Learning and Nonlinear Models - Journal of the Brazilian Society on Computational Intelligence (SBIC), Vol. 23, Iss. 2, pp. 23–45, 2025
https://doi.org/10.21528/lnlm-vol23-no2-art3

Statistical Approaches to Calculate Weighted Principal Components with Applications in Human Face Image Analysis

Laura Costa Pereira Miranda
Laboratório Nacional de Computação Científica - LNCC
lauracpmiranda@gmail.com

Carlos Eduardo Thomaz
Centro Universitário da FEI - FEI
cet@fei.edu.br

Diego Barreto Haddad
Centro Federal de Educação Tecnológica Celso Suckow da Fonseca - CEFET-RJ
diego.haddad@cefet-rj.br

Ana Cláudia Souza Vidal de Negreiros
Universidade Federal Rural do Semi-Árido - UFERSA
anaclaudia.negreiros@ufersa.edu.br

Gilson Antonio Giraldi
Laboratório Nacional de Computação Científica - LNCC
gilson@lncc.br

Resumo – O principal objetivo deste trabalho é investigar a eficiência de técnicas de ponderação dos dados para o cálculo de componentes principais ponderadas em experimentos de gênero e expressão facial, considerando problemas de classificação e reconstrução. Especificamente, a metodologia consiste em gerar pesos espaciais (mapas estatísticos) para ponderar os pixels das imagens de entrada da análise de componentes principais (PCA). Estaremos considerando as seguintes técnicas para o cálculo dos pesos espaciais: (a) Estatística T de Student; (b) Hiperplanos computados usando SVM (*Support Vector Machine*); (c) Entropia de Shannon calculada pixel-a-pixel; (d) Inverso da Entropia de Shannon calculada pixel-a-pixel; (e) Divergência Jensen-Shannon. A aplicação das técnicas (a),(c),(d), e (e) para cálculo de pesos espaciais em reconhecimento de faces é a principal contribuição deste trabalho. A avaliação da eficiência das diferentes componentes principais obtidas com cada técnica é feita analisando os resultados da reconstrução de imagens e da classificação usando as bases de imagens FEI e FERET. Para classificação, foram utilizados os seguintes classificadores: K-Vizinhos Mais Próximos e a Distância de Mahalanobis. Foi utilizada também uma metodologia baseada em *quadtree* para reduzir pequenas variações locais dos mapas estatísticos, para testar seu efeito na reconstrução e classificação.

Palavras-chave – Reconhecimento de faces, PCA, Métodos de Ponderação, Classificação, Reconstrução.

Abstract – The main objective of this work is to investigate the efficiency of data weighting techniques for calculating weighted principal components in gender and facial expression experiments, considering classification and reconstruction problems. Specifically, the methodology consists of generating spatial weights (statistical maps) to weight the pixels of the input images. Then, the weighted data are used as input for the principal component analysis (PCA) algorithm. We will consider the following techniques for calculating spatial weights: (a) Student's t-test; (b) Hyperplanes computed using SVM (Support Vector Machine); (c) Shannon entropy calculated pixel-by-pixel; (d) Inverse Shannon entropy calculated pixel-by-pixel; (e) Jensen-Shannon divergence. The application of techniques (a), (c), (d), (e) for calculating spatial weights in face recognition is the main contribution of this work. The evaluation of the efficiency of the different principal components obtained with each technique is done by analyzing the results of image reconstruction and classification over FEI and FERET databases. For classification, the following classifiers were used: K-Nearest Neighbors and Mahalanobis Distance. A quadtree-based methodology was also used to reduce small local variations in the statistical maps, to test its effect on reconstruction and classification.

Keywords – Face Recognition, PCA, Weighting Methods, Classification, Reconstruction.

1. INTRODUÇÃO

Muitas áreas que utilizam técnicas de reconhecimento de padrões, tais como visão computacional, processamento de sinais e análise de imagens, requerem o tratamento de conjuntos de dados com um grande número de características ou dimensões.

https://doi.org/10.21528/lnlm-vol23-no2-art3

Portanto, devemos considerar métodos de redução de dimensionalidade para obter uma representação compacta dos dados capaz de descartar redundâncias, ruídos e reduzir o custo computacional de operações posteriores [1, 2].

Nesse trabalho, nosso foco em relação ao tipo de dados será o processamento de imagens em tons de cinza, as quais podem ser representadas por matrizes do tipo $\mathbb{R}^{M\times N}$. Contudo, a metodologia pode ser facilmente adaptada para imagens coloridas.

Considerando o estado da arte em reconhecimento de padrões, a redução de dimensionalidade pode ser realizada dentro de um contexto de engenharia de características [3, 4], aprendizado de variedades [5–7] e utilizando métodos de aprendizado de máquina, tais como metodologias envolvendo redes neurais profundas [8]. Em todos esses casos, do ponto de vista matemático, a ideia é construir uma função $f : \mathbb{R}^{M\times N} \longrightarrow \mathbb{R}^m$, onde $m \ll M \cdot N$, que preserva as informações de interesse nos dados originais.

Cada subárea descrita acima tem suas vantagens e desvantagens. A engenharia de características envolve métodos lineares, multilineares e técnicas procedurais baseadas em extratores de características, tais como momentos, características de hitogramas e matriz de co-ocorrência, para gerar a função f. Os métodos lineares/multilineares são baseados em conceitos de álgebra linear/multilinear, gerando mapeamentos para os quais a pré-imagem $f^{-1}(\mathbb{R}^m)$ pode ser calculada, embora potencialmente com perda de detalhes, processo denominado reconstrução. As técnicas procedurais, em geral, não admitem tal pré-imagem, sendo uma desvantagem destes métodos. Considerando a vasta quantidade de métodos procedurais para fazer a extração de características [9], em geral, fica-se em um processo de tentativa e erro para encontrar a função (procedimento) f, marcado pela busca de uma combinação eficiente de espaços de características. Essa é outra desvantagem dessa metodologia. Contudo, apesar dessas potenciais desvantagens, os métodos de extração não automática de características também possuem o benefício de, em muitos casos, fornecer uma compreensão mais clara acerca das informações contidas nos dados, o que pode melhorar a compreensão do fenômeno de interesse e da respectiva tarefa envolvida.

As técnicas de aprendizado de variedades são fundamentadas em elementos geométricos que permitem tratar a natureza não linear dos dados [10]. Neste caso, o problema central pode ser expresso como encontrar uma aplicação $f : \mathbb{R}^{M\times N} \longrightarrow \mathcal{M}^m \subset \mathbb{R}^n$, onde $n = M \cdot N$ e $\mathcal{M}^m$ é uma variedade diferenciável de dimensão $m \ll n$ (simplificadamente, hipersuperfície suave contida em $\mathbb{R}^n$). Dentre esse grupo de técnicas, alguns métodos clássicos são o ISOMAP [11], tendo sido um dos primeiros a considerar a preservação das distâncias geodésicas no processo de redução de dimensionalidade, e o LLE [12] que reconstrói a estrutura não-linear global com partes localmente lineares. Um pouco mais recente, o método t-SNE foi proposto com a vantagem de preservar a estrutura local dos dados no espaço reduzido, bem como a estrutura global, como a presença de agrupamentos em diversas escalas [13]. Porém, não há uma metodologia unificada para a construção da função $f : \mathbb{R}^{M\times N} \longrightarrow \mathcal{M}^m$ nesta área, o que é uma desvantagem da mesma.

As técnicas de aprendizado profundo resolvem o problema da computação da função f através do treinamento de redes neurais. Porém, essas técnicas necessitam de grandes volumes de dados para que o treinamento ocorra eficientemente, o que é um problema principalmente no caso do aprendizado supervisionado. Esse problema persiste na área, mesmo com o desenvolvimento de métodos de aumento de dados [14, 15], transferência de aprendizado [16, 17] e aprendizado fracamente supervisionado [18] e semi-supervisionado [19]. Além disso, o treinamento costuma ser muito caro computacionalmente, demandando longos períodos de processamento, mesmo em GPUs mais modernas.

Considerando o contexto acima, esse trabalho trata de métodos lineares para redução de dimensionalidade com aplicação para reconhecimento de padrões em faces humanas. As metodologias lineares neste contexto incluem análise de componentes principais (PCA), análise fatorial, [20], *Multidimensional scaling*, [21], *Projection pursuit* [22], *random projection* [23], PCA ponderado (*weighted PCA* - WPCA) [24], dentre outros. Uma generalização dos métodos lineares é dada pelos métodos multilineares (tensoriais) [25], os quais não necessitam da etapa de vetorização das imagens, a qual é utilizada pelos demais métodos, respeitando assim as relações espaciais entre os pixels das mesmas.

O PCA é a técnica mais popular nesta área, sendo uma abordagem não-supervisionada, bem-sucedida para o problema de representar amostras em espaços de dimensão reduzida. Desde o trabalho pioneiro de Sirovich e Kirby, [26], vários trabalhos subsequentes usaram PCA para extração de características com vistas à classificação e análise de imagens [20, 27].

Em aplicações de análise de imagens, foco desse trabalho, todos os métodos lineares mencionados concatenam as intensidades de pixels da imagem original, suposta em escala de cinza, em um vetor de alta dimensão. Com as imagens assim representadas, as técnicas lineares recebem um banco de dados como entrada e *aprendem* matrizes de projeção que transformam os dados originais de alta dimensão para um espaço de dimensão reduzida. Em seguida, a reconstrução é realizada para visualizar os dados reduzidos no espaço da imagem original.

O desenvolvimento de técnicas lineares que reúnem redução de dimensionalidade e conhecimento prévio pode ser realizado no âmbito de métodos de aprendizagem supervisionada. Um exemplo conhecido nesta área é a Análise Discriminante Linear (LDA) [28]. Em [24] é proposto um outro método supervisionado que permite incorporar explicitamente conhecimento prévio do domínio ao PCA, gerando um novo espaço linear eficiente para classificação que herda também propriedades de otimalidade para redução de dimensionalidade. O método depende de pesos discriminantes dados por hiperplanos de separação para gerar um PCA espacialmente ponderado, chamado de WPCA, ou PCA ponderado. Esses dois métodos para redução de dimensionalidade, LDA e WPCA, são supervisonados e lineares. Contudo, é possível a utilização de métodos supervisionados não lineares, como o KLDA (*Kernel Linear Discriminant Analysis*) [29], que consiste na utilização do artifício da função kernel no LDA. Mais recentemente, técnicas supervisionadas baseadas em redes neurais artificiais vêm sendo usadas para redução de dimensionalidade [30].

O ponto de partida deste trabalho está na observação de que, apesar da eficiência do PCA como um método para redução de dimensionalidade, seu algoritmo não incorpora informações prévias extraídas do domínio específico correspondente aos da-

dos. Uma forma de tratar essa limitação do PCA consiste na aplicação de pesos espaciais computados utilizando os próprios dados. Neste contexto, o principal objetivo deste trabalho é investigar a eficiência de novas técnicas de ponderação espacial em experimentos de gênero e expressão facial, considerando problemas de classificação e reconstrução. Assim, o trabalho está fundamentado na referência [24], a qual apresenta um algoritmo para incorporar conhecimento a priori no PCA tradicional via aplicação de pesos espaciais, computados por hiperplanos de separação, como o $SSVM$ (*Smooth Support Vector Machine*), descrito em [31], para gerar o WPCA. Outros hiperplanos de separação, como o LDA, podem ser utilizados para extrair os pesos para computar o WPCA. Porém, a formulação do LDA, baseada no critério de Fisher gera uma direção discriminante que considera a distribuição total dos dados, uma vez que depende da média global, das médias das classes e das matrizes de espalhamento inter e intra-classes, enquanto o SSVM vai se concentrar nas amostras mais difíceis para a classificação (vetores suporte), podendo assim ser mais eficiente quando características mais sutis, como relacionadas à expressão facial, estão envolvidas [32]. Além disso, considerando as características do LDA é possível identificar certa relação entre sua direção discriminante e o mapa gerado pela Estatística T [33] utilizada neste trabalho, uma vez que essa também considera as médias das classes e a distribuição das classes em torno das médias.

Estaremos considerando as seguintes técnicas para o cálculo dos pesos espaciais: (a) Estatística T de Student, denotada por T [33]; (b) Hiperplanos computados usando o SSVM [31]; (c) Entropia H de Shannon [34], calculada pixel-a-pixel; (d) Inverso da Entropia de Shannon calculada pixel-a-pixel ($\widehat{H}$); (e) Divergência de Jensen-Shannon, denotada por JS [35]. A técnica (b) é conhecida na literatura para o método em foco [24], sendo escolhida como uma referência para comparação com as demais. Os mapas obtidos pelo método (a) são utilizados em análise de faces, porém, em metodologias distintas daquela abordada neste trabalho [36–39]. A aplicação da Estatística T de Student, entropia de Shannon e seu inverso, e da divergência de Jensen-Shannon para cálculo de pesos espaciais para o método descrito em [24], para reconhecimento de faces, é a principal contribuição deste trabalho.

A motivação para a utilização da entropia vem da interpretação desta quantidade como uma medida da complexidade dos padrões de intensidade em um banco de imagens. A divergência de Jensen-Shannon (JS) e a Estatística T de Student quantificam diferenças entre as classes consideradas: pixels com diferenças mais significativas devem ter maior *peso* que os demais.

A avaliação da eficiência das diferentes componentes principais obtidas com cada técnica é feita analisando os resultados da reconstrução de imagens, da visualização de componentes principais e da classificação. Para esta última, foram utilizados classificadores K-Vizinhos mais próximos (KNN) com $K = 1$ e o classificador baseado na distância de Mahalanobis, denotado por DM [40]. A motivação para essa escolha está na comparação entre o desempenho de um classificador não paramétrico (KNN) e um classificador paramétrico (DM).

Foi utilizada também uma metodologia baseada em *quadtree* para reduzir pequenas variações locais dos mapas de ponderação, para testar seu efeito na reconstrução e classificação, baseada na referência [38]. A aplicação da estrutura espacial de *quadtree* para processamento de mapas espaciais para calcular versões ponderadas do PCA é outra contribuição deste trabalho.

Os experimentos computacionais foram realizados utilizando imagens do banco de imagens de faces da FEI [41] e da FERET [42]. Como medida de desempenho dos classificadores, utilizamos a acurácia média calculada com o uso do $K-$fold, para $K = 10$. No caso da reconstrução, além da análise visual, utilizamos a medida $SSIM$ (*Structural Similarity Index*) [43].

Os experimentos de classificação e reconstrução foram executados duas vezes: sem aplicação da *quadtree* e com a aplicação da *quadtree* para processar cada mapa de ponderação calculado. Uma análise dos experimentos computacionais leva à conclusão de que os melhores métodos para reconstrução são: PCA e JS. Para a classificação usando o classificador DM, os melhores métodos foram PCA, JS_{qt} e $\widehat{H}_{qt}$ (estes últimos sendo os métodos JS e $\widehat{H}$ com aplicação da *quadtree*). No caso do classificador KNN os melhores métodos foram H_{qt}, T, $\widehat{H}_{qt}$ e $\widehat{H}$. Esses resultados mostram a influência do classificador nestes experimentos e vantagens da aplicação da *quadtree*.

Assim, podemos sintetizar as contribuições deste trabalho nos seguintes itens:

1. A aplicação da Estatística T de Student, entropia de Shannon e seu inverso, e divergência de Jensen-Shannon para cálculo de pesos espaciais para computar o PCA ponderado;

2. Aplicação da técnica de *quadtree* para processar mapas de ponderação para cálculo do PCA ponderado;

3. Testes exaustivos, além da análise visual, envolvendo reconstrução e classificação em cenários envolvendo gênero e expressão facial .

O texto está organizado da seguinte forma: a seção 2 apresenta a abordagem proposta, a qual tem como ponto central o cálculo do PCA ponderado, apresentado na seção 2.3. A seção 3 sumariza o processo de reconstrução e discute teoricamente os efeitos da aplicação da *quadtree*. Os resultados computacionais e a discussão dos mesmos são apresentados nas seções 4 e 5, respectivamente. As conclusões e trabalhos futuros são comentados na seção 6.

2. ABORDAGEM PROPOSTA

Nesta seção descrevemos as principais contribuições deste trabalho, o qual está baseado no WPCA [24], sendo um sumário da dissertação [44]. Especificamente, estaremos propondo as seguintes técnicas para o cálculo dos mapas de ponderação: (a) Entropia de Shannon [34]; (b) Inverso da Entropia de Shannon; (c) Divergência Jensen-Shannon [35] entre as distribuições associadas a cada uma das classes consideradas; (d) Estatística T de Student [33]. Além destas técnicas, usaremos também

pesos espaciais calculados usando hiperplanos computados via $SSVM$ [31], para fins de comparação com a literatura da área. A Tabela 1 centraliza a notação e os principais símbolos usados no texto.

Símbolo	significado
P, U, etc.	Letras maiúsculas normais representam matrizes
$\mathbf{x}$, $\mathbf{y}$	Letras minúsculas em negrito representam vetores
λ, α, etc.	Símbolos gregos em minúsculo representam escalares
$\mathcal{D}$, etc.	Conjuntos de dados e subespaços são representados usando fonte caligráfica
$\|\| \cdot \|\|$	Norma de vetor ou matriz
$E(\cdot)$	Esperança matemática
$\widehat{H}$	Mapa gerado usando o Inverso da Entropia
H	Mapa gerado usando a Entropia
JS	Mapa gerado usando a Divergência de Jensen-Shannon
T	Mapa gerado usando a Estatística T de Student
W	Mapa gerado usando o Smooth Support Vector Machine
$\lambda(R)$	Espectro da matriz R
$SSIM$	Índice de Similaridade Estrutural (Structural Similarity Index)
PCA	Análise de Componentes Principais (Principal Component Analysis)
$\min$	Valor mínimo
$\max$	Valor máximo
$diag(d_1, d_2, \ldots, d_n)$	Matriz diagonal com elementos diagonais $d_1, d_2, \ldots, d_n$
$WPCA$	Weighted PCA

Tabela 1: Principais símbolos usados e seus significados.

A principal motivação para esse trabalho vem da análise apresentada em [45], a qual mostra que componentes principais com alta variância não necessariamente são eficientes para reconhecimento de padrões. Isso motiva a inclusão de informação *a priori* no algoritmo do PCA.

2.1 Modelagem do Banco de Imagens

Primeiramente, cada imagem do banco de dados original é convertida em tons de cinza e redimensionada para uma forma quadrada, 128×128 no caso deste trabalho, usando recursos de bibliotecas como Scikit-Image ou mesmo do MatLab, de modo a facilitar os cálculos e o processamento relativos à *quadtree*. Em seguida, a matriz obtida é vetorizada, gerando um vetor em $\mathbb{R}^n$. Desta forma, representamos o conjunto de imagens na forma:

$$\mathcal{D} = \{(\mathbf{x}_1, l_1), (\mathbf{x}_2, l_2), \ldots, (\mathbf{x}_N, l_N)\} \subset \mathbb{R}^n \times \{-1, 1\}, \tag{1}$$

onde $\mathbf{x}_i$ é o vetor que representa a imagem i e l_i seu rótulo, o qual define a classe correspondente. A Figura 1 mostra duas amostras de cada banco de dados, fornecendo exemplos das características principais de cada um. São imagens em tons de cinza, com pouco fundo.

Vamos modelar o banco de dados $\mathcal{D}$ usando um vetor de variáveis aleatórias:

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)^T, \tag{2}$$

onde $(\cdot)^T$ significa transposta, x_i é uma variável aleatória discreta real tomando valores no conjunto $I = \{0, 1, ..., 255\}$, já que estamos considerando imagens em tons de cinza. Agora, devemos estimar as distribuições de probabilidade p_j, $j = 1, 2, \ldots, n$. Isso pode ser feito calculando histogramas a partir da representação matricial do conjunto $\mathcal{D}$, dada por:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdot & \cdot & x_{1n} \\ x_{21} & x_{22} & \cdot & \cdot & x_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{N1} & x_{N2} & \cdot & \cdot & x_{Nn} \end{bmatrix}. \tag{3}$$

Assim, podemos definir os contadores:

$$C(j, k) = \sum_{i=1}^{N} \delta(i, j, k), \quad 1 \leq j \leq n, \quad 0 \leq k \leq 255, \tag{4}$$

onde:

$$\delta(i, j, k) = \begin{cases} 1, & \text{se } x_{ij} = k, \\ 0, & \text{caso contrário.} \end{cases} \tag{5}$$

https://doi.org/10.21528/lnlm-vol23-no2-art3

(a) (b) (c) (d)

Figura 1: (a)-(b) Exemplos de faces neutras da FEI. (c)-(d)Imagens do banco de dados FERET.

Com isso, podemos estimar a distribuição de probabilidades associada a cada variável aleatória x_j do vetor $\mathbf{x}$ como:

$$p_{jk} = \frac{C\left(j,k\right)}{N}, \quad 1 \leq j \leq n, \quad 0 \leq k \leq 255. \tag{6}$$

2.2 Cálculo dos mapas de ponderação

Neste trabalho, estaremos considerando problemas de classificação com duas classes. Desta forma, vamos representar as classes pelos conjuntos:

$$X_l = \{\mathbf{x}_{1l}, \mathbf{x}_{2l}, \ldots \mathbf{x}_{n_l}\} \subset \mathbb{R}^n, \quad l = 1, 2, \tag{7}$$

onde $\mathbf{x}_{il}$ representa a imagem i da classe l.

- Estatística T de Student: No contexto acima, primeiramente calculamos as médias dadas pela expressão:

$$\overline{\mathbf{x}}_l = \sum_{i=1}^{n_l} \frac{\mathbf{x}_{il}}{n_l}, \quad l = 1, 2, \tag{8}$$

e variâncias:

$$S^2_{X_l}\left(k,l\right) = \sum_{i=1}^{n_l} \frac{\left(x_{ikl} - \overline{x}_{kl}\right)^2}{n_l - 1}, \quad 1 \leq k \leq n, \quad l = 1, 2. \tag{9}$$

Então a Estatística T de Student é computada a partir da expressão:

$$\widehat{T}_k = \frac{\left|\overline{x}_{k1} - \overline{x}_{k2}\right|}{\left(\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right) S_{X_1X_2}\left(k\right)}, \quad 1 \leq k \leq n, \tag{10}$$

onde:

$$S_{X_1X_2}\left(k\right) = \sqrt{\frac{\left(n_1 - 1\right) S^2_{X_1}\left(k,1\right) + \left(n_2 - 1\right) S^2_{X_2}\left(k,2\right)}{n_1 + n_2 - 2}}, \quad 1 \leq k \leq n. \tag{11}$$

Em seguida, vetorizamos $\widehat{T}$ e, seguindo a metodologia do WPCA [24], normalizamos os valores de $\widehat{T}$ para obter o mapa de ponderação T dado por:

$$T_k = \frac{\left|\widehat{T}_k\right|}{\sum_{j=1}^{n}\left|\widehat{T}_j\right|}; \quad 1 \leq k \leq n. \tag{12}$$

- SSVM: Neste caso, o vetor normal $\mathbf{w} \in \mathbb{R}^n$ ao hiperplano de separação das classes X_1 e X_2, gerado pelo SSVM [31], é utilizado para gerar o mapa W via normalização:

$$W_i = \frac{\left|\mathbf{w}_i\right|}{\sum_{j=1}^{n}\left|\mathbf{w}_j\right|}; \quad 1 \leq i \leq n. \tag{13}$$

- Entropia: De acordo com o desenvolvimento da Seção 2.1, cada variável aleatória x_j do vetor (2) possui uma distribuição de probabilidade p_j dada pela expressão (6), para $1 \leq j \leq n$. Assim, podemos calcular o vetor de entropias:

$$H\left(x_j\right) = \frac{1}{8} \sum_{i=0}^{255} p_{ji} \left(-\log p_{ji}\right), 1 \leq j \leq n. \tag{14}$$

Uma vez que $0 \leq H \leq 1$, para gerar o mapa de ponderação neste caso basta realizar a normalização:

$$H\left(x_j\right) \leftarrow \frac{H\left(x_j\right)}{\sum_{j=1}^{n} H\left(x_j\right)}. \tag{15}$$

- Inverso da Entropia: Dado por:

$$\widehat{H}(j) = \frac{1}{H_j}, \quad j = 1, \cdots, n, \tag{16}$$

com H_j sendo os elementos do vetor de entropias H calculado na expressão (15). Caso $H_j = 0$, então fazemos a substituição $H_j \leftarrow 10^{-5}$.

- Divergência de Jensen-Shannon: Agora temos dois vetores de variáveis aleatórias discretas $\mathbf{x}_1 = \left(x_{11}, x_{21}, \ldots, x_{n1}\right)^T$ e $\mathbf{x}_2 = \left(x_{12}, x_{22}, \ldots, x_{n2}\right)^T$, associados às classes de imagens 1 e 2, respectivamente. Por sua vez, cada variável aleatória x_{jl} toma valores no conjunto $\{0, 1, 2, \ldots, 255\}$ com probabilidades:

$$p_{jl}\left(k\right) = p_{jlk}, \quad 1 \leq j \leq n,\ 0 \leq k \leq 255,\ l = 1, 2,$$

as quais são obtidas computando a expressão (6) para cada classe separadamente. Nestas condições, a divergência JS associada ao pixel j é dada por:

$$JS(p_{j1}||p_{j2}) = KL\left(p_{j1}(x)||\left(\frac{p_{j1}(x) + p_{j2}(x)}{2}\right)\right) + KL\left(p_{j2}(x)||\left(\frac{p_{j1}(x) + p_{j2}(x)}{2}\right)\right), \tag{17}$$

com a divergência de KL calculada pela equação:

$$KL(p_1||p_2) = \sum_{i=1}^{k} p_{i1} \log\left(\frac{p_{i1}}{p_{i2}}\right). \tag{18}$$

Assim como no caso anterior, o mapa JS é calculado pela normalização dos resultados obtidos pela expressão (17):

$$JS\left(j\right) \leftarrow \frac{JS(p_{j1}||p_{j2})}{\sum_{j=1}^{n} JS(p_{j1}||p_{j2})}, \quad 1 \leq j \leq n. \tag{19}$$

A principal motivação para a utilização da entropia neste contexto vem da apresentação da teoria de informação de Shannon [34], a qual mostra que a entropia H normalizada fornece um limite inferior para compressão da informação associada a uma sequência de variáveis aleatórias idênticas e igualmente distribuídas. No caso das imagens, valores mais altos da entropia indicam maior complexidade de padrões de intensidade, fazendo com que a taxa R do esquema de compressão seja mais elevada. Assim, considerando um pixel i e seus valores x_{ij}, $j = 1, 2, \ldots, N$ no banco de dados, *maior complexidade* indica distribuição de probabilidade mais próxima da distribuição uniforme, onde a entropia tem valor máximo. Assim, nesta interpretação, estaremos dando maior peso aos pixels com distribuição de intensidades mais dispersa em relação à média, os quais esperamos serem mais relevantes para as tarefas de reconhecimento facial nos bancos de dados considerados.

Por outro lado, tanto a Divergência de Jensen-Shannon (JS) quanto a Estatística T de Student quantificam diferenças entre as classes consideradas. A primeira considerando as distribuições de probabilidades associadas às classes, ponderando com maior valor os pixeis cujas distribuições de intensidade são diferentes para as duas classes. No caso do Estatística T, olhando para diferenças entre médias das distribuições de intensidade (numerador da expressão (10)) das classes, normalizadas em função de $S_{X_1X_2}$, que por sua vez depende das variâncias pixel-a-pixel ($S^2_{X_1}$ e $S^2_{X_2}$) e da cardinalidade relativa de cada classe (expressão (11)), de forma a dar maior peso aos pixels cujas intensidades possuem alta variação inter-classe (numerador da equação (10)) e baixa variância intra-classe, quantificada pelo denominador da equação (10).

Com relação ao mapa W, proposto em [24], esse mapa é inspirado na seguinte observação: dada uma amostra centralizada $\mathbf{z} = \mathbf{x} - \overline{\mathbf{x}}$ do conjunto de teste, a classe desta amostra segundo o SSVM será definida pela expressão:

$$f\left(z\right) = signal\left(\sum_{i=1}^{n} z_i w_i + b\right), \tag{20}$$

onde b é o viés do modelo e $signal$ retorna $+1$ se o resultado for positivo e -1 caso contrario. Assim, os elementos do vetor $\mathbf{w} = (w_1, w_2, ..., w_n)$ com maior valor em módulo serão os mais significativos para a classificação via SSVM. Consequentemente, os pixeis correspondentes devem receber as maiores ponderações.

Os pesos computados podem formar um mapa de ponderação ruidoso, o que pode trazer artefatos na reconstrução e interferir na classificação, o que é justificado a seguir (seções 3 e 5). Com o objetivo de reduzir tais efeitos, seguimos [38] e utilizamos uma metodologia baseada em *quadtree* para reduzir pequenas variações locais. Durante a construção da *quadtree* um quadrante é subdividido em quatro subquadrantes, caso atenda um critério, baseado em um limiar t de intensidade, definido pela condição: se $q \times q$ é a dimensão de um quadrante Q da *quadtree* então há uma subdivisão desse quadrante se

$$\max(Q_{ij}) - \min(Q_{ij}) > t, \quad 1 \leq i \leq q \;\; e \;\; 1 \leq j \leq q. \tag{21}$$

Esse processo pode ser executado até gerar quadrantes de tamanho 1×1 pixels. Porém, em nossa implementação (feita em MATLAB), optamos por limitar inferiormente o tamanho dos quadrantes para até 8×8 pixels. Utilizamos limiares $t = 0.2, 0.4, 0, 6, 0.8$. Calculada a *quadtree* sobre o mapa, cada quadrante assume valor igual à média de pixels do quadrante. Os mapas obtidos são normalizados, seguindo o mesmo processo feito para os mapas originais.

2.3 Cálculo do PCA Ponderado

Depois de computarmos os mapas estatísticos da seção 2.2, precisamos calcular o PCA ponderado para então prosseguir para a etapa de reconstrução e classificação.

Nossa metodologia para calcular as versões ponderadas do PCA é baseada no WPCA [24]. Ela é constituída pelas mesmas etapas do WPCA, com uma mudança na etapa 4, como descrito abaixo:

1. Cálculo do vetor de ponderação espacial bruto $\widehat{\mathbf{w}} = (\widehat{w}_1, \widehat{w}_2, \ldots, \widehat{w}_n)^T$ usando dados rotulados e alguma técnica de ponderação da seção 2.2;
2. Normalização de $\widehat{\mathbf{w}}$ para gerar o vetor de ponderação espacial $\mathbf{w} = (w_1, w_2, \ldots, w_n)^T$, onde:
$$w_i = \frac{|\widehat{w}_i|}{\sum_{j=1}^{n} |\widehat{w}_j|}; \tag{22}$$
3. Cálculo da média global $\overline{\mathbf{x}}$ dado por:
$$\overline{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_i = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n)\,; \tag{23}$$
4. Centralização dos dados, organizados na matriz de dados de treinamento $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]^T \in \mathbb{R}^{N\times n}$, com respeito à média, gerando uma nova matriz de treinamento Z computada pela equação:
$$z_{ij} = x_{ij} - \overline{x}_j,\ 1 \leq i \leq N,\ 1 \leq j \leq n; \tag{24}$$
5. Ponderação das variáveis centralizadas z_{ij} usando os pesos espaciais calculados no passo 2 e geração de uma nova matriz $Z^* = \left\{z^*_{ij}\right\}$ onde:
$$z^*_{ij} = z_{ij}\sqrt{w_j}; \tag{25}$$
6. Cálculo das m componentes principais ponderadas $P^*_{wpca} = [\mathbf{p}^*_1, \mathbf{p}^*_2, \ldots, \mathbf{p}^*_m]^T$ selecionando os autovetores da matriz $(Z^*)^T Z^*$ com os m maiores autovalores.

Neste trabalho, as técnicas de ponderação utilizadas estão resumidas na Seção 2.2. Porém, outras técnicas podem ser utilizadas, dependendo da aplicação. Por outro lado, suprimimos a normalização em relação ao desvio padrão que é realizada no $WPCA$ [24], com o objetivo de evitar problemas com tal normalização, visto que a representação assim obtida pode não ser eficiente para propósitos de classificação e de reconstrução, como observado em [44], seção 4.6.

3 PROCESSO DE RECONSTRUÇÃO E *QUADTREE*

Dada uma imagem $\mathbf{x}$ do conjunto de dados, o processo de reconstrução pode ser sintetizado em duas etapas:

1. Projeção de $\mathbf{x}$ num espaço de dimensão m reduzida, ou seja, $m < n$;
2. Representação da projeção no espaço original das imagens.

https://doi.org/10.21528/lnlm-vol23-no2-art3

A etapa 1 é realizada segundo a expressão:

$$\mathbf{x}_p = P_{pca}(\mathbf{x} - \overline{\mathbf{x}}), \tag{26}$$

na qual P_{pca} é a matriz do PCA, tradicional ou ponderado, $\overline{\mathbf{x}}$ é a média global do conjunto de dados, computada pela expressão (23).

O passo 2 é realizado através da equação:

$$\mathbf{x}_r = P_{pca}^T \mathbf{x}_p + \overline{\mathbf{x}}, \tag{27}$$

onde $\mathbf{x}_r$ representa o vetor $\mathbf{x}$ reconstruído, ou seja, o vetor $\mathbf{x}_p$ representado no espaço original. Tomamos então o valor absoluto em cada elemento do vetor $\mathbf{x}_r$ para que possamos convertê-lo em uma imagem.

Com relação ao efeito da *quadtree* no resultado da reconstrução, seja $P^*_{wpca} \in \mathbb{R}^{m\times n}$ a matriz do WPCA gerada na etapa (6) da seção 2.3. Então, dado $\mathbf{z} = \mathbf{x} - \overline{\mathbf{x}} \in \mathbb{R}^n$, o problema de reconstrução formalizado na expressão (27) pode ser escrito como:

$$\mathbf{x}_r = \left(P^*_{wpca}\right)^T P^*_{wpca}\mathbf{z} + \overline{\mathbf{x}}.$$

Renomeando $A \equiv P^*_{wpca}\left(P^*_{wpca}\right)^T$, consideremos agora uma variação da expressão acima na forma:

$$\mathbf{x}_r(\epsilon) = (A + \epsilon F)\,\mathbf{z} + \overline{\mathbf{x}}, \tag{28}$$

onde $F \in \mathbb{R}^{n\times n}$ representa a perturbação na matriz A, suposta com efeito aditivo por simplicidade, gerada por artefatos no mapa de pesos e $\epsilon \in \mathbb{R}$. Então:

$$\mathbf{x}_r(\epsilon) = A\mathbf{z} + \overline{\mathbf{x}} + \epsilon F\mathbf{z}, \tag{29}$$

logo:

$$\mathbf{x}_r(\epsilon) - \mathbf{x}_r(0) = \epsilon F\mathbf{z}. \tag{30}$$

permitindo concluir que:

$$\frac{\|\mathbf{x}_r(\epsilon) - \mathbf{x}_r(0)\|}{\|\mathbf{z}\|} = \frac{\|\epsilon F\mathbf{z}\|}{\|\mathbf{z}\|} \leq |\epsilon|\,\|F\|, \tag{31}$$

onde $\|\cdot\|$ representa a norma de Frobenius. Neste caso, o efeito da *quadtree* é $\|F_{qtree}\| \leq \|F\|$, pois a aplicação da *quadtree*, gerando F_{qtree}, remove pequenas variações locais das intensidades dos pixeis. Portanto:

$$\frac{\|\mathbf{x}_r(\epsilon) - \mathbf{x}_r(0)\|}{\|\mathbf{z}\|} \leq |\epsilon|\,\|F_{qtree}\| \leq |\epsilon|\,\|F\|, \tag{32}$$

e, consequentemente, o erro relativo pode diminuir, indicando um resultado mais *suave* para a reconstrução.

4. RESULTADOS

Foram conduzidos experimentos computacionais objetivando verificar a eficácia do PCA ponderado para a reconstrução e classificação de grupos de dados, seguindo a metodologia dos trabalhos [24,38,45] e utilizando as abordagens propostas na seção 2. Trata-se de uma síntese dos principais resultados apresentados na dissertação [44].

Para os experimentos a seguir, foram utilizadas imagens do banco de faces da FEI [41], onde cada pessoa neste banco de imagens possui uma foto de seu rosto com a expressão facial neutra e outra com a expressão facial sorrindo, como exemplificam as Figuras 1-(a),(b). No total, há 400 imagens, sendo 200 de homens e 200 de mulheres. Estas imagens originalmente possuem tamanho de 250×300 pixels, centralizadas, utilizando os olhos e nariz como referência de localização, para diminuir a variância amostral. Outra característica dessas imagens é que todas estão em tons de cinza, com intensidade de pixel variando de 0 a 255.

Nos experimentos desta seção também foram utilizadas imagens do banco de faces FERET [42] com características análogas ao banco da FEI. Porém, neste banco, as imagens apresentam uma porção menor de fundo, se comparadas com as imagens da FEI, como pode ser observado nas Figuras 1-(c),(d). Portanto, proporcionalmente, há mais pixeis com informações potencialmente relevantes para o reconhecimento facial nas imagens FERET. Também há 400 imagens nesta base, das quais 214 são de homens e 186 são de mulheres (em cada caso, com ambas as expressões faciais neutra e sorrindo). As imagens em questão possuem resolução de 128×128 pixels. Assim, inicialmente, foi feito um redimensionamento para 128×128 pixels para as imagens do banco de dados da FEI. Ademais, os experimentos foram separados em dois tipos: reconstrução e classificação, sem e com aplicação da *quadtree*, para ambas as bases de dados FEI e FERET, envolvendo homens e mulheres com as duas expressões faciais.

No caso da reconstrução, foram realizados experimentos com o objetivo de analisar o efeito da ponderação nas características faciais guardadas pelas direções principais quando variamos o número de componentes principais. Para a classificação, foram utilizados os classificadores DM e KNN. Objetivamos aqui verificar o efeito da ponderação para distinguir automaticamente diferentes grupos faciais. As estatísticas da classificação foram obtidas usando o método $K-$fold com $K = 10$.

Todos os experimentos abordados nesta seção para as imagens da FEI foram conduzidos de forma que separamos 40 imagens de teste, das 400 imagens disponíveis, dispondo então de 360 imagens para a etapa de treinamento e, consequentemente, de 359 componentes principais, uma vez que nos testes realizados não foram identificados autovalores nulos para essas componente. No caso da FERET, separamos também 40 imagens de teste, das 400 imagens do banco de dados, para o experimento de expressão facial. Porém, pela quantidade de imagens existentes envolvendo gênero (214 imagens de rostos masculinos e 186 imagens de rostos femininos), para o experimento de gênero foram separadas 20 imagens de homens e 19 imagens de mulheres para teste, a fim de manter as mesmas proporções do caso da FEI, totalizando 39 imagens para teste. Assim, teremos 360 e 361 imagens de treinamento em cada experimento para a FERET. Para fins de comparação utilizaremos 359 componentes principais em ambos os experimentos, considerando que não foram encontrados autovalores nulos para essas componentes nos experimentos com as bases de imagens consideradas. Os resultados obtidos podem ser visualizados nos gráficos das Figuras 6 e 7, do Apêndice A. As seções 4.1-4.3 descrevem os principais achados obtidos a partir destes gráficos.

4.1 Mapas de Ponderação

Num primeiro momento, apresentamos as imagens dos mapas de ponderação para a base de dados da FEI e FERET na Figura 2. Esses mapas foram gerados usando os métodos descritos na seção 2.2, computados sobre o conjunto de treinamento da FEI e sobre o conjunto de treinamento da FERET. Primeiramente, considerando as duas primeiras linhas, notamos que os mapas $\widehat{H}$, H e JS são bem ruidosos quando comparados com os demais, enquanto o mapa W é sempre o mapa com mais artefatos. A correlação entre os mapas $\widehat{H}$ e H ($\widehat{H} = 1/H$) justifica o fato do mapa $\widehat{H}$ apresentar tons de cinza mais elevados nas regiões onde o mapa H é mais escuro. Em todos os casos, os mapas $\widehat{H}$ e H são mais limitados, do ponto de vista de guardar características faciais mais definidas, quando comparados com os demais.

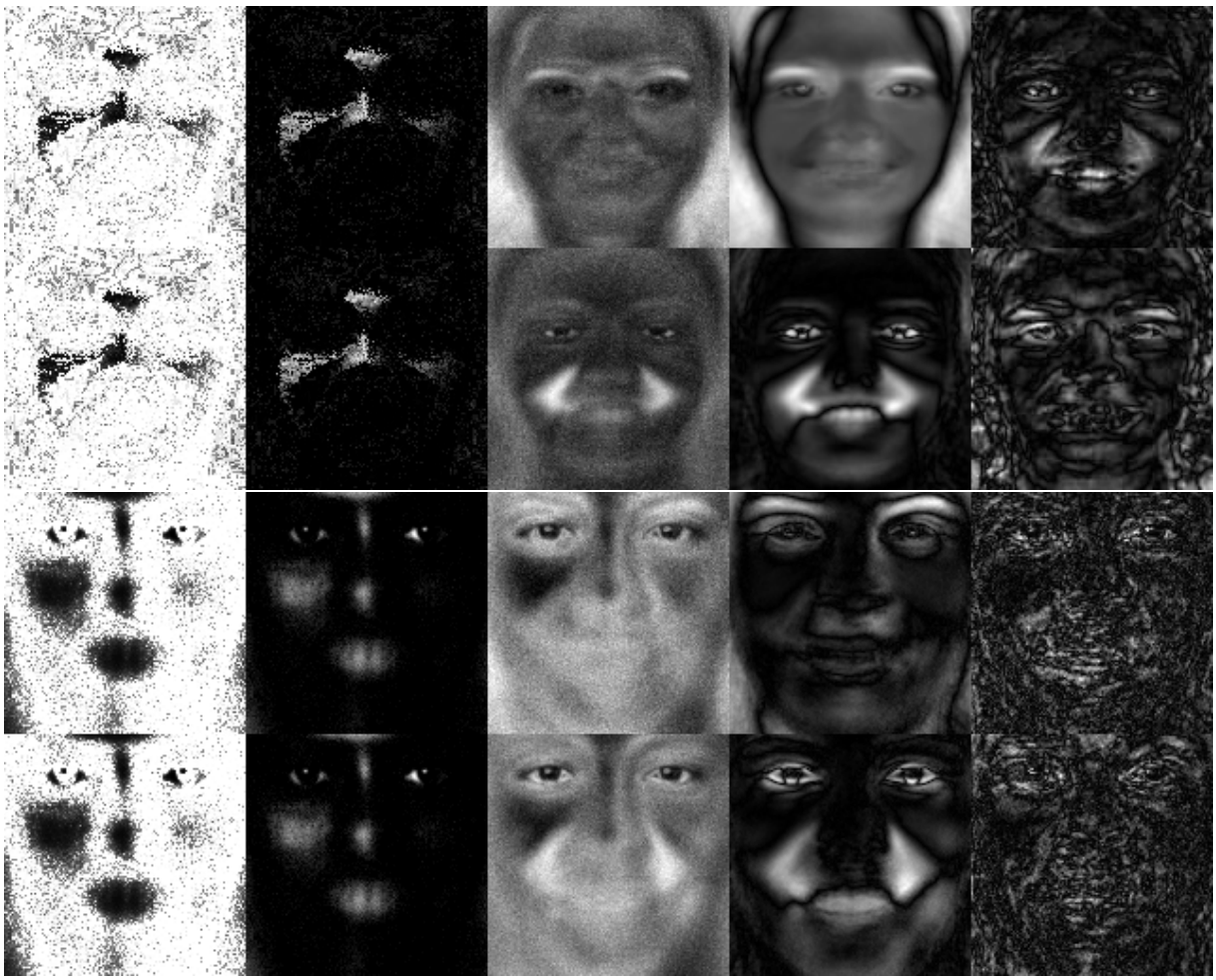


Figura 2: Reprodução dos mapas de ponderação utilizados para calcular o PCA ponderado (da esquerda para a direita): $\widehat{H}$, H, JS, T e W. $1^{\underline{a}}$ linha: separação por gênero para FEI; $2^{\underline{a}}$ linha: separação por expressão facial para FEI, $3^{\underline{a}}$ linha: experimento com separação por gênero para FERET; $4^{\underline{a}}$ linha: experimento com separação por expressão facial para FERET

Comparando os mapas relativos aos experimentos com a base de dados FERET, na $3^{\underline{a}}$ e $4^{\underline{a}}$ linhas da Figura 2, com os correspondentes das duas primeiras linhas, notamos bastante diferença nos mapas $\widehat{H}$, principalmente nas regiões da boca, bochechas e olhos. Em termos de percepção visual, as imagens do $\widehat{H}$ nas duas últimas linhas permitem reconhecer caricaturas de faces humanas, o que não é tão óbvio nas imagens dos mapas $\widehat{H}$ da FEI. Esse fato é consequência do que observamos para o mapa H,

https://doi.org/10.21528/lnlm-vol23-no2-art3

onde as diferenças que se destacam quando comparamos as imagens dos mapas da FEI e FERET estão nas regiões dos olhos e da boca.

4.2 Reconstrução

Com respeito às reconstruções, a pergunta fundamental é: qual foi o melhor método nos cenários considerados? Além da análise visual, vamos considerar o $SSIM$ para responder essa pergunta. Pelas Figuras 3, 4 e 5, percebe-se que, usando $d > 5$ componentes principais, já é possível identificar a expressão facial das pessoas das imagens originais reproduzidas na base das Figuras. Porém, é possível verificar que sem a aplicação da *quadtree* há perda de nitidez nas imagens resultantes, mesmo com todas as 359 componentes principais, devido à presença de ruídos e artefatos causados pelos métodos, em especial pelos métodos T e W. O ruído causado pelos métodos $\widehat{H}$, H e JS, apesar de existir, não se destaca tanto nas imagens resultantes por ocorrer de forma mais dispersa nos rostos das imagens. Contudo, é possível notar que, ao aplicarmos a *quadtree* (Figura 4) esses ruídos e artefatos são suavizados, fato justificado no final da seção 3, e as imagens resultantes tornam-se mais nítidas, particularmente no caso dos métodos T e W. Isso vai se refletir no desempenho de alguns métodos com respeito ao $SSIM$, como será verificado na Tabela 2.

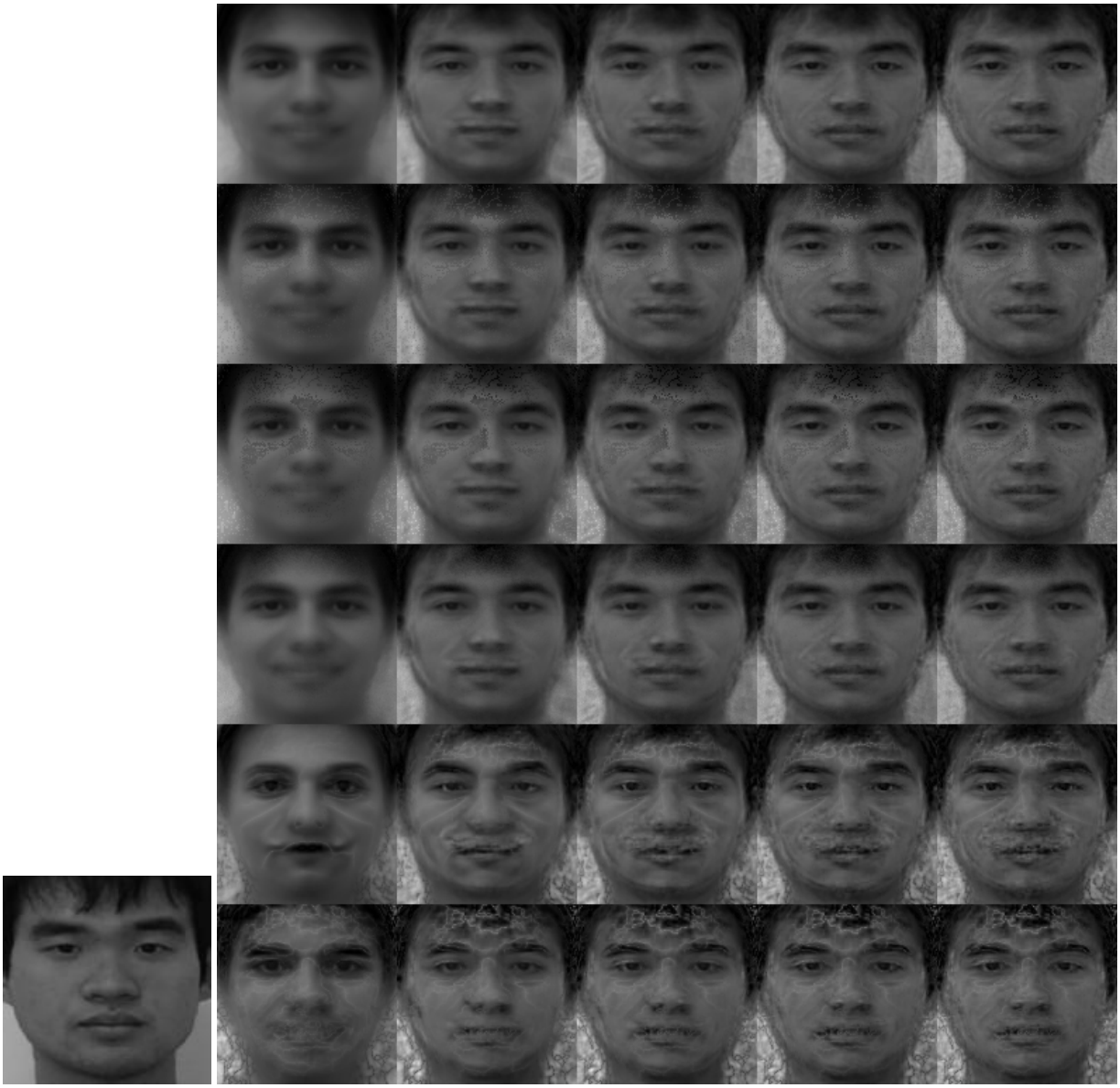


Figura 3: Experimento FEI com separação de expressão facial sem uso de quadtree: Reconstrução de uma das imagens de teste, reproduzida na base, através dos métodos (de cima para baixo): PCA, $\widehat{H}$, H, JS, T e W, com projeção nos subespaços (da esquerda para a direita) 5, 80, 160, 250 e 359.

A Tabela 2 mostra, para cada subespaço considerado, qual foi o método com maior $SSIM$ médio, quando comparamos as imagens de teste (40 para a FEI e 40 para FERET) com suas versões reconstruídas. Além disso, essa tabela também mostra, para esses métodos, seus respectivos valores de $SSIM$. Os gráficos da Figura 6 (Apêndice A) permitem analisar a evolução do $SSIM$ para cada método ao longo dos diferentes subespaços.

https://doi.org/10.21528/lnlm-vol23-no2-art3

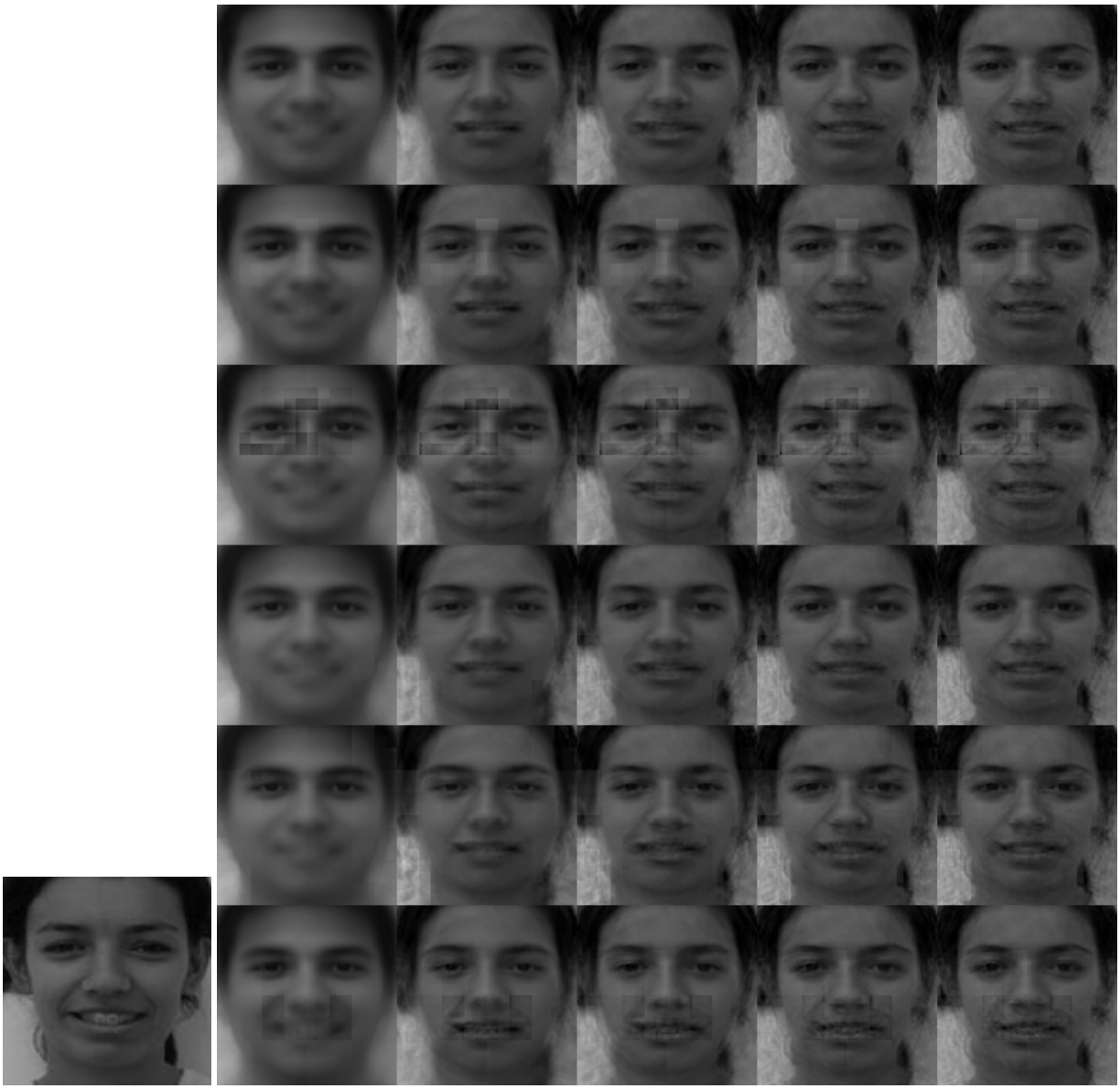


Figura 4: Experimento FEI com separação de gênero usando quadtree: Reconstrução de uma das imagens de teste, reproduzida na base, através dos métodos (de cima para baixo): PCA, $\widehat{H}_{qt}$, H_{qt}, JS_{qt}, T_{qt} e W_{qt}, com $t = 0.8$, com projeção nos subespaços (da esquerda para a direita) 5, 80, 160, 250 e 359.

d	5	10	20	40	80	160	250	320	359
(i) Gênero FEI	JS_{qt} (0.6348)	PCA (0.6191)	PCA (0.6970)	PCA (0.7159)	PCA (0.7340)	PCA (0.7509)	PCA (0.7591)	PCA (0.7620)	**PCA** (**0.7630**)
(ii) Exp. Facial FEI	W_{qt} (0.6699)	W_{qt} (0.6875)	W_{qt} (0.7014)	W_{qt} (0.7198)	W_{qt} (0.7227)	**PCA** **0.7271**	$\{PCA, W_{qt}\}$ (0.7255)	PCA (0.7248)	PCA (0.7241)
(iii) Gênero FERET	PCA (0.3402)	JS_{qt} (0.3515)	JS_{qt} (0.3666)	JS (0.3746)	JS (0.3794)	JS (0.3827)	JS (0.3846)	JS (0.3867)	**JS** (**0.3878**)
(iv) Exp. Facial FERET	W_{qt} (0.3165)	JS_{qt} (0.3290)	W_{qt} (0.3386)	JS (0.3469)	JS (0.3508)	JS (0.3519)	JS (0.3539)	JS (0.3553)	**JS** (**0.3564**)

Tabela 2: Tabela com melhor método para reconstrução, segundo o $SSIM$ que faz a comparação entre a reconstrução e a imagem original. Em negrito estão destacados os melhores métodos para cada linha.

https://doi.org/10.21528/lnlm-vol23-no2-art3

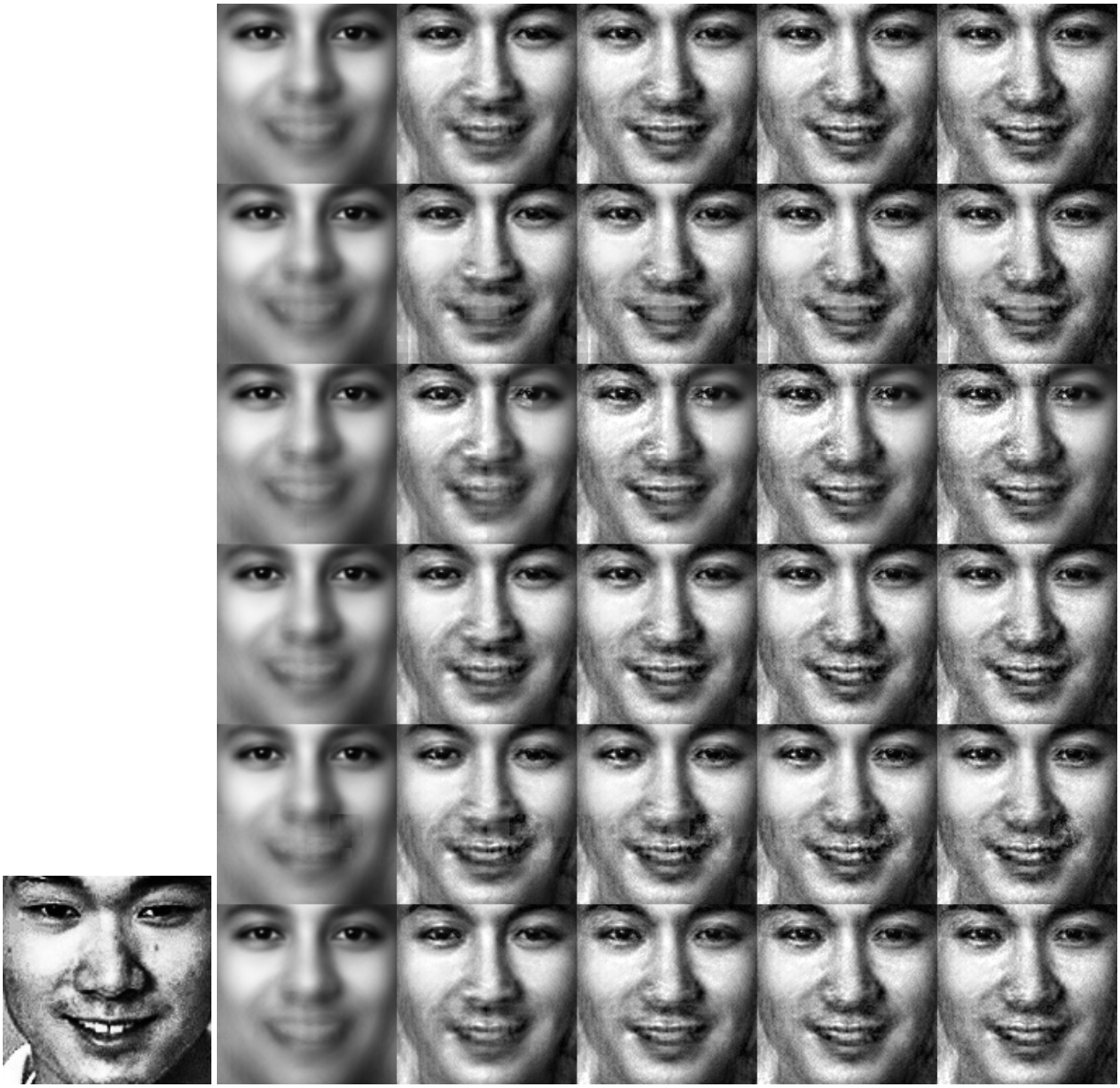


Figura 5: Experimento FERET com separação de expressão facial usando quadtree: Reconstrução de uma das imagens de teste, reproduzida no topo, através dos métodos (de cima para baixo): PCA, $\widehat{H}_{qt}$, H_{qt}, JS_{qt}, T_{qt} e W_{qt}, com $t = 0.8$, com projeção nos subespaços (da esquerda para a direita) 5, 80, 160, 250 e 359.

Ao considerarmos os dados da linha (i) da Tabela 2, percebemos que, com exceção de $d = 5$, o PCA é o método que melhor reconstrói as imagens de teste deste experimento de gênero com os dados da FEI, alcançando maior $SSIM$ igual a 0.7363, com $d = 359$. Porém, notemos nesta mesma linha que com $d = 320$ o PCA atinge $SSIM = 0.7620$, bem próximo do máximo.

Considerando variação de expressão facial, observando a linha (ii) da Tabela 2, vemos que o melhor método para a reconstrução com poucas componentes principais é o W_{qt} atingindo $SSIM = 0.7227$ para $d = 80$. No entanto, aumentando o número de componentes principais, o método que melhor reconstruiu as imagens foi o PCA, com 160 componentes principais, atingindo neste caso $SSIM = 0.7271$.

Abordando agora o caso dos experimentos com o banco de imagens FERET, na linha (iii) da Tabela 2, podemos ver que referente ao experimento de gênero com a base de dados FERET, percebe-se que, com exceção de $d = 5$, os métodos JS_{qt} e JS são os que melhor reconstroem as imagens de teste. O melhor método nesta seção foi JS com $d = 359$ e $SSIM = 0.3878$. Porém, notamos que para $d = 320$ temos $SSIM = 0.3867$ bem próximo do valor máximo. Assim, ao contrário do experimento análogo para as imagens da FEI (linha (i)), a alternativa mais eficiente neste caso é o método JS, que gera reconstruções mais próximas das imagens originais.

Analisando as informações contidas na linha (iv) da Tabela 2, considerando o experimento com variação de expressão facial com as imagens da base FERET, as imagens foram melhor reconstruídas, com poucas componentes principais, pelo método W_{qt}. Utilizando 40 componentes principais ou mais, vemos que o melhor $SSIM$ é alcançado pelo método JS, o qual melhor reconstrói as imagens neste experimento. Para 40 componentes observamos $SSIM = 0.3469$ indicando um resultado satisfatório para a reconstrução, o que pode ser verificado nas imagens da Figura 5. Assim, no caso das imagens da FERET, os dois experimentos de reconstrução realizados envolvendo separação de gênero (linha (iii)) e expressão facial (linha (iv)) indicaram superioridade do método JS.

4.3 Classificação

Nas Tabelas 3 e 4 fazemos um sumário dos melhores resultados obtidos nos experimentos de classificação. Para cada experimento, considerando os cenários sem e com *quadtree*, selecionamos o melhor método; ou seja, aquele que obteve a melhor acurácia média, sendo que no caso de empate, escolhemos o método correspondente ao subespaço de menor dimensão d. Os gráficos da Figura 7 (Apêndice A) permitem analisar a evolução da acurácia média, calculada sobre o conjunto de teste, para cada método ao longo dos diferentes subespaços escolhidos. A seguir, descrevemos os principais achados retirados a partir dos resultados destes gráficos.

Na Tabela 3, considerando o classificador DM, notamos que o PCA obteve melhores resultados tanto em situações de classificação de gênero quanto de expressão facial, obtendo acurácias sempre acima de 0.82, mesmo com apenas 10 componentes principais, como é o caso do experimento com expressão facial da base FERET. Por outro lado, quando consideramos o classificador KNN na Tabela 3 notamos que os melhores métodos foram T e $\widehat{H}$, mostrando a influência do método de classificação nos resultados. Neste caso, pelo classificador KNN e de acordo com a Tabela 3, no caso das imagens da FEI, o método $\widehat{H}$ destacou-se na classificação de gênero, atingindo acurácia de 0.8925 para 40 componentes principais. Já no caso das imagens do banco FERET, este mesmo método foi o melhor na classificação de expressão facial, obtendo acurácia 0.7899 com apenas 10 componentes. Por outro lado, o método T possui destaque no caso de classificação de expressão facial com os dados da FEI, chamando a atenção por obter boa acurácia (0.7699) para um subespaço de dimensão $d = 5$. Além disso, foi o mais eficiente também na classificação de gênero com a base FERET, com acurácia 0.7923 no subespaço de dimensão $d = 160$.

Sem *Quadtree*	DM	KNN
Experimento	Melhor;Acurácia;d	Melhor;Acurácia;d
FEI Gênero	$PCA; 0.8874; 320$	$\widehat{H}; 0.8925; 40$
FEI Exp. Facial	$PCA; 0.8574; 160$	$T; 0.7699; 5$
FERET Gênero	$PCA; 0.8282; 359$	$T; 0.7923; 160$
FERET Exp. Facial	$PCA; 0.8225; 10$	$\widehat{H}; 0.7899; 10$

Tabela 3: Tabela com melhor resultado para cada experimento de classificação sem uso de *quadtree*

Levando em consideração o uso de *quadtree* nos experimentos, a Tabela 4 resume os melhores desempenhos na classificação obtidos pelos métodos abordados. O critério para a exibição do melhor método nesta etapa é o mesmo critério utilizado para os experimentos sem *quadtree*.

Segundo a Tabela 4, e de acordo com o classificador DM, percebemos que, com a base de dados da FEI, os melhores resultados pertencem aos métodos: PCA, para o experimento de gênero, com acurácia 0.8874 e $d = 320$; e $\widehat{H}_{qt}$, com limiar $t = 0.6$, para o experimento de expressão facial, obtendo taxa de reconhecimento de 0.8574 com $d = 80$. Nota-se que, neste caso, o uso da *quadtree* mostrou melhora significativa na classificação de expressão facial para o método H, visto que H_{qt} aparece na

https://doi.org/10.21528/lnlm-vol23-no2-art3

Uso de *quadtree*	DM	KNN
Experimento	Melhor;Acurácia;d;t	Melhor;Acurácia;d;t
FEI Gênero	$PCA; 0.8874; 320$	$H_{qt}; 0.8949; 40; t = 0.4$
FEI Exp. Facial	$\widehat{H}_{qt}; 0.8574; 80; t = 0.6$	$T; 0.7699; 5$
FERET Gênero	$PCA; 0.8282; 359$	$\widehat{H}_{qt}; 0.7930; 250; t = 0.8$
FERET Exp. Facial	$JS_{qt}; 0.8299; 80; t = 0.8$	$\widehat{H}; 0.7899; 10$

Tabela 4: Tabela com o melhor resultado para cada experimento de classificação com o uso de *quadtree*.

Tabela 4, com taxa $0,8574$ com $d = 80$, enquanto H não aparece na Tabela 3. Já para os experimentos com a base de dados FERET, os métodos mais eficientes são: PCA, para o experimento de classificação de gênero, com acurácia 0.8282 e $d = 359$; e JS_{qt}, com limiar $t = 0.8$, obtendo taxa de 0.8299 com $d = 80$, para a classificação de expressão facial.

Ainda de acordo com a Tabela 4, mas agora observando o classificador KNN, no caso dos dados da FEI, vemos que o melhor método para classificar gênero é o H_{qt}, com limiar $t = 0.4$, obtendo acurácia de 0.8949 com $d = 40$. Para classificar expressão facial, o melhor método é o T, com acurácia 0.7699 com apenas 5 componentes principais. Por outro lado, no caso dos dados da base FERET, percebemos que o melhor método para classificar gênero é o $\widehat{H}_{qt}$, com limiar $t = 0.8$, atingindo a taxa de reconhecimento de 0.7930 com 250 componentes principais, e o método mais eficaz para classificar expressão facial é o $\widehat{H}$, com taxa de 0.7899 com apenas 10 componentes principais.

Observando a Tabela 3 vemos que o mapa H não aparece como melhor método em nenhuma linha da coluna do KNN. Porém, quando utilizamos sua versão com *quadtree*, a Tabela 4 nos mostra que o método H_{qt} figura como melhor método em um cenário na coluna do KNN. Fazendo estatísticas para os autovalores correspondentes a H e H_{qt} notamos um aumento da faixa de variação dos autovalores observado para H_{qt}, o que foi positivo para a classificação. Isso é um indicativo da influência da dispersão dos dados para a classificação.

Com relação ao efeito dos artefatos observados para método W para classificação, notamos que esse método não aparece em nenhuma linha das Tabelas 3 e 4, apontando para a influência negativa deste fato na classificação, o que será discutido na seção 5.

Outro ponto importante a se destacar são as conclusões quando consideramos um cenário em que são necessárias uma boa reconstrução e uma boa taxa de classificação. A Tabela 5 sumariza os resultados apresentados em [44], alguns deles reproduzidos nas Tabelas 2-4 em relação a este item. Especificamente, pelas Tabelas 2-4 fica claro que, no caso do experimento de gênero para a base da FEI com o classificador DM, o PCA é o melhor método tanto para reconstrução quanto para classificação. A comparação das Tabelas 2-4 justifica também os resultados da Tabela 5 para FERET com expressão facial e classificador DM. Para os demais resultados reportados na Tabela 5, a justificativa está na comparação dos gráficos correspondentes do Apêndice A. Por exemplo, quando usamos o classificador KNN e consideramos o experimento de gênero com a base FEI, não há um consenso nas Tabelas 3-4. Porém, observando as Figuras 6.(a) e 7.(b) do Apêndice A, fica claro que, novamente, o PCA é a melhor escolha tanto para a classificação quanto para a reconstrução.

Com e sem *quadtree*	DM	KNN
Experimento	Reconstrução+Classificação	Reconstrução+Classificação
FEI Gênero	PCA	PCA
FEI Exp. Facial	W_{qt}	T
FERET Gênero	JS	JS
FERET Exp. Facial	JS_{qt}	$\widehat{H}$

Tabela 5: Tabela com melhor método para cada experimento, considerando cenário onde necessitamos de boa classificação e boa reconstrução.

Novamente, na Tabela 5, notamos a influência do classificador no resultado. Nos experimentos referentes à base de dados da FEI, notamos que, para o classificador DM, para o experimento de classificação de gênero, o PCA é o melhor método, e para o experimento de classificação de expressão facial, o método W_{qt} com limiar $t = 0.8$ é o mais eficaz. Os resultados do classificador

KNN coincidem com o DM para os experimentos e gênero nas duas bases. Porém, o KNN mostra que os melhores métodos para os experimentos de expressão facial para FEI e FERET foram, respectivamente, os métodos T e $\widehat{H}$.

5. DISCUSSÃO

De maneira geral, é possível perceber que há diferenças entre os resultados obtidos através da base da FEI e os resultados obtidos através da base FERET. Apesar de ambas as bases envolverem imagens de faces humanas com os dois gêneros e as mesmas expressões faciais, foi observado nos resultados experimentais que os mapas computacionais gerados são diferentes nos dois casos. Consequentemente, o PCA ponderado gera resultados distintos para cada base tanto para a classificação quanto para a reconstrução.

Especificamente, com relação às variações de desempenho observadas nas Tabelas 2-5 entre as imagens da FEI e FERET, é importante ter em mente que o desempenho do PCA ponderado, na forma proposta neste trabalho, depende das propriedades dos mapas de ponderação, mostrados na Figura 2. Assim, o aprofundamento da análise destas diferenças passa pela análise de cada técnica em conjunto com as características de cada banco de imagens (Figura 1).

A análise da Tabela 5 mostra preponderância do método JS, ou sua versão JS_{qt}, para as imagens FERET quando se deseja boa taxa de reconstrução e boa acurácia. Porém, no caso da FEI, JS ou JS_{qt} nunca aparece como o melhor método nesta tabela. O método JS quantifica diferenças entre as classes considerando as distribuições de probabilidades associadas às classes, pixel-a-pixel. Essas diferenças, destacadas nas regiões mais claras dos mapas JS da Figura 2, foram significativas para a reconstrução das imagens FERET, como destacado na Tabela 2. Porém, no caso do mapa JS para gênero da FEI, a observação dos mapas correspondentes na Figura 2 mostra menor relevância dos pesos na região da face, prejudicando a performance do método. No caso do mapa JS para separação de expressão facial da FEI, as regiões da bochecha, próximas do nariz, boca e fundo dos olhos têm maiores intensidades. Porém, estes fatos não foram significativos para a classificação, visto que o JS não aparece na Tabela 3.

De acordo com a apresentação da seção 2.2, no mapa H, computado via entropia de Shannon, estaremos dando maior peso para pixels com distribuição de intensidades com mais dispersão em relação à média. A hipótese por trás disso é que esses pixeis sejam mais relevantes para as tarefas de reconhecimento facial. O mapa $\widehat{H}$ considera o inverso; ou seja, dará mais peso para os pixeis cujas distribuições de intensidade sejam menos dispersas em relação à média. No caso das imagens da FEI (Figuras 1.(a)-(b)), os pixeis do fundo da imagem têm intensidades próximas nas diferentes imagens, gerando a região mais clara (embora ruidosa) próxima da borda da imagem no mapa $\widehat{H}$ mostrado na Figura 2. Essas regiões avançam para o interior da face devido proximidades de tons de pele, uma vez que não há variação considerável de tons de pele, havendo também controle da iluminação durante a aquisição. No caso do FERET, algo análogo acontece, embora nestas imagens a região do fundo seja mais reduzida, como pode ser verificado nas Figuras 1.(c)-(d). Esse funcionamento realçou regiões no mapa $\widehat{H}$ da FEI que, para o classificador KNN foram mais importantes para a classificação de gênero, visto que o mapa $\widehat{H}$ foi o melhor método neste caso, segundo a Tabela 3, o mesmo acontecendo para a expressão facial para as imagens FERET. Porém, no caso do classificador DM, $\widehat{H}$ nunca aparece como melhor método, embora para o subespaço com dimensão $d = 40$, os resultados apresentados na dissertação [44] mostram que o mapa $\widehat{H}$ forneceu resultados melhores para a classificação do que o mapa H.

Ainda pela Tabela 3, no caso da classificação de gênero para FERET pelo KNN, os realces obtidos para H e $\widehat{H}$ não foram preponderantes, mas sim aqueles gerados pelo mapa T (Figura 2) o qual considera diferenças entre médias locais, normalizadas em função das variâncias intra-classe das distribuições de intensidade de cada pixel. Como observado na Figura 2, o mapa correspondente destaca pixeis em regiões como sobrancelha, proximidade dos olhos e bigode, que são importantes para a classificação de gênero. O realce nessas regiões foi importante para o classificador KNN mas não se destacou para classificação de gênero no caso do classificador DM para nenhum dos subespaços do PCA considerados, o que pode ser observado na Tabela 3, bem como na Figura 7.(a).

Essas análises mostram que o desempenho de cada método depende não somente das características do mapa gerado, as quais estão diretamente relacionadas com as características de cada banco de dados, mas também do classificador utilizado. Com os testes realizados até o momento, os padrões encontrados ocorreram na primeira linha da Tabela 5, onde o PCA supera os demais para os dois classificadores. Além disso, pela Tabela 4 há preponderância do PCA para classificação de gênero nas duas bases, usando o classificador DM. Na reconstrução, o PCA é o melhor método para a FEI e o JS é o melhor para a FERET (Tabela 2).

Outras propostas de ponderação do PCA foram aplicadas em contextos como detecção de falhas [46], expressão genética [47], classificação de arritmia cardíaca [48] , dentre outros [49,50], indicando que a abordagem proposta pode ser promissora em outras aplicações. Porém, as observações abaixo devem ser consideradas para explorar o PCA ponderado de forma eficiente:

- O mapa H não depende de classes pré-definidas para seu cálculo, podendo então ser aplicado a problemas não-supervisionados.
- Os mapas T, JS e W dependem de classes previamente definidas para seu cálculo, ficando assim restritos a aplicações com supervisão.
- Analogamente ao caso do PCA, o nível de redução de dimensionalidade depende da dispersão dos dados na base definida pelos autovetores da matriz $(Z^*)^T Z^*$. Por exemplo, se considerarmos imagens de faces adquiridas sem controle de iluminação e fundo, a taxa de redução tanto para reconstrução quanto para classificação poderá ficar baixa.

- Em [44] foi demonstrado que as relações entre o maior (λ_{max}) e o menor autovalor (λ_{min}) do PCA tradicional e os correspondentes do PCA ponderado (λ^*_{max} e λ^*_{min}, respectivamente) são: $\lambda^*_{max} \leq \lambda_{max}$ e $\lambda^*_{min} \leq \lambda_{min}$. Assim, a ponderação dada pela expressão (25) pode reduzir efeitos indesejáveis decorrentes da escala de variação dos atributos envolvidos no problema. Porém, dependendo do tamanho do intervalo $[\lambda^*_{min}, \lambda^*_{max}]$ pode também ocorrer mais dificuldade para selecionar as componentes principais.

Um outro ponto importante é uma justificativa para o efeito dos artefatos observados para o método W na classificação. Por exemplo, no caso da Distância de Mahalanobis, seja $\mathbf{y}$ um dado no conjunto de teste, representado no espaço do $WPCA$. O classificador baseado na Distância de Mahalanobis, denotado por DM, classifica o dado de teste $\mathbf{y}$ como pertencente a classe $c(\mathbf{y})$ que satisfaz:

$$c(\mathbf{y}) = \begin{cases} +1, & \text{se} \quad \arg\min\{d_1(\mathbf{y}), d_2(\mathbf{y})\} = 1, \\ -1, & \text{se} \quad \arg\min\{d_1(\mathbf{y}), d_2(\mathbf{y})\} = 2. \end{cases} \tag{33}$$

onde:

$$d_i(\mathbf{y}) = \sqrt{\sum_{j=1}^{m} \frac{1}{\lambda_j}(y_j - \overline{y}_{ij})^2}, \quad i = 1, 2, \tag{34}$$

com λ_j sendo o autovalor correspondente à componente principal j, m a quantidade de componentes principais utilizadas, e $\overline{\mathbf{y}}_i = (\overline{y}_{i1}, \overline{y}_{i2}, \ldots, \overline{y}_{im}) \in \mathbb{R}^m$ a média das amostras da classe i, no espaço do $WPCA$. A equação (34) define a distância de Mahalanobis, computada no espaço de características do $WPCA$, entre o ponto $\mathbf{y}$ e a média $\overline{\mathbf{y}}_i$. Podemos modelar os artefatos observados nas imagens do método W como perturbações no espaço de características e nos autovalores, decorrentes de ruídos ou artefatos nos dados ponderados, gerados pela ponderação com o mapa W, que podem influenciar o resultado da expressão (33) alterando a classificação. Algo análogo ocorre para o KNN. Formalmente, seja $\widetilde{R} = R + \Delta$ a matriz de covariância do $WPCA$ computada sobre os dados ponderados, definida por uma perturbação Δ da matriz de covariância R, e P a matriz contendo todos os autovetores de $\widetilde{R}$. Definamos também $D = diag(d_1, d_2, \ldots, d_n)$ como a matriz diagonal dos autovalores de $\widetilde{R}$. Então:

$$P^T \widetilde{R} P = D \quad \Longleftrightarrow \quad P^T (R + \Delta) P = D, \tag{35}$$

e consequentemente:

$$P^T R P = D - P^T \Delta P \equiv \widehat{D} + F, \tag{36}$$

onde $\widehat{D} = D - diag\left(P^T \Delta P\right)$ e $F \equiv \{f_{ij}\} = diag\left(P^T \Delta P\right) - P^T \Delta P$. Consequentemente, $diag(F) = \mathbf{0}$. Então, pela teoria de perturbação para autovalores [51], $\lambda(R) \subset \cup_i^n [d_i - r_i, d_i + r_i]$, $r_i = \sum_{j=1}^n |f_{ij}|$. Esse resultado mostra com que intensidade os autovalores da matriz R podem ser deslocados em função da perturbação Δ, podendo influir significativamente no resultado da expressão (34) e na classificação dada pela expressão (33), dependendo do valor de $\max_{1 \leq i \leq n} r_i$.

O sucesso do PCA em tarefas de classificação relatadas nas Tabelas 3-4 chama a atenção bem como o sucesso do método JS na reconstrução da FERET (Tabela 2). As componente principais associadas aos maiores autovalores do PCA, estão relacionadas com características de maior variância no conjunto de dados, que são as mais importantes para reconstrução. É esperado que o subespaço correspondente seja eficiente para reconstrução, podendo também ser bom para tarefas de classificação relacionadas a essas características. Isso explica o sucesso do PCA para reconstrução no caso da FEI. Porém, não explica por que o PCA foi o melhor método para a classificação nas situações observadas, visto que está competindo com métodos direcionados para realçar pixeis importantes para separação entre classes. Por outro lado, o sucesso do JS para reconstrução também surpreende, visto que sua formulação prioriza pixeis relevantes para as características que têm maior importância para tarefas de classificação, ponderando com valor menor pixeis de outras regiões. Mais testes precisam ser feitos para responder essas dúvidas de forma conclusiva.

6. CONCLUSÃO

O principal objetivo deste trabalho foi investigar a eficiência de novas técnicas de ponderação para calcular o PCA ponderado, em experimentos de gênero e expressão facial, considerando problemas de classificação e reconstrução. Como principais conclusões dos experimentos realizados com a base da FEI e com a base FERET podemos destacar:

1. As reconstruções tiveram os artefatos das imagens resultantes suavizados pela aplicação da *quadtree*. Isto se deve ao potencial da quadtree para reduzir pequenas variações locais dos mapas estatísticos o que se reflete em mais suavidade para as reconstruções, um efeito justificado teoricamente no final da seção 3;

2. Considerando a Tabela 2, os melhores métodos para reconstrução são: PCA e JS, com $SSIM$ acima de 0.72 e 0.35, respectivamente. O resultado para o método JS não é esperado, visto que esse método visa priorizar apenas pixeis de regiões importantes para a classificação. Mais testes são necessários para responder essa dúvida.

3. Comparando as Tabelas 3 e 4 vemos que a utilização da *quadtree* pode melhorar o resultado da classificação, visto que o método H não aparece como melhor método em nenhuma linha da primeira tabela, mas sua versão com *quadtree* (H_{qt})

aparece na primeira linha do classificador KNN na Tabela 4. Esse fato pode ser justificado pelo efeito de suavização dos mapas de características proporcionado pela aplicação da *quadtree*. Esse efeito pode diminuir a intensidade da perturbação Δ que aparece nas expressões (35)-(36) melhorando a classificação.

4. Para o classificador KNN, para o experimento de gênero da FEI o melhor método é o $\widehat{H}_{qt}$. Para a expressão facial o método T é o melhor para a FEI. No caso da FERET, ainda para o KNN, o método $\widehat{H}_{qt}$ é superior para gênero, enquanto que para expressão facial o melhor método é o $\widehat{H}$. A menor taxa de reconhecimento nestes casos foi para o método T, com acurácia de 0.7699 com apenas $d = 5$ componentes principais. Chama a atenção a influência da aplicação da *quadtree* observada quando comparamos as Tabelas 3 e 4. A análise da influência de artefatos na classificação feita no final da seção 5 traz uma justificativa para esse fato.

5. Quando consideramos um cenário onde necessitamos de boa classificação e boa reconstrução, a Tabela 5 mostra que, para o classificador DM os melhores métodos são PCA, JS, JS_{qt} e W_{qt} (com $t = 0.8$), enquanto que no caso do classificador KNN destacam-se os métodos PCA, T, JS e $\widehat{H}$. Novamente, observa-se a influência do classificador na lista dos melhores métodos.

6. Do ponto de vista de custo computacional, o método W é mais custoso, uma vez que demanda o treinamento do SSVM enquanto as abordagens propostas não dependem de nenhum tipo de treinamento, sendo computadas diretamente a partir dos dados. Isso é demonstrado pela análise realizada em [44], onde verificamos a seguinte ordem de complexidade: $\widehat{H}$ (Complexidade $O\left(n \cdot N\right)$); H (Complexidade $O\left(n \cdot N\right)$), JS (Complexidade $O\left(n \cdot N\right)$); T (Complexidade $O\left(n \cdot N\right)$); W (Complexidade $O\left(n \cdot N^2\right)$), considerando N imagens representadas por pontos em $\mathbb{R}^n$.

7. Apesar da justificativa apresentada após a equação (19) para a utilização da entropia H ser bem fundamentada, observa-se apenas um caso de sucesso desta técnica (sua versão *quadtree*, Tabela 4), enquanto que seu inverso (mapa $\widehat{H}$ ou $\widehat{H}_{qt}$) apresentou melhor desempenho em cinco cenários relatados nas Tabelas 3-4. Uma vez que as imagens nos dois bancos de dados foram adquiridas em ambientes controlados, passando por processos de normalização, existem pixeis fora das regiões destacadas nos mapas H da Figura 2 importantes para a classificação, embora as distribuições de intensidades correspondentes sejam mais concentradas em torno da média. Essas regiões são realçadas nos mapas $\widehat{H}$ influenciando positivamente a classificação. Isso é uma justificativa para o sucesso do mapa $\widehat{H}$ observado.

Por fim, para trabalhos futuros, pretendemos testar outras bases de imagens de faces, testar outros classificadores, para responder às questões levantadas no último parágrafo da seção 5. Serão também testados outros tipos de abordagens (como redes neurais convolucionais profundas) e comparação dos resultados procurando o melhor *setup*. Em particular, pretendemos treinar redes neurais profundas usando a representação das imagens nos diferentes subespaços das versões ponderadas do PCA. Se tivermos resultados satisfatórios, poderemos combinar os vetores de características do PCA ponderado com vetores de características gerados por redes neurais convolucionais e verificar a influência desse processo na classificação. Essa combinação pode acontecer por meio do empilhamento e/ou da concatenação das características. Os vetores de características assim combinados serviriam de entrada para uma rede *multi-layer perceptron* que faria a fusão e classificação final. A ideia é também comparar com resultados advindos das redes neurais sem essa combinação de características. Um procedimento análogo foi realizado em trabalhos envolvendo outros espaços de características com resultados promissores [52]. Por fim, podemos explorar o *Multi-Class Discriminant Analysis* [53] para computar pesos espaciais e estender o $WPCA$ para problemas multi-classe. Essa técnica combina classificadores binários para formar uma função discriminante global, que permite ordenar as componentes do espaço de acordo com a importância de cada característica para a classificação. Se executada no espaço dos pixeis, poderia gerar pesos espaciais que podem ser usados no WPCA proposto.

REFERÊNCIAS

[1] A. Soni, A. Rasool, A. Dubey and N. Khare. "Data Mining based Dimensionality Reduction Techniques". *2022 International Conference for Advancement in Technology (ICONAT)*, pp. 1–8, 2022.

[2] G. T. Reddy, M. P. K. Reddy, K. Lakshmanna, R. Kaluri, D. S. Rajput, G. Srivastava and T. Baker. "Analysis of Dimensionality Reduction Techniques on Big Data". *IEEE Access*, vol. 8, pp. 54776–54788, 2020.

[3] A. Chahid, T. N. Alotaiby, S. A. Alshebeili and T. Laleg-Kirati. "Feature Generation and Dimensionality Reduction using the Discrete Spectrum of the Schrödinger Operator for Epileptic Spikes Detection". *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2373–2376, 2019.

[4] I. Guyon, S. Gunn, M. Nikravesh and L. A. Zadeh, editors. *Feature Extraction: Fundations and Applications*. 2006.

[5] J. A. Lee and M. Verleysen. *Nonlinear Dimensionality Reduction*. Springer Publishing Company, Incorporated, first edition, 2007.

[6] Y. Ma and Y. Fu. *Manifold Learning Theory and Applications*. Taylor & Francis Group, 2012.

[7] B. Ghojogh, M. Crowley, F. Karray and A. Ghodsi. *Elements of Dimensionality Reduction and Manifold Learning*. Springer International Publishing, 2023.

[8] P. Li, Y. Pei and J. Li. "A comprehensive survey on design and application of autoencoder in deep learning". *Applied Soft Computing*, vol. 138, pp. 110176, 2023.

[9] A. K. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, Inc., 1989.

[10] G. F. Miranda, C. E. Thomaz and G. A. Giraldi. "Geometric Data Analysis Based on Manifold Learning with Applications for Image Understanding". *2017 30th SIBGRAPI Conference on Graphics, Patterns and Images - Tutorials*, pp. 42–62, 2017.

[11] J. Teenbaum, D. Silva and J. Langford. "global geometric framework for nonlinear dimensionality reduction [J]". *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[12] S. T. Roweis and L. K. Saul. "Nonlinear dimensionality reduction by locally linear embedding". *science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[13] L. v. d. Maaten and G. Hinton. "Visualizing data using t-SNE". *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[14] A. G. Mumuni and F. Mumuni. "Data augmentation: A comprehensive survey of modern approaches". *Array*, vol. 16, pp. 100258, 2022.

[15] S. Yang, W.-T. Xiao, M. Zhang, S. Guo, J. Zhao and S. Furao. "Image Data Augmentation for Deep Learning: A Survey". *ArXiv*, vol. abs/2204.08610, 2022.

[16] F. Yu, X. Xiu and Y. Li. "A Survey on Deep Transfer Learning and Beyond". *Mathematics*, 2022.

[17] H. mohammed Essa and A. M. Murshid. "Optimizing Image Processing with CNNs through Transfer Learning: Survey". *Al-Kitab Journal for Pure Sciences*, 2023.

[18] C. Fasana, S. Pasini, F. Milani and P. Fraternali. "Weakly Supervised Object Detection for Remote Sensing Images: A Survey". *Remote. Sens.*, vol. 14, pp. 5362, 2022.

[19] X. Yang, Z. Song, I. King and Z. Xu. "A survey on deep semi-supervised learning". *IEEE transactions on knowledge and data engineering*, vol. 35, no. 9, pp. 8934–8954, 2022.

[20] I. Fodor. "A Survey of Dimension Reduction Techniques". Technical report, 2002.

[21] T. Hastie, R. Tibshirani and J. Friedman. "The Elements of Statistical Learning". *Springer*, 2001.

[22] H. Safavi and C.-I. Chang. "Projection pursuit-based dimensionality reduction". *Proc. SPIE*, vol. 6966, pp. 69661H–69661H–11, 2008.

[23] E. Bingham and H. Mannila. "Random projection in dimensionality reduction: applications to image and text data". In *Knowledge Discovery and Data Mining*, 2001.

[24] C. E. Thomaz. "A Simple and Efficient Supervised Method for Spatially Weighted PCA in Face Image Analysis". Technical Report TR-2010-13, Department of Computing, Imperial College, London, UK, 2010.

[25] S. Yuan, X. Mao and L. Chen. "Multilinear Spatial Discriminant Analysis for Dimensionality Reduction". *IEEE Transactions on Image Processing*, vol. 26, pp. 2669–2681, 2017.

[26] L. Sirovich and M. Kirby. "Low-dimensional procedure for the characterization of human faces". *J. Opt. Soc. Am. A*, vol. 4, no. 3, pp. 519–524, Mar 1987.

[27] A. Franc. "Linear Dimensionality Reduction". *ArXiv*, vol. abs/2209.13597, 2022.

[28] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, New York, 1990.

[29] F. Lu and H. Li. "KLDA - An Iterative Approach to Fisher Discriminant Analysis". In *2007 IEEE International Conference on Image Processing*, volume 2, pp. II – 201–II – 204, 2007.

[30] Y. Tian, L. Wang, S. Yang, J. Ding, Y. Jin and X. Zhang. "Neural Network-Based Dimensionality Reduction for Large-Scale Binary Optimization With Millions of Variables". *IEEE Transactions on Evolutionary Computation*, pp. 1–1, 2024.

[31] Y.-J. Lee and O. L. Mangasarian. "SSVM: A Smooth Support Vector Machine for Classification". *Computational Optimization and Applications*, vol. 20, pp. 5–22, 2001.

[32] C. E. Thomaz, P. S. Rodrigues and G. A. Giraldi. "Using face images to investigate the differences between lda and svm separating hyper-planes". In *II workshop de visao computacional*, 2006.

[33] M. Spiegel and L. Stephens. *Schaum's Outline of Statistics*. Schaum's Outline Series. McGraw-Hill Education, 2007.

[34] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press., December 2000.

[35] R. Connor, F. A. Cardillo, R. Moss and F. Rabitti. "Evaluation of Jensen-Shannon Distance over Sparse Data". In *Similarity Search and Applications*, edited by N. Brisaboa, O. Pedreira and P. Zezula, pp. 163–168, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[36] V. Amaral, G. A. Giraldi and C. E. Thomaz. "LBP Estatístico Aplicado ao Reconhecimento de Expressões Faciais". In *proceedings of the 10th Encontro Nacional de Inteligencia Artificial e Computacional, ENIAC*, Fortaleza, Ceara, Brazil, 2013.

[37] V. Amaral, G. A. Giraldi and C. E. Thomaz. "Segmentacao espacial nao uniforme aplicada ao reconhecimento de genero e expressoes faciais". In *proceedings of the 11th Encontro Nacional de Inteligencia Artificial, ENIAC*, pp. 383–388, Sao Carlos, Sao Paulo, Brazil, 2014.

[38] V. do Amaral, G. A. Giraldi and C. E. Thomaz. "A Statistical Quadtree Decomposition to Improve Face Analysis". In *International Conference on Pattern Recognition Applications and Methods*, 2016.

[39] V. Amaral, G. A. Giraldi and C. E. Thomaz. "Statistical and cognitive spatial mapping applied to face analysis". In *Proceedings of the 28th SIBGRAPI, Conference on Graphics, Patterns and Images-Workshop of Works in Progress, Salvador, Bahia, Brazil*, 2015.

[40] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[41] C. Thomaz. "FEI Face Database". `https://fei.edu.br/~cet/facedatabase.html`, 2012. Acesso em: 2022-09-26.

[42] N. I. of Standards Technology. "color FERET Database". `https://www.nist.gov/itl/products-and-services/color-feret-database`, 2019. Acesso em: 2023-09-25.

[43] Z. Wang, A. Bovik, H. Sheikh and E. Simoncelli. "Image quality assessment: from error visibility to structural similarity". *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[44] L. C. P. Miranda. "Abordagens computacionais para calcular componentes principais ponderadas com aplicações em análise de imagens de faces humanas". Master's thesis, Laboratório Nacional de Computação Científica, Petrópolis, RJ, 2023. Orientador: Gilson Antonio Giraldi. Co-orientador Diego Barreto Haddad.

[45] C. E. Thomaz and G. A. Giraldi. "A new ranking method for principal components analysis and its application to face image analysis". *Image Vision Comput.*, vol. 28, no. 6, pp. 902–913, 2010.

[46] H. Yue and M. Tomoyasu. "Weighted principal component analysis and its applications to improve FDC performance". In *2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, volume 4, pp. 4262–4267 Vol.4, 2004.

[47] J. F. Pinto da Costa, H. Alonso and L. Roque. "A Weighted Principal Component Analysis and Its Application to Gene Expression Data". *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 1, pp. 246–252, 2011.

[48] J. Rodríguez-Sotelo, E. Delgado-Trejos, D. Peluffo-Ordóñez, D. Cuesta-Frau and G. Castellanos-Domínguez. "Weighted-PCA for unsupervised classification of cardiac arrhythmias". In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pp. 1906–1909. IEEE, 2010.

[49] P. Harris, C. Brunsdon and M. Charlton. "Geographically weighted principal components analysis". *International Journal of Geographical Information Science*, vol. 25, no. 10, pp. 1717–1736, 2011.

[50] J. T. Vogelstein, E. W. Bridgeford, M. Tang, D. Zheng, C. Douville, R. Burns and M. Maggioni. "Supervised dimensionality reduction for big data". *Nature communications*, vol. 12, no. 1, pp. 2872, 2021.

[51] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, third edition, 1996.

[52] M. Hajij, G. Zamzmi and F. Batayneh. "TDA-Net: fusion of persistent homology and deep learning features for COVID-19 detection from chest X-Ray images". In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 4115–4119. IEEE, 2021.

[53] T. A. Filisbino, G. A. Giraldi and C. E. Thomaz. "Nested AdaBoost procedure for classification and multi-class nonlinear discriminant analysis". *Soft Computing*, vol. 24, no. 23, pp. 17969–17990, 2020.

Learning and Nonlinear Models - Journal of the Brazilian Society on Computational Intelligence (SBIC), Vol. 23, Iss. 2, pp. 23–45, 2025
https://doi.org/10.21528/lnlm-vol23-no2-art3

A GRÁFICOS REFERENTES ÀS TABELAS 2-4

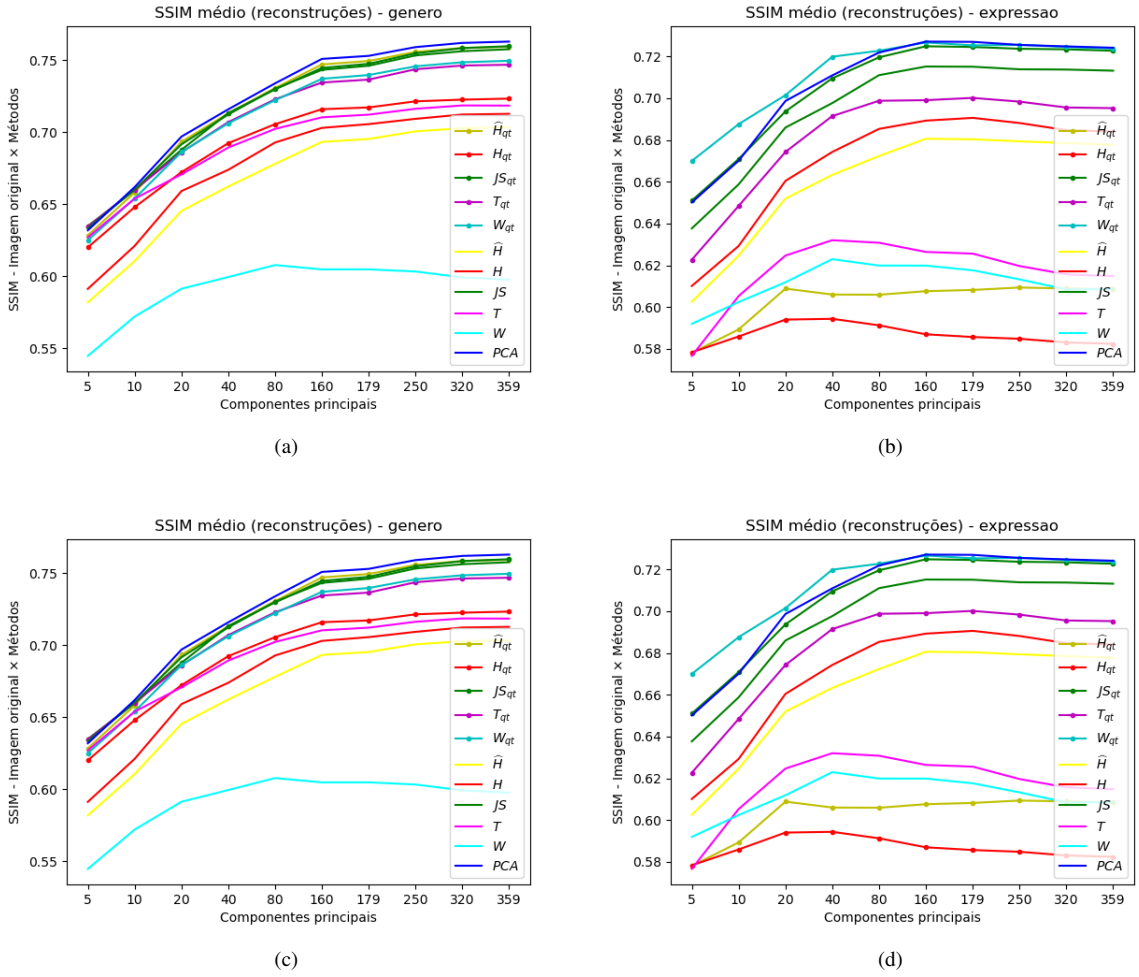
Este apêndice apresenta os gráficos dos quais foram retirados os dados que constam nas Tabelas 2-5. Esses gráficos permitem analisar a evolução do desempenho de cada método ao longo dos diferentes subespaços, para as tarefas de reconstrução (seção A.1) e classificação (seção A.2).

Os experimentos abordados nesta seção com a base de dados da FEI, foram conduzidos de forma que separamos 40 imagens de teste, das 400 imagens disponíveis, ficando então 360 imagens para a etapa de treinamento e, consequentemente, gerando 359 componentes principais, uma vez que nos testes realizados não foram identificados autovalores nulos para essas componentes. Um procedimento análogo de separação treinamento/teste foi realizado para os experimentos envolvendo expressão facial da FERET. Porém, pela quantidade de imagens existentes envolvendo gênero, foram separadas 20 imagens de homens e 19 imagens de mulheres para tarefas envolvendo gênero, a fim de manter as mesmas proporções do caso da FEI, como descrito na introdução da seção 4. Neste caso, foram utilizadas também 359 componentes principais para reconstrução e classificação.

https://doi.org/10.21528/lnlm-vol23-no2-art3

A.1 Gráficos para reconstrução

Para gerar os gráficos abaixo, projetamos os dados em subespaços com 5, 10, 20, 40, 80, 160, 179, 250, 320 e 359 componentes principais obtidas com os métodos abordados. Em seguida, é efetuada a reconstrução através da expressão (27) e calculado o $SSIM$ entre as reconstruções geradas e as imagens originais. Exibiremos aqui a média dos cálculos realizados sobre o conjunto de teste. Os experimentos realizados foram feitos sem e com o uso da *quadtree*. A seção 4.2 descreve as principais conclusões obtidas a partir dos resultados destes gráficos.



(a) (b)

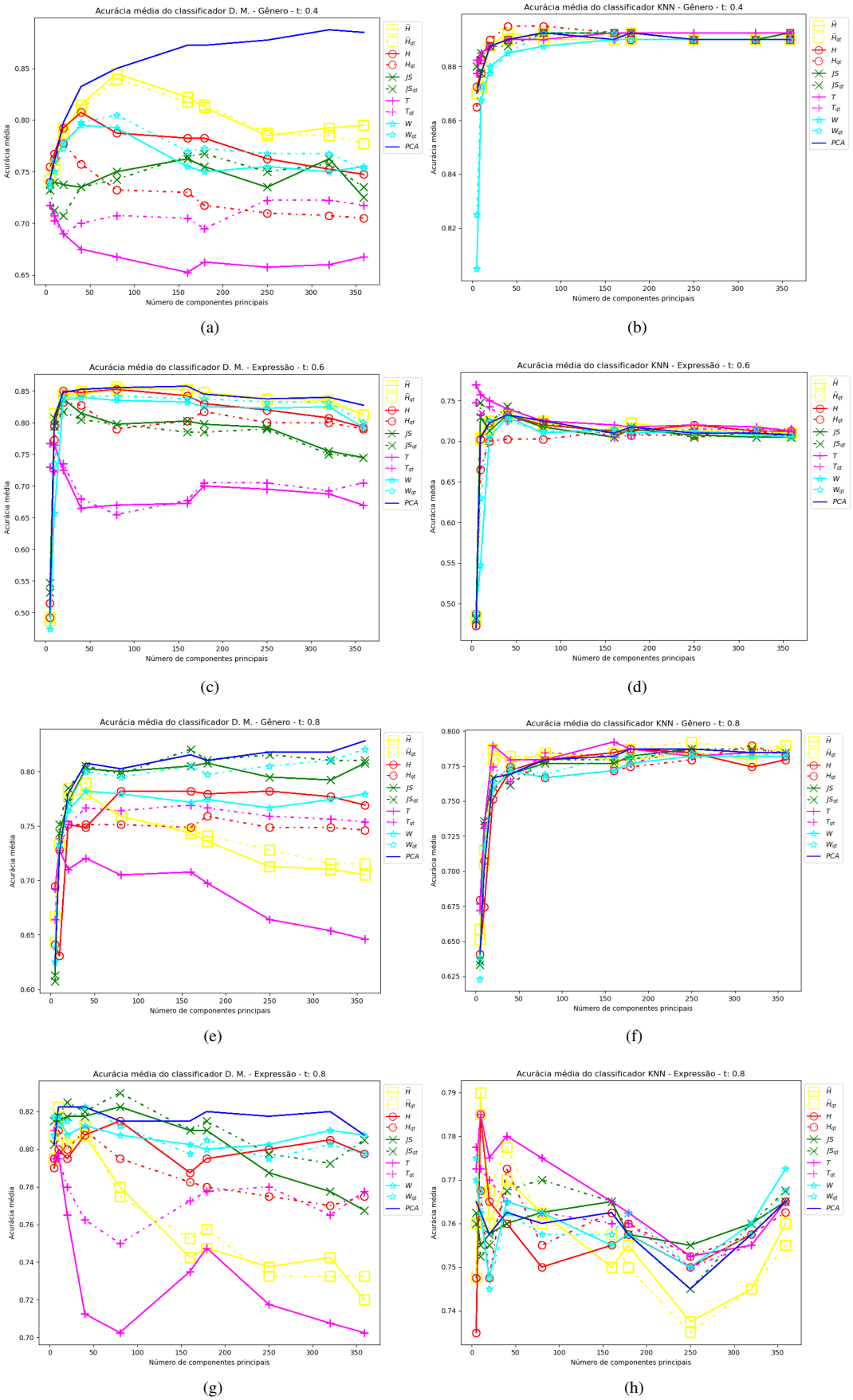
(c) (d)

Figura 6: Curvas do $SSIM$ referentes às reconstruções das 40 imagens de teste, com o uso de *quadtree*: (a) Gênero FEI: Linha (i) da Tabela 2, limiar $t = 0.8$. (b) Expressão Facial FEI: Linha (ii) da Tabela 2, $t = 0.8$. (c) Gênero FERET: Linha (iii) da Tabela 2, com limiar $t = 0.8$: (d) Expressão Facial FERET: Linha (iv) da Tabela 2, com limiar $t = 0.8$.

A.2 Gráficos da Classificação

Os gráficos abaixo foram gerados computando as acurácias médias sobre o conjunto de teste para subespaços de dimensão 5, 10, 20, 40, 80, 160, 179, 250, 320 e 359, computados com os métodos considerados, usando $K-$fold para $K = 10$. A seção 4.3 descreve as principais conclusões retiradas a partir dos resultados destes gráficos.

https://doi.org/10.21528/lnlm-vol23-no2-art3



(a) (b) (c) (d) (e) (f) (g) (h)

Figura 7: Gráficos das acurácias médias dos classificadores, referentes à Tabela 4, com *quadtree* (linhas tracejadas) e sem *quadtree* (linhas contínuas): FEI Gênero: (a) DM com $t = 0.4$. (b) KNN com $t = 0.4$. FEI Expressão Facial: (c) DM com $t = 0.6$. (d) KNN com $t = 0.6$. FERET Gênero: (e) DM com $t = 0.8$. (f) KNN com $t = 0.8$. FERET Expressão Facial: (g) DM com $t = 0.8$. (h) KNN com $t = 0.8$.