face anti-spoofing: A review. *IEEE Access*, IEEE, v. 9, p. 24838–24867, 2021.

65 SHAO, J.; ZHENG, Y.; ZHANG, P.; WU, Z. Deep learning for face anti-spoofing: A review. *Artificial Intelligence Review*, Springer, v. 50, p. 1–26, 2017.

66 YANG, J.; LIU, Y.; SONG, W.; ZHANG, J.; ZHENG, Y.; YANG, K.; WANG, X.; LIU, H.; ZHANG, P.; SONG, J. et al. Face anti-spoofing with learning on the high-frequency features. In: IEEE. *2017 IEEE International Joint Conference on Biometrics (IJCB)*. [S.l.], 2017. p. 130–139.

67 KIM, J.; SEO, B.; LEE, H. S.; CHOI, H. S.; KIM, J. C. Feature pyramid network for face anti-spoofing. In: IEEE. *2019 IEEE International Conference on Image Processing (ICIP)*. [S.l.], 2019. p. 4169–4173.

68 WANG, P.; CHEN, G.; WANG, X.; ZHANG, J. Deep texture and structure anti-spoofing for face recognition. In: IEEE. *2018 24th International Conference on Pattern Recognition (ICPR)*. [S.l.], 2018. p. 2501–2506.

69 ATOUM, Y.; LIU, Z.; LI, X.; JOURNET, T.; LI, X.; YANG, Y.; REN, X. Face anti-spoofing using deep multi-spectral networks. In: *European Conference on Computer Vision (ECCV) Workshops*. [S.l.: s.n.], 2018. p. 0–0.

70 JOURNET, T. D. H.; LE, T. D. H.; DUGELAY, J.-L. Learning a generalizable face presentation attack detector through multi-task learning. In: IEEE. *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. [S.l.], 2018. p. 1–8.

71 YU, Z.; YANG, K.; LYU, S.; TIAN, Y. Searching for efficient face anti-spoofing models. In: IEEE. *2020 IEEE International Joint Conference on Biometrics (IJCB)*. [S.l.], 2020. p. 1–8.

72 GARG, H.; BHAMBRI, A.; KUMAR, R.; YADAV, V. K. Unsupervised domain adaptation for face anti-spoofing. In: IEEE. *2019 International Conference on Biometrics (ICB)*. [S.l.], 2019. p. 1–6.

73 WEN, S.; LI, Y.; CAI, Y.; JIANG, Z.; JIANG, B.; LI, S.; JAIN, A. K. Meta-learning for face anti-spoofing. In: IEEE. *2020 IEEE International Joint Conference on Biometrics (IJCB)*. [S.l.], 2020. p. 1–8.

74 XIANG, J.; ZHU, G. Joint face detection and facial expression recognition with mtcnn. In: IEEE. *2017 4th international conference on information science and control engineering (ICISCE)*. [S.l.], 2017. p. 424–427.

75 OSELEDETS, I. V. Tensor-train decomposition. *SIAM Journal on Scientific Computing*, SIAM, v. 33, n. 5, p. 2295–2317, 2011.

76 ZHAO, Q.; ZHOU, G.; XIE, S.; ZHANG, L.; CICHOCKI, A. Tensor ring decomposition. *arXiv preprint arXiv:1606.05535*, 2016.

77 KOSSAIFI, J.; LIPTON, Z. C.; KOLBEINSSON, A.; KHANNA, A.; FURLANELLO, T.; ANANDKUMAR, A. Tensor regression networks. *Journal of Machine Learning Research*, v. 21, n. 123, p. 1–21, 2020.

78 ESTER, M.; KRIEGEL, H.-P.; SANDER, J.; XU, X. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: *kdd*. [S.l.: s.n.], 1996. v. 96, n. 34, p. 226–231.

79 REYNOLDS, D. Gaussian mixture models. In: *Encyclopedia of biometrics*. [S.l.]: Springer, 2015. p. 827–832.

80 LUXBURG, U. V. A tutorial on spectral clustering. *Statistics and computing*, Springer, v. 17, p. 395–416, 2007.

81 YANG, B. Projection approximation subspace tracking. *IEEE Transactions on Signal processing*, IEEE, v. 43, n. 1, p. 95–107, 1995.

82 BALZANO, L.; NOWAK, R.; RECHT, B. Online identification and tracking of subspaces from highly incomplete information. In: IEEE. *2010 48th Annual allerton conference on communication, control, and computing (Allerton)*. [S.l.], 2010. p. 704–711.

83 LI, Y. On incremental and robust subspace learning. *Pattern recognition*, Elsevier, v. 37, n. 7, p. 1509–1518, 2004.

84 ROEMER, F. *Advanced algebraic concepts for efficient multi-channel signal processing*. Tese (Doutorado) — Ilmenau, Technische Universität Ilmenau, Diss., 2012, 2013.

85 NION, D.; SIDIROPOULOS, N. D. Adaptive algorithms to track the parafac decomposition of a third-order tensor. *IEEE Transactions on Signal Processing*, IEEE, v. 57, n. 6, p. 2299–2310, 2009.

APÊNDICES

I.1 NOTAÇÃO

Com o objetivo de simplificar a distinção entre escalares, vetores, matrizes e tensores de ordem superior, este trabalho adota a seguinte notação:

- Escalares são indicados por letras minúsculas:

$$a, b, c, \ldots, \alpha, \beta, \ldots \in \mathbb{C}^{1\times 1}$$

- Vetores de dimensão R são representados por letras minúsculas em negrito:

$$\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots, \alpha, \beta, \ldots \in \mathbb{C}^{R\times 1}$$

- Matrizes $I \times J$ são representadas por letras maiúsculas em negrito:

$$\mathbf{A}, \mathbf{B}, \ldots, \mathbf{\Sigma}, \mathbf{\Gamma}, \ldots \in \mathbb{C}^{I\times J}$$

  O elemento da matriz $\mathbf{A}$ com índice de linha i e índice de coluna j, ou seja, $\mathbf{A}ij$, é simbolizado por a_{ij}.

- Para um dado $Q < N$, dada uma matriz $\mathbf{A} \in \mathbb{C}^{M\times N}$, tal que:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{\mathrm{Q}} & \cdots & \mathbf{a}_n & \cdots & \mathbf{a}_{\mathrm{N}} \end{bmatrix} \in \mathbb{C}^{M\times N}$$

- Tensores de ordem superior são representados por letras caligráficas:

$$\mathcal{A}, \mathcal{B}, \ldots \in \mathbb{C}^{I\times J\times K\times \ldots}$$

I.2 OPERAÇÕES COM MATRIZES

Nesta seção, são apresentadas as operações com matrizes empregadas na concepção do modelo de dados derivado no Capítulo (2).

I.2.1 O Produto de Kronecker

O produto de Kronecker de duas matrizes $\mathbf{A} \in \mathbb{C}^{I\times J}$ e $\mathbf{B} \in \mathbb{C}^{K\times L}$ é uma matriz de dimensões $IK \times JL$. O produto é definido como:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \cdots & a_{1,J}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \cdots & a_{2,J}\mathbf{B} \\ \vdots & \vdots & \vdots & \vdots \\ a_{I,1}\mathbf{B} & a_{I,2}\mathbf{B} & \ddots & a_{I,J}\mathbf{B} \end{bmatrix} \in \mathbb{C}^{IK\times JL} \tag{1}$$

I.2.2 O Produto de Khatri-Rao

O produto de Khatri-Rao de duas matrizes $\mathbf{A} \in \mathbb{C}^{I\times J}$ e $\mathbf{B} \in \mathbb{C}^{K\times L}$ é uma matriz de dimensões $IK \times L$. O produto é definido como:

$$\mathbf{A} \diamond \mathbf{B} = \begin{bmatrix} a_{1,1}\mathbf{b}_1 & a_{1,2}\mathbf{b}_2 & \cdots & a_{1,J}\mathbf{b}_L \\ a_{2,1}\mathbf{b}_1 & a_{2,2}\mathbf{b}_2 & \cdots & a_{1,J}\mathbf{b}_L \\ \vdots & \vdots & \vdots & \vdots \\ a_{I,1}\mathbf{b}_1 & a_{I,2}\mathbf{b}_2 & \cdots & a_{I,J}\mathbf{b}_L \end{bmatrix} \tag{2}$$

$$= \begin{bmatrix} \mathbf{a}_1 \otimes \mathbf{b}_1 & \mathbf{a}_2 \otimes \mathbf{b}_2 & \cdots & \mathbf{a}_L \otimes \mathbf{b}_L \end{bmatrix} \in \mathbb{C}^{IK\times L} \tag{3}$$

I.2.3 Produto Externo

O produto externo de dois vetores $\mathbf{a} \in \mathbb{C}^{I\times 1}$ e $\mathbf{b} \in \mathbb{C}^{J\times 1}$ resulta em uma matriz de dimensões $I \times J$. O produto é definido como:

$$\mathbf{a} \circ \mathbf{b} = \mathbf{a}\mathbf{b}^{\mathrm{T}} = \begin{bmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_J \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_J \\ \vdots & \vdots & \ddots & \vdots \\ a_I b_1 & a_I b_2 & \cdots & a_I b_J \end{bmatrix} \in \mathbb{C}^{I\times J} \tag{4}$$

Note que o produto externo de três vetores resulta em um tensor, tal que:

$$\mathcal{D} = \mathbf{a} \circ \mathbf{b} \circ \mathbf{c} \in \mathbb{C}^{I\times J\times K} \tag{5}$$

Onde $\mathbf{c} \in \mathbb{C}^{K\times 1}$.

I.2.4 O Operador $\mathrm{vec}\{\ \}$

Denotando $\mathbf{a}_i$ como a i-ésima coluna de uma matriz $\mathbf{A} \in \mathbb{C}^{I\times J}$. O operador vec empilha as colunas de $\mathbf{A}$ em um único vetor, resultando em:

$$\mathrm{vec}\{\mathbf{A}\} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_J \end{bmatrix} \in \mathbb{C}^{IJ\times 1} \tag{6}$$

Outra propriedade importante do operador vec é para $\mathbf{X} = \mathbf{ABC}$, onde $\mathbf{A} \in \mathbb{C}^{I\times J}$, $\mathbf{B} \in \mathbb{C}^{I\times I}$ é uma matriz diagonal e $\mathbf{C} \in \mathbb{C}^{I\times K}$.

$$\mathrm{vec}\{\mathbf{ABC}\} = \left(\mathbf{C}^{\mathrm{T}} \diamond \mathbf{A}\right) \mathrm{diag}\{\mathbf{B}\} \in \mathbb{C}^{IK} \tag{7}$$

I.2.5 O Operador $\mathrm{unvec}\{\ \}$

O operador unvec concatena uma quantidade escolhida de colunas de um vetor $\mathbf{a} = \begin{bmatrix}\mathbf{a}_1, \cdots, \mathbf{a}_J\end{bmatrix} \in \mathbb{C}^{IJ\times 1}$ formando uma matriz $\mathbf{A} \in \mathbb{C}^{I\times J}$. Onde $\mathbf{a}_i \in \mathbb{C}^{1\times J}$ denota um grupo de J elementos no vetor $\mathbf{a}$.

O resultado do operador unvec é dado por:

$$\mathrm{unvec}\{\mathbf{a}\} = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_3\end{bmatrix} \in \mathbb{C}^{I\times K} \tag{8}$$

I.2.6 A Decomposição em Valores Singulares

Toda matriz $\mathbf{M}$ cujas entradas são do domínio real ou complexo, admite uma fatoração chamada: *Decomposição em Valores Singulares (SVD)* de tal forma que para $\mathbf{M} \in \mathbb{C}^{M\times N}$:

$$\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{H}} \tag{9}$$

Onde $\mathbf{U} \in \mathbb{C}^{M\times M}$ é uma matriz unitária, $\mathbf{\Sigma} \in \mathbb{C}^{N\times N}$ é uma matriz diagonal e $\mathbf{V} \in \mathbb{C}^{N\times N}$ é uma matriz unitária.

I.2.7 A Fatoração de Khatri-Rao

Dado o produto de Khatri-Rao: $(\mathbf{A} \diamond \mathbf{B}) = \begin{bmatrix}\mathbf{a}_1 \otimes \mathbf{b}_1 & \cdots & \mathbf{a}_{\mathrm{R}} \otimes \mathbf{b}_{\mathrm{R}}\end{bmatrix} \in \mathbb{C}^{IJ\times R}$ entre as matrizes $\mathbf{A} \in \mathbb{C}^{I\times R}$ $\mathbf{B} \in \mathbb{C}^{J\times R}$, cada elemento $\mathbf{a}_{\mathrm{r}} \otimes \mathbf{b}_{\mathrm{r}}$ pode ser rearranjado de tal forma que:

$$\underset{J\times I}{\mathrm{unvec}}\{\mathbf{a}_{\mathrm{r}} \otimes \mathbf{b}_{\mathrm{r}}\} = \begin{bmatrix} a_{1,\mathrm{r}}b_{1,\mathrm{r}} & \cdots & a_{\mathrm{I},\mathrm{r}}b_{1,\mathrm{r}} \\ \vdots & \vdots & \vdots \\ a_{1,\mathrm{r}}b_{J,\mathrm{r}} & \cdots & a_{\mathrm{I},\mathrm{r}}b_{\mathrm{J},\mathrm{r}} \end{bmatrix} \in \mathbb{C}^{J\times I} \tag{10}$$

A estrutura da matriz obtida na Equação (10) é equivalente à seguinte operação:

$$\underset{J\times I}{\text{unvec}}\{\mathbf{a}_\text{r} \otimes \mathbf{b}_\text{r}\} = \mathbf{b_r}\, \mathbf{a}_\text{r}^\text{T} \tag{11}$$

Onde cada $\mathbf{a}_\text{r}$ e $\mathbf{b}_\text{r}$ podem ser estimados usando uma decomposição em valores singulares de posto 1:

$$\mathbf{b}_\text{r}\mathbf{a}_\text{r}^\text{T} = \mathbf{u}_\text{r}\, s_\text{r}\, \left(\mathbf{v}_\text{r}{}^*\right)^\text{T} \tag{12}$$

As estimativas são então obtidas tomando a raiz quadrada do valor singular para distribuir ao vetor singular, de tal forma que:

$$\hat{\mathbf{a}}_\text{r} = \mathbf{v}_\text{r}{}^*\sqrt{s_\text{r}} \tag{13}$$

$$\hat{\mathbf{b}}_\text{r} = \mathbf{u}_\text{r}\sqrt{s_\text{r}} \tag{14}$$

I.3 CONCEITOS DE CÁLCULO TENSORIAL

Nesta seção, são apresentadas as operações gerais baseadas em tensores necessárias para a derivação das técnicas abordadas nos Capítulos (2) e (4).

I.3.1 Tensores

Da mesma forma que vetores podem ser vistos como um conjunto de escalares, e matrizes como um conjunto de vetores, tensores podem ser considerados um conjunto de matrizes. No entanto, enquanto as matrizes são limitadas a apenas duas dimensões, os tensores podem ter um número ilimitado de dimensões. Nesse contexto, um tensor de N-dimensões é denotado como $\mathcal{X} \in \mathbb{C}^{M_1\times M_2\cdots\times M_n\cdots\times M_R}$. Além disso, tensores podem ser vistos como uma coleção de *vetores de modo-n*, que são obtidos se um dado índice n é variado e todos os outros índices são mantidos fixos. Por exemplo, dado um tensor de terceira ordem $\mathcal{X} \in \mathbb{C}^{I\times J\times K}$, seus vetores de modo-n são dados por: $\{\mathbf{x}_{(i,:,:)}\ \ \mathbf{x}_{(:,j,:)}\ \ \mathbf{x}_{(:,:,k)}\}$.

I.3.1.1 O Desdobramento de Modo-n (n-mode unfolding)

O desdobramento de modo-n é um processo que reordena os elementos de um tensor de ordem N em uma matriz. Isso é feito rearranjando os vetores de modo-n do tensor para que se tornem as colunas da matriz resultante. Diferentes trabalhos por vezes utilizam diferentes ordenamentos. A notação adotada neste trabalho é consistente com [84].

A ordenação de colunas escolhida neste trabalho é referida como o desdobramento cíclico reverso, proposto em [54], que começa com o índice $(n-1)$-ésimo e prossegue para trás, até o índice $(n+1)$.

Propriedades dos desdobramentos de modo-n são discutidas nas Seções (I.3.1.3) e (I.3.1.2).

I.3.1.2 O Produto de Modo-n (n-mode product)

Um tensor pode ser multiplicado por uma matriz através do produto de modo-n. Para um tensor de ordem N, $\mathcal{X} \in \mathbb{C}^{M_1 \times M_2 \cdots \times M_n \cdots \times M_R}$, e uma matriz $\mathbf{B} \in \mathbb{C}^{M \times M_n}$, o produto de modo-n desses dois termos é dado por:

$$\mathcal{X} \times_n \mathbf{B} = \mathbf{B}\,[\mathcal{X}]_{(n)} \tag{15}$$

Onde $[\mathcal{X}]_{(n)} \in \mathbb{C}^{M_n \times M_1 M_2 \ldots M_N}$ é o desdobramento de modo-n de $\mathcal{X}$. O tensor de saída é o produto matricial da Equação (15) dobrado de volta em um tensor de tamanho $M_1 \times M_2 \cdots \times M \cdots \times M_R$.

I.3.1.3 A Decomposição PARAFAC

A Análise Fatorial Paralela (PARAFAC) [16] é uma forma poliádica de um tensor. Ela expressa um dado tensor como a soma de um número finito de tensores de posto um. Por exemplo, dado um tensor de terceira ordem de posto R, $\mathcal{X} \in \mathbb{C}^{I \times J \times K}$, a seguinte expressão:

$$\mathcal{X} \approx \sum_{r=1}^{R} \mathbf{a} \circ \mathbf{b} \circ \mathbf{c} \in \mathbb{C}^{I \times J \times K} \tag{16}$$

Fatora o tensor $\mathcal{X}$ em uma soma de tensores componentes de posto um. As matrizes de fatores são referidas como a combinação dos vetores de cada um dos componentes de posto um. Portanto, as matrizes de fatores associadas à decomposição PARAFAC do tensor $\mathcal{X}$ são escritas como:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_R \end{bmatrix} \in \mathbb{C}^{I \times R} \tag{17}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_R \end{bmatrix} \in \mathbb{C}^{J \times R} \tag{18}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_R \end{bmatrix} \in \mathbb{C}^{K \times R} \tag{19}$$

Sob a condição de unicidade [85] a decomposição PARAFAC de um tensor pode ser expressa em termos de suas matrizes de fatores, de tal forma que:

$$\mathcal{X} = \mathcal{I}_{3,R} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} \in \mathbb{C}^{I \times J \times K} \tag{20}$$

Uma propriedade é que para um dado tensor de terceira ordem que segue a Equação (20), seus desdobramentos podem ser expressos por:

$$[\mathcal{X}]_{(1)} = \mathbf{A}\,(\mathbf{B} \diamond \mathbf{C})^{\mathrm{T}} \in \mathbb{C}^{I \times JK} \tag{21}$$

$$[\mathcal{X}]_{(2)} = \mathbf{B}\,(\mathbf{C} \diamond \mathbf{A})^{\mathrm{T}} \in \mathbb{C}^{J \times KI} \tag{22}$$

$$[\mathcal{X}]_{(3)} = \mathbf{C}\,(\mathbf{A} \diamond \mathbf{B})^{\mathrm{T}} \in \mathbb{C}^{K \times IJ} \tag{23}$$