



UNIVERSIDADE ESTADUAL DO CEARÁ
CENTRO DE CIÊNCIAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
MESTRADO ACADÊMICO EM CIÊNCIA DA COMPUTAÇÃO

FRANCISCO EDVAN CHAVES

**ANÁLISE COMPARATIVA DAS CARACTERÍSTICAS DE FILTROS DE GABOR,
LBP E DCT PARA O RECONHECIMENTO AUTOMÁTICO DE EXPRESSÕES
FACIAIS**

FORTALEZA – CEARÁ

2015

FRANCISCO EDVAN CHAVES

ANÁLISE COMPARATIVA DAS CARACTERÍSTICAS DE FILTROS DE GABOR, LBP E
DCT PARA O RECONHECIMENTO AUTOMÁTICO DE EXPRESSÕES FACIAIS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências e Tecnologia da Universidade Estadual do Ceará, como requisito parcial à obtenção do título de mestre em Ciência da Computação. Área de Concentração: Ciência da Computação

Orientadores: Thelmo Pontes de Araújo e José Everardo Bessa Maia

FORTALEZA – CEARÁ

2015

Dados Internacionais de Catalogação na Publicação

Universidade Estadual do Ceará

Sistema de Bibliotecas

Chaves, Francisco Edvan.

Análise comparativa das características de filtros de Gabor, LBP e DCT para o reconhecimento automático de expressões faciais [recurso eletrônico] /

Francisco Edvan Chaves. – 2015.

1 CD-ROM: il.; 4 ¾ pol.

CD-ROM contendo o arquivo no formato PDF do trabalho acadêmico com 88 folhas, acondicionado em caixa de DVD Slim (19 x 14 cm x 7 mm).

Dissertação (mestrado acadêmico) – Universidade Estadual do Ceará, Centro de Ciências e Tecnologia, Mestrado Acadêmico em Ciência da Computação, Fortaleza, 2015.

Área de concentração: Ciência da Computação.

Orientação: Prof. Dr. Thelmo Pontes de Araújo.

Coorientação: Prof. Dr. José Everardo Bessa Maia.

1. Expressões faciais. 2. Reconhecimento de padrões. 3. Classificação. 4. Redução de dimensionalidade. I. Título.

FRANCISCO EDVAN CHAVES

ANÁLISE COMPARATIVA DAS CARACTERÍSTICAS DE FILTROS DE GABOR, LBP E
DCT PARA O RECONHECIMENTO AUTOMÁTICO DE EXPRESSÕES FACIAIS

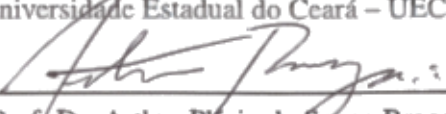
Dissertação apresentada ao Curso de Mestrado Acadêmico em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências e Tecnologia da Universidade Estadual do Ceará, como requisito parcial à obtenção do título de mestre em Ciência da Computação. Área de Concentração: Ciência da Computação

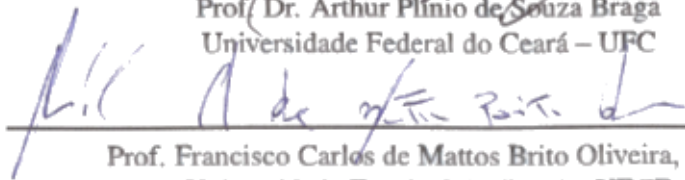
Aprovada em: 24 de agosto de 2015.

BANCA EXAMINADORA


Prof. Thelmo Pontes de Araújo, Ph.D. (Orientador)
Universidade Estadual do Ceará – UECE


Prof. Dr. José Everardo Bessa Maia (Orientador)
Universidade Estadual do Ceará – UECE


Prof. Dr. Arthur Plínio de Souza Braga
Universidade Federal do Ceará – UFC


Prof. Francisco Carlos de Mattos Brito Oliveira, Ph.D.
Universidade Estadual do Ceará – UECE

À Iza, à Bia, ao Miguel e à Malu.

AGRADECIMENTOS

A Deus, por ter me dado saúde (principalmente), coragem, perseverança e dedicação, por nunca ter desistido deste objetivo.

À minha esposa Izautina, pelo amor, companheirismo, apoio e incentivo.

Aos meus filhos Ana Beatriz, João Miguel e Maria Luíza, pelo amor e por serem minhas fontes de inspiração.

Aos meus pais, Francisco e Raimunda, ao meu irmão Elcivan e à minha tia Rosalba, por terem me ajudado a seguir o caminho da educação.

Aos professores orientadores Thelmo e Everardo, pelo conhecimento compartilhado e pela paciência. Aos professores Arthur e Francisco Carlos (Fran) por aceitarem participar da banca e pelas recomendações de melhorias deste trabalho.

Aos meus sogros, Francisco e Socorro, pelo apoio, principalmente com o auxílio com as crianças, quando eu precisei me ausentar delas.

À Universidade Estadual do Ceará (UECE), pela possibilidade de prosseguir essa caminhada. A todos os professores e colegas do MACC.

Ao Banco do Nordeste do Brasil (BNB) e à Faculdade Integrada da Grande Fortaleza (FGF), pelo incentivo. Aos colegas/amigos do BNB, em especial, à equipe do *Internet Banking*, George e Dênio, e ao Grande Moisés, pelo imenso apoio.

A todos amigos que sempre me apoiaram e torceram por essa conquista. Em especial, aos amigos da turma de Ciência da Computação da UFC 2001.

Aos que estiveram na platéia: compadres Claudina e Bruno, compadres Lucilane e Dante, Dante Júnior, comadre Mônica, comadre Kellyane, “compadres” José e Raimunda, Dona Socorro, amigos Roberto, Humberto e George, minha esposa Iza e nossa pequena Bia.

E que venham mais caminhos e desafios a serem trilhados!

*“No fim tudo dá certo, e se não deu é porque ainda
não chegou ao fim.”*

Fernando Sabino

RESUMO

O reconhecimento automático de expressões faciais é tratado como um problema de reconhecimento de padrões que consiste em determinar a expressão facial ou o estado emocional expresso (alegre, triste, cansado, dentre outras) pela imagem frontal de uma face humana. Entre as características da imagem extraídas para esta tarefa, as mais promissoras encontradas na literatura são filtros de Gabor, Padrões Binários Locais (LBP) e Transformada Discreta do Cosseno (DCT). Tipicamente, as bases de dados utilizadas para avaliar o desempenho são regionalizadas, sendo, por exemplo, uma base composta somente de faces asiáticas, outra de faces ocidentais, outra de afro-descendentes e assim por diante. O objetivo deste estudo é avaliar o desempenho destas características em diferentes bases de imagens frontais de faces (intra-bases) e avaliar a universalidade delas entre diferentes bases (inter-bases e multibases) considerando os casos de indivíduos conhecidos e desconhecidos. Os resultados obtidos demonstram a não universalidade das características estudadas, pois nos experimentos inter-bases, a melhor acurácia média foi 56,74%. Porém, em alguns experimentos intra-bases e multibases, a acurácia foi superior a 90%. Por fim, algumas recomendações de uso são extraídas dos resultados.

Palavras-chave: Expressões faciais. Reconhecimento de padrões. Classificação. Redução de dimensionalidade.

ABSTRACT

Automatic recognition of facial expressions is treated as a pattern recognition problem is to determine the facial expression or expressed emotional state (happy, sad, tired, among others) the front image of a human face. Between the image features extracted for this task the most promising in the literature are Gabor filters, Binary Local Patterns (LBP) and Discrete Cosine Transform (DCT). Typically, databases used to evaluate the performance is regionalized, being, for example, a base made up of only Asian faces other Western faces another african descent, and so forth. The objective of this study is to evaluate the performance of these features in different bases of frontal images of faces (intra-bases) and assess the universality of them between different bases (inter-bases) considering cases of known and unknown individuals. The results obtained didn't demonstrate the universality of the features, because in the inter-base experiments, the best average accuracy was 56.74 %. However, in some intra-bases and multibase experiments, the accuracy was more than 90 %. Finally, some recommendations of use are drawn from the results.

Keywords: Facial expressions. Patterns recognition. Classification. Reducing dimensionality.

LISTA DE ILUSTRAÇÕES

Figura 1 – Etapas do processo de Reconhecimento Automático de Expressões Faciais. .	19
Figura 2 – Imagens de filtros de Gabor - parte real.	27
Figura 3 – Imagens de filtros de Gabor - parte imaginária.	27
Figura 4 – Imagem original (canto superior esquerdo) e suas respectivas convoluções com 40 filtros de Gabor.	28
Figura 5 – Exemplo de um LBP original.	30
Figura 6 – Imagem original (à esquerda) da base JAFFE e o resultado da aplicação do LBP (à direita).	30
Figura 7 – Imagem dividida em regiões.	31
Figura 8 – Imagem original (à esquerda) e sua transformada DCT (à direita).	33
Figura 9 – PCA - ideia geral.	35
Figura 10 – LDA - ideia geral.	36
Figura 11 – À esquerda, duas primeiras componentes principais obtidas com PCA. À esquerda, com LDA.	38
Figura 12 – Estrutura de uma rede RBF.	40
Figura 13 – Representação do SVM.	40
Figura 14 – FCNN - exemplo.	43
Figura 15 – Imagens originais da base JAFFE. Da esquerda para a direita: zangado, desgosto, medo, feliz, neutro, triste e surpreso.	44
Figura 16 – Imagens originais da base YALE. Da esquerda para a direita: neutro, feliz, triste e surpreso.	45
Figura 17 – Imagens originais da base CK+. Da esquerda para a direita: zangado, desprezo, desgosto, medo, feliz, triste e surpreso.	45
Figura 18 – Imagens recortadas com o método Viola-Jones (VIOLA; JONES, 2001). . .	47
Figura 19 – Comparativo entre os experimentos realizados com a base JAFFE.	58
Figura 20 – Comparativo entre os experimentos realizados com a base Yale.	65
Figura 21 – Comparativo entre os experimentos realizados com a base CK.	73

LISTA DE TABELAS

Tabela 1 – Resumo dos trabalhos relacionados.	23
Tabela 2 – Comparação entre os banco de faces utilizados nos experimentos.	46
Tabela 3 – Comparação entre os classificadores.	48
Tabela 4 – Base JAFFE: acurácias obtidas (%) com filtros de Gabor usando <i>leave-one-out</i>	51
Tabela 5 – Base JAFFE: acurácias obtidas (%) com filtros de Gabor usando <i>leave-one-subject-out</i>	51
Tabela 6 – Base JAFFE: melhor resultado obtido com filtros de Gabor para indivíduos conhecidos (<i>leave-one-out</i>): SVM sem seleção de protótipos com acurácia média de 96,71%.	52
Tabela 7 – Base JAFFE: melhor resultado obtido com filtros de Gabor para indivíduos desconhecidos (<i>leave-one-subject-out</i>): SVM sem seleção de protótipos com acurácia média de 69,48%.	52
Tabela 8 – Base JAFFE: acurácias obtidas (%) usando LBP aplicado a base JAFFE.	53
Tabela 9 – Base JAFFE: melhor resultado obtido com LBP para indivíduos conhecidos (<i>leave-one-out</i>): KNN (sem FCNN) com acurácia média de 87,78%.	54
Tabela 10 – Base JAFFE: melhor resultado obtido com LBP para indivíduos desconhecidos (<i>leave-one-subject-out</i>): SVM (sem FCNN) com acurácia média de 63,38%.	54
Tabela 11 – Base JAFFE: resultados usando DCT aplicado às imagens inteiras.	55
Tabela 12 – Base JAFFE: resultados usando DCT aplicado em blocos.	56
Tabela 13 – Base JAFFE: melhor resultado obtido com DCT (16 × 16) para indivíduos conhecidos (<i>leave-one-out</i>): SVM (sem FCNN) com acurácia média de 91,55%.	57
Tabela 14 – Base JAFFE: melhor resultado obtido com DCT (8 × 8) para indivíduos desconhecidos (<i>leave-one-subject-out</i>): NB (sem FCNN) com acurácia média de 57,75%.	57
Tabela 15 – Resultados obtidos com filtros de Gabor usando <i>leave-one-out</i>	59
Tabela 16 – Resultados obtidos com filtros de Gabor usando <i>leave-one-group-out</i>	59
Tabela 17 – Base Yale: melhor resultado obtido com Gabor para indivíduos desconhecidos (<i>leave-one-subject-out</i>): SVM (com FCNN) com acurácia média de 70,00%.	60
Tabela 18 – Resultados usando LBP aplicado às imagens divididas em blocos.	60

Tabela 19 – Base Yale: melhor resultado obtido com LBP para indivíduos conhecidos (<i>leave-one-out</i>): SVM (sem FCNN) com acurácia média de 51,67%.	61
Tabela 20 – Base Yale: melhor resultado obtido com LBP para indivíduos desconhecidos (<i>leave-one-group-out</i>): SVM (sem FCNN) com acurácia média de 78,33%.	61
Tabela 21 – Resultados usando DCT aplicado às imagens inteiras.	62
Tabela 22 – Resultados usando DCT aplicado em blocos.	63
Tabela 23 – Base Yale: melhor resultado obtido com DCT para indivíduos conhecidos (<i>leave-one-out</i>): SVM (sem FCNN) com acurácia média de 45,00%.	64
Tabela 24 – Base Yale: melhor resultado obtido com DCT (8×8) para indivíduos desconhecidos (<i>leave-one-group-out</i>): SVM (sem FCNN) com acurácia média de 58,33%.	64
Tabela 25 – Base CK+: resultados obtidos com filtros de Gabor usando <i>leave-one-out</i>	66
Tabela 26 – Base CK+: resultados obtidos com filtros de Gabor usando <i>leave-one-group-out</i>	66
Tabela 27 – Base CK+: melhor resultado obtido com filtros de Gabor para indivíduos conhecidos (<i>leave-one-out</i>): SVM sem seleção de protótipos com acurácia média de 92,05%.	67
Tabela 28 – Base CK+: melhor resultado obtido com filtros de Gabor para indivíduos desconhecidos (<i>leave-one-subject-out</i>): SVM sem seleção de protótipos com acurácia média de 93,27%.	67
Tabela 29 – Base CK+: resultados usando LBP aplicado às imagens divididas em blocos.	68
Tabela 30 – Base CK+: melhor resultado obtido com LBP (16×16) para indivíduos conhecidos (<i>leave-one-out</i>): SVM sem seleção de protótipos com acurácia média de 88,69%.	68
Tabela 31 – Base CK+: melhor resultado obtido com LBP (16×16) para indivíduos desconhecidos (<i>leave-one-subject-out</i>): SVM com seleção de protótipos com acurácia média de 91,13%.	69
Tabela 32 – Base CK+: resultados usando DCT aplicado às imagens inteiras.	70
Tabela 33 – Base CK+: resultados usando DCT aplicado em blocos.	71
Tabela 34 – Base CK+: melhor resultado obtido com DCT para indivíduos desconhecidos (<i>leave-one-subject-out</i>): bloco 8×8 , SVM, sem seleção de protótipos e acurácia média de 82,87%.	72

Tabela 35 – Base CK+: melhor resultado obtido com DCT para indivíduos desconhecidos (<i>leave-one-subject-out</i>): bloco 32×32 , SVM, sem seleção de protótipos e acurácia média de 86,54%.	72
Tabela 36 – Experimentos com bases cruzadas usando JAFFE como treinamento.	73
Tabela 37 – Experimentos com bases cruzadas usando Yale como treinamento.	74
Tabela 38 – Experimentos com bases cruzadas usando CK+ como treinamento.	74
Tabela 39 – Matriz de confusão referente ao experimento com melhor acurácia (Treinamento: Yale; teste: CK+).	75
Tabela 40 – Teste de Hipóteses - Experimentos inter-bases.	75
Tabela 41 – Experimentos multibases.	76
Tabela 42 – Teste de Hipóteses - Experimentos multibases.	76
Tabela 43 – Tempo de Extração/Classificação (em <i>ms</i>).	77

LISTA DE ALGORITMOS

1	OBTERCARACTERISTICASGABOR	29
2	OBTERCARACTERISTICASLBP	31
3	OBTERCARACTERISTICASDCT	33
4	FCNN	42

LISTA DE ABREVIATURAS E SIGLAS

CK+	<i>Cohn-Kanade</i> Estendida.
DCT	<i>Discret Cosine Transform</i> (Transformada Discreta do Cosseno).
FACS	<i>Face Action Coding System</i> (Sistema de Codificação de Ações Faciais).
FCNN	<i>Fast Condensed Nearest Neighbor</i> .
IHC	Interface Homem-Máquina.
JAFFE	<i>Japanese Female Facial Expression</i> .
KNN	<i>K-Nearest Neighbors</i> .
MLP	<i>MultiLayer Perceptron</i> (Perceptron Multi-Camadas).
PCA	Análise de Componentes Principais.
PID	Processamento de Imagens Digitais.
LBP	<i>Local Binary Pattern</i> (Padrão Binário Local).
LDA	Análise de Discriminantes Lineares.
RBF	<i>Radial Base Function</i> (Função de Base Radial).
REAF	Reconhecimento Automático de Expressões Faciais.
RP	Reconhecimento de Padrões.
RFLD	<i>Recursive Linear Discriminant Analysis</i> .
SMO	<i>Sequential Minimal Optimization</i> .
SPAM	<i>Supervised Partitioning Around Medoids</i> .
SVM	<i>Support Vector Machine</i> (Máquinas de Vetores de Suporte).

SUMÁRIO

1	INTRODUÇÃO	17
1.1	OBJETIVOS	18
1.2	METODOLOGIA	18
1.3	VISÃO GERAL DO RAEF	19
1.4	TRABALHOS RELACIONADOS	20
1.5	CONTRIBUIÇÕES	24
1.6	ORGANIZAÇÃO DO TRABALHO	24
2	EXTRAÇÃO DE CARACTERÍSTICAS E CLASSIFICAÇÃO	25
2.1	EXTRAÇÃO DE CARACTERÍSTICAS	25
2.1.1	Extração de Características com filtros de Gabor	25
2.1.2	Extração de Características com LBP	29
2.1.3	Extração de Características com DCT	32
2.2	REDUÇÃO DE DIMENSIONALIDADE	34
2.2.1	Análise de Componentes Principais	34
2.2.2	Análise de Discriminantes Lineares	36
2.3	CLASSIFICAÇÃO	38
2.3.1	<i>K-Nearest Neighbors (KNN)</i>	38
2.3.2	Naive Bayes	39
2.3.3	Redes RBF	39
2.3.4	Máquinas de Vetores de Suporte	40
2.4	SELEÇÃO DE PROTÓTIPOS	41
3	RESULTADOS E DISCUSSÃO	44
3.1	BASE DE DADOS CONSIDERADAS	44
3.1.1	JAFFE	44
3.1.2	Yale	44
3.1.3	CK+	45
3.1.4	Comparativo entre os bancos de faces	45
3.2	PROCEDIMENTOS EXPERIMENTAIS	46
3.3	RESULTADOS	49
3.3.1	Experimentos com a base JAFFE	49

3.3.2	Experimentos com a base Yale	58
3.3.3	Experimentos com a base CK+	65
3.3.4	Experimentos com bases cruzadas (inter-bases)	73
3.3.5	Experimentos multibases	75
3.3.6	Comparativo entre os tempos de processamento	76
3.4	DISCUSSÃO	77
4	CONCLUSÃO E TRABALHOS FUTUROS	79
4.1	CONCLUSÕES	79
4.2	TRABALHOS FUTUROS	80
	REFERÊNCIAS	81
	APÊNDICES	84
	APÊNDICE A – Implementação Gabor	85
	APÊNDICE B – Implementação PCA e LDA	87
	APÊNDICE C – Implementação Classificadores	88

1 INTRODUÇÃO

O Reconhecimento Automático de Expressões Faciais (RAEF) consiste na utilização de um sistema computacional que lê uma imagem digital de uma face humana e atribui a ela uma expressão facial como, por exemplo, triste, alegre, cansado ou sonolento. Trata-se de um problema de reconhecimento de padrões extremamente complexo e vem sendo bastante estudado na literatura (PANTIC; ROTHKRANTZ, 2000) (FASEL; LUETTIN, 2003) (TIAN *et al.*, 2011) (BETTADAPURA, 2012). O RAEF ocupa atualmente cada vez mais espaço nas áreas de Reconhecimento de Padrões (RP) e Processamento de Imagens Digitais (PID). Tal crescimento deve-se ao fato do avanço de pesquisas em detecção e reconhecimento de faces (FASEL; LUETTIN, 2003).

Diversos métodos têm sido propostos na literatura para esse problema. No entanto, a atividade de reconhecer expressões faciais automaticamente é extremamente complexa, pois a fisionomia facial varia bastante entre pessoas distintas. Além disso, há a interferência de outros fatores como identidade, idade, etnia, gênero, cabelos, produtos cosméticos, acessórios (óculos, brincos, piercings), dentre outros. Há ainda a influência de aspectos como pose e iluminação (FASEL; LUETTIN, 2003).

Segundo Gu e colegas (GU *et al.*, 2012), há uma correlação da expressão facial com o indivíduo. Assim, indivíduos desconhecidos nem sempre têm suas expressões faciais reconhecidas corretamente, fazendo-se necessário, o desenvolvimento de metodologias mais eficientes. Neste trabalho, abordaremos o problema de reconhecimento de expressões faciais tanto para indivíduos desconhecidos ou como para indivíduos já conhecidos pelo sistema.

Há uma diversidade de áreas de atuação que podem ser beneficiadas com a utilização de um sistema automático de reconhecimento de expressões faciais. Inicialmente utilizado em Psicologia, o RAEF pode ser usado também como meio de interação homem-máquina (IHC), mais especificamente em aplicações de computação pervasiva (ubíqua), na qual a interação dos humanos com as máquinas é feita de modo natural. Porém, convém enfatizar que a expressão facial não é o único meio utilizado. Pode-se utilizar outros meios para IHC, como reconhecimento de voz, emoções, gestos, dentre outros (PANTIC; ROTHKRANTZ, 2000). Reconhecimento de expressões pode ser extremamente útil também em sistemas de segurança, principalmente em aeroportos, onde pode-se detectar a presença de terroristas (CHUNG, 2010), e em sistemas de monitoramento usados em sala de aula, onde poderia identificar os alunos sonolentos ou desatentos. Diversas técnicas têm sido aplicadas comercialmente, como detecção

de sorrisos em câmeras digitais. Tem surgido também diversas ferramentas de marketing como *CrowdEmotion*, *RealEyes*, *Affectiva* (VALSTAR, 2015).

1.1 OBJETIVOS

Objetivo Geral

Tendo em vista o problema apresentado, a presente dissertação tem como objetivo principal a realização de uma análise comparativa entre filtros de Gabor, LBP (do inglês, *Local Binary Pattern*) e DCT (do inglês, *Discret Cossine Transform*) como métodos de representação de características para serem utilizados em classificadores de expressões faciais.

Objetivos Específicos

Dentre os objetivos específicos, destacam-se os seguintes:

- a) Implementar diversos métodos para representação e extração de características aplicados ao problema em estudo;
- b) Implementar diversos métodos para classificação aplicados ao problema em estudo;
- c) Realizar experimentos com os principais bancos de expressões faciais referenciados na literatura, que apresentem variações de etnias, gênero, efeitos de iluminação e presença de itens de oclusão;
- d) Realizar experimentos com o algoritmo de seleção de protótipos FCNN (do inglês *Fast Condensed Nearest Neighbor*) aplicado aos conjuntos de treinamentos;
- e) Desenvolver uma ferramenta que possa ser facilmente ajustada, através de parâmetros, e acoplada a outros sistemas que necessitem classificar expressões faciais;
- f) Propor itens de evolução e continuidade deste trabalho.

1.2 METODOLOGIA

Visando atingir aos objetivos deste trabalho, serão implementados os algoritmos para extração de características com filtros de Gabor, LBP e DCT. Em seguida, serão realizados experimentos com três bases de dados disponibilizadas pela comunidade científica, a saber: JAFFE,

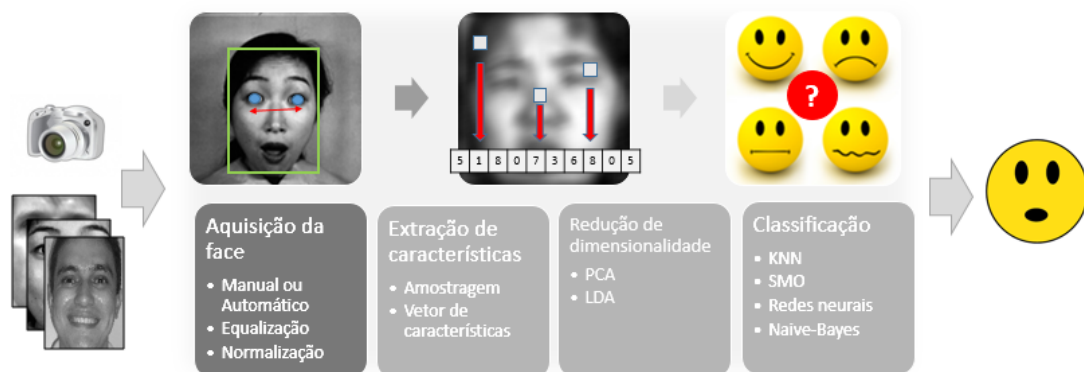
Yale e CK+. Para estimação das classes de expressões faciais, serão realizados experimentos com os classificadores KNN (do inglês, *K-Nearest Neighbors*), SVM (do inglês, *Support Vector Machine*), Naive Bayes e redes RBF (do inglês, *Radial Basis Functions*).

Serão realizados experimentos intra, inter e multibases bem como avaliadas algumas medidas de performance como acurácia, sensibilidade, precisão, especificidade e *F-Measure*. Além dessas medidas, serão comparados os tempos de processamento para extração de características de cada método bem como para realização da classificação. Por fim, será avaliada a universalidade de cada método de extração, comparando experimentos nos quais os indivíduos são ou não previamente conhecidos pelos algoritmos.

1.3 VISÃO GERAL DO RAEF

O processo de reconhecimento de expressões faciais consiste em quatro etapas: aquisição da face ou pré-processamento, representação/extração de características faciais, redução de dimensionalidade das características e classificação da expressão facial (FASEL; LUETTIN, 2003) (TIAN *et al.*, 2011). Essas fases podem ser vistas no esquema da Figura 1.

Figura 1 – Etapas do processo de Reconhecimento Automático de Expressões Faciais.



Fonte: Elaborado pelo autor.

A etapa de aquisição consiste em localizar uma região facial em uma imagem ou sequência de imagens (vídeo). No caso de um vídeo, a região facial precisa ser localizada em cada um de seus quadros (*frames*) ou apenas no primeiro e ser rastreada nos demais quadros. O pré-processamento consiste em extrair a região de interesse da imagem da face adquirida e aplicar algumas operações de PID, como redimensionamento e equalização.

A representação/extração de características consiste em descrever as mudanças faciais causadas por uma expressão facial. Por exemplo, quando um ser humano está sorrindo, as dimensões de sua boca e dos seus olhos são diferentes de quando está com uma expressão neutra. A forma como essas mudanças são descritas são denominadas de características (em inglês *features*) e são representadas por meio de vetores. Há duas categorias de métodos para representação de características: os geométricos e os baseados em aparência, também conhecido como holísticos (TIAN *et al.*, 2011). Os geométricos usam como características a localização ou as dimensões de parte das faces como, por exemplo, coordenadas geométricas referentes à localização dos olhos ou dos cantos da boca, medidas como dimensão da boca ou sobrancelhas, dentre outros. Os holísticos trabalham apenas com os níveis de intensidade dos pixels da imagem. Estes podem ser divididos em globais e locais. Os globais trabalham com a imagem inteira, enquanto os locais com partes dela, que contenham algumas partes específicas da face, como boca e olhos (FASEL; LUETTIN, 2003). Os métodos holísticos usam técnicas de processamento de imagens, como filtros ou transformadas, de forma a realçar informações significativas. Neste trabalho, abordar-se apenas métodos holísticos.

Em geral, o número de características, principalmente nos métodos holísticos, é muito elevado. Então, faz-se uso de métodos de redução de dimensionalidade que consistem em projetar os vetores correspondentes às características obtidas em uma nova base vetorial de forma que, as primeiras componentes contenhas as maiores variâncias. Assim, para efeitos de redução de dimensionalidade, seleciona-se apenas um subconjunto das características projetadas.

A classificação pode ser feita através dos códigos de ações faciais (FACS, do inglês *Face Action Coding System*) ou dos protótipos de expressões (FASEL; LUETTIN, 2003). Pode ser baseada em quadros ou em sequência deles. Neste trabalho, aborda-se apenas a classificação através de protótipos e baseada em quadros.

1.4 TRABALHOS RELACIONADOS

O presente trabalho apresenta um estudo comparativo entre três metodologias de extração de características para o reconhecimento de expressões faciais: filtros de Gabor (SHEN; BAI, 2006), LBP (SHAN *et al.*, 2009) e DCT (PEDRINI; SCHWARTZ, 2008). Com isso, apresenta-se nesta seção alguns trabalhos relacionados, que serviram como base para esta análise comparativa bem como para a escolha de algumas parametrizações de alguns métodos.

Deng e colegas (DENG *et al.*, 2005) apresentaram uma metodologia usando um

banco de filtros locais de Gabor. A ideia principal foi usar um subconjunto dos 40 filtros que são comumente usados, com 5 frequências e 8 orientações. Dessa forma, os autores reduziram o tempo de extração de características bem como a quantidade de memória necessária para seu armazenamento. Durante o pré-processamento, foi usada uma metodologia para obtenção da face a partir da distância intraocular. Os autores usaram PCA (do inglês *Principal Component Analysis*) (JOLLIFFE, 2002) e LDA (do inglês *Linear Discriminant Analysis*) (WEBB; COPSEY, 2011) para redução de dimensionalidade e as características obtidas foram então testadas num classificador baseado nas distâncias euclidiana e *cityblock*. Os experimentos foram realizados com a base de dados JAFFE (LYONS *et al.*, 1998), usando imagens dos mesmos indivíduos para treinamento e teste. O melhor resultado obtido forneceu uma acurácia de 97,33 %, usando um subconjunto com 12 filtros de Gabor (denominado pelo autores de LG3), PCA e LDA para redução de dimensionalidade e classificação com a distância euclidiana.

Em 2012, Gu e colegas (GU *et al.*, 2012) criaram uma metodologia para extração de características baseada no funcionamento do córtex visual humano e em classificadores hierárquicos. Durante o pré-processamento, os olhos e a boca foram marcados manualmente e as posições foram usadas para obtenção da face. Uma vez que face tenha sido localizada, esta foi dividida em blocos com sobreposição. Para cada um desses blocos, os autores aplicaram 24 filtros de Gabor (3 escalas e 8 orientações). Os blocos filtrados foram mapeados para um *grid* que simula o córtex visual humano para seleção das características. Um diferencial desse trabalho é a forma como os blocos filtrados foram concatenados para a geração do vetor de características. Geralmente, agrupa-se todos em um único vetor de característica. Nesse trabalho, o agrupamento foi feito por escalas. Dessa forma, cada imagem gerou 3 vetores de características. Para redução de dimensionalidade, os autores usaram RFLD (do inglês, *Recursive Linear Discriminant Analysis*), que foi proposto por Xiang e colegas. (XIANG *et al.*, 2006). As características locais foram classificadas usando o classificador KNN (do inglês, *K-Nearest Neighbors*) e combinadas para a decisão global usando *Borda count* e *Decision template*. Os autores obtiveram excelentes resultados para expressões de indivíduos não conhecidos usando as bases JAFFE e CK (LUCEY *et al.*, 2010). Com a primeira base, a acurácia média foi de 89,67% e com a segunda, 91,51 %. Os autores realizaram experimentos com as bases cruzadas. Usando JAFFE para treinamento, obtiveram uma acurácia de 54,05 % usando CK para teste. Fazendo o experimento contrário, a acurácia foi de 55,87 %.

Outro trabalho relacionado, Owusua e colegas (OWUSU *et al.*, 2014) utilizaram o

tradicional método Viola-Jones (VIOLA; JONES, 2001) para detecção das faces. Em seguida, aplicaram funções de Bessel para reduzir as imagens das faces que foram detectadas. As imagens foram reduzidas para um tamanho padronizado de 20×20 pixels. Com isso, não houve necessidade de realização de amostragem (*downsampling*) nas etapas seguintes. As características foram extraídas com a utilização de filtros de Gabor (5 escalas e 8 orientações). Para redução de dimensionalidade, os autores criaram um procedimento baseado no algoritmo *AdaBoost* (NORVIG; RUSSELL, 2014). Foram utilizadas redes neurais MLP (do inglês, *MultiLayer Perceptron*), como arquitetura *feedforward* e aprendizagem por *backpropagation*, para classificação. Foram realizados experimentos com a base JAFFE e Yale (YALE, 2000). Os autores obtiveram uma acurácia de 96,83 % com a base JAFFE, porém foram usados apenas indivíduos conhecidos para testes.

Em (SHAN *et al.*, 2009), os autores fazem um estudo do Padrão Binário Local (LBP, do inglês, *Local Binary Pattern*) como método para extração e representação de características para expressões faciais. Os autores usaram o histograma para geração do vetor de características considerando apenas os padrões uniformes e propuseram um método, denominado *Boosted-LBP*, para selecionar as características mais relevantes. Foram apresentados resultados comparando diversos classificadores, destacando as Máquinas de Vetores de Suporte (SVM, do inglês, *Support Vector Machine*). A acurácia obtida com a base CK foi 92,6 % usando SVM com núcleo de Função de Base Radial (RBF, do inglês *Radial Basis Function*). Nesse mesmo trabalho, os autores fizeram testes do uso do LBP para imagens com baixa resolução e obtiveram resultados aceitáveis, mesmo para indivíduos desconhecidos. Em (ABDULRAHMAN *et al.*, 2014), os autores realizaram experimentos com a base JAFFE combinando filtros de Gabor e LBP e obtiveram acurácia de 90 %.

Outros trabalhos têm usado Transformada Discreta do Cosseno (DCT, do inglês, *Discrete Cossine Transform*) para extração de características, por exemplo (MA; KHORASANI, 2004). Nesse trabalho, os autores utilizam DCT para extração das características. DCT foi aplicado sobre a diferença entre a face neutra e a face com expressão. Em seguida, foram selecionados um subconjunto dos coeficientes obtidos e alimentado a um classificador de rede neural. Os testes foram realizados com um banco próprio, utilizando 40 indivíduos para treinamento e 20 para teste. A acurácia obtida foi de 93,75 %. Em um outro trabalho dos mesmos autores (MA *et al.*, 2004), foi usados DCT e *K-Means*. No trabalho de Badhu e colegas (BHADU *et al.*, 2012) também foi usado DCT, que foi aplicado nas imagens originais e extraídos, em

zigzague, os coeficientes mais significativos. A acurácia nesse trabalho com a base JAFFE e o classificador *Adaboost* foi de 71 %, porém não foi citado como a base foi particionada em treinamento e teste. Nesse mesmo trabalho, os autores usaram ainda filtros de Gabor e wavelets de Daubechies com o classificador *Adaboost*, obtendo uma acurácia de 65,33 % e 62,5 %, respectivamente.

Motivado pelos trabalhos de (LABOREIRO, 2014) e (MAIA *et al.*, 2013), embora não diretamente relacionados a RAEF, resolveu-se investigar a eficiência da aplicação de algoritmos de seleção de protótipos nos conjuntos de treinamento. No primeiro trabalho, o autor usou o algoritmo de seleção de protótipos FCNN (do inglês, *Fast Condensed Nearest Neighbor*) (ANGIULLI, 2005) para o problema de segmentação de imagens. No segundo trabalho, os autores investigaram a utilização do SPAM (do inglês, *Supervised Partitioning Around Medoids*) (KAUFMAN; ROUSSEEUW, 2009) para problemas gerais de classificação e regressão.

Além do PCA e LDA, outros métodos vêm sendo utilizados para redução de dimensionalidade para o reconhecimento de expressões faciais. Dentre esses, podemos citar o K-PCA (do inglês, *Kernel - Principal Components Analysis*) e GDA (do inglês, *General Discriminant Analysis*) (CHUNG, 2010).

Resumo dos trabalhos relacionados

A Tabela 1 apresenta uma síntese dos trabalhos analisados indicando os métodos utilizados em cada etapa do RAEF, banco de faces usados para validação e resultados obtidos. Não foi encontrado indicação de utilização de seleção de protótipos nos trabalhos pesquisados.

Tabela 1 – Resumo dos trabalhos relacionados.

Referência	PP	Características	RD	Classificador	BF	Resultados	ID
(DENG <i>et al.</i> , 2005)	Manual	Gabor (locais)	PCA/LDA	KNN	JAFFE	93,00%	Não
(GU <i>et al.</i> , 2012)	Manual	Gabor	RLDA	Classificadores hierárquicos	JAFFE, CK	91,51 (CK) 55,87 (inter-bases)	Sim
(OWUSU <i>et al.</i> , 2014)	Viola-Jones	Gabor	<i>AdaBoost</i> adaptado	MLP	JAFFE	96,83%	Não
(SHAN <i>et al.</i> , 2009)	Manual	LBP	Boosted-LBP	SVM	CK	92,6%	Sim
(ABDULRAHMAN <i>et al.</i> , 2014)	Nenhum	Gabor, LBP	PCA	KNN	JAFFE	>90%	Não
(MA; KHORASANI, 2004)	Não	DCT	Não	MLP	Própria	93,75%	Não
(BHADU <i>et al.</i> , 2012)	Não	DCT	PCA	Adaboost	JAFFE	71%	Não

Fonte: Elaborado pelo autor. PP = pré-processamento; RD = redução de dimensionalidade; BF = Banco de faces; ID = indicador de indivíduos desconhecidos.

1.5 CONTRIBUIÇÕES

As principais contribuições deste trabalho são as seguintes:

- a) Análise comparativa entre as características de Gabor, LBP e DCT para reconhecimento automático de expressões faciais;
- b) Avaliação da utilização do algoritmo de seleção de protótipos FCNN aplicado aos conjuntos de treinamentos;
- c) Artigo **Análise de Alguns Discriminantes para o Problema de Reconhecimento de Expressões Faciais** a ser submetido a uma revista.

1.6 ORGANIZAÇÃO DO TRABALHO

O restante do trabalho está organizado da seguinte forma. O Capítulo 2 apresenta a fundamentação teórica descrevendo os principais tópicos para a compreensão deste trabalho. O capítulo seguinte apresenta o detalhamento dos experimentos, os banco de expressões faciais utilizados, análise dos resultados e discussão. Por fim, o Capítulo 4 expõe as conclusões obtidas com a realização deste trabalho, bem como as recomendações para trabalhos futuros.

2 EXTRAÇÃO DE CARACTERÍSTICAS E CLASSIFICAÇÃO

Este capítulo apresenta inicialmente uma descrição dos três métodos de extração de características utilizados. Em seguida, apresenta-se uma visão geral dos classificadores utilizados. Por fim, apresenta-se o algoritmo FCNN para seleção de protótipos.

2.1 EXTRAÇÃO DE CARACTERÍSTICAS

Em geral, sistemas de classificação de padrões atuam sobre conjuntos de dados com milhares de informações que precisam ser tratadas ou processadas de forma a serem melhor representadas. As informações processadas ou extraídas são denominadas características. Como exemplo de características obtidas dos dados amostrados, podemos citar média aritmética, desvio padrão, máximo e mínimo.

Como citado na Seção 1.3, este trabalho aborda apenas métodos holísticos para extração de características. Porém, utilizar somente as intensidades dos *pixels* não traz bons resultados para reconhecimento de expressões faciais. Com isso, faz-se necessário utilizar algumas técnicas de processamento de imagens, como filtragem ou transformadas, para extração das características a serem utilizadas pelo classificador. Neste trabalho, optou-se por implementar filtros de Gabor, LBP e DCT para extração de características faciais, por serem os mais utilizados na literatura.

As subseções seguintes apresentam uma visão geral sobre as três metodologias usadas para representação de características: filtros de Gabor, Padrões Binários Locais e Transformada Discreta do Cosseno. Junto com essa visão geral, descreve-se como os vetores de características são gerados com cada um desses métodos.

2.1.1 Extração de Características com filtros de Gabor

Os filtros bidimensionais de Gabor, também conhecidos como *Gabor wavelets*, vêm sendo bastante utilizado para extração de características e são compostos por uma função gaussiana modulada por uma *wavelet* complexa (SHEN *et al.*, 2007). Os núcleos dos filtros de Gabor, compostos por uma função *wavelet*, simulam o comportamento das células do córtex cerebral dos mamíferos. Têm sido muito utilizados em Visão Computacional e Análise de Imagens devido a suas propriedades de localização, tanto no domínio espacial quanto no de frequência (DENG *et al.*, 2005). Há várias formulações desses filtros, porém neste trabalho

usou-se a mesma formulação que (DENG *et al.*, 2005), que os define por:

$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x')^2+(y')^2}{2\sigma^2}} \left(e^{i\omega x'} - e^{-\frac{\omega^2\sigma^2}{2}} \right), \quad (2.1)$$

sendo

$$x' = x \cos \theta + y \sin \theta, \quad y' = -x \sin \theta + y \cos \theta,$$

ω o centro de frequência radial, σ o desvio padrão da função gaussiana e i a unidade imaginária.

Geralmente, utiliza-se 5 variações de frequência (escalas) e 8 orientações, totalizando 40 filtros (DENG *et al.*, 2005) (OWUSU *et al.*, 2014). A frequência é definida como:

$$\omega_m = \omega_{max} \lambda^{-(m-1)},$$

onde $\omega_{max} = \pi/2$, $\lambda = \sqrt{2}$ e $m = 1, 2, 3, 4, 5$. A orientação é dada por

$$\theta_n = (n-1)/8,$$

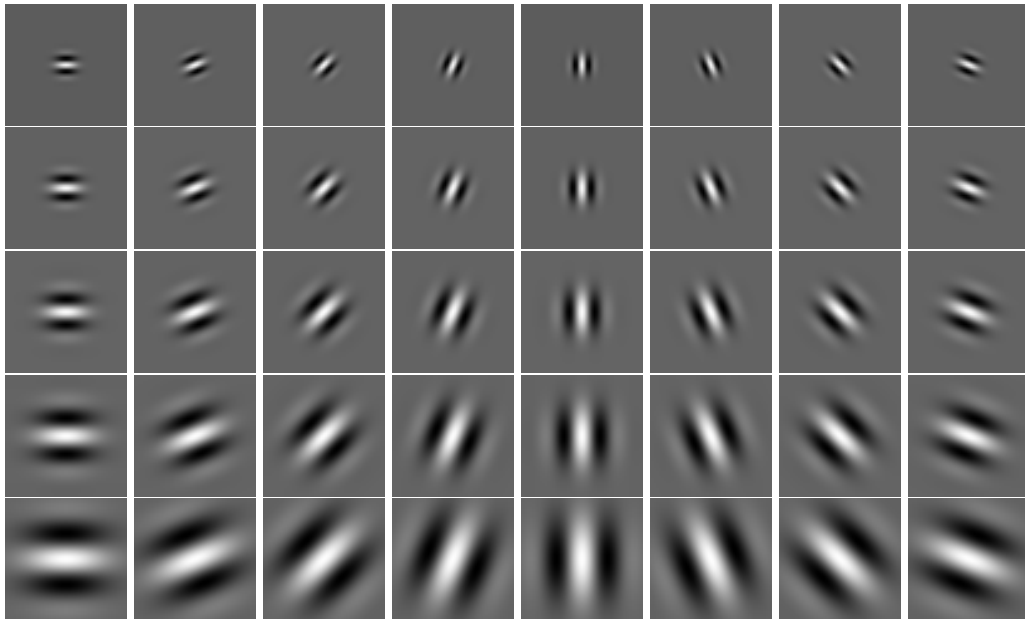
com $n = 1, 2, \dots, 8$. Assim como em (DENG *et al.*, 2005) e (LEE, 1996), o valor do desvio padrão σ usado nos experimentos foi $\frac{\pi}{\omega}$. Cada filtro é convoluído com a imagem original e os resultados são concatenados para a geração do vetor de características.

A aplicação de um filtro de Gabor em uma imagem pode ser feita de três formas: usar somente a parte real, usar somente a parte imaginária ou a magnitude. Neste trabalho, adotou-se a mais comumente utilizado: a magnitude. As Figuras 2 e 3 representam imagens contendo, respectivamente, a parte real e a parte imaginária de filtros com cinco variações de escala ω_m e oito orientações θ_μ . Já a Figura 4 representa a magnitude da imagem de um indivíduo convoluída com os 40 filtros de Gabor.

Para a extração das características de uma imagem usando filtros de Gabor, determina-se uma quantidade de variações de escala e orientações e gera-se um filtro para cada par de escala e orientação. Cada filtro é convoluído com a imagem original, sendo que as imagens filtradas são concatenadas para a geração do vetor de características. Geralmente, utiliza-se 5 escalas e 8 orientações, totalizando 40 filtros (DENG *et al.*, 2005) (OWUSU *et al.*, 2014). Dependendo das dimensões da imagem original, o número de características torna-se muito elevado, sendo necessária a utilização de métodos de amostragem e redução de dimensionalidade.

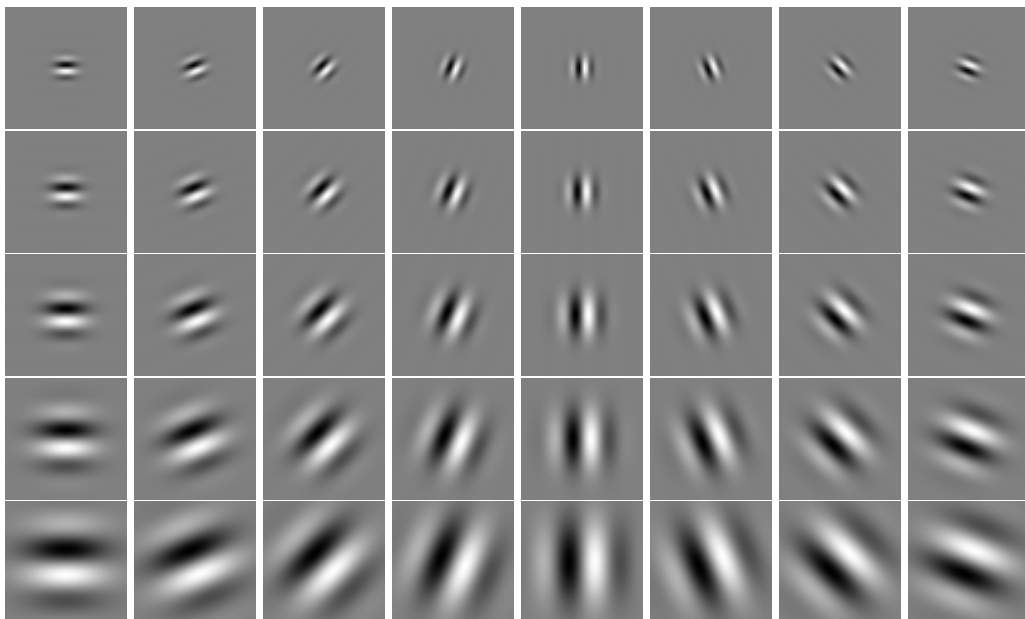
Dessa forma, são feitos os seguintes procedimentos para extração de características de uma imagem I usando filtros de Gabor:

Figura 2 – Imagens de filtros de Gabor - parte real.



Fonte: Elaborado pelo autor.

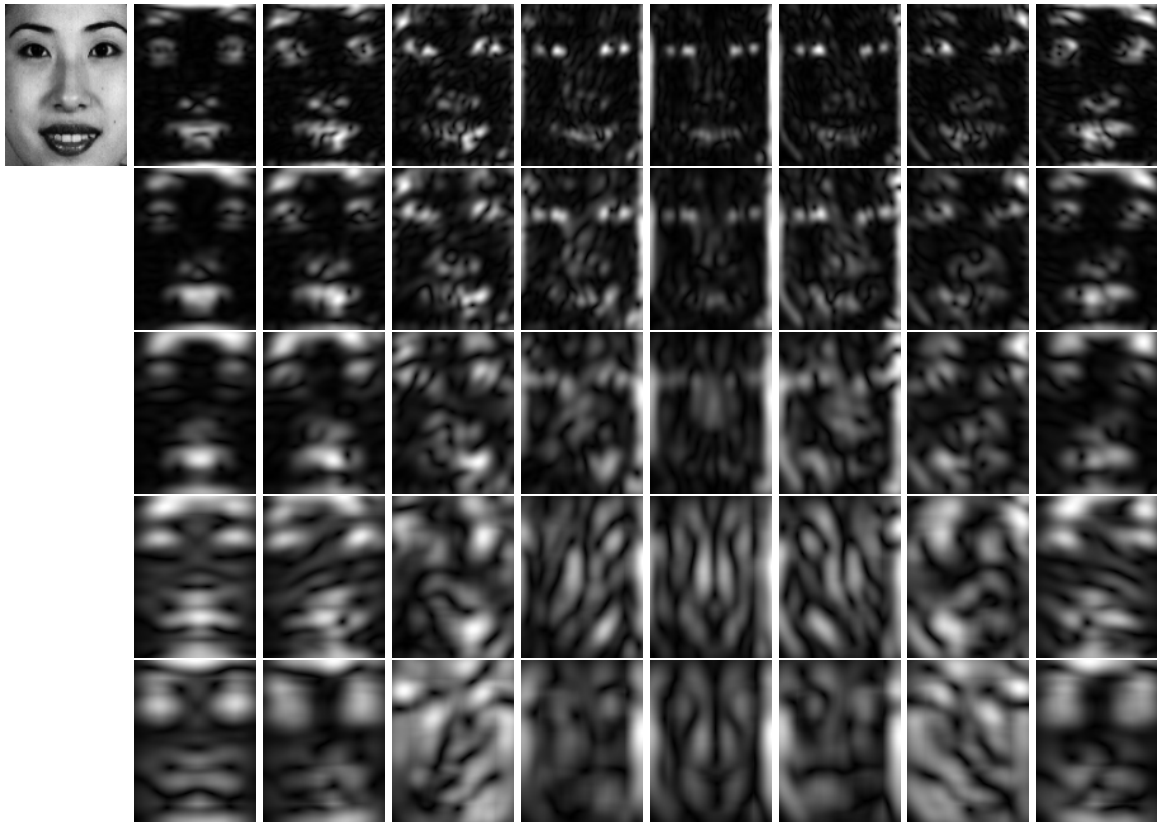
Figura 3 – Imagens de filtros de Gabor - parte imaginária.



Fonte: Elaborado pelo autor.

- a) Obtém-se os filtros de Gabor, cuja quantidade é determinada pela variações de escala e orientações. Por exemplo, para 5 escalas e 8 orientações, são obtidos 40 filtros;
- b) A imagem I é individualmente convoluída com cada filtro;
- c) Obtém-se a magnitude da convolução feita com a parte real e parte imaginária e,

Figura 4 – Imagem original (canto superior esquerdo) e suas respectivas convoluções com 40 filtros de Gabor.



Fonte: Elaborado pelo autor.

em seguida, aplica-se um fator de amostragem;

- d) Os valores restantes são normalizados;
- e) Concatena-se todas as magnitudes amostradas para geração do vetor de características.

O Algoritmo 1 ilustra o processo de obtenção do vetor de características de uma imagem.

Algoritmo 1: OBTERCARACTERISTICASGABOR

Entrada: Imagem I , **inteiro** $numEscalas$, **inteiro** $numOrientacoes$
Saída: Vetor de Características X

```

inicio
   $X = 0$ 
  para  $i = 0$  até  $numEscalas$  faça
    para  $j = 0$  até  $numOrientacoes$  faça
       $\omega = i + 1$ 
       $\theta = j(\pi/numOrientacoes)$ 
       $G = ObterGabor(\omega, \theta)$ 
       $I_r = Convolucao(I, G)$ 
       $I_m = Magnitude(I_r)$ 
      Concatena( $X$ , Amostragem( $I_m$ ))
    fim
  fim
fim

retorna  $X$ 

```

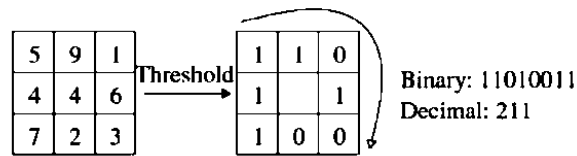
Fonte: Elaborado pelo autor.

2.1.2 Extração de Características com LBP

O Padrão Binário Local (LBP, do inglês *Local Binary Pattern*) foi originalmente definido como um operador que, atuando em regiões 3×3 de uma imagem, associa a cada um dos oito vizinhos do pixel central o bit 0 ou o bit 1, de acordo com a seguinte regra: se o vizinho tiver intensidade menor que a do pixel central, associa-se o bit 0; caso contrário, associa-se o bit 1. Os oito bits resultantes geram uma nova intensidade que substituirá o pixel central. A Figura 5 ilustra um exemplo da aplicação do LBP original, também conhecido como LBP(8, 1) (PEDRINI; SCHWARTZ, 2008).

O padrão gerado é considerado uniforme se tiver no máximo duas mudanças de 0 para 1 ou de 1 para 0. Por exemplo, o padrão 00011100 é considerado uniforme, enquanto que o 01100011 é não uniforme.

Figura 5 – Exemplo de um LBP original.



Fonte: (SHAN *et al.*, 2009).

O LBP tem sido usado em processamento de imagens, principalmente em análise de texturas. Ao ser usado em uma imagem, o LBP é aplicado em cada submatriz 3×3 da imagem, resultando em uma nova imagem. A Figura 6 apresenta uma imagem original e sua resultante da aplicação do operador LBP.

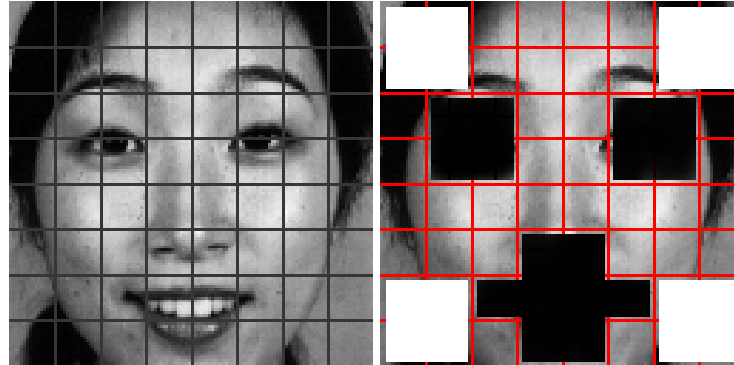
Figura 6 – Imagem original (à esquerda) da base JAFFE e o resultado da aplicação do LBP (à direita).



Fonte: Figura original (LYONS *et al.*, 1998); figura modificada (produzido pelo autor).

Para a extração das características de uma imagem resultante da aplicação do LBP, utiliza-se geralmente o histograma 256-bin da imagem. Neste trabalho, usou-se uma metodologia similar a (SHAN *et al.*, 2009). Cada imagem foi dividida em blocos de mesmo tamanho, sem sobreposição. Para cada um desses blocos, aplica-se LBP (uniforme, não uniforme, outras variações, etc.) e obtém o histograma. Os histogramas foram concatenados para gerar o vetor de características da imagem analisada. A Figura 7 apresenta uma imagem dividida em blocos. LBP é aplicada em cada uma das regiões e os histogramas correspondentes (ponderados ou não) são concatenados para geração de características. A imagem à direita mostra a utilização de pesos, sendo que os blocos em branco tiveram peso 0 e os em preto, peso 3.

Figura 7 – Imagem dividida em regiões.



Fonte: Elaborado pelo autor.

O Algoritmo 2 ilustra o processo de obtenção do vetor de características de uma imagem. A função *DivideImagemEmBlocos* divide a imagem em blocos do mesmo tamanho. Nos testes com aplicação de LBP à imagem inteira, o tamanho do bloco (*tamBloco*), deve ser igual às dimensões da imagem original. A função *Histograma* retorna o histograma considerando o número de níveis de intensidade. Por exemplo, considerando imagens com dimensão 96×128 divididas em 48 blocos de tamanho 16×16 , o vetor de características gerado possui dimensão 12.288, usando o histograma com 256 níveis.

Algoritmo 2: OBTENCARACTERISTICASLBP

Entrada: Imagem I , **inteiro** $tamBloco$, **inteiro** $tamBin$

Saída: Vetor de Características X

inicio

$X = 0$

$B = \text{DivideImagemEmBlocos}(I, tamBlocos)$

para cada $B_i \in B$ **faça**

$I_{LBP} = \text{AplicarLBP}(B_i)$

$H_i = \text{Histograma}(I_{LBP}, tamBin)$

 Concatena(X, H_i)

fim

fim

retorna X

Fonte: Elaborado pelo autor.

2.1.3 Extração de Características com DCT

A transformada discreta do cosseno (DCT, do inglês, *Discret Cossine Transform*), surgida em 1974, é uma das transformadas comumente utilizadas em processamento de imagens, processamento de sinais e compressão de dados. Há quatro variações da DCT: DCT-I, DCT-II, DCT-III e DCT-IV. A DCT-II é mais usada em processamento de imagens por possuir uma forte capacidade de compactação de energia, sendo definida pela função a seguir (PEDRINI; SCHWARTZ, 2008):

$$F(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} f(x, y) \cos \frac{(2x+1)u\pi}{2w} \cos \frac{(2y+1)v\pi}{2h},$$

sendo

$$\alpha(u), \alpha(v) = \begin{cases} \sqrt{1/n}, & \text{se } u = v = 0; \\ \sqrt{2/n}, & \text{caso contrário;} \end{cases}$$

$$u = 0, 1, \dots, w-1,$$

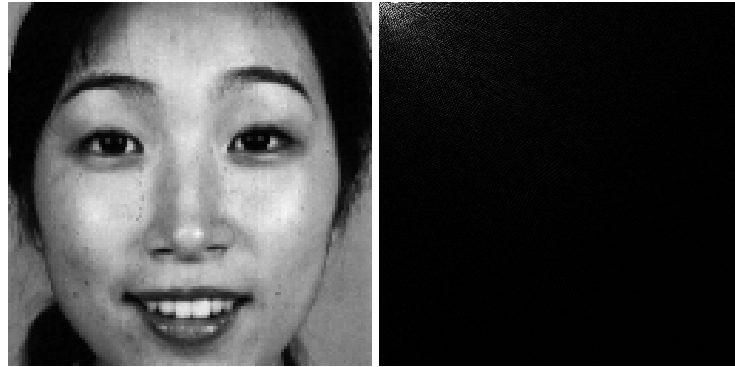
$$v = 0, 1, \dots, h-1,$$

e w e h as dimensões da imagem.

Uma das principais propriedades da DCT é concentrar a maior parte das informações em regiões próximas da origem. Assim, a maioria dos coeficientes, por serem próximos a zero, pode ser descartada sem perda de informações significativas da imagem transformada. A DCT tem sido bastante utilizada em compressão, filtragem e extração de características em imagens (PEDRINI; SCHWARTZ, 2008) (GONZALEZ; WOODS, 2009).

Para extração das características de uma imagem resultante da aplicação da DCT, usa-se apenas os primeiros coeficientes e descarta-se os demais. Neste trabalho, usou-se a mesma metodologia de (BHADU *et al.*, 2012), que aplica a transformada DCT à imagem inteira. Porém, resolveu-se investigar a aplicação da DCT em blocos. Com isso, a imagem é dividida em blocos, por exemplo, de tamanho 16 x 16. Para cada bloco, obtém-se a DCT e armazena-se os coeficientes mais significativos. Os coeficientes foram concatenados para gerar o vetor de características da imagem analisada. O Algoritmo 3 ilustra o processo de obtenção do vetor de características de uma imagem. A função *DivideImagemEmBlocos* divide a imagem em blocos

Figura 8 – Imagem original (à esquerda) e sua transformada DCT (à direita).



Fonte: Figura original (LYONS *et al.*, 1998); figura modificada (produzido pelo autor).

do mesmo tamanho. Nos testes com aplicação de DCT à imagem inteira, o tamanho do bloco (*tamBloco*), deve ser igual às dimensões da imagem original. A função *ObterCoeficientes* retorna os coeficientes correspondentes à submatriz do canto superior esquerdo do bloco transformado, cuja dimensão é $\sqrt{numCoefs} \times \sqrt{numCoefs}$. A dimensão do vetor de características é dada pelo produto do número de blocos pelo número de coeficientes.

Algoritmo 3: OBTERCARACTERISTICASDCT

Entrada: Imagem I , inteiro $tamBloco$, inteiro $numCoefs$

Saída: Vetor de Características X

inicio

$X = 0$

$B = \text{DivideImagemEmBlocos}(I, tamBloco)$

para cada $B_i \in B$ **faça**

$I_{DCT} = \text{AplicarDCT}(B_i)$

$C_i = \text{ObterCoeficientes}(I_{DCT}, numCoefs)$

 Concatena(X, C_i)

fim

fim

retorna X

Fonte: Elaborado pelo autor.

2.2 REDUÇÃO DE DIMENSIONALIDADE

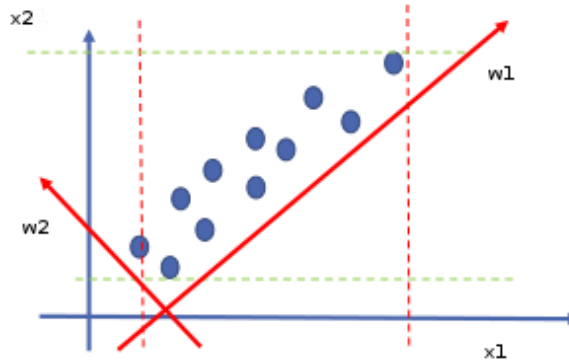
Em geral, sistemas de classificação de padrões consistem na manipulação de conjunto de dados com milhares de características, o que pode reduzir a performance do classificador, tornando-se inviável. Para viabilizar essa tarefa de classificação é necessário simplificar os dados trabalhados de forma a reduzir o número de características, mantendo as mais significativas. Neste trabalho, foram usadas duas metodologias de redução de dimensionalidade: Análise de Componentes Principais (PCA) e Análise de Discriminantes Lineares (LDA).

2.2.1 Análise de Componentes Principais

A Análise de Componentes Principais (PCA, do inglês *Principal Component Analysis*) considera os dados originais do problema como variáveis aleatórias correlacionadas e que formam uma base de um espaço vetorial de elevada dimensão. A PCA encontra uma base de vetores não correlacionados de um subespaço do espaço original de modo a preservar o máximo possível de variabilidade. A projeção dos dados nesse subespaço de menor dimensão é o que garante a redução de dimensionalidade do problema original (JOLLIFFE, 2002).

Consideremos um conjunto $X = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^d$. A PCA consiste em projetar os elementos de X em novas direções $w_1, w_2, \dots, w_d$, de forma que a primeira componente contenha a maior variância dentre os dados projetados, a segunda componente a segunda maior variância e assim, sucessivamente. Por exemplo, a Figura 9 apresenta um conjunto de dados bidimensionais com componentes x_1 e x_2 . Observemos que a variância dos dados nas duas direções originais é praticamente a mesma. Porém, quando projetamos esses dados nas direções w_1 e w_2 percebemos que a variância dos dados na primeira direção é bem maior que na segunda. Assim, w_1 é denominada de primeira componente principal e w_2 a segunda.

Figura 9 – PCA - ideia geral.



Fonte: Elaborado pelo autor.

A projeção de um elemento $x \in X$ sobre uma direção unitária w_i é dada por $y = w_i^T x$. Assim, a variância de y é dada por:

$$\begin{aligned}
 \text{Var}(w_i^T x) &= \text{Var}(y) \\
 &= E[(w_i^T x - w_i^T \mu)^2] \\
 &= E[(w_i^T x - w_i^T \mu)(w_i^T x - w_i^T \mu)] \\
 &= E[w_i^T (x - \mu)(x - \mu)^T w_i] \\
 &= w_i^T \Sigma w_i,
 \end{aligned}$$

sendo E o valor esperado, Σ a matriz de covariância e μ a média.

Assim, a primeira componente principal maximiza $w_1^T \Sigma w_1$ tal que $w_1^T w_1 = 1$. Segundo (JOLLIFFE, 2002), essa direção é dada por:

$$L(w, \lambda) = \max_w \{w^T \Sigma w + \lambda(1 - w^T w)\}.$$

Derivando a equação anterior em relação a w , obtém-se:

$$\Sigma w = \lambda w.$$

Assim sendo, a primeira componente principal é determinada pelo autovetor de Σ correspondente ao seu maior autovalor; a segunda componente, pelo autovetor correspondente ao segundo maior autovalor e, assim, sucessivamente.

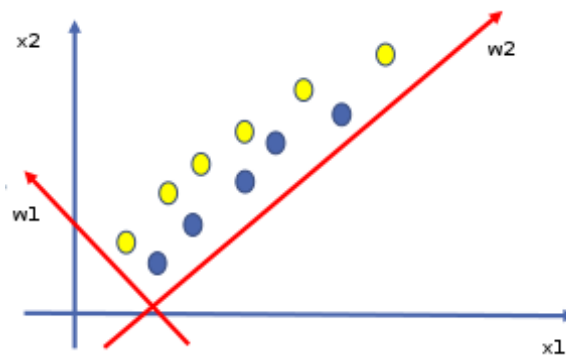
Dessa forma, para obtermos as $k (\ll d)$ componentes principais de um elemento $z \in \mathbb{R}^d$, devemos projetá-lo nas direções de cada um dos k autovetores selecionados.

2.2.2 Análise de Discriminantes Lineares

A Análise de Discriminantes Lineares (LDA, do inglês *Linear Discriminant Analysis*) consiste na redução de dimensionalidade de um conjunto de dados pertencentes a classes distintas de forma que cada um dos elementos fique mais próximos dos pertencentes à mesma classe e mais distante dos que pertencem a outras classes. A PCA, vista na seção anterior, não considera as classes durante a obtenção das componentes principais. A Figura 10 ilustra essa idéia. Ao aplicarmos PCA, a primeira componente obtida seria a indicada por w_2 e não forneceria uma separação entre as classes. Porém, ao aplicar LDA, a primeira componente obtida é a indicada por w_1 , que separa claramente as classes.

Consideremos novamente um conjunto de dados $X = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^d$, distribuídos em c classes diferentes. Seja D_i um subconjunto de X cujo elementos pertencem a classe i . A LDA consiste em projetar os elementos de X em uma direção w que maximiza a distância entre as classes e minimiza a distância entre os elementos de uma mesma classe.

Figura 10 – LDA - ideia geral.



Fonte: Elaborado pelo autor.

Esse objetivo é atingido quando a função abaixo é maximizada (XIANG *et al.*, 2006):

$$J(w) = \frac{w^T S_B w}{w^T S_W w},$$

sendo S_B a matriz de espalhamento entre classes e S_W a matriz de espalhamento dentro das classes. Ou seja, a idéia é maximizar S_B e minimizar S_W .

Para obter S_B , calcula-se a média dos elementos de X ,

$$m = \frac{1}{n} \sum_{k=1}^n x_k$$

e a média de cada conjunto (classe) D_i ,

$$m_i = \frac{1}{n_i} \sum_{x \in D_i} x.$$

Assim, temos:

$$S_B = \sum_{i=1}^c n_i (m_i - m)(m_i - m)^T.$$

Para calcular S_W , obtém-se a matriz de espalhamento S_i de cada classe:

$$S_i = \sum_{x \in D_i} (x - m_i)(x - m_i)^T, \quad i = 1, \dots, c.$$

Assim, temos:

$$S_W = \sum_{i=1}^c S_i.$$

Segundo (XIANG *et al.*, 2006), o vetor w que maximiza $J(w)$ satisfaz:

$$S_B w = \lambda S_W w.$$

Se S_W não for singular, a equação acima pode ser transformada em um problema de autovalores.

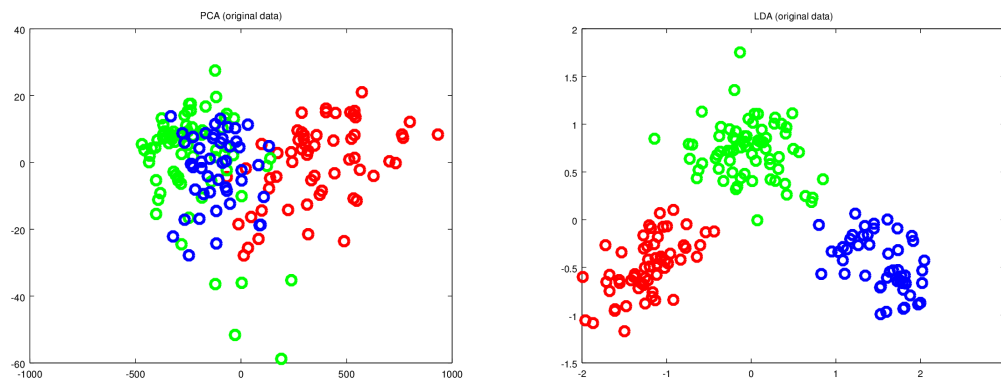
$$(S_W^{-1} S_B) w = \lambda w.$$

Segundo (XIANG *et al.*, 2006), como o posto de S_B é no máximo $c - 1$, é possível extrair $c - 1$ componentes principais, cujas direções correspondem aos autovetores associados aos $c - 1$ autovalores de $(S_W^{-1} S_B)$, ordenados de forma decrescente. Dessa forma, para obtermos as $c - 1$ componentes principais de um elemento $z \in \mathbb{R}^d$, devemos projetá-lo nas direções de cada um dos autovetores obtidos.

A Figura 11, baseada em (WAGNER, 2011), mostra exemplos de redução de dimensionalidade usando PCA e LDA. Percebe-se facilmente que as componentes principais obtidas

por LDA separam claramente as classes do problema ilustrado, enquanto que com PCA, essa separação não ocorre.

Figura 11 – À esquerda, duas primeiras componentes principais obtidas com PCA. À esquerda, com LDA.



Fonte: Elaborado pelo autor.

Geralmente, para utilização da LDA, primeiramente aplica-se PCA para extrair os $n - c$ componentes principais, sendo n o número de padrões e c o número de classes. Os dados originais são então projetados nas $n - c$ direções das componentes principais e, em seguida, aplica-se LDA para extração $c - 1$ componentes.

2.3 CLASSIFICAÇÃO

A classificação pode ser feita através dos códigos de ações faciais (FACS, do inglês *Face Action Coding System*) ou dos protótipos de expressões (FASEL; LUETTIN, 2003). Neste trabalho, aborda-se apenas a classificação por meio de protótipos. Assim, a identificação da expressão é feita por métricas de comparação com cada um dos protótipos previamente determinados.

Os classificadores usados foram o KNN, Naive-Bayes, Redes RBF e Máquinas de Vetores de Suporte, os quais estão descritos a seguir.

2.3.1 *K-Nearest Neighbors (KNN)*

O KNN (do inglês, *K-Nearest Neighbors*) é um modelo de classificação não paramétrico baseado apenas no conjunto de treinamento. Para classificar um elemento x de classe

desconhecida, o KNN seleciona dentre o conjunto de treinamento os k padrões mais próximos de x , usando alguma métrica de distância. Dentre esses k elementos selecionados, verifica-se qual a classe predominante entre eles e a atribui a x . Em problemas de classificação binária, para evitar empate, escolhe-se sempre o valor de k como sendo ímpar (NORVIG; RUSSELL, 2014).

A distância mais comumente utilizada é a euclidiana, porém outras métricas vêm sendo aplicadas, como a distância de Mahalanobis. Neste trabalho, utilizou-se a distância euclidiana e $k = 1$. Foram feitos experimentos com outros valores de k , porém não houve melhoria considerável na taxa de acertos do classificador.

2.3.2 Naive Bayes

O classificador Naive Bayes é um classificador probabilístico baseado no teorema de Bayes com forte independência entre as características. Também conhecido como Bayes ingênuo ou idiota (NORVIG; RUSSELL, 2014), pois assumem que as características são independentes (WEBB; COPSEY, 2011).

Esse classificador vem sendo estudado desde os anos 1950 e é muito usado em categorização de informações, como por exemplo, na identificação se um documento trata-se de economia ou de esportes (NORVIG; RUSSELL, 2014).

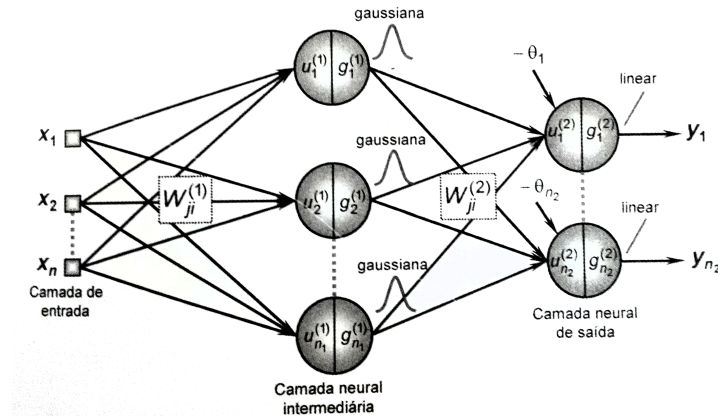
2.3.3 Redes RBF

A Rede Neural de Função de Base Radial (RBF, do inglês *Radial Basis Functions*) corresponde a um tipo de rede neural artificial com uma única camada intermediária, cuja função de ativação é do tipo radial. A arquitetura de uma rede RBF é do tipo *feed-forward* não-linear e o treinamento é realizado de forma supervisionada (SILVA *et al.*, 2010).

Uma rede RBF é composta de três camadas. A primeira camada é responsável pela captação dos dados de entrada. A segunda, a camada intermediária, realiza um mapeamento de um espaço de entrada para um espaço intermediário e usa funções de ativação do tipo radial, como as gaussianas. A terceira camada, a de saída, é linear (HAYKIN, 2009). A Figura 12 ilustra uma rede RBF.

As redes RBF são bastante utilizados em aproximação de funções e classificação de padrões complexos. Para os problemas de classificação, as fronteiras de decisão são delimitadas por regiões hipersféricas (SILVA *et al.*, 2010).

Figura 12 – Estrutura de uma rede RBF.

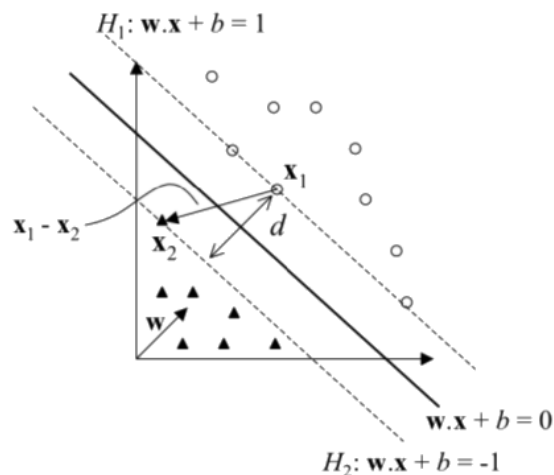


Fonte: (SILVA *et al.*, 2010).

2.3.4 Máquinas de Vetores de Suporte

As Máquinas de Vetores de Suporte (SVM, do inglês *Support Vector Machine*) também é um modelo de classificação paramétrico. SVM constrói um separador de margem máxima com a maior distância possível de todas as classes. Trabalha com dados de qualquer dimensão, que são mapeados para um outro espaço para que se tornem separáveis (NORVIG; RUSSELL, 2014) (HAYKIN, 2009). A Figura 13 ilustra a superfície de separação máxima entre as classes *triângulo* e *círculo* representada pela reta $w \cdot x + b = 0$, onde w e b são obtidos por uma problema de otimização de programação quadrática (PQ).

Figura 13 – Representação do SVM.



Fonte: (LORENA; CARVALHO, 2007).

Para resolver o problema do SVM de forma mais rápida, Platt e colegas (PLATT *et al.*, 1999) propuseram um algoritmo denominado SMO (Sequential Minimal Optimization). Esse algoritmo é usado para treinamento de máquinas de vetores de suporte e resolve o problema de programação quadrática (PQ) do SVM sem o armazenamento de matrizes extras e sem o uso de métodos de solução numérica. O problema de PQ global é decomposto em uma série de subproblemas e então escolhe-se o menor problema de otimização possível para resolver cada etapa. Esses pequenos subproblemas são então solucionados de forma analítica, evitando assim o uso de métodos numéricos de otimização (PLATT *et al.*, 1999).

A quantidade de espaço em memória necessária para a execução do SMO é proporcional ao tamanho do conjunto de treinamento. Com isso, o SMO trabalha facilmente com conjuntos de treinamento com elevada cardinalidade (PLATT *et al.*, 1999). Utiliza geralmente núcleos polinomiais ou de Gauss.

2.4 SELEÇÃO DE PROTÓTIPOS

Alguns classificadores, como o KNN, precisam conhecer todos os padrões de treinamento durante a predição de um novo padrão. Se o conjunto de treinamento possuir muitos elementos, a classificação consumirá tempo e exigirá muita memória para processamento. Visando a reduzir a cardinalidade do conjunto de treinamento, alguns métodos, denominados seleção de protótipos, têm sido propostos (ANGIULLI, 2005).

Uma dessas técnicas obtém um subconjunto consistente do conjunto de padrões originais. Um subconjunto S de T é dito consistente se, a partir dele, todos os elementos de $T - S$ são classificados corretamente usando a regra do vizinho mais próximo. Um algoritmo usado para implementar essa idéia é o FCNN (do inglês *Fast Condensed Nearest Neighbor*) (ANGIULLI, 2005). O Algoritmo 4 resume o funcionamento do FCNN. A função *ObterCentroides* retorna um conjunto contendo todos os centroides das diferentes classes contidas em T . Já a função *ObterMaisProximoComDiferenteClasse* retorna um elemento q de $T - S$ tal que p é o elemento de S mais próximo de q e cuja classe seja distinta da classe de p . De forma matemática, temos que: $ObterMaisProximoComDiferenteClasse(p, S, T) = \{q \mid distancia(q, p) \leq distancia(q, p'), \forall p' \in S, \text{ e } classe(p) \neq classe(q)\}$. Dessa forma, a função *ObterMaisProximoComDiferenteClasse* retorna o elemento de $T - S$ que é erroneamente classificado quando comparado ao elemento p .

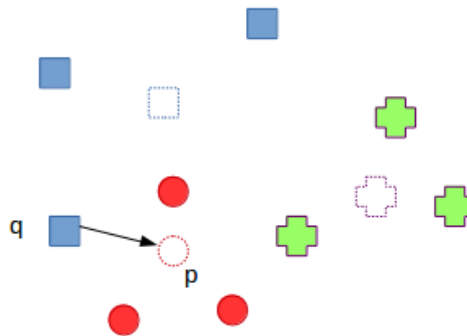
Algoritmo 4: FCNN

Entrada: Conjunto T
Saída: Subconjunto consistente S de T
inicio
 $S = \emptyset;$
 $\Delta S = \text{ObterCentroides}(T);$
repita
 $S = S \cup \Delta S;$
 $\Delta S = \emptyset;$
para cada $p \in S$ **faça**
 $q = \text{ObterMaisProximoComDiferenteClasse}(p, S, T);$
se $q \neq \text{NULL}$ **então**
 $\Delta S = \Delta S \cup q;$
fim
fim
até $\Delta S = \emptyset;$
fim
retorna S

Fonte: Elaborado pelo autor.

Como visto na representação algorítmica, o conjunto S é inicializado com os centróides de cada classe distinta em T . Em seguida, para cada elemento $p \in S$, verifica se existe algum elemento $q \in (T - S)$, tal que q é o mais próximo de p e a classe de q é diferente de p . Os elementos q que atendam a essa verificação são adicionados a S . Esse processo é repetido até que não existam mais elementos em $T - S$ que atendam a essa condição. A Figura 14 apresenta um exemplo de um ponto q que deve ser adicionado a S , que está representado pelo elementos tracejados e que correspondem aos centróides das classes *quadrado*, *círculo* e *cruz*. Ao verificar o elemento p de S , a função *ObterMaisProximoComDiferenteClasse* identifica q como sendo o mais próximo de p e com classe distinta. Logo, q é adicionado S . Nesse caso, o conjunto formado pelos centróides e o elemento q corresponde a um subconjunto consistente de T .

Figura 14 – FCNN - exemplo.



Fonte: Elaborado pelo autor.

CONSIDERAÇÕES FINAIS

Este capítulo apresentou um detalhamento de cada um dos métodos de extração de características (filtros de Gabor, LBP e DCT), dos classificadores (KNN, Naive-Bayes, RBF e SVM-SMO) e do algoritmo de seleção de protótipos FCNN.

O próximo capítulo apresenta o detalhamento dos experimentos realizados, combinando os métodos de extração de características e os classificadores apresentados. Detalha ainda os banco de expressões faciais utilizados, análise dos resultados e discussão.

O último capítulo expõe as conclusões obtidas com a realização destes experimentos, enfatizando as características estudadas, bem como algumas recomendações para trabalhos futuros.

3 RESULTADOS E DISCUSSÃO

O presente capítulo visa a detalhar os experimentos realizados a fim de comparar os métodos de extração de características. Inicialmente, descrevem-se os bancos de faces utilizados. Em seguida, apresentam-se os detalhes de implementação, os índices de desempenho avaliados, os softwares e o hardware utilizados na realização dos experimentos. Por fim, os resultados dos experimentos realizados são analisados.

3.1 BASE DE DADOS CONSIDERADAS

Os bancos selecionados para realização dos experimentos são os mais utilizados na literatura, a saber:

3.1.1 JAFFE

A base JAFFE (*Japanese Female Facial Expression*) (LYONS *et al.*, 1998) contém 213 imagens de 10 indivíduos japoneses do sexo feminino. Cada uma das pessoas possui pelo menos 2 imagens em cada uma das 7 expressões faciais: feliz (*happy*), triste (*sad*), zangado (*angry*), surpreso (*surprised*), desgosto ou repugnado (*disgusted*), medo ou assustado (*scared*) e neutro. Essas expressões estão ilustradas na Figura 15. As imagens originais possuem dimensões 256×256 e podem ser obtidas em <<http://www.kasrl.org/jaffe.html>>.

Figura 15 – Imagens originais da base JAFFE. Da esquerda para a direita: zangado, desgosto, medo, feliz, neutro, triste e surpreso.



Fonte: (LYONS *et al.*, 1998).

3.1.2 Yale

A base Yale (YALE, 2000) contém imagens de 15 indivíduos, cada um deles com imagens com as seguintes expressões faciais: neutro, feliz (*happy*), surpreso (*surprised*), triste (*sad*), sonolento (*sleep*) e piscando olho (*wink*), sendo as 4 primeiras utilizadas neste trabalho. Há ainda imagens com variações de iluminação e com oclusão (óculos, barba, cabelos na

face etc). As imagens originais possuem dimensões 320×243 e podem ser obtida em <<http://vision.ucsd.edu/content/yale-face-database>>. A Figura 16 apresenta algumas imagens dessa base.

Figura 16 – Imagens originais da base YALE. Da esquerda para a direita: neutro, feliz, triste e surpreso.



Fonte: (YALE, 2000).

3.1.3 CK+

A base CK+ (Cohn-Kanade estendida) (LUCEY *et al.*, 2010) contém 593 sequências de 123 pessoas distintas. A quantidade de imagens por sequência varia de 10 a 60. A sequência de cada pessoa/expressão varia desde o *onset* (imagem neutra) até o pico da expressão (intensidade máxima). É o banco de face mais utilizado para testes de reconhecimento de expressões faciais. As imagens possuem dimensão 640×490 .

Do total de sequências, somente 327 são rotuladas com uma das sete expressões apresentadas na Figura 17. Porém, para facilitar os testes inter-bases, foram selecionadas aleatoriamente 33 imagens correspondentes à expressão neutra.

Figura 17 – Imagens originais da base CK+. Da esquerda para a direita: zangado, desprezo, desgosto, medo, feliz, triste e surpreso.



Fonte: (LUCEY *et al.*, 2010).

3.1.4 Comparativo entre os bancos de faces

Apresenta-se na Tabela 2, um resumo comparativo entre os 3 banco de dados utilizados. Consideremos I.R. - número de imagens rotuladas; IND. - número de indivíduos distintos; EXP. - número de expressões; IND./EXP. - número de expressões de um mesmo tipo para um

mesmo indivíduo; ZA-DE-ME-FE-NE-TR-SU-DP - números de expressões dos tipos zangado, desgosto, medo, feliz, neutro, triste, surpreso e desprezo, respectivamente.

Tabela 2 – Comparação entre os banco de faces utilizados nos experimentos.

Base	I.R.	IND.	EXP.	IND./EXP.	ZA	DE	ME	FE	NE	TR	SU	DP
JAFFE	213	10	7	2-4	30	28	32	32	30	31	30	0
YALE	60	15	4	1	0	0	0	15	15	15	15	0
CK+	360	118	7	0-1	45	59	25	69	33	28	83	18

Fonte: Elaboração própria.

3.2 PROCEDIMENTOS EXPERIMENTAIS

Para a realização dos procedimentos experimentais, foi desenvolvido um sistema contendo três módulos: pré-processamento, extração de características e classificação. Os dois primeiros foram implementados em C++ com a biblioteca OpenCV. Para o módulo de classificação, foi implementado em Java uma funcionalidade de integração com os métodos disponibilizados na biblioteca WEKA. Nos experimentos com os filtros de Gabor, como os vetores de características possuem dimensões elevadas, realizou-se testes com métodos de redução de dimensionalidade PCA e LDA. Foi utilizada a implementação da PCA disponibilizada na OpenCV e a implementação da LDA da linguagem R.

Os experimentos foram realizados em um computador *desktop* com processador Intel® Core™i5-2410M (2.30GHz até 2.90GHz com Turbo Boost, 4Threads, 3M cache), com 6GB de memória RAM, placa de vídeo AMD Radeon™ (HD 6470M 64-bit, 512MB) com sistema operacional Ubuntu 14.04.

Inicialmente foram realizados alguns experimentos visando a comparar os métodos Viola-Jones (VIOLA; JONES, 2001) e Deng (DENG *et al.*, 2005) para a etapa de pré-processamento. Para a implementação de ambos, foram utilizadas as funcionalidades de localização de face (para o primeiro método) e olhos (para o segundo) da biblioteca OpenCV. Verificou-se que o segundo apresentou melhor acurácia, verificou-se a necessidade de marcação manual dos olhos, principalmente com a base CK+¹. Dessa forma, optou-se por utilizar o método Viola-Jones em todos os demais experimentos. Como o método Viola-Jones recorta boa

¹ Em algumas imagens, não foi possível localizar corretamente os olhos de forma automática. Nesses casos, foi necessário marcar os olhos manualmente para o sistema fazer o recorte da face corretamente.

parte do fundo da imagem contendo a face, adotou-se um recorte usando a escala $0,7L \times 0,9A$, sendo L a largura da face e A a altura da face recortada por Viola-Jones. A Figura 18 apresenta uma imagem de uma face recortada pelo Viola-Jones original (esquerda) e outra com o método ajustado (direita).

Figura 18 – Imagens recortadas com o método Viola-Jones (VIOLA; JONES, 2001).



Fonte: Adaptadas de (LUCEY *et al.*, 2010).

Como foi utilizada uma abordagem supervisionada para o problema em estudo, os conjuntos de dados foram divididos em treinamento e testes. Em todos os experimentos, foram usados métodos de validação cruzada *leave-one-out* e *leave-one-group-out*. Para o primeiro caso, retirou-se uma imagem de cada conjunto para teste, ficando o restante para treinamento. Esse procedimento foi repetido para cada imagem de cada banco de dados. Dessa forma, o número de *folds* corresponde ao número de imagens da base. Já no segundo caso, retirou-se todas as imagens de uma pessoa ou de um grupo de pessoas para teste, ficando as imagens das pessoas restantes para treinamento. Neste caso, o número de *folds* corresponde ao número de grupos. Inicialmente, os experimentos foram realizados individualmente com cada um dos banco de faces. Todas as imagens foram pré-processadas e normalizadas para o tamanho 96×128 pixels. Geralmente, trabalha-se com dimensões maiores na vertical pelo fato de face ser mais alongada nessa direção, principalmente em imagens relacionada à expressão de surpresa.

Com isso, os experimentos para análise das características foram agrupados conforme a seguir:

- a) Gabor;
- b) Gabor e PCA;
- c) Gabor, PCA e LDA;
- d) LBP;

e) DCT.

Cada um dos conjuntos de características foi avaliado com os classificadores KNN, Naive-Bayes, Redes RBF e SVM, cujos parâmetros utilizados em cada um são mostrados na Tabela 3. Os parâmetros dos classificadores KNN ($K = 1$), SVM ($C = 0,6$) e RBF ($B = 2$) foram obtidos através de uma funcionalidade da biblioteca WEKA (Classe *CVParameterSelection*) (WEKA, 2014). A escolha destes deu-se por conta de serem bastante referenciados na literatura.

Tabela 3 – Comparação entre os classificadores.

CLASSIFICADOR	PARÂMETROS
KNN	$K = 1$, distância euclidiana
SVM	$C = 0,6$, núcleo linear
RBF	$B = 2$
NB	Sem parâmetros

Fonte: Produzido pelo autor.

Outro experimento realizado foi a avaliação de um método de seleção de protótipo, no caso o FCNN. Em cada um dos grupos de características, o FCNN foi aplicado. Os protótipos selecionados foram testados com cada um dos classificadores.

A análise dos experimentos avaliou os seguintes itens gerais:

- Acurácia média** - foi calculada a acurácia média de cada um dos experimentos, que é obtida pela média aritmética das acurácias obtidas em cada *fold*. Também foi verificado o *desvio-padrão*.
- Tempo** - em cada experimento, calculou-se o tempo médio gasto para a extração das características e para a classificação.
- Relação entre indivíduos conhecidos x desconhecidos** - foram realizados testes considerando indivíduos conhecidos e desconhecidos durante a etapa de treinamento. Dessa forma, comparou-se as acurácias obtidas nesses dois cenários.

Para alguns experimentos, foram avaliados índices de desempenho por classes, que foram:

- Sensitividade (SENS.)** - também conhecida como taxa de verdadeiros positivos ou acurácia média de uma classe, corresponde ao percentual de acertos de todos

os padrões de uma mesma classe. Supondo que em um determinado experimento com 40 imagens de pessoas com a expressão sorriso, 30 tenham sido classificadas corretamente, o valor da sensibilidade para essa classe é $\frac{30}{40} = 0,75$.

- b) **Precisão (PREC.)** - indica o percentual de padrões verdadeiros positivos dentre o total de todos os classificados para uma determinada classe. Considerando o mesmo exemplo anterior, suponhamos que todas as 40 imagens da expressão feliz tenham sido classificadas corretamente, mas que 10 imagens da expressão surpresa tenham sido classificadas erroneamente como feliz. Nesse caso, a precisão da classe feliz é $\frac{40}{40+10} = 0,80$.
- c) **Especificidade (SPEC.)** - também conhecida como taxa de verdadeiros negativos, indica para uma determinada classe, o percentual de acertos das demais classes.
- d) **F Measure** - fornecida pela expressão

$$\frac{2}{1/PREC + 1/(SENS)},$$

sendo *PREC* a precisão e *SENS* a sensibilidade (WEBB; COPSEY, 2011).

3.3 RESULTADOS

Nesta seção, apresenta-se uma análise dos experimentos realizados, cujos resultados estão nas subseções seguintes, organizadas por banco de faces. Em cada subseção, apresentam-se os resultados combinando os métodos de extração de características com os classificadores. Mostram-se também resultados com variações de parâmetros dos métodos de extração de características, com ou sem redução de dimensionalidade e utilização do FCNN como seleção de protótipos. Ao final de cada subseção é apresentado um resumo dos resultados obtidos. Para finalizar esta seção, mostra-se uma análise comparativa entre os métodos utilizados.

3.3.1 Experimentos com a base JAFFE

Os experimentos com a base JAFFE foram realizados usando os métodos *leave-one-out* e *leave-one-group-out*. Para o primeiro método, selecionou-se uma imagem da base e manteve-se as demais para treinamento, repetindo-se 213 vezes. Quanto ao segundo, a base foi dividida em 10 partições da seguinte forma: selecionou-se um dos indivíduos e todas as suas imagens foram colocadas no conjunto de teste; as imagens dos demais indivíduos, no conjunto

de treinamento. Em cada grupo, selecionou-se um indivíduo distinto para teste, totalizando 10 *folds*, que nesse caso também é denominado de *leave-one-subject-out*.

Filtros de Gabor

As imagens filtradas com Gabor também possuem a mesma dimensão da imagem original, que é 96×128 . Utilizando os 40 filtros comumente usados, o vetor de características gerado teria dimensão 491.520 ($40 \times 96 \times 128$), que inviabilizaria o seu uso em diversas aplicações. Em função disso, usou-se o fator de amostragem 4 aplicado às imagens filtradas em todos os experimentos com filtros de Gabor. Assim, as dimensões das imagens filtradas foram reduzidas para 24×32 *pixels* e a dimensão do vetor de características para 30720 (40 filtros $\times 24 \times 32$).

Os primeiros testes foram realizados usando somente as características extraídas com filtros de Gabor, sem aplicação de métodos para redução de dimensionalidade. Para os experimentos com PCA, foi feito um procedimento empírico para determinar o número de componentes principais que fornecesse a melhor acurácia com esse banco de dados. Dessa forma, o procedimento empírico consistiu em variar o número de componentes principais selecionadas de 5 a 200, com fator de variação 5. Para cada quantidade selecionada, realizou-se o experimento *leave-one-out* e calculou-se a acurácia média. Verificou-se que o valor 80 produziu a melhor acurácia. Assim, optou-se por utilizar esse número de componentes principais.

Foram realizados ainda, testes com uso de LDA para redução de dimensionalidade. Como o número de componentes principais extraídas pelo LDA é de $c - 1$, sendo c o número de classes distintas, o vetor de características usados nesses experimentos possui dimensão 6.

Além dos experimentos com redução de dimensionalidade, realizou-se testes aplicando o seletor de protótipos FCNN em todos os conjuntos de treinamento.

As Tabelas 4 e 5 apresentam as acurácias médias obtidas utilizando cada um dos conjuntos de características de Gabor, com e sem seleção de protótipos (FCNN)². A segunda e terceira colunas apresentam os resultados obtidos com somente filtros de Gabor. A quarta e quinta colunas apresentam as acurácias obtidas com filtros de Gabor e PCA. A sexta e sétima colunas apresentam os resultados obtidos com filtros de Gabor e PCA/LDA. As colunas 3, 5 e 7 apresentam as acurácias obtidas com o uso do FCNN aplicado aos conjuntos de treinamentos.

² Em função da dimensão e da cardinalidade reduzidas dos conjuntos de treinamentos, não foi possível realizar alguns experimentos com redes RBF e características PCA/LDA/FCNN.

Tabela 4 – Base JAFFE: acurácias obtidas (%) com filtros de Gabor usando *leave-one-out*.

CLASSIFICADOR	GABOR		GABOR/PCA		GABOR/PCA/LDA	
	AC	AC/FCNN	AC	AC/FCNN	AC	AC/FCNN
KNN	92,49	91,55	90,14	92,02	93,90	91,55
SVM	96,71	91,55	93,43	81,22	94,84	17,37
RBF	80,75	70,89	86,38	54,46	96,24	-
NB	81,22	53,05	79,81	56,34	94,84	65,73

Fonte: Elaborado pelo autor. AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

Tabela 5 – Base JAFFE: acurácias obtidas (%) com filtros de Gabor usando *leave-one-subject-out*.

CLASSIFICADOR	GABOR		GABOR/PCA		GABOR/PCA/LDA	
	AC	AC/FCNN	AC	AC/FCNN	AC	AC/FCNN
KNN	49,77	67,61	49,30	69,01	66,67	66,20
SVM	69,48	65,26	69,01	56,81	66,67	37,56
RBF	59,15	48,36	53,05	32,39	68,08	-
NB	61,97	31,92	53,52	31,92	67,61	33,33

Fonte: Elaborado pelo autor. AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

Em ambos os casos, o classificador SVM com características de Gabor gerou as melhores acurácias: 96,71% para indivíduos conhecidos e 69,48% para desconhecidos. As Tabelas 6 e 7 apresentam as matrizes de confusão desses dois resultados. Apesar de uma taxa de acerto menor para desconhecidos, observa-se que nesses experimentos, as expressões de felicidade (FE), surpresa (SU) e raiva (RA) apresentam boas precisões. Observa-se também uma melhoria significativa com o uso do FCNN com o classificador KNN, nos quais houve uma melhoria de 19,71% na acurácia média com as características Gabor/PCA (Tabela 5).

Tabela 6 – Base JAFFE: melhor resultado obtido com filtros de Gabor para indivíduos conhecidos (*leave-one-out*): SVM sem seleção de protótipos com acurácia média de 96,71%.

MATRIZ DE CONFUSÃO								MÉTRICAS			
CLASSES	FELIZ	TRISTE	SURPRESO	DESGOSTO	RAIVA	MEDO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	96,77	03,23	00,00	00,00	00,00	00,00	00,00	0,94	0,97	0,99	0,95
TRISTE	03,23	93,55	00,00	00,00	00,00	03,23	00,00	0,97	0,94	0,99	0,95
SURPRESO	03,33	00,00	93,33	00,00	00,00	03,33	00,00	1,00	0,93	1,00	0,97
DESGOSTO	00,00	00,00	00,00	96,55	03,45	00,00	00,00	0,97	0,97	0,99	0,97
RAIVA	00,00	00,00	00,00	03,33	96,67	00,00	00,00	0,97	0,97	0,99	0,97
MEDO	00,00	00,00	00,00	00,00	00,00	100,00	00,00	0,94	1,00	0,99	0,97
NEUTRO	00,00	00,00	00,00	00,00	00,00	00,00	100,00	1,00	1,00	1,00	1,00

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Tabela 7 – Base JAFFE: melhor resultado obtido com filtros de Gabor para indivíduos desconhecidos (*leave-one-subject-out*): SVM sem seleção de protótipos com acurácia média de 69,48%.

MATRIZ DE CONFUSÃO								MÉTRICAS			
CLASSES	FELIZ	TRISTE	SURPRESO	DESGOSTO	RAIVA	MEDO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	70,97	06,45	12,90	00,00	00,00	00,00	09,68	0,88	0,71	0,98	0,79
TRISTE	03,23	61,29	03,23	16,13	03,23	09,68	03,23	0,59	0,61	0,93	0,60
SURPRESO	03,33	00,00	83,33	00,00	00,00	06,67	06,67	0,69	0,83	0,94	0,76
DESGOSTO	00,00	06,90	00,00	75,86	17,24	00,00	00,00	0,65	0,76	0,93	0,70
RAIVA	03,33	10,00	00,00	03,33	80,00	00,00	03,33	0,69	0,80	0,94	0,74
MEDO	00,00	18,75	15,62	18,75	00,00	37,50	09,38	0,71	0,38	0,97	0,49
NEUTRO	00,00	00,00	03,33	00,00	16,67	00,00	80,00	0,71	0,80	0,95	0,75

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Padrões Binários Locais (LBP)

Os experimentos com LBP foram realizados com o padrão LBP (8, 1) aplicado em blocos nas imagens. Foram feitos experimentos utilizando os seguintes tamanhos de blocos: 8×8 (192 blocos), 16×16 (48 blocos) e 32×32 (12 blocos). Usou-se o histograma 58-*bin*, que considera somente os padrões uniformes, de cada bloco para geração do vetor de características. Por exemplo, para os blocos 16×16 , a dimensão do vetor de características gerado foi 2784 (48×58).

A Tabela 8 apresenta um resumo dos experimentos realizados usando LBP como método de extração de características. Assim como em filtros de Gabor, usou-se as metodologias *leave-one-out* e *leave-one-group-out*. A primeira verifica somente indivíduos conhecidos

enquanto que a segunda, desconhecidos. As colunas 4 e 5 (IND. CONHEC.) apresentam os resultados dos experimentos *leave-one-out* enquanto que as colunas 6 e 7 (IND. DESCON.), dos experimentos *leave-one-group-out*.

Tabela 8 – Base JAFFE: acurácias obtidas (%) usando LBP aplicado a base JAFFE.

TAM. BL.	NUM. BL.	CLASSIF	IND. CONHEC.		IND. DESCON.	
			AC	AC/FCNN	AC	AC/FCNN
8 × 8	192	KNN	87,78	46,95	37,56	41,78
8 × 8	192	SVM	86,85	59,62	62,44	42,25
8 × 8	192	RBF	35,68	19,25	25,35	15,96
8 × 8	192	NB	77,93	15,02	51,17	14,55
16 × 16	48	KNN	87,32	60,56	49,30	56,34
16 × 16	48	SVM	86,38	77,46	63,38	60,09
16 × 16	48	RBF	62,91	27,23	50,70	21,13
16 × 16	48	NB	67,14	29,11	52,58	20,66
32 × 32	12	KNN	86,85	60,56	45,07	59,62
32 × 32	12	SVM	79,34	76,06	54,93	59,62
32 × 32	12	RBF	65,26	16,43	55,87	26,76
32 × 32	12	NB	65,73	31,92	55,87	25,82

Fonte: Elaborado pelo autor. TAM. BL. = tamanho bloco; NUM. BL. = número de blocos; CLASSIF = classificador; AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

As acurácias obtidas foram relativamente baixas, principalmente para *leave-one-subject-out*, que correspondem às colunas 6 e 7 da Tabela 8. O melhor resultado para esse grupo foi 63,38%, com bloco 16 × 16 e classificador SVM. Já para indivíduos conhecidos, o melhor resultado foi de 87,78%, com bloco 8 × 8 e classificador KNN. As Tabelas 9 e 10 apresentam os detalhes desses 2 melhores experimentos.

Tabela 9 – Base JAFFE: melhor resultado obtido com LBP para indivíduos conhecidos
(*leave-one-out*): KNN (sem FCNN) com acurácia média de 87,78%.

MATRIZ DE CONFUSÃO								MÉTRICAS			
CLASSES	FELIZ	TRISTE	SURPRESO	DESGOSTO	RAIVA	MEDO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	83,87	00,00	00,00	00,00	00,00	00,00	16,13	0,93	0,84	0,99	0,88
TRISTE	03,23	80,65	00,00	00,00	00,00	03,23	12,90	0,96	0,81	0,99	0,88
SURPRESO	03,33	00,00	86,67	00,00	00,00	03,33	06,67	0,96	0,87	0,99	0,91
DESGOSTO	00,00	00,00	00,00	89,66	06,90	03,45	00,00	0,93	0,90	0,99	0,91
RAIVA	00,00	03,33	00,00	00,00	90,00	00,00	06,67	0,93	0,90	0,99	0,92
MEDO	00,00	00,00	03,12	06,25	00,00	84,38	06,25	0,90	0,84	0,98	0,87
NEUTRO	00,00	00,00	00,00	00,00	00,00	00,00	100,00	0,67	1,00	0,92	0,80

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Tabela 10 – Base JAFFE: melhor resultado obtido com LBP para indivíduos desconhecidos
(*leave-one-subject-out*): SVM (sem FCNN) com acurácia média de 63,38%.

MATRIZ DE CONFUSÃO								MÉTRICAS			
CLASSES	FELIZ	TRISTE	SURPRESO	DESGOSTO	RAIVA	MEDO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	74,19	00,00	12,90	00,00	00,00	00,00	12,90	0,92	0,74	0,99	0,82
TRISTE	00,00	35,48	03,23	22,58	06,45	22,58	09,68	0,33	0,35	0,88	0,34
SURPRESO	03,33	03,33	76,67	00,00	00,00	16,67	00,00	0,68	0,77	0,94	0,72
DESGOSTO	00,00	31,03	00,00	55,17	10,34	03,45	00,00	0,52	0,55	0,92	0,53
RAIVA	00,00	10,00	00,00	03,33	86,67	00,00	00,00	0,79	0,87	0,96	0,83
MEDO	03,12	21,88	09,38	15,62	00,00	46,88	03,12	0,54	0,47	0,93	0,50
NEUTRO	00,00	06,67	10,00	06,67	06,67	00,00	70,00	0,72	0,70	0,96	0,71

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Transformada Discreta do Cosseno (DCT)

Os experimentos com o uso das características extraídas com DCT foram realizados de duas maneiras. Primeiramente, a transformada DCT foi aplicada à imagem inteira e um subconjunto dos coeficientes foram selecionados para a formação do vetor de características. No segundo experimento, a imagem foi dividida em subimagens e a transformada DCT foi aplicada em cada uma delas. Em cada subimagem transformada, selecionou-se um subconjunto dos coeficientes que concatenados geraram o vetor de características. Em ambos os casos, os n coeficientes foram selecionados do canto superior esquerdo da imagem transformada, que correspondem a submatriz $\sqrt{n} \times \sqrt{n}$.

A Tabela 11 apresenta os resultados dos experimentos com DCT aplicada às imagens inteiras. Foram testadas as seguintes quantidades de coeficientes selecionados: 64, 256, 1024,

e 4096 que correspondem, respectivamente, a 0,52%, 2,08%, 8,33% e 33,33% do total de coeficientes, que é 12.228 (96×128). As colunas agrupadas como “IND. CONHEC.” apresentam os resultados dos experimentos *leave-one-out* enquanto que as colunas “IND. DESCON.”, dos experimentos *leave-one-group-out*.

Observa-se que a melhor acurácia para indivíduos conhecidos, 89,67%, foi obtida com o classificador SVM utilizado 256 características. Já para indivíduos desconhecidos, a melhor acurácia, 57,75 %, também foi obtida com o classificador SVM utilizando 256 características.

Tabela 11 – Base JAFFE: resultados usando DCT aplicado às imagens inteiras.

CARAC	COEFS(%)	CLASSIF	IND. CONHEC.		IND. DESCON.	
			AC	AC/FCNN	AC	AC/FCNN
64	00,52	KNN	87,79	84,51	38,50	51,64
64	00,52	SVM	75,59	70,89	51,17	46,48
64	00,52	RBF	62,91	59,15	37,56	25,35
64	00,52	NB	55,40	58,69	46,01	43,66
256	02,08	KNN	87,32	82,16	40,85	53,52
256	02,08	SVM	89,67	84,98	57,75	57,75
256	02,08	RBF	69,95	67,14	52,11	31,46
256	02,08	NB	62,91	66,67	53,52	40,85
1024	08,33	KNN	83,57	77,93	34,27	47,42
1024	08,33	SVM	88,26	79,81	48,36	50,23
1024	08,33	RBF	71,36	54,93	45,54	23,00
1024	08,33	NB	71,36	57,75	51,17	30,99
4096	33,33	KNN	64,79	60,09	19,72	37,09
4096	33,33	SVM	67,14	65,73	37,56	39,44
4096	33,33	RBF	62,44	26,76	41,31	16,43
4096	33,33	NB	66,20	32,39	44,13	17,84

Fonte: Elaborado pelo autor. CARAC = número de características; COEFS(%) = porcentagem de coeficientes; CLASSIF = classificador; AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

A Tabela 12 apresenta os resultados da aplicação da DCT em blocos das imagens. Foram feitos experimentos utilizando os seguintes tamanhos de blocos: 8×8 (192 blocos), 16×16 (48 blocos) e 32×32 (12 blocos). A melhor acurácia para indivíduos conhecidos (*leave-one-out*), 91,55%, foi obtida com o classificador SVM usando blocos 16×16 . Já para indivíduos desconhecidos (*leave-one-subject-out*), a melhor acurácia, 57,75%, foi obtida com

o Naive-Bayes usando blocos de tamanho 8×8 . O classificador SVM também forneceu um acurácia equivalente de 57,28%. As Tabelas 13 e 14 apresentam a matriz de confusão juntamente com outras métricas referente aos estes dois experimentos. Observa-se que, para indivíduos desconhecidos, o uso do seletor de protótipos FCNN melhorou significativamente (até 22%) a acurácia do classificador KNN. Observa-se também um pequena melhoria com o classificador SVM.

Tabela 12 – Base JAFFE: resultados usando DCT aplicado em blocos.

TAM. BL.	CARAC	COEFS(%)	CLASSIF	IND. CONHEC.		IND. DESCON.	
				AC	AC/FCNN	AC	AC/FCNN
8×8	4	25,00	KNN	80,75	69,95	31,92	53,05
8×8	4	25,00	SVM	87,32	80,28	54,93	53,05
8×8	4	25,00	RBF	65,73	37,56	56,34	21,60
8×8	4	25,00	NB	69,48	45,07	57,75	24,41
16×16	4	06,25	KNN	88,73	84,51	39,44	53,52
16×16	4	06,25	SVM	91,55	86,38	55,87	57,28
16×16	4	06,25	RBF	73,24	55,87	50,23	27,23
16×16	4	06,25	NB	69,48	62,44	52,58	30,99
16×16	8	25,00	KNN	77,00	67,61	25,82	51,64
16×16	8	25,00	SVM	81,69	77,06	48,83	50,70
16×16	8	25,00	RBF	73,71	37,56	53,05	19,25
16×16	8	25,00	NB	68,08	43,19	53,05	21,60
32×32	4	01,56	KNN	89,20	88,73	45,54	51,64
32×32	4	01,56	SVM	88,26	87,32	50,70	49,77
32×32	4	01,56	RBF	70,42	74,65	44,60	29,58
32×32	4	01,56	NB	61,97	68,08	49,77	38,50
32×32	8	06,25	KNN	87,32	83,57	36,15	50,23
32×32	8	06,25	SVM	89,20	81,69	51,17	51,64
32×32	8	06,25	RBF	68,54	55,87	46,95	25,35
32×32	8	06,25	NB	65,73	66,67	49,77	30,05
32×32	16	25,00	KNN	75,12	66,20	23,00	48,83
32×32	16	25,00	SVM	78,40	73,71	45,07	46,01
32×32	16	25,00	RBF	70,89	-	46,95	17,37
32×32	16	25,00	NB	69,48	41,78	46,95	23,00

Fonte: Elaborado pelo autor. TAM. BL. = tamanho bloco; CARAC = número de características; COEFS(%) = porcentagem de coeficientes; CLASSIF = classificador; AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

Tabela 13 – Base JAFFE: melhor resultado obtido com DCT (16×16) para indivíduos conhecidos (*leave-one-out*): SVM (sem FCNN) com acurácia média de 91,55%.

MATRIZ DE CONFUSÃO								MÉTRICAS			
CLASSES	FELIZ	TRISTE	SURPRESO	DESGOSTO	RAIVA	MEDO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	96,77	00,00	00,00	00,00	00,00	00,00	03,23	0,94	0,97	0,99	0,95
TRISTE	03,23	90,32	00,00	03,23	00,00	03,23	00,00	0,90	0,90	0,98	0,90
SURPRESO	03,33	00,00	93,33	00,00	00,00	03,33	00,00	0,90	0,93	0,98	0,92
DESGOSTO	00,00	06,90	00,00	82,76	10,34	00,00	00,00	0,86	0,83	0,98	0,84
RAIVA	00,00	00,00	00,00	06,67	93,33	00,00	00,00	0,90	0,93	0,98	0,92
MEDO	00,00	03,12	09,38	03,12	00,00	84,38	00,00	0,93	0,84	0,99	0,89
NEUTRO	00,00	00,00	00,00	00,00	00,00	00,00	100,00	0,97	1,00	0,99	0,98

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Tabela 14 – Base JAFFE: melhor resultado obtido com DCT (8×8) para indivíduos desconhecidos (*leave-one-subject-out*): NB (sem FCNN) com acurácia média de 57,75%.

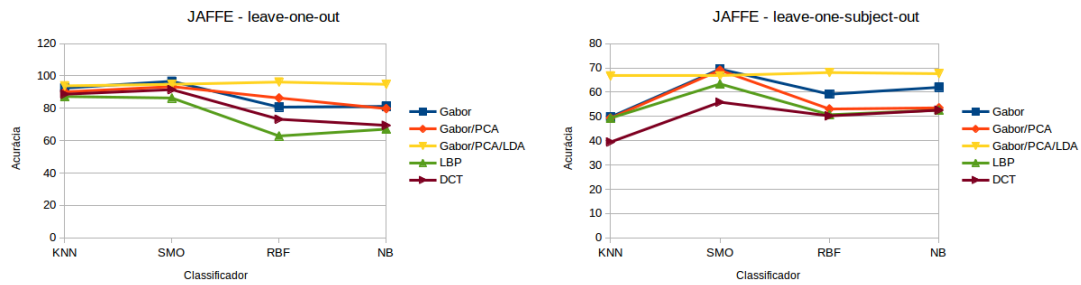
MATRIZ DE CONFUSÃO								MÉTRICAS			
CLASSES	FELIZ	TRISTE	SURPRESO	DESGOSTO	RAIVA	MEDO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	74,19	03,23	06,45	00,00	00,00	16,13	00,00	0,70	0,74	0,95	0,72
TRISTE	00,00	35,48	09,68	25,81	06,45	22,58	00,00	0,46	0,35	0,93	0,40
SURPRESO	03,33	03,33	66,67	00,00	00,00	20,00	06,67	0,65	0,67	0,94	0,66
DESGOSTO	10,34	10,34	00,00	44,83	20,69	13,79	00,00	0,50	0,45	0,93	0,47
RAIVA	06,67	06,67	00,00	16,67	66,67	03,33	00,00	0,59	0,67	0,92	0,62
MEDO	09,38	09,38	06,25	00,00	03,12	71,88	00,00	0,46	0,72	0,85	0,56
NEUTRO	03,33	10,00	13,33	00,00	16,67	13,33	43,33	0,87	0,43	0,99	0,58

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Resumo dos experimentos realizados com a base JAFFE

Os gráficos da Figura 21 apresentam o resumo dos experimentos com os 3 grupos de características aplicados à base JAFFE. O gráfico à esquerda resume os experimentos *leave-one-out* e à direita, os experimentos *leave-one-subject-out*. Destaca-se o uso das características de Gabor selecionadas com PCA/LDA que tiveram acurácia uniforme com os 4 classificadores testados.

Figura 19 – Comparativo entre os experimentos realizados com a base JAFFE.



Fonte: Elaborado pelo autor.

3.3.2 Experimentos com a base Yale

Os experimentos com a base Yale foram realizados tanto com a metodologia *leave-one-out* quanto a *leave-one-group-out*. Para o primeiro método, uma imagem foi selecionada para teste e as demais foram mantidas para treinamento. Esse procedimento foi repetido 60 vezes de forma que em cada ciclo, uma imagem distinta foi usada para teste. Quanto ao segundo método, a base foi dividida em 5 partições, que foram geradas da seguinte forma: três indivíduos distintos foram selecionados e todas as suas imagens foram colocadas no conjunto de teste. As imagens dos demais indivíduos foram mantidas no conjunto de treinamento. Os indivíduos escolhidos para testes em uma das partições não são mais testados em outra.

Filtros de Gabor

Os experimentos com filtros de Gabor foram realizados da mesma forma que com a base JAFFE. Em função do número menor de imagens presentes na base Yale (60 imagens), foram utilizadas as 30 primeiras componentes principais nos testes com PCA e LDA. As Tabelas 15 e 16 apresentam, respectivamente, os resultados obtidos com os experimentos *leave-one-out* e *leave-one-group-out*.

Tabela 15 – Resultados obtidos com filtros de Gabor usando *leave-one-out*.

CLASSIFICADOR	GABOR		GABOR PCA		GABOR PCA/LDA	
	AC	AC/FCNN	AC	AC/FCNN	AC	AC/FCNN
KNN	1,67	31,67	1,67	33,33	65,00	55,00
SVM	46,67	41,67	40,00	58,33	58,33	21,67
RBF	30,00	35,00	30,00	21,67	70,00	36,67
NB	31,67	38,33	21,67	26,67	68,33	26,67

Fonte: Elaborado pelo autor.

Tabela 16 – Resultados obtidos com filtros de Gabor usando *leave-one-group-out*.

CLASSIFICADOR	GABOR		GABOR PCA		GABOR PCA/LDA	
	AC	AC/FCNN	AC	AC/FCNN	AC	AC/FCNN
KNN	61,67	63,33	48,33	65,00	60,00	60,00
SVM	63,33	70,00	66,67	63,33	66,67	28,33
RBF	63,33	40,00	60,00	41,67	66,67	45,00
NB	61,67	38,33	61,67	46,67	61,67	38,33

Fonte: Elaborado pelo autor.

Observa-se com esses experimentos, que houve uma inversão de acertos em relação aos experimentos com a base JAFFE. As acurácias obtidas com indivíduos desconhecidos foram melhores do que indivíduos conhecidos. Em (GU *et al.*, 2012), os autores demonstraram que as características de Gabor puras, como extraídas aqui, estão mais relacionadas aos indivíduos do que sua expressão. Nos experimentos com a base JAFFE, como há mais de uma expressão por indivíduo, ao retirarmos uma imagem para teste, provavelmente ela será classificada com a mesma expressão de uma das duas similares. No caso da Yale, que há somente uma imagem de expressão por indivíduo, ao retirarmos essa imagem para teste, é provável que ela seja associada a uma das outras expressões correspondente ao mesmo indivíduo da imagem de teste. Por isso, a acurácia com os experimentos com *leave-one-out* foi muito baixa.

Com relação aos experimentos *leave-one-group-out*, destaca-se o resultado produzido com o classificador SVM utilizando as características de Gabor dos padrões selecionados pelo FCNN: 70%. A Tabela 17 exhibe o detalhamento desse experimento.

Tabela 17 – Base Yale: melhor resultado obtido com Gabor para indivíduos desconhecidos (*leave-one-subject-out*): SVM (com FCNN) com acurácia média de 70,00%.

CLASSES	MATRIZ DE CONFUSÃO				MÉTRICAS			
	FELIZ	TRISTE	SURPRESO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	100,00	00,00	00,00	00,00	0,94	1,00	0,98	0,97
TRISTE	06,67	53,33	00,00	40,00	0,47	0,53	0,80	0,50
SURPRESO	00,00	20,00	73,33	06,67	0,92	0,73	0,98	0,81
NEUTRO	00,00	40,00	06,67	53,33	0,53	0,53	0,84	0,53

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Padrões Binários Locais (LBP)

Da mesma forma que com a base JAFFE, os experimentos com LBP foram realizados com o padrão LBP (8, 1) utilizando blocos de tamanho 8×8 (192 blocos), 16×16 (48 blocos) e 32×32 (12 blocos) com o histograma 58-bin. A Tabela 18 apresenta um resumo dos experimentos realizados usando LBP.

Tabela 18 – Resultados usando LBP aplicado às imagens divididas em blocos.

TAM. BL.	NUM. BL.	CLASSIF	IND. CONHEC.		IND. DESCON.	
			AC	AC/FCNN	AC	AC/FCNN
8×8	192	KNN	8,33	45,00	48,33	50,00
8×8	192	SVM	45,00	23,33	76,67	63,33
8×8	192	RBF	26,67	-	40,00	-
8×8	192	NB	21,67	23,33	58,33	26,67
16×16	48	KNN	1,67	23,33	58,33	63,33
16×16	48	SVM	51,67	33,33	78,33	65,00
16×16	48	RBF	38,33	28,33	45,00	-
16×16	48	NB	25,00	25,00	63,33	33,33
32×32	12	KNN	3,33	28,33	66,67	70,00
32×32	12	SVM	50,00	40,00	70,00	61,67
32×32	12	RBF	33,33	26,67	55,00	35,00
32×32	12	NB	25,00	33,33	53,33	36,67

Fonte: Elaborado pelo autor. TAM. BL. = tamanho bloco; NUM. BL. = número de blocos; CLASSIF = classificador; AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

Assim como os resultados com a base JAFFE, as acurácias obtidas foram relati-

vamente baixas, principalmente para *leave-one-subject-out*, que correspondem às colunas 6 e 7 da Tabela 18. O melhor resultado para esse grupo foi 51,67%, com bloco 16×16 e classificador SVM. Já para indivíduos conhecidos, o melhor resultado foi de 78,33%, com bloco 16×16 e classificador SVM. As Tabelas 19 e 20 apresentam os detalhes desses 2 melhores experimentos.

Tabela 19 – Base Yale: melhor resultado obtido com LBP para indivíduos conhecidos (*leave-one-out*): SVM (sem FCNN) com acurácia média de 51,67%.

CLASSES	MATRIZ DE CONFUSÃO				MÉTRICAS			
	FELIZ	TRISTE	SURPRESO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	93,33	06,67	00,00	00,00	0,93	0,93	0,98	0,93
TRISTE	06,67	13,33	00,00	80,00	0,13	0,13	0,71	0,13
SURPRESO	00,00	00,00	80,00	20,00	1,00	0,80	1,00	0,89
NEUTRO	00,00	80,00	00,00	20,00	0,17	0,20	0,67	0,18

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Tabela 20 – Base Yale: melhor resultado obtido com LBP para indivíduos desconhecidos (*leave-one-group-out*): SVM (sem FCNN) com acurácia média de 78,33%.

CLASSES	MATRIZ DE CONFUSÃO				MÉTRICAS			
	FELIZ	TRISTE	SURPRESO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	93,33	00,00	00,00	06,67	0,93	0,93	0,98	0,93
TRISTE	06,67	60,00	00,00	33,33	0,60	0,60	0,87	0,60
SURPRESO	00,00	00,00	100,00	00,00	1,00	1,00	1,00	1,00
NEUTRO	00,00	40,00	00,00	60,00	0,60	0,60	0,87	0,60

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Transformada Discreta do Cosseno (DCT)

Da mesma forma que com a base JAFFE, os experimentos com a base Yale foram realizados de duas maneiras: aplicação de DCT à imagem inteira e à imagem dividida em blocos. Os vetores de características também foram gerados como no experimento anterior (Subseção 3.3.1).

A Tabela 21 apresenta os resultados dos experimentos com DCT aplicada às imagens inteiras. Também foram testadas as seguintes quantidades de coeficientes selecionados: 64, 256,

1024, e 4096 que correspondem, respectivamente, a 0,52%, 2,08%, 8,33% e 33,33% do total de coeficientes, que é 12.228 (96×128). Observa-se que a melhor acurácia para indivíduos conhecidos, 46,67%, foi obtida com o classificador SVM utilizando 64 características obtidas com os padrões de treinamento selecionados pelo FCNN. Já para indivíduos desconhecidos, a melhor acurácia, 55,00 %, também foi obtida com o classificador SVM utilizando 256 características.

Tabela 21 – Resultados usando DCT aplicado às imagens inteiras.

CARAC	COEFS(%)	CLASSIF	IND. CONHEC.		IND. DESCON.	
			AC	AC/FCNN	AC	AC/FCNN
64	00,52	KNN	08,33	15,00	45,00	43,33
64	00,52	SVM	41,67	46,67	51,67	53,33
64	00,52	RBF	18,33	23,33	35,00	40,00
64	00,52	NB	35,00	31,67	46,67	48,33
256	02,08	KNN	06,67	15,00	41,67	56,67
256	02,08	SVM	40,00	41,67	55,00	53,33
256	02,08	RBF	15,00	15,00	48,33	38,33
256	02,08	NB	28,33	21,67	53,33	53,33
1024	08,33	KNN	06,67	20,00	28,33	45,00
1024	08,33	SVM	30,00	25,00	50,00	41,67
1024	08,33	RBF	20,00	23,33	48,33	28,33
1024	08,33	NB	23,33	20,00	51,67	25,00
4096	33,33	KNN	13,33	16,67	25,00	28,33
4096	33,33	SVM	20,00	16,67	38,33	33,33
4096	33,33	RBF	28,33	21,67	41,67	31,67
4096	33,33	NB	18,33	25,00	36,67	31,67

Fonte: Elaborado pelo autor. CARAC = número de características; COEFS(%) = porcentagem de coeficientes; CLASSIF = classificador; AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

A Tabela 22 apresenta os resultados da aplicação da DCT em blocos das imagens. Foram feitos experimentos utilizando os seguintes tamanhos de blocos: 8×8 (192 blocos), 16×16 (48 blocos) e 32×32 (12 blocos). A melhor acurácia para indivíduos conhecidos (*leave-one-out*), 58,33%, foi obtida com o classificador SVM usando blocos 8×8 e 16×16 . Já para indivíduos desconhecidos (*leave-one-subject-out*), a melhor acurácia, 45,00%, foi obtida com o SVM (com FCNN) usando blocos de tamanho 32×32 . As Tabelas 23 e 24 apresentam a matriz de confusão juntamente com outras métricas referente aos esses dois experimentos.

Tabela 22 – Resultados usando DCT aplicado em blocos.

TAM. BL.	CARAC	COEFS(%)	CLASSIF	IND. CONHEC.		IND. DESCON.	
				AC	AC/FCNN	AC	AC/FCNN
8 × 8	4	25,00	KNN	10,00	18,33	31,67	50,00
8 × 8	4	25,00	SVM	38,33	25,00	58,33	56,67
8 × 8	4	25,00	RBF	31,67	31,67	50,00	38,33
8 × 8	4	25,00	NB	30,00	36,67	53,33	43,33
16 × 16	4	06,25	KNN	10,00	20,00	46,67	61,67
16 × 16	4	06,25	SVM	40,00	33,33	58,33	60,00
16 × 16	4	06,25	RBF	21,67	31,67	48,33	41,67
16 × 16	4	06,25	NB	33,33	18,33	51,67	30,00
16 × 16	8	25,00	KNN	11,67	26,67	26,67	50,00
16 × 16	8	25,00	SVM	36,67	23,33	50,00	48,33
16 × 16	8	25,00	RBF	23,33	18,33	60,00	35,00
16 × 16	8	25,00	NB	35,00	33,33	60,00	36,67
32 × 32	4	01,56	KNN	03,33	18,33	45,00	48,33
32 × 32	4	01,56	SVM	43,33	45,00	60,00	55,00
32 × 32	4	01,56	RBF	25,00	20,00	45,00	43,33
32 × 32	4	01,56	NB	35,00	31,67	51,67	51,67
32 × 32	8	06,25	KNN	10,00	23,33	35,00	48,33
32 × 32	8	06,25	SVM	35,00	26,67	51,67	46,67
32 × 32	8	06,25	RBF	23,33	28,33	53,33	40,00
32 × 32	8	06,25	NB	28,33	31,67	56,67	43,33
32 × 32	16	25,00	KNN	15,00	15,00	23,33	40,00
32 × 32	16	25,00	SVM	30,00	20,00	48,33	41,67
32 × 32	16	25,00	RBF	35,00	20,00	55,00	28,33
32 × 32	16	25,00	NB	26,67	40,00	53,33	40,00

Fonte: Elaborado pelo autor. TAM. BL. = tamanho bloco; CARAC = número de características; COEFS(%) = porcentagem de coeficientes; CLASSIF = classificador; AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

Tabela 23 – Base Yale: melhor resultado obtido com DCT para indivíduos conhecidos
(*leave-one-out*): SVM (sem FCNN) com acurácia média de 45,00%.

CLASSES	MATRIZ DE CONFUSÃO				MÉTRICAS			
	FELIZ	TRISTE	SURPRESO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	86,67	13,33	00,00	00,00	0,81	0,87	0,93	0,84
TRISTE	20,00	20,00	06,67	53,33	0,17	0,20	0,67	0,18
SURPRESO	00,00	13,33	46,67	40,00	0,88	0,47	0,98	0,61
NEUTRO	00,00	73,33	00,00	26,67	0,22	0,27	0,69	0,24

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Tabela 24 – Base Yale: melhor resultado obtido com DCT (8×8) para indivíduos desconhecidos
(*leave-one-group-out*): SVM (sem FCNN) com acurácia média de 58,33%.

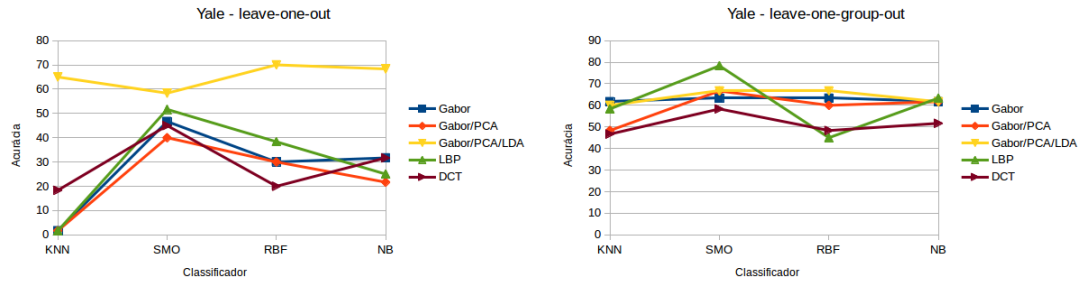
CLASSES	MATRIZ DE CONFUSÃO				MÉTRICAS			
	FELIZ	TRISTE	SURPRESO	NEUTRO	PREC.	SENS.	SPEC.	F-MEASURE
FELIZ	86,67	00,00	06,67	06,67	0,81	0,87	0,93	0,84
TRISTE	13,33	46,67	13,33	26,67	0,41	0,47	0,78	0,44
SURPRESO	00,00	26,67	66,67	06,67	0,62	0,67	0,87	0,65
NEUTRO	06,67	40,00	20,00	33,33	0,45	0,33	0,87	0,38

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Resumo dos experimentos realizados com a base Yale

Os gráficos da Figura 20 apresentam o resumo dos experimentos com os 3 grupos de características aplicados à base Yale. O gráfico à esquerda resume os experimentos *leave-one-out* e à direita, os experimentos *leave-one-subject-out*. Destaca-se o uso das características de Gabor selecionadas com PCA/LDA que tiveram acurácia uniforme com os 4 classificadores testados.

Figura 20 – Comparativo entre os experimentos realizados com a base Yale.



Fonte: Elaborado pelo autor.

3.3.3 Experimentos com a base CK+

Da mesma forma que com a base Yale, os experimentos com a base CK+ foram realizados tanto com a metodologia *leave-one-out* quanto com a *leave-one-group-out*. Para o primeiro método, uma imagem foi selecionada para teste e as demais foram mantidas para treinamento, repetindo-se 327 vezes. Quanto à segunda, a base foi dividida em 10 partições, que foram geradas da seguinte forma: selecionou-se um grupo dos indivíduos da base e todas as suas imagens foram colocadas no conjunto de teste, enquanto que as imagens dos demais indivíduos foram mantidas no conjunto de treinamento. Os indivíduos escolhidos para testes em uma partição não são mais testados em outra. Nos experimentos apresentados nesta seção, foram usadas apenas as expressões rotuladas disponíveis na base CK+, sendo excluída a expressão neutra.

Filtros de Gabor

As características de Gabor foram extraídas da mesma forma que na base JAFFE: somente filtros de Gabor e com redução de dimensionalidade com PCA e LDA. No caso do LDA, o número de componentes selecionados foi 6, visto que nesses experimentos foram utilizadas 7 classes distintas.

As Tabelas 25 e 26 apresentam, respectivamente, um resumo dos experimentos realizados com *leave-one-out* (considerando indivíduos conhecidos) e com *leave-one-group-out* (considerando indivíduos desconhecidos).

Tabela 25 – Base CK+: resultados obtidos com filtros de Gabor usando *leave-one-out*.

CLASSIFICADOR	GABOR		GABOR PCA		GABOR PCA/LDA	
	AC.	AC/FCNN	AC.	AC/FCNN	AC.	AC/FCNN
KNN	48,62	83,18	37,92	85,02	89,60	86,24
SVM	91,44	88,69	89,60	74,31	92,05	30,89
RBF	80,43	18,35	80,12	32,72	90,83	51,99
NB	74,92	34,56	83,18	46,79	90,83	48,01

Fonte: Elaborado pelo autor.

Tabela 26 – Base CK+: resultados obtidos com filtros de Gabor usando *leave-one-group-out*.

CLASSIFICADOR	GABOR		GABOR PCA		GABOR PCA/LDA	
	AC.	AC/FCNN	AC.	AC/FCNN	AC.	AC/FCNN
KNN	83,49	88,38	76,15	91,44	92,05	88,69
SVM	93,27	91,13	91,74	72,78	91,74	32,11
RBF	82,57	30,58	84,10	37,00	91,74	-
NB	80,12	34,56	88,07	48,93	92,97	47,71

Fonte: Elaborado pelo autor.

Observa-se que as acurácias para indivíduos desconhecidos são bem melhores que para conhecidos. Acredita-se que isso acontece pelo fato de a base CK+ ser bem heterogênea.

O classificador SVM juntamente com as características de Gabor apresentaram os melhores resultados: 66,97% para *leave-one-out* e 93,27% para *leave-one-group-out*. As Tabelas 27 e 28 apresentam as matrizes de confusão juntamente com outras métricas referentes a esses dois experimentos.

Para efeitos comparativos, foram realizados testes com a base CK+ incluindo a expressão neutra, usando o classificador SVM. Os resultados indicaram uma redução na acurácia média que foi de 91,11%, 87,50% e 90,28%, respectivamente, com Gabor, Gabor PCA e Gabor PCA/LDA. A inclusão da expressão neutra reduziu o número de acerto de algumas classes, principalmente triste e desprezo.

Tabela 27 – Base CK+: melhor resultado obtido com filtros de Gabor para indivíduos conhecidos (*leave-one-out*): SVM sem seleção de protótipos com acurácia média de 92,05%.

CLASSES	MATRIZ DE CONFUSÃO							MÉTRICAS			
	RAIVA	DESPREZO	DESGOSTO	MEDO	FELIZ	TRISTE	SURPRESO	PREC	SENS	SPEC	F-MEASURE
RAIVA	93,33	00,00	02,22	00,00	00,00	04,44	00,00	0,72	0,93	0,94	0,82
DESPREZO	22,22	72,22	00,00	00,00	00,00	05,56	00,00	0,87	0,72	0,99	0,79
DESGOSTO	00,00	00,00	98,31	00,00	01,69	00,00	00,00	0,98	0,98	1,00	0,98
MEDO	08,00	00,00	00,00	84,00	04,00	00,00	04,00	1,00	0,84	1,00	0,91
FELIZ	00,00	01,45	00,00	00,00	98,55	00,00	00,00	0,97	0,99	0,99	0,98
TRISTE	35,71	00,00	00,00	00,00	00,00	60,71	03,57	0,85	0,61	0,99	0,71
SURPRESO	00,00	01,20	00,00	00,00	00,00	00,00	98,80	0,98	0,99	0,99	0,98

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Tabela 28 – Base CK+: melhor resultado obtido com filtros de Gabor para indivíduos desconhecidos (*leave-one-subject-out*): SVM sem seleção de protótipos com acurácia média de 93,27%.

CLASSES	MATRIZ DE CONFUSÃO							MÉTRICAS			
	RAIVA	DESPREZO	DESGOSTO	MEDO	FELIZ	TRISTE	SURPRESO	PREC	SENS	SPEC	F-MEASURE
RAIVA	95,56	02,22	02,22	00,00	00,00	00,00	00,00	0,81	0,96	0,96	0,88
DESPREZO	11,11	72,22	00,00	00,00	00,00	05,56	11,11	0,87	0,72	0,99	0,79
DESGOSTO	01,69	00,00	96,61	00,00	01,69	00,00	00,00	0,98	0,97	1,00	0,97
MEDO	08,00	00,00	00,00	84,00	00,00	00,00	08,00	1,00	0,84	1,00	0,91
FELIZ	00,00	00,00	00,00	00,00	100,00	00,00	00,00	0,99	1,00	1,00	0,99
TRISTE	17,86	00,00	00,00	00,00	00,00	71,43	10,71	0,95	0,71	1,00	0,82
SURPRESO	00,00	01,20	00,00	00,00	00,00	00,00	98,80	0,92	0,99	0,97	0,95

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Padrões Binários Locais (LBP)

Os experimentos com LBP foram realizados com o padrão LBP (8, 1) aplicado em blocos nas imagens. Também foram feitos experimentos utilizando os seguintes tamanhos de blocos: 8×8 (192 blocos), 16×16 (48 blocos) e 32×32 (12 blocos), ambos utilizando o histograma 58 -bin.

A Tabela 29 apresenta um resumo dos experimentos realizados usando LBP como método de extração de características. Assim como em filtros de Gabor, usou-se as metodologias *leave-one-out* e *leave-one-group-out*.

Tabela 29 – Base CK+: resultados usando LBP aplicado às imagens divididas em blocos.

TAM. BL.	NUM. BL.	CLASSIF	IND. CONHEC.		IND. DESCON.	
			AC	AC/FCNN	AC	AC/FCNN
8 × 8	192	KNN	38,84	82,26	74,01	85,02
8 × 8	192	SVM	88,38	84,71	89,60	85,32
8 × 8	192	RBF	57,19	45,87	57,19	45,26
8 × 8	192	NB	71,56	44,34	75,84	48,62
16 × 16	48	KNN	38,23	81,04	78,29	83,49
16 × 16	48	SVM	88,69	88,69	90,83	91,13
16 × 16	48	RBF	74,62	66,67	75,84	69,11
16 × 16	48	NB	75,23	69,72	81,65	73,70
32 × 32	12	KNN	44,04	67,28	74,01	78,29
32 × 32	12	SVM	87,46	89,91	90,21	89,91
32 × 32	12	RBF	69,11	66,06	73,39	68,50
32 × 32	12	NB	69,72	69,11	76,15	73,70

Fonte: Elaborado pelo autor. TAM. BL. = tamanho bloco; NUM. BL. = número de blocos; CLASSIF = classificador; AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

O melhor resultado para os experimentos considerando indivíduos conhecidos (*leave-one-out*), a melhor acurácia obtida com LBP foi 89,91%, com bloco 32 × 32. Já para indivíduos desconhecidos (*leave-one-group-out*), o melhor resultado foi de 91,13%, com bloco 16 × 16. Em ambos os casos, foi utilizado o classificador SVM e os testes foram realizados com os padrões de treinamento selecionados pelo FCNN. As Tabelas 30 e 31 apresentam os detalhes desses 2 melhores experimentos.

Tabela 30 – Base CK+: melhor resultado obtido com LBP (16 × 16) para indivíduos conhecidos (*leave-one-out*): SVM sem seleção de protótipos com acurácia média de 88,69%.

CLASSES	MATRIZ DE CONFUSÃO							MÉTRICAS			
	RAIVA	DESPREZO	DESGOSTO	MEDO	FELIZ	TRISTE	SURPRESO	PREC	SENS	SPEC	F-MEASURE
RAIVA	86,67	00,00	00,00	00,00	00,00	13,33	00,00	0,75	0,87	0,95	0,80
DESPREZO	16,67	77,78	00,00	00,00	00,00	05,56	00,00	0,74	0,78	0,98	0,76
DESGOSTO	00,00	00,00	100,00	00,00	00,00	00,00	00,00	0,94	1,00	0,99	0,97
MEDO	04,00	12,00	08,00	52,00	12,00	00,00	12,00	0,93	0,52	1,00	0,67
FELIZ	00,00	00,00	00,00	00,00	100,00	00,00	00,00	0,96	1,00	0,99	0,98
TRISTE	32,14	03,57	07,14	00,00	00,00	53,57	03,57	0,68	0,54	0,98	0,60
SURPRESO	00,00	01,20	00,00	01,20	00,00	00,00	97,59	0,95	0,98	0,98	0,96

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Tabela 31 – Base CK+: melhor resultado obtido com LBP (16×16) para indivíduos desconhecidos (*leave-one-subject-out*): SVM com seleção de protótipos com acurácia média de 91,13%.

CLASSES	MATRIZ DE CONFUSÃO							MÉTRICAS			
	RAIVA	DESPREZO	DESGOSTO	MEDO	FELIZ	TRISTE	SURPRESO	PREC	SENS	SPEC	F-MEASURE
RAIVA	88,89	00,00	00,00	00,00	00,00	11,11	00,00	0,80	0,89	0,96	0,84
DESPREZO	11,11	88,89	00,00	00,00	00,00	00,00	00,00	0,73	0,89	0,98	0,80
DESGOSTO	00,00	00,00	100,00	00,00	00,00	00,00	00,00	1,00	1,00	1,00	1,00
MEDO	00,00	12,00	00,00	64,00	24,00	00,00	00,00	0,89	0,64	0,99	0,74
FELIZ	01,45	00,00	00,00	02,90	95,65	00,00	00,00	0,92	0,96	0,98	0,94
TRISTE	25,00	07,14	00,00	00,00	00,00	67,86	00,00	0,79	0,68	0,98	0,73
SURPRESO	00,00	01,20	00,00	00,00	00,00	00,00	98,80	1,00	0,99	1,00	0,99

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Transformada Discreta do Cosseno (DCT)

Da mesma forma que com as bases anteriores, os experimentos com a base CK+ foram realizados de duas maneiras: aplicação de DCT às imagens inteiras e às imagens divididas em blocos. Os vetores de características também foram gerados como nos experimentos anteriores (Subseções 3.3.1 e 3.3.2).

A Tabela 32 apresenta os resultados da aplicação da DCT às imagens inteiras, cujas dimensões são 96×128 , variando-se o número de coeficientes selecionados para a geração do vetor de características. Foram testados os primeiros 64, 256, 1024 e 4096 coeficientes. Observa-se que os resultados para indivíduos desconhecidos foram ligeiramente melhores que para conhecidos.

Tabela 32 – Base CK+: resultados usando DCT aplicado às imagens inteiras.

CARAC	COEFS(%)	CLASSIF	IND. CONHEC.		IND. DESCON.	
			AC	AC/FCNN	AC	AC/FCNN
64	00,52	KNN	33,94	42,20	66,97	66,67
64	00,52	SVM	74,92	71,87	76,15	75,23
64	00,52	RBF	63,00	54,13	66,36	65,75
64	00,52	NB	65,75	67,28	69,42	69,42
256	02,08	KNN	48,32	67,89	76,76	81,96
256	02,08	SVM	82,26	76,76	84,10	79,51
256	02,08	RBF	74,31	63,61	78,90	65,50
256	02,08	NB	78,29	69,72	82,26	71,25
1024	08,33	KNN	54,13	67,89	65,44	72,48
1024	08,33	SVM	75,23	71,56	78,59	73,70
1024	08,33	RBF	68,81	46,48	70,03	44,65
1024	08,33	NB	70,95	52,60	74,92	59,63
4096	33,33	KNN	27,83	65,44	32,42	66,67
4096	33,33	SVM	70,03	66,06	71,87	68,50
4096	33,33	RBF	57,49	26,91	56,27	27,22
4096	33,33	NB	57,19	27,52	59,02	32,72

Fonte: Elaborado pelo autor. CARAC = número de características; COEFS(%) = porcentagem de coeficientes; CLASSIF = classificador; AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

A Tabela 33 apresenta os resultados da aplicação da DCT em blocos das imagens. Considerando apenas indivíduos conhecidos, a melhor acurácia, 82,87%, foi obtida com o SVM usando blocos de tamanho 8×8 e 16×16 . Já considerando os indivíduos desconhecidos (*leave-one-group-out*), o melhor resultado (86,54%) foi gerado também pelo classificador SVM, porém usando blocos 32×32 . Em relação ao uso do seletor de protótipos FCNN, observa-se que apenas o classificador melhora sua performance com sua utilização.

Tabela 33 – Base CK+: resultados usando DCT aplicado em blocos.

TAM. BL.	CARAC	COEFS(%)	CLASSIF	IND. CONHEC.		IND. DESCON.	
				AC	AC/FCNN	AC	AC/FCNN
8 × 8	4	25,00	KNN	55,05	76,76	68,50	79,51
8 × 8	4	25,00	SVM	82,87	80,73	84,71	84,40
8 × 8	4	25,00	RBF	-	-	77,37	46,79
8 × 8	4	25,00	NB	70,03	55,96	74,92	61,77
16 × 16	4	06,25	KNN	53,52	75,54	74,92	81,65
16 × 16	4	06,25	SVM	82,87	80,43	85,32	83,18
16 × 16	4	06,25	RBF	76,76	66,06	78,29	59,94
16 × 16	4	06,25	NB	74,92	68,50	78,90	73,39
16 × 16	8	25,00	KNN	56,57	77,68	62,08	81,65
16 × 16	8	25,00	SVM	80,43	77,37	82,26	79,51
16 × 16	8	25,00	RBF	74,92	52,91	77,06	48,32
16 × 16	8	25,00	NB	72,48	55,96	76,45	61,47
32 × 32	4	01,56	KNN	52,60	68,50	77,06	81,04
32 × 32	4	01,56	SVM	85,32	77,98	86,54	78,90
32 × 32	4	01,56	RBF	76,76	68,20	79,51	74,62
32 × 32	4	01,56	NB	77,98	74,62	82,57	78,29
32 × 32	8	06,25	KNN	57,49	73,39	73,09	79,51
32 × 32	8	06,25	SVM	82,57	80,43	85,32	82,87
32 × 32	8	06,25	RBF	77,06	68,20	79,51	63,91
32 × 32	8	06,25	NB	79,51	73,09	82,87	77,37
32 × 32	16	25,00	KNN	51,38	73,39	57,49	74,62
32 × 32	16	25,00	SVM	77,98	72,78	78,90	74,92
32 × 32	16	25,00	RBF	77,68	51,07	75,84	51,38
32 × 32	16	25,00	NB	74,01	56,88	76,45	63,91

Fonte: Elaborado pelo autor. TAM. BL. = tamanho bloco; CARAC = número de características; COEFS(%) = porcentagem de coeficientes; CLASSIF = classificador; AC = acurácia; AC/FCNN = acurácia com uso do FCNN.

As Tabelas 34 e 35 detalham os dois experimentos com melhores acurácias para *leave-one-out* e *leave-one-group-out*. As expressões FELIZ e SURPRESO tiveram as melhores acurácias e precisão, em ambos os experimentos.

Tabela 34 – Base CK+: melhor resultado obtido com DCT para indivíduos desconhecidos (*leave-one-subject-out*): bloco 8×8 , SVM, sem seleção de protótipos e acurácia média de 82,87%.

CLASSES	MATRIZ DE CONFUSÃO							MÉTRICAS			
	RAIVA	DESPREZO	DESGOSTO	MEDO	FELIZ	TRISTE	SURPRESO	PREC.	SENS.	SPEC.	F-MEASURE
RAIVA	77,78	00,00	04,44	00,00	00,00	17,78	00,00	0,62	0,78	0,93	0,69
DESPREZO	38,89	33,33	00,00	11,11	00,00	05,56	11,11	0,60	0,33	0,99	0,43
DESGOSTO	03,39	00,00	94,92	00,00	01,69	00,00	00,00	0,95	0,95	0,99	0,95
MEDO	08,00	08,00	00,00	52,00	16,00	08,00	08,00	0,87	0,52	0,99	0,65
FELIZ	00,00	00,00	01,45	00,00	98,55	00,00	00,00	0,93	0,99	0,98	0,96
TRISTE	32,14	03,57	00,00	00,00	00,00	42,86	21,43	0,52	0,43	0,96	0,47
SURPRESO	01,20	01,20	00,00	00,00	00,00	00,00	97,59	0,89	0,98	0,96	0,93

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Tabela 35 – Base CK+: melhor resultado obtido com DCT para indivíduos desconhecidos (*leave-one-subject-out*): bloco 32×32 , SVM, sem seleção de protótipos e acurácia média de 86,54%.

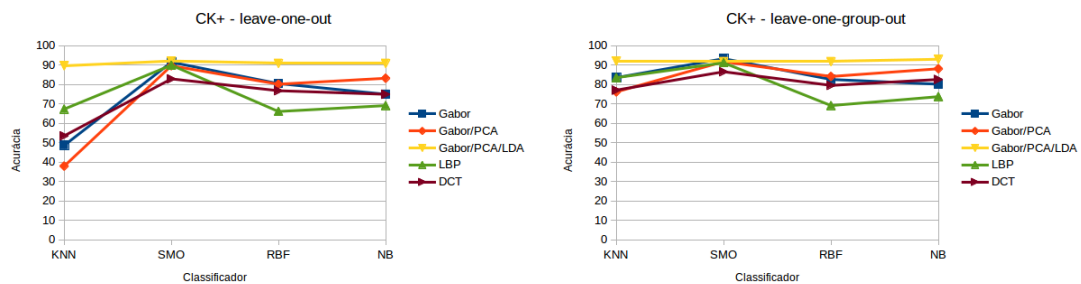
CLASSES	MATRIZ DE CONFUSÃO							MÉTRICAS			
	RAIVA	DESPREZO	DESGOSTO	MEDO	FELIZ	TRISTE	SURPRESO	PREC.	SENS.	SPEC.	F-MEASURE
RAIVA	82.22	02.22	02.22	00.00	00.00	13.33	00.00	0.70	0.82	0.94	0.76
DESPREZO	16.67	50.00	00.00	16.67	00.00	05.56	11.11	0.60	0.50	0.98	0.55
DESGOSTO	05.08	00.00	93.22	01.69	00.00	00.00	00.00	0.96	0.93	0.99	0.95
MEDO	04.00	16.00	00.00	64.00	16.00	00.00	00.00	0.76	0.64	0.98	0.70
FELIZ	00.00	00.00	01.45	01.45	97.10	00.00	00.00	0.94	0.97	0.98	0.96
TRISTE	28.57	00.00	00.00	00.00	00.00	64.29	07.14	0.72	0.64	0.98	0.68
SURPRESO	01.20	01.20	00.00	00.00	00.00	00.00	97.59	0.95	0.98	0.98	0.96

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Resumo dos experimentos realizados com a base CK+

Os gráficos da Figura 21 apresentam o resumo dos experimentos com os 3 grupos de características aplicados à base CK. O gráfico à esquerda resume os experimentos *leave-one-out* e à direita, os experimentos *leave-one-subject-out*. Destaca-se o uso das características de Gabor selecionadas com PCA/LDA que tiveram acurácia em torno de 90% com os 4 classificadores testados.

Figura 21 – Comparativo entre os experimentos realizados com a base CK.



Fonte: Elaborado pelo autor.

3.3.4 Experimentos com bases cruzadas (inter-bases)

Esta subseção apresenta os resultados dos experimentos com bases cruzadas, no qual uma das bases foi selecionada para treinamento. Os testes foram realizados individualmente com cada uma das duas bases restantes. Como a base Yale possui apenas 4 classes distintas, esses experimentos usaram as mesmas expressões da base JAFFE e CK+, que foram: feliz, triste, surpreso e neutro.

A Tabela 36 apresenta a acurácia média dos experimentos nos quais a base JAFFE foi usada como treinamento. Foram testadas cinco grupos de características (Gabor, Gabor/PCA, Gabor/LDA e LBP e DCT) e quatro classificadores que foram utilizados nos experimentos intra-bases. As colunas 2 a 5 apresentam os resultados obtidos com a base Yale enquanto que as colunas 6 a 9, os resultados com a base CK+. Analisando os resultados, observa-se que as acurácias obtidas com a base CK+ foram ligeiramente superiores às obtidas com a base Yale.

Tabela 36 – Experimentos com bases cruzadas usando JAFFE como treinamento.

	Yale				CK+			
	KNN	SVM	RBF	NB	KNN	SVM	RBF	NB
Gabor	46,67	41,67	43,33	38,33	64,32	74,18	72,77	52,58
Gabor/PCA	41,67	41,67	38,33	35,00	48,36	66,67	53,99	66,67
Gabor/PCA/LDA	40,00	28,33	41,67	43,33	46,48	44,13	39,91	47,89
LBP	40,00	45,00	46,67	45,00	59,62	76,06	54,46	66,20
DCT	33,33	38,33	25,00	28,33	43,19	42,72	48,83	47,42

Fonte: Elaborado pelo autor.

As Tabelas 37 e 38 apresentam, respectivamente, os resultados dos experimentos com as bases Yale e CK+ como treinamento.

Tabela 37 – Experimentos com bases cruzadas usando Yale como treinamento.

	JAFPE				CK+			
	KNN	SVM	RBF	NB	KNN	SVM	RBF	NB
Gabor	55,74	36,89	56,56	34,43	80,75	88,73	78,87	76,53
Gabor/PCA	53,28	34,43	24,59	25,41	75,59	82,16	66,20	74,65
Gabor/PCA/LDA	33,61	37,70	25,41	44,26	77,93	80,75	61,03	77,00
LBP	25,41	24,59	25,41	26,23	65,26	80,75	53,52	60,09
DCT	49,18	24,59	37,70	39,34	55,87	80,28	76,06	79,81

Fonte: Elaborado pelo autor.

Tabela 38 – Experimentos com bases cruzadas usando CK+ como treinamento.

	JAFPE				Yale			
	KNN	SVM	RBF	NB	KNN	SVM	RBF	NB
Gabor	40,16	45,08	43,44	39,34	55,00	73,33	56,67	56,67
Gabor/PCA	45,08	40,16	43,44	40,98	46,67	63,33	38,33	48,33
Gabor/PCA/LDA	41,80	42,62	43,44	44,26	65,00	61,67	58,33	68,33
LBP	39,34	45,90	40,16	28,69	56,67	66,67	68,33	63,33
DCT	28,69	31,15	31,15	35,25	41,67	65,00	56,67	56,67

Fonte: Elaborado pelo autor.

Com os resultados apresentados nas Tabelas 36, 37 e 38, observa-se uma pequena predominância das características extraídas somente com filtros de Gabor com o classificador SVM. Porém, como as acurácias são relativamente baixa, não é possível garantir a universalidade de nenhum dos três métodos estudados para extração de características. A Tabela 39 apresenta a matriz de confusão referente ao melhor resultado obtido nos experimentos inter-bases, que foi obtido com as características de Gabor e classificador SVM. A acurácia desse experimento foi de 88,73%.

Tabela 39 – Matriz de confusão referente ao experimento com melhor acurácia (Treinamento: Yale; teste: CK+).

CLASSES	MATRIZ DE CONFUSÃO				MÉTRICAS			
	FELIZ	TRISTE	SURPRESO	NEUTRO	PREC	SENS	SPEC	F-MEASURE
FELIZ	100,00	00,00	00,00	00,00	1,00	1,00	1,00	1,00
TRISTE	00,00	35,71	21,43	42,86	1,00	0,35	1,00	0,74
SURPRESO	00,00	00,00	98,80	01,20	0,88	0,98	0,85	0,95
NEUTRO	00,00	00,00	15,15	84,85	0,62	0,84	0,97	0,90

Fonte: Elaborado pelo autor. PREC = Precisão; SENS = Sensitividade; SPEC = Especificidade.

Foram realizados testes de hipótese *TESTE – T* com nível de significância 5%, usando as características de Gabor como referência. Considerando que os valores *P* obtidos, todos maiores que 0,05, conforme Tabela 40, conclui-se que esses experimentos são equivalentes. Ou seja, os dados são estaticamente válidos, confirmando-se a hipótese de que os dados não foram randomicamente gerados.

Tabela 40 – Teste de Hipóteses - Experimentos inter-bases.

CARACTERÍSTICAS	MÉDIA	DESVIO PADRÃO	P-VALUE
Gabor	56,34	16,18	-
Gabor/PCA	46,79	15,76	0,16
Gabor/PCA/LDA	46,79	15,37	1,00
LBP	50,14	16,74	0,94
DCT	45,68	16,63	0,36

Fonte: Elaborado pelo autor.

3.3.5 Experimentos multibases

Esta subseção apresenta os resultados dos experimentos multibases, no qual juntou-se todas as imagens das expressões comuns às três bases (JAFPE, Yale e CK+) para formar uma única base. Dessa forma, a base gerada possui 395 imagens e 4 expressões distintas, a saber: alegre - 115, triste - 71, surpreso - 128 e neutro - 78. Os experimentos foram realizados usando validação cruzada com 10 *folds* e os resultados são apresentados na Tabela 41.

Tabela 41 – Experimentos multibases.

CARACTERÍSTICAS	KNN	SVM	RBF	NB
Gabor	61,27	88,61	73,92	76,71
Gabor/PCA	54,68	85,82	74,18	73,16
Gabor/PCA/LDA	96,46	96,46	96,71	96,71
LBP	57,47	83,78	69,62	59,49
DCT	62,78	81,01	74,43	73,67

Fonte: Elaborado pelo autor.

Assim como nos experimentos inter-bases, foram realizados testes de hipótese *TESTE – T* com nível de significância 5%, usando as características de Gabor como referência. Considerando que os valores *P* obtidos, todos maiores que 0,05, conforme Tabela 42, conclui-se que esses experimentos são equivalentes.

Tabela 42 – Teste de Hipóteses - Experimentos multibases.

CARACTERÍSTICAS	MÉDIA	DESVIO PADRÃO	P-VALUE
Gabor	75,13	11,22	-
Gabor/PCA	71,96	12,87	0,72
Gabor/PCA/LDA	96,59	0,14	0,03
LBP	67,59	12,03	0,40
DCT	72,97	7,55	0,76

Fonte: Elaborado pelo autor.

3.3.6 Comparativo entre os tempos de processamento

A Tabela 43 compara o tempo de processamento, extração de características e classificação, de cada um dos métodos. DCT apresentou um tempo de extração de características aproximadamente 30 vezes menor que o tempo gasto com filtros de Gabor e 8 vezes menor que LBP. Dessa forma, DCT é mais indicada para aplicações com resposta em tempo real. Em relação à classificação, o tempo gasto foi praticamente o mesmo, porém observou-se que o tempo está diretamente relacionado à dimensão do vetor de características. O classificador KNN consumiu mais tempo com características de Gabor, cujas dimensões são bem mais elevadas. Deve-se destacar o SVM, que consumiu um tempo médio menor que os demais classificadores.

Tabela 43 – Tempo de Extração/Classificação (em *ms*).

Método de Extração	Extração	KNN	SVM	RBF	NB
Gabor	120,37	192,90	7,68	19,61	89,57
Gabor/PCA	145,06	0,58	0,08	0,11	0,50
Gabor/PCA/LDA	148,02	0,19	0,07	0,07	0,11
LBP	32,00	9,49	0,69	1,95	7,18
DCT	4,00	3,60	0,70	0,50	2,50
Tempo médio	89,89	41,35	1,84	4,45	19,97

Fonte: Elaborado pelo autor.

3.4 DISCUSSÃO

Analizando os resultados de todos os experimentos, destaca-se a acurácia obtida com a utilização de filtros de Gabor. Porém, a acurácia foi relativamente baixa, principalmente para indivíduos desconhecidos, não garantindo a universalidade dessa metodologia para extração de características para o reconhecimento de expressões faciais.

Porém, os resultados com indivíduos conhecidos foram mais precisos. A melhor acurácia foi 96,71%, com os experimentos com a base JAFFE e o classificador SVM. Nesse mesmo experimento, as expressões neutra e de medo tiveram precisão de 100% enquanto as de raiva e felicidade, 96,67% e 96,77%. Dessa forma, se uma área de aplicação utilizar apenas indivíduos conhecidos, a combinação Gabor e SVM é bastante satisfatória.

Os resultados dos experimentos inter-bases foram menos precisos. A melhor acurácia foi 88,73%, usando a base Yale como treinamento e CK+ como teste. Nesse mesmo experimento, com características de Gabor e classificador SVM, as expressões de felicidade e surpresa tiveram precisão de 100% e 98,80%, respectivamente. O experimento usando a base JAFFE como treinamento e CK+ como teste gerou uma acurácia de 76,06% no melhor resultado (Tabela 36). Invertendo-se as bases de treinamento e teste, o melhor resultado foi de 45,90%. Em (GU *et al.*, 2012), os autores fizeram um procedimento similar e obtiveram acurácia de 54,05% e 55,87%, respectivamente.

Ainda sobre os experimentos inter-bases, observa-se ainda que a utilização da base Yale para treinamento produziu a melhor acurácia. Apesar do número de imagens reduzido, a base Yale é mais heterogênea do que a base JAFFE. Acredita-se que essa diferença tenha contribuído para esse resultado. Esperava-se que a base CK+ tivesse produzido melhores resultados. Sobre os experimentos multibases, as características de Gabor com PCA/LDA obteve uma acurácia

acima de 96% com todos os classificadores. A média obtida foi de 96,59% e desvio de 0,14%. Esse resultado reforça a não universalidade das características, pois os resultados foram bem melhores que os obtidos com os experimentos inter-bases.

Por outro lado, a extração de filtros de Gabor consome bem mais tempo de processamento do que LBP e DCT. A combinação DCT e SVM apresentou um tempo total de 4,7ms, que corresponde a um tempo 27 vezes menor do que Gabor e SVM. Dessa forma, DCT é extremamente recomendado para aplicações em tempo real.

4 CONCLUSÃO E TRABALHOS FUTUROS

4.1 CONCLUSÕES

Os desempenhos das características LPB, DCT e filtros de Gabor foram analisados na tarefa de reconhecimento automático de expressões faciais considerando indivíduos conhecidos e não conhecidos. O objetivo foi obter recomendações de uso considerando o contexto da aplicação: se circunscrita a uma mesma base de dados ou na construção de reconhecedores pretensamente universais.

As principais conclusões derivadas dos resultados são as seguintes:

- a) Que mesmo circunscrita a uma base de dados, não há uma característica que sobreponha as outras em todos os casos. Assim é que LPB apresentou melhor resultado na base de faces ocidentais enquanto filtros de Gabor foi melhor na base de faces asiática, tanto para indivíduos conhecidos quanto para indivíduos desconhecidos.
- b) Por outro lado, no treinamento e testes entre bases, não houve superioridade de nenhuma das características.
- c) Nenhuma das características apresentou superioridade com conjuntos de dados independentes de identidades. Porém, se a área de aplicação analisar apenas indivíduos conhecidos, alguns resultados são satisfatórios.
- d) Não foi possível garantir eficiência do uso do seletor de protótipos FCNN. Embora o seletor tenha melhorado a classificação em alguns casos do KNN, não deve ser considerado para uso prático.
- e) Uma das dificuldades encontradas durante a realização deste trabalho foi encontrar uma forma de recortar a região de uma imagem contendo uma face. Foi verificado a eficiência do método Viola-Jones, porém esse método fornece algumas características desnecessárias para a classificação da expressão. Dessa forma, foi necessário realizar alguns ajustes nas imagens recortadas por esse método.

4.2 TRABALHOS FUTUROS

Entretanto, em todos os casos os resultados não podem ser considerados satisfatórios. A acurácia média de 55%, considerando todos os experimentos, significa que há um amplo campo de pesquisa ainda a ser explorado na solução deste problema.

Por outro lado, estes resultados apontam para algumas direções de pesquisa, as principais sendo: avaliar a utilização de um conjunto misto de características, avaliar novas combinações de pares característica-classificador e investigar novas características com maior poder de discriminação, independente da identidade do indivíduo ser conhecida ou não.

Recomenda-se ainda investigar a utilização das características LBP e DCT aplicadas a blocos das imagens com sobreposição.

REFERÊNCIAS

- ABDULRAHMAN, M.; GWADABE, T. R.; ABDU, F. J.; ELEYAN, A. Gabor wavelet transform based facial expression recognition using pca and lbp. In: IEEE. **Signal Processing and Communications Applications Conference (SIU), 2014 22nd**. [S.l.], 2014. p. 2265–2268.
- ANGIULLI, F. Fast condensed nearest neighbor rule. In: ACM. **Proceedings of the 22nd international conference on Machine learning**. [S.l.], 2005. p. 25–32.
- BETTADAPURA, V. Face expression recognition and analysis: the state of the art. 2012. Disponível em: <<http://arxiv.org/pdf/1203.6722v1.pdf>>. Acesso em: 1 jul. 2015.
- BHADU, A.; TOKAS, R.; KUMAR, D. V. Facial expression recognition using dct, gabor and wavelet feature extraction techniques. **International Journal of Engineering and Innovative Technology**, v. 2, n. 1, 2012.
- CHUNG, K. Y. C. **Facial Expression Recognition by Using Class Mean Gabor Responses with Kernel Principal Component Analysis**. 69 f. Dissertação (Master of Science) — Ohio University, Computer Science (Engineering and Technology), 2010.
- DENG, H.-B.; JIN, L.-W.; ZHEN, L.-X.; HUANG, J.-C. A new facial expression recognition method based on local gabor filter bank and pca plus lda. **International Journal of Information Technology**, v. 11, n. 11, p. 86–96, 2005.
- FASEL, B.; LUETTIN, J. Automatic facial expression analysis: a survey. **Pattern Recognition**, Elsevier, v. 36, n. 1, p. 259–275, 2003.
- GONZALEZ, R.; WOODS, R. **Processamento de imagens digitais**. [S.l.]: Pearson, 2009. ISBN 9788576054016.
- GU, W.; XIANG, C.; VENKATESH, Y.; HUANG, D.; LIN, H. Facial expression recognition using radial encoding of local gabor features and classifier synthesis. **Pattern Recognition**, Pergamon, v. 45, n. 1, p. 80–91, 2012.
- HAYKIN, S. **Neural Networks and Learning Machines**. [S.l.]: Pearson Education Upper Saddle River, 2009.
- JOLLIFFE, I. **Principal component analysis**. [S.l.]: Wiley Online Library, 2002.
- KAUFMAN, L.; ROUSSEEUW, P. J. **Finding groups in data: an introduction to cluster analysis**. [S.l.]: John Wiley & Sons, 2009.
- LABOREIRO, V. R. S. **Uma abordagem baseada em Análise de Texturas para Segmentação Supervisionada de Faces**. 46 f. Dissertação (Mestrado em Ciência da Computação) — Universidade Estadual do Ceará, Fortaleza, 2014.
- LEE, T. S. Image representation using 2d gabor wavelets. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 18, n. 10, p. 959–971, 1996.
- LORENA, A. C.; CARVALHO, A. C. de. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43–67, 2007.

- LUCEY, P.; COHN, J. F.; KANADE, T.; SARAGIH, J.; AMBADAR, Z.; MATTHEWS, I. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In: IEEE. **Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on**. [S.l.], 2010. p. 94–101.
- LYONS, M.; AKAMATSU, S.; KAMACHI, M.; GYOKA, J. Coding facial expressions with gabor wavelets. In: IEEE. **Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on**. [S.l.], 1998. p. 200–205.
- MA, L.; KHORASANI, K. Facial expression recognition using constructive feedforward neural networks. **Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on**, IEEE, v. 34, n. 3, p. 1588–1595, 2004.
- MA, L.; XIAO, Y.; KHORASANI, K.; WARD, R. K. A new facial expression recognition technique using 2d dct and k-means algorithm. In: IEEE. **Image Processing, 2004. ICIP'04. 2004 International Conference on**. [S.l.], 2004. v. 2, p. 1269–1272.
- MAIA, J.; LABOREIRO, V.; CHAVES, F.; MAIA, F.; SILVA, T.; FERREIRA, T. Performance comparison between edited knn and mq-rbfn for regression and classification tasks. In: SBIC. **XI Congresso Brasileiro de Inteligência Computacional**. Ipojuca, Pernambuco, Brazil, 2013.
- NORVIG, P.; RUSSELL, S. **Inteligência Artificial, 3ª Edição**. [S.l.]: Elsevier Brasil, 2014.
- OWUSU, E.; ZHAN, Y.; MAO, Q. R. A neural-adaboost based facial expression recognition system. **Expert Systems with Applications**, Pergamon, v. 41, n. 7, p. 3383–3390, 2014.
- PANTIC, M.; ROTHKRANTZ, L. J. M. Automatic analysis of facial expressions: The state of the art. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 22, n. 12, p. 1424–1445, 2000.
- PEDRINI, H.; SCHWARTZ, W. R. **Análise de imagens digitais: princípios, algoritmos e aplicações**. [S.l.]: Thomson Learning, 2008.
- PLATT, J. *et al.* Fast training of support vector machines using sequential minimal optimization. **Advances in kernel methods—support vector learning**, Cambridge, MA, v. 3, 1999.
- SHAN, C.; GONG, S.; MCOWAN, P. W. Facial expression recognition based on local binary patterns: A comprehensive study. **Image and Vision Computing**, Elsevier, v. 27, n. 6, p. 803–816, 2009.
- SHEN, L.; BAI, L. A review on gabor wavelets for face recognition. **Pattern analysis and applications**, Springer, v. 9, n. 2-3, p. 273–292, 2006.
- SHEN, L.; BAI, L.; FAIRHURST, M. Gabor wavelets and general discriminant analysis for face identification and verification. **Image and Vision Computing**, Elsevier, v. 25, n. 5, p. 553–563, 2007.
- SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais para engenharia e ciências aplicadas curso prático**. [S.l.]: Editora Artliber, 2010.
- TIAN, Y.; KANADE, T.; COHN, J. F. Facial expression recognition. **Handbook of face recognition**, Springer London, p. 487–519, 2011.

VALSTAR, M. Automatic facial expression analysis. In: MANDAL, M. K.; AWASTHI, A. (Ed.). **Understanding Facial Expressions in Communication**. Springer India, 2015. p. 143–172. ISBN 978-81-322-1933-0. Disponível em: <http://dx.doi.org/10.1007/978-81-322-1934-7_8>.

VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: IEEE. **Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on**. [S.l.], 2001. v. 1, p. I–511.

WAGNER, P. Principal component analysis and linear discriminant analysis with gnu octave. 2011. Disponível em: <http://www.bytefish.de/blog/pca_lda_with_gnu_octave/>.

WEBB, A. R.; COPSEY, K. D. **Statical Pattern Recognition**. [S.l.]: Wiley, 2011.

WEKA. Weka 3: Data mining software in java. 2014. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em: 1 jul. 2015.

XIANG, C.; FAN, X.; LEE, T. H. Face recognition using recursive fisher linear discriminant. **Image Processing, IEEE Transactions on**, IEEE, v. 15, n. 8, p. 2097–2105, 2006.

YALE. Yale face database. 2000. Disponível em: <<http://vision.ucsd.edu/content/yale-face-database>>. Acesso em: 1 jul. 2015.

APÊNDICES

APÊNDICE A – Implementação Gabor

Implementação: Gabor

```

Mat getFiltroGaborDeng(float ptheta, float pomega, int pnx, int pny, int ptipo)
{
    float theta = ptheta;
    float escala = pomega;
    int nx = floor(pnx / 2);
    int ny = floor(pny / 2);
    int dimx = 2*nx + 1;
    int dimy = 2*ny + 1;

    Mat real = cv::Mat::zeros(dimx, dimy, CV_32F);
    Mat imag = cv::Mat::zeros(dimx, dimy, CV_32F);
    Mat mag = cv::Mat::zeros(dimx, dimy, CV_32F);

    int x, y;

    float wmax = PI/2.0;
    float wm = wmax * pow(sqrt(2.0), -escala);
    float sigma = PI/wm;

    for(x = -nx; x <= nx; x++)
    {
        for(y = -ny; y <= ny; y++)
        {
            float x1 = x*cos(theta) + y*sin(theta);
            float y1 = -x*sin(theta) + y*cos(theta);

            float A = (1.0/(2*PI*sigma*sigma)) * exp(-((x1*x1 + y1*y1)/(2*sigma*sigma)));
            float Br = cos(wm*x1);
            float Bi = sin(wm*x1);
            float C = exp(-(wm*wm*sigma*sigma)/2.0);

            real.at<float>(x + nx, y + ny) = A*(Br - C);
            imag.at<float>(x + nx, y + ny) = A*Bi;
        }
    }

    int i, j;

    for(i = 0; i < dimx; i++)
    {
        for(j = 0; j < dimy; j++)
        {

```

```
float re = real.at<float>(i, j);
float im = imag.at<float>(i, j);

float mg = sqrt(re*re + im*im);

mag.at<float>(i, j) = mg;
}
}

if (ptipo == 0)
    return real;
else if (ptipo == 1)
    return imag;
else
    return mag;
}
```

APÊNDICE B – Implementação PCA e LDA

Implementação: PCA

```
// Obtendo o modelo PCA
PCA pca = PCA(featuresGabor, media, CV_PCA_DATA_AS_ROW, numPCA);

// Projetando os dados de treinamento
Mat treinamentoProjetado = pca.project(featuresGabor);

//Projetando os dados de teste
Mat testeProjetadoTemp = pca.project(featuresGaborTeste);
```

Implementação: LDA

```
library(caret)
library(MASS)

#lendo os arquivos de treinamento e teste
treino <- read.table("/home/chaves/Dropbox/Mestrado/R/treina2.csv",
header = TRUE, sep=",")

teste <- read.table("/home/chaves/Dropbox/Mestrado/R/teste2.csv" ,
header = TRUE, sep=",")

set.seed(32343)

#gerando o modelo
modelo <- lda(CLASSE ~., data=treino)

#obtendo os autovetores
base <- modelo$scaling
n <- ncol(treino) - 1

#projetando os dados de treinamento e teste na nova base
treinoMatrix <- matrix(as.numeric(unlist(treino[,1:n])),nrow=nrow(treino[,1:n]))
treinoProj <- treinoMatrix %*% base

testeMatrix <- matrix(as.numeric(unlist(teste[,1:n])),nrow=nrow(teste[,1:n]))
testeProj <- testeMatrix %*% base

#salvando os arquivos projetados na nova base
write.table(treinoProj, file = "/home/chaves/Dropbox/Mestrado/R/treinaLDA.csv",
sep = ";", col.names = FALSE, row.names=FALSE, qmethod = "double")

write.table(testeProj, file = "/home/chaves/Dropbox/Mestrado/R/testeLDA.csv",
sep = ";", col.names = FALSE, row.names=FALSE, qmethod = "double")
```

APÊNDICE C – Implementação Classificadores

Implementação: Classificador SMO

```

for(int i = 1; i <= numFolds; i++)
{
    // Lendo arquivo ARFF de treinamento
    BufferedReader readerTreino = new BufferedReader(new FileReader("treino.arff"));
    Instances treinoARFF = new Instances(readerTreino);

    // Lendo arquivo ARFF de teste
    BufferedReader readerTeste = new BufferedReader(new FileReader("teste.arff"));
    Instances testeARFF = new Instances(readerTeste);

    treinoARFF.setClassIndex(treinoARFF.numAttributes() - 1);
    testeARFF.setClassIndex(testeARFF.numAttributes() - 1)

    // Instanciando o classificador SMO
    weka.classifiers.functions.SMO scheme = new weka.classifiers.functions.SMO();

    // Setando parametros, construindo o classificador e testando
    try
    {
        scheme.setOptions(
            weka.core.Utils.splitOptions("-C0.6-L0.001-P1.0E-12
UU-NU0U-VU-1U-WU1
UU-KU\"weka.classifiers.functions.supportVector.PolyKernel-U-CU250007-U-EU1.0\""));

        scheme.buildClassifier(treinoARFF);

        Evaluation evaluation = new Evaluation(treinoARFF);

        evaluation.evaluateModel(scheme, testeARFF);

        System.out.println("-----");
        System.out.println("FOLD" + i + " - Percentual: " + evaluation.pctCorrect());
        System.out.println("-----");
        System.out.println(evaluation.toSummaryString());
        System.out.println(evaluation.toMatrixString());

        System.out.println("-----");
    }
    catch (Exception e)
    {
        e.printStackTrace();
    }
} // for i

```