

UNB – UNIVERSIDADE DE BRASÍLIA

FGA – FACULDADE GAMA

PROGRAMA DE PÓS-GRADUAÇÃO

CLASSIFICAÇÃO DE IMAGENS DE RETINA PARA AUXÍLIO A
DIAGNÓSTICO DE RETINOPATIA DIABÉTICA VIA DETECÇÃO DE
EXSUDATOS DUROS, COM BASE EM CLASSIFICADORES ROBUSTOS,
MÁQUINAS DE VETORES DE SUPORTE E ALGORITMOS COM
TREINAMENTO DE DICIONÁRIOS

CHARLES CORRÊA MENDES

ORIENTADOR: PROF. CRISTIANO JACQUES MIOSSO, PHD

DISSERTAÇÃO DE MESTRADO EM
ENGENHARIA BIOMÉDICA

PUBLICAÇÃO: 068A/2017

BRASÍLIA/DF, JUNHO DE 2017

UNIVERSIDADE DE BRASÍLIA
FACULDADE UNB GAMA
ENGENHARIA BIOMÉDICA

CLASSIFICAÇÃO DE IMAGENS DE RETINA PARA AUXÍLIO A
DIAGNÓSTICO DE RETINOPATIA DIABÉTICA VIA DETECÇÃO
DE EXSUDATOS DUROS, COM BASE EM CLASSIFICADORES
ROBUSTOS, MÁQUINAS DE VETORES DE SUPORTE E
ALGORITMOS COM TREINAMENTO DE DICIONÁRIOS

CHARLES CORREA MENDES

DISSERTAÇÃO DE MESTRADO SUBMETIDA À FACULDADE UNB GAMA DA
UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO TÍTULO DE MESTRE EM ENGENHARIA BIOMÉDICA.

APROVADA POR:



CRISTIANO JACQUES MIOSSO RODRIGUES MENDES, DR. – FGA / UNB
(ORIENTADOR)



MARCUS VINICIUS CHAFFIM COSTA, DR. - FGA / UNB
(EXAMINADOR INTERNO)



ALEXANDRE RICARDO SOARES ROMARIZ, DR. – ENE / UNB
(EXAMINADOR EXTERNO)

BRASÍLIA, 13 DE JUNHO DE 2017

BRASÍLIA/DF, JUNHO DE 2017

FICHA CATALOGRÁFICA

CORRÊA MENDES, CHARLES

Classificação de imagens de retina para auxílio a diagnóstico de retinopatia diabética via detecção de exsudatos duros, com base em classificadores robustos, máquinas de vetores de suporte e algoritmos com treinamento de dicionários [Distrito Federal], 2017.

68p., 210 × 297 mm (FGA/UnB Gama, mestrado em engenharia biomédica, 2017).

Dissertação de mestrado em engenharia biomédica, Faculdade UnB Gama, programa de pós-graduação em engenharia biomédica.

- | | |
|--|-----------------------------|
| 1. Retinopatia Diabética | 2. Classificadores |
| 3. <i>Principal Component Analysis</i> | 4. <i>Fourier Transform</i> |
| I. FGA UnB/UnB. | II. Título (série) |

REFERÊNCIA

CORRÊA MENDES, CHARLES (2017). Classificação de imagens de retina para auxílio a diagnóstico de retinopatia diabética via detecção de exsudatos duros, com base em classificadores robustos, máquinas de vetores de suporte e algoritmos com treinamento de dicionários. Dissertação de mestrado em engenharia biomédica, Publicação 068A/2017, Programa de Pós-Graduação, Faculdade UnB Gama, Universidade de Brasília, Brasília, DF, 68p.

CESSÃO DE DIREITOS

AUTOR: Charles Corrêa Mendes

TÍTULO: Classificação de imagens de retina para auxílio a diagnóstico de retinopatia diabética via detecção de exsudatos duros, com base em classificadores robustos, máquinas de vetores de suporte e algoritmos com treinamento de dicionários

GRAU: Mestre

ANO: 2017

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor.



Charlesmendes8@hotmail.com

Brasília, DF – Brasil

RESUMO

Esta pesquisa aborda o desenvolvimento e avaliação de um sistema para classificação de imagens de retina como forma de auxílio a diagnóstico da retinopatia diabética. Atualmente, caso não haja tratamento, a retinopatia é a maior causa de cegueira no mundo para as pessoas com Diabetes Mellitus. Sendo assim, como o diagnóstico por imagem é a forma de detecção da doença, o trabalho tem por finalidade desenvolver um sistema com melhor desempenho e praticidade e que possa ser utilizado, por exemplo, em ambientes carentes de serviços de saúde universais, como ferramenta de apoio ao diagnóstico para os profissionais de saúde.

O método tradicional de classificação de imagens médicas utiliza extração de características, porém muitas vezes pode ocorrer a perda de informações importantes na imagem que serviriam para um diagnóstico mais preciso. Por isso, o intuito é comparar algoritmos de classificação de imagens com algumas características propostas e também com algoritmos que dispensam extração de características por meio de treinamento de dicionários. Para isso, a ideia é implementar a classificação utilizando-se classificadores frequentemente usados, como *Support Vector Machine* (SVM) e do tipo Comitês, além de um método mais recente que explora o treinamento de dicionários simultaneamente ao treinamento do classificador, o chamado *Learning Algorithm for Soft-thresholding Classifier* (LAST). Além disso, avalia-se a utilização de algumas ferramentas que possam melhorar o desempenho da classificação como análise de componentes principais (PCA, do inglês *Principal Component Analysis*) e o uso da transformada discreta de Fourier para extração de características espectrais.

Testes de hipóteses realizados após avaliação sistemática do sistema mostraram que a classificação usando o método do LAST com PCA teve melhor desempenho nas comparações com o classificador do tipo Comitê com o PCA com nível de significância de 98% e com o SVM com PCA com nível de significância de 99.99%. Já na comparação entre o classificador do tipo Comitês com PCA e SVM com PCA, o método do tipo Comitês obteve resultado melhor com nível de significância de 99.99%. Outros testes de hipóteses mostraram que as três classificações possuem acurácia superior a 0.8 com nível de significância de 91% para o LAST, 99.99% para o classificador do tipo Comitês e 99% para o SVM. Em experimentos com outras variáveis, tendo a Medida-F como parâmetro obteve-se a seguinte ordem de desempenho: LAST com PCA e o módulo da Transformada de Fourier (0.88), LAST (0.75), LAST com módulo de Fourier (0.70), LAST com PCA e fase de Fourier (0.65) e LAST com fase de Fourier (0.60). A análise no plano *Receiver Operating Characteristic* mostrou que as classificações da imagens com SVM com PCA, do tipo Comitês com PCA, LAST com PCA, LAST e LAST com o módulo da transformada de Fourier e PCA possuem boa relação entre verdadeiros positivos e falsos negativos, o mesmo não acontece nos casos do LAST módulo da transformada de Fourier, LAST com fase da transformada de Fourier e PCA e no LAST com a fase da transformada de Fourier.

Ao final, conclui-se que no treinamento com o uso de dicionários, sem a extração prévia de características, o LAST obteve melhor acurácia média (0.84) do que a acurácia comumente encontrada na literatura (0.8) e com as acurácias médias encontradas na pesquisa usando o classificador do tipo Comitês (0.81) e o SVM (0.79). Estes resultados sugerem que algumas configurações propostas, sobretudo a do LAST aplicado diretamente a imagens de retina, podem auxiliar o diagnóstico de retinopatia diabética, dados os índices de verdadeiros positivos encontrados e os baixos índices de falsos negativos. Em trabalho futuro, o autor pretende combinar técnicas de aprendizagem profunda com as abordagens aqui empregadas, para avaliar o impacto sobre as métricas de desempenho, com base em bancos de imagens maiores a serem obtidos em parcerias de pesquisa.

ABSTRACT

This research is the development and evaluation of a system to classify retinal images as a way to assist in the diagnosis of diabetic retinopathy. Currently, if there is no treatment, the retinopathy is the largest cause of blindness in the world for people with Diabetes Mellitus. Thus, as imaging diagnosis is the form of detection of the disease, the purpose of the work is to develop a system with better performance and that can be used, for example, in environments lacking universal health services, as a tool for support for health professionals.

The traditional method of medical imaging uses extractions of features, but often there is a loss of important information in the image that would serve a more accurate diagnosis. Therefore, the aim is to compare algorithms of image classification with some characteristics proposed and also with algorithms that do not require extraction of features through the training of dictionaries. For this, the idea is to implement classification using frequently used classifiers, such as Support Vector Machine (SVM) and Ensemble, as well as a more recent method that explores the training of dictionaries simultaneously to training Classifier, the Learning Algorithm for Soft-thresholding Classifier (LAST). In addition, it is evaluated the use of some tools that can improve the classification performance as main component analysis (PCA) and use of the discrete fourier transform to extract spectral features.

Hypothesis tests performed after systematic evaluation showed that the classification using the LAST method with PCA performed better in the comparisons with the Ensemble with the PCA with a level of significance 98% and with the SVM with PCA With a level of significance 99.99%. In the comparison between Ensemble with PCA and SVM with PCA, the Ensemble method obtained a better result with a level of significance 99.99%. Other hypothesis tests showed that all three classifications had accuracy above 0.8 with a level of significance 91% for LAST, 99.99% for Ensemble and 99% for SVM. In the experiments with other variables, the performance of F-measure order was obtained: LAST with PCA and Fourier transform (0.88), LAST (0.75), LAST with Fourier module (0.70), LAST with PCA and Fourier phase (0.65) and LAST with Fourier phase (0.60). The analysis in the Receiver Operating Characteristic plane showed that the classifications of SVM images with PCA, Ensemble with PCA, LAST with PCA, LAST and LAST with the Fourier transform module and PCA have a good relationship between true positives and false negatives, the same does not happen in the cases of LAST Fourier transform module, LAST with Fourier transform phase and PCA, and LAST with the Fourier transform phase.

Finally, we conclude that dictionary training without the previous extraction of features with the LAST obtained a better average accuracy (0.84) than the accuracy commonly found in the literature (0.8) and with the average accuracy Found in the search using Ensemble (0.81) and SVM (0.79). The research shows that the classification using LAST would help in the diagnosis of diabetic retinopathy, because the system presented good index of true positives and low index of false negatives. In future work, the authors intend to combine deep learning techniques with the approaches used here to evaluate the impact on performance metrics, based on larger database to be obtained in research partnerships.

SUMÁRIO

1	Histórico, consequência, diagnóstico e limitações	1
1.1	Contextualização	1
1.2	Definição do problema e Proposta	2
1.3	Objetivos	4
1.3.1	Objetivo geral	4
1.3.2	Objetivos específicos	4
1.4	Justificativa	5
1.5	Estrutura da dissertação	6
2	Fundamentação teórica	7
2.1	Retinopatia	7
2.1.1	Retinopatia Diabética	7
2.2	Extração de características	11
2.3	Classificadores	12
2.3.1	Máquinas de Vetores de Suporte	12
2.3.2	Classificadores do tipo Comitês	18
2.3.3	Algoritmo de Aprendizagem por Limiarização Suave	23
2.4	Análise de Componentes Principais	25
2.5	Transformada de Fourier	29

2.6	Avaliação de desempenho para classificadores	34
2.6.1	Métricas de desempenho	34
2.6.2	Plano ROC	36
2.6.3	Testes de hipótese	38
3	Metodologia	41
3.1	Desenvolvimento dos algoritmos de classificação das imagens de retina . .	42
3.1.1	Detecção do disco óptico	47
3.1.2	Rotação imagens com exsudato duro	48
3.2	Experimentos de avaliação dos métodos implementados de classificação de imagens de retina	49
3.3	Metodologia de Análise dos Resultados de Classificação – Validação e Com- paração dos Classificadores	51
3.3.1	Métricas para avaliação de desempenho dos classificadores	51
3.3.2	Testes de Hipóteses	52
3.3.3	Medida de desempenho no plano ROC	53
4	Resultados e Discussão	54
4.1	Histogramas de desempenho dos métodos testados sistematicamente	54
4.2	Medidas de desempenho para os testes realizados uma vez	56
4.3	Testes de Hipóteses para comparação entre os testes sistemáticos	58
4.4	Resultados de desempenho de todos os testes no plano ROC	59
5	Conclusão	61

LISTA DE TABELAS

2.1	Principais complicações oftalmológicas causadas pela diabetes mellitus. . .	8
2.2	Exemplo de dados para Análise de Componentes Principais.	26
2.3	Exemplo de uma matriz de confusão.	34
2.4	Possibilidades envolvidas em um teste de hipóteses.	40
3.1	Resultados de experimentos preliminares com os métodos do tipo Comitês.	50
4.1	Matrizes de confusão para os resultados da classificação com testes feitos uma vez.	57
4.2	Resultados de experimentos feitos com LAST e ferramentas como PCA e Transformada de Fourier.	57
4.3	Valores dos níveis de significância os testes de hipótese entre SVM com PCA, Comitês com PCA e LAST com PCA.	58
4.4	Valores dos níveis de significância os testes de hipótese das médias das acurácias do SVM com PCA, do tipo Comitês com PCA e LAST com PCA com acurácia de referência de 0.8.	59

LISTA DE FIGURAS

1.1	Comparação entre os campos visuais de uma pessoa com (a) visão normal e (b) com retinopatia severa.	2
2.1	Principais partes anatômicas da retina.	9
2.2	Retinógrafo responsável pelo exame de fundo de retina.	10
2.3	Fases Retinopatia Diabética.	10
2.4	Principais características da imagem de retina.	11
2.5	Detecção manual de exsudatos duros em imagens retinianas.	11
2.6	Ilustração da separação de classes SVM.	13
2.7	Interpretação geométrica de w e b sobre um hiperplano.	14
2.8	Ilustração de uma distribuição dos dados em duas classes perfeitamente separadas.	15
2.9	Cálculo da distância d entre os hiperplanos H_1 e H_2	17
2.10	Arquitetura comum para o classificador do tipo Comitês.	18
2.11	Modelo árvore de decisão.	22
2.12	Classificação em tempo de teste realizados pelo LAST.	25
2.13	Algumas funções bidimensionais e seus espectros de Fourier.	31
2.14	Espaço ROC.	37
2.15	Diferentes pontos no espaço ROC.	38
3.1	Semelhança entre o disco óptico e região com exsudatos duros.	42

3.2	Diagrama esquemático com a metodologia implementada para o treinamento e classificação usando SVMs, Comitês e LAST.	46
3.3	Exemplificação das posições diferentes do disco óptico.	47
3.4	Exemplo de marcação do disco óptico.	48
3.5	Exemplo da obtenção de imagem com exsudatos duros.	48
3.6	Exemplo de rotações em uma imagem com exsudato duro.	49
3.7	Exemplos de imagens do banco de dados MESSIDOR.	50
4.1	Histograma dos testes do classificador SVM com PCA.	55
4.2	Histograma dos testes do classificador do tipo Comitês com PCA.	55
4.3	Histograma dos testes do classificador LAST com PCA.	56
4.4	Metricas de desempenho no plano ROC para todos metodos testados.	60

LISTA DE SÍMBOLOS, NOMENCLATURAS E ABREVIACÕES

RD	–	Retinopatia Diabética
RDNP	–	Retinopatia Diabética não-proliferativa
RDP	–	Retinopatia Diabética proliferativa
DM	–	Diabetes Mellitus
OMS	–	Organização Mundial da Saúde
DO	–	disco óptico
CP	–	componente principal
RBF	–	<i>Radius Basis Function</i>
SLFN	–	<i>Single Layer Feedforward Network</i>
ROC	–	<i>Receiver Operating Characteristics</i>
AdaBoost	–	<i>Adaptative Boosting</i>
FT	–	<i>Fourier Transform</i>
FFT	–	<i>Fast Fourier Transform</i>
DFT	–	<i>Discrete fourier transform</i>
LAST	–	<i>Learning Algorithm for Soft-Thresholding</i>
PCA	–	<i>Principal Component Analysis</i>
SVD	–	<i>Singular Value Decomposition</i>
SVM	–	<i>Support Vector Machine</i>
VP	–	verdadeiro positivo
VN	–	verdadeiro negativo
FP	–	falso positivo
FN	–	falso negativo

1 HISTÓRICO, CONSEQUÊNCIA, DIAGNÓSTICO E LIMITAÇÕES

1.1 CONTEXTUALIZAÇÃO

Ao longo das duas últimas décadas, a Diabetes Mellitus (DM) tornou-se um dos problemas de saúde mais desafiadores. Estima-se que havia 366 milhões de pessoas com diabetes em 2011, e projeta-se que o número aumente para 552 milhões em 2030 [1]. As taxas de ocorrência estimadas de retinopatia diabética (RD) e o risco de perda da visão são de 40,3% e 8,2%, respectivamente, em adultos no Estados Unidos acima dos 40 anos [1]. Estima-se que metade dos pacientes com retinopatia no estágio mais avançado, chamada de proliferativa, perde a visão dentro de cinco anos se não recebem o tratamento adequado [1].

A RD constitui a principal causa de deficiência visual e perda de visão entre adultos e afeta mais de 90% dos pacientes diabéticos em algum grau. Aproximadamente 10% dos pacientes diabéticos com mais de 15 anos desenvolvem deficiência visual grave, sendo que cerca de 10% chegam a cegueira [2]. A Figura 1.1 ilustra o efeito da retinopatia diabética avançada sobre o campo visual.

A retinopatia é uma doença que causa danos microvasculares e macrovasculares na retina e que pode ser causada pela Diabetes Mellitus (retinopatia diabética) e pela hipertensão arterial (retinopatia hipertensiva). A DM é uma doença comum em torno do mundo e é grande causadora de cegueira. Diversos efeitos negativos sobre a retina se desenvolvem em consequência da hiperglicemia crônica [4].

A RD é a mais comum causa de perda de visão em pacientes com diabetes no mundo, segundo a Organização Mundial da Saúde (OMS). Cerca de 80% dos indivíduos com diabetes vivem em países em desenvolvimento, como China, Índia e Brasil [5]. No Brasil, no final da década de 1980, estimou-se a prevalência de DM na população adulta em 7.6%. Em 2014, estimou-se que existiriam 11,9 milhões de pessoas, na faixa etária de 20 a 79 anos, com diabetes no Brasil, podendo alcançar 19,2 milhões em 2035 [5]. Além



Figura 1.1. Comparação entre os campos visuais de uma pessoa com (a) visão normal e (b) retinopatia severa. Fonte: [3].

disso, segundo o censo realizado pelo Conselho Brasileiro de Oftalmologia em 2014 havia um médico especialista em saúde ocular para cada grupo de 11.604 pessoas. A OMS estabelece como ideal a proporção de um para 17.000. O dado é animador, porém apenas 848 dos 5.570 municípios brasileiros, ou 15%, contam com atendimento de especialistas em saúde ocular. Ou seja, 85% das cidades não possuem tais especialistas, resultando na carência de médicos especializados em saúde ocular em pequenas cidades pelo país [6].

Assim como a diabetes, a hipertensão arterial acomete grande parte da população mundial causando elevado risco de doenças cardiovasculares e eventos cerebrovasculares, tais como infarto e acidente vascular cerebral. Além disso, a hipertensão causa danos em órgãos específicos como os rins (insuficiência renal) e a retina (retinopatia hipertensiva) [7].

1.2 DEFINIÇÃO DO PROBLEMA E PROPOSTA

A detecção precoce da RD é de fundamental importância para a eficácia dos tratamentos [8]. O exame oftalmológico completo incluindo o exame do fundo do olho e biomicroscopia é necessário para a detecção e determinação da fase e gravidade da retinopatia. A documentação fotográfica (retinografia) também é importante para a detecção e avaliação da progressão da doença e dos resultados do tratamento [8, 9].

Alguns métodos para classificação de imagens de retina com RD estão presentes na literatura. Em 2007, *Sanchez et al.* propuseram um algoritmo baseado na Análise Linear Discriminante de Fisher e de intensidade de cores para classificar retinas com exsudatos [10]. Em 2009, *Welfera et al.* fizeram uso do método de morfologia matemática

para detectar exsudatos em imagens de fundo de olho de diabéticos [11]. Em 2011, *Kumaria e Maruthib* usaram a técnica de Equalização de histograma para segmentar exsudatos em imagens de retina [12]. Em 2012, *Youssef e Salouma* usaram métodos para segmentação de bordas para detectar os vasos sanguíneos e classificar automaticamente imagens com retinopatia [13]. Também no ano de 2012, *Ranamuka e Meegama* usaram a morfologia matemática para detectar exsudatos, então por meio de um algoritmo de lógica *Fuzzy* com funções de associação no espaço de cor RGB [14]. Já em 2015, *Rokade e Manza* usaram, no pré-processamento, cálculos de mapas de probabilidade para exsudatos e então fizeram uso de uma Transformada Wavelet para classificação da retinopatia diabética.

Apesar de todos os avanços no entendimento de complicações diabéticas crônicas, a retinopatia diabética continua a ser um problema clínico grave devido, muitas vezes, à falta de intervenção terapêutica eficaz [15]. Tendo em vista que em muitos lugares há carência de profissionais capacitados para diagnosticar pacientes que desenvolvem a RD, a proposta desse trabalho é desenvolver um sistema automatizado para classificação de imagens com RD. O sistema visa detectar as retinas que apresentam depósitos de lipídios, chamados de os exsudatos duros, que estão presentes no estágio inicial da retinopatia o que poderá servir como uma ferramenta de apoio ao diagnóstico. Posteriormente, após os pacientes diabéticos passarem por esse exame eles poderão ser encaminhados a oftalmologistas.

O sistema tradicional automatizado utilizado para o diagnóstico de algumas doenças por imagem é gerado pela extração das principais características. Elas são usadas no classificador de forma separada. A dificuldade encontrada, geralmente, é saber com precisão quais características devem ser empregadas no classificador gerando, muitas vezes, perda de informações importantes no diagnóstico. O treinamento de dicionários substitui a extração prévia de características. Apesar dos novos métodos já utilizados na classificação de imagens retinianas, não foram encontrados estudos sistemáticos comparando diversos classificadores, e sobretudo classificadores que explorem treinamento de dicionários, isoladamente ou em conjunto com a utilização de características extraídas previamente como em análises de componentes principais (PCA, do inglês *Principal Component Analysis*). Houve muitos avanços em classificadores como do tipo Comitês [16], Aprendizagem Profunda [17], *Extreme Learning Machine* [18] e *Dictionary learning for fast classification based on soft-thresholding* (LAST) [19], porém estes não foram comparados sistematicamente para retinopatia.

Na presente pesquisa, as informações extraídas por meio da PCA serão utilizadas para a classificação utilizando classificadores clássicos como a máquina de vetores de suporte (SVM) [20, 21] e classificadores do tipo Comitês [16, 22]. Além disso, os resultados obtidos com a SVM e do tipo Comitês serão comparados com um classificador implementado

mais recentemente, o algoritmo de aprendizagem por limiarização suave (LAST, do inglês *Dictionary learning for fast classification based on soft-thresholding*). Nele a intenção é integrar a extração de características e o classificador em um único sistema de forma que o classificador esteja sempre atualizado com as novas informações relevantes retiradas da imagem. Teoricamente, é possível extrair um conjunto de características para uma classificação desejada junto com o treinamento do classificador, o que poderá resultar em melhor desempenho.

A proposta é implementar, avaliar e validar um sistema para imagens de retina especificamente para o pré-diagnóstico de retinopatia diabética. A intenção é criar e avaliar um classificador automático eficiente que sirva como uma ferramenta de diagnóstico mais rápida e estabeleça com precisão os pacientes com maior suspeita de retinopatia diabética.

1.3 OBJETIVOS

1.3.1 Objetivo geral

O objetivo desta pesquisa é implementar e analisar um sistema de classificação de imagens de retina baseado em técnicas para redução de dimensionalidade aplicadas a imagens e em técnicas baseadas em treinamento de dicionários. O sistema deve servir como ferramenta de apoio ao diagnóstico em pacientes diabéticos a fim de identificar a presença de exsudatos duros, por se tratar de um dos principais sintomas da retinopatia. O sistema passará por testes de validação e comparação com técnicas de aprendizagem de máquina e otimização numérica como SVM, do tipo Comitês e LAST para confirmar se a classificação de imagens de retina sem a extração prévia de características aumentará a precisão, acurácia e sensibilidade na tarefa de detectar exsudatos.

1.3.2 Objetivos específicos

Para a comparação sistemática das estratégias de classificação propostas, é necessário o cumprimento de alguns objetivos específicos, a saber:

- Implementação de um sistema de detecção de disco óptico, para que esta região imagem não seja classificada erroneamente como exsudato;
- Implementação de um sistema para treinamento e validação de imagens de retina com base em treinamento de dicionários e usando a metodologia LAST [19];

- Implementação de um sistema para treinamento e validação de imagens de retina usando SVM, do tipo Comitês e LAST com as características extraídas pelo PCA;
- Implementação de um sistema para treinamento e validação de imagens de retina usando LAST com as características extraídas pelo PCA e informações dadas pelo módulo e fase da Transformada de Fourier;
- Implementação de um sistema para treinamento e validação de imagens de retina usando LAST com informações dadas pelo módulo e fase da Transformada de Fourier;
- Realização de testes com diferentes conjuntos de testes e validação para avaliação estatística das métricas de desempenho;
- Comparação dos classificadores SVM, do tipo Comitês e LAST por meio Testes de Hipóteses;
- Comparação do desempenho do classificador LAST usando métricas estatísticas da acurácia, precisão sensibilidade e medida-F.
- Análise da hipótese de que as técnicas mais recentes de aprendizagem de máquinas (LAST) comparadas com a abordagem clássica (extração de características seguida de classificação) podem levar a um aumento da precisão, acurácia, sensibilidade na detecção de exsudatos duros.

1.4 JUSTIFICATIVA

O sistema poderá ser usado em lugares com maior carência como em alguns lugares da África, Índia e no Brasil, onde não se tem acesso facilitado a exames de diagnóstico para doenças na retina. Cabe enfatizar que, por falta de um diagnóstico precoce, a retinopatia diabética é um dos maiores causadores de cegueira em todo o mundo. Portanto, a falta de compreensão do dano que a diabetes e a hipertensão causam, sem o desenvolvimento de novos e eficazes diagnósticos preventivos, é motivo de significativa perda de visão [1].

A intenção é que em lugares onde há carência de profissionais especializados em oftalmologia o sistema possa ser utilizado por técnicos treinados. Eles realizarão o exame de forma mais precisa e o sistema implementado servirá como uma ferramenta automática de apoio ao diagnóstico para atender a pessoas com hiperglicemia suspeitas ou portadoras de retinopatia. A ideia é que os casos apontados como suspeitos de retinopatia sejam encaminhados a profissionais de saúde para análise mais detalhada e tratamento.

1.5 ESTRUTURA DA DISSERTAÇÃO

A dissertação está organizada da seguinte forma. O Capítulo 2 aborda conceitos básicos de diabetes mellitus e retinopatia diabética, sobretudo a necessidade do diagnóstico rápido da retinopatia por meio da classificação automatizada. Uma seção em particular aborda os principais pontos de extração de características na imagens de retina. Em seguida são abordados os classificadores que serão utilizados para comparação dividido com subsessões específicas para o máquina de vetores de suporte, algoritmo de aprendizagem por limiarização suave e do tipo Comitês. Finalmente, são apresentados os princípios de análise de componentes principais e da Transformada de Fourier.

O Capítulo 3 apresenta toda a metodologia utilizada na pesquisa. O desenvolvimento do sistema é detalhado nessa seção, incluindo a detecção do disco óptico e a rotação das imagens com exsudatos duros para composição dos conjuntos de treinamento dos classificadores. Também são detalhados os experimentos de validação dos classificadores.

O Capítulo 4 mostra com histogramas os resultados obtidos nos testes das classificações, além de avaliar e comparar os resultados por meio testes de hipóteses e métricas de desempenho como precisão, acurácia, sensibilidade e Medida-F.

Finalmente, o Capítulo 5 apresenta as conclusões desta pesquisa, com ênfase nos ganhos das novas técnicas de classificação avaliadas com respeito aos métodos tradicionais de extração de características e classificação, e nos trabalhos futuros que se pretende desenvolver a partir destes resultados.

2 FUNDAMENTAÇÃO TEÓRICA

A retinopatia diabética é a maior causa de cegueira entre diabéticos e o diagnóstico ocorre por meio de exame de imagens do fundo da retina. Neste capítulo são apresentados conceitos e um breve contexto histórico da doença, com uma descrição de suas causas e das principais características encontradas em imagens de retina. Além disso, descrevem-se os principais conceitos relacionados a extração de características e alguns dos principais classificadores estudados e utilizados como LAST, SVM e do tipo Comitês. É apresentada a técnica de análise de componentes principais e sua aplicação em redução de dimensionalidade. Finalmente, detalham-se os principais conceitos da Transformada de Fourier em processamento de imagens.

2.1 RETINOPATIA

2.1.1 Retinopatia Diabética

A Diabetes Mellitus é uma enfermidade reconhecida desde o século II AC, contudo somente foi associada a complicações oftalmológicas em meados do século XIX. A primeira descrição da Retinopatia Diabética Proliferativa foi apresentada por Wilhelm Manz (1833-1911) em 1886 com a publicação “Retinis proliferans”. A primeira classificação da retinopatia diabética foi usada em 1890 por Julius Hirschberg (1843-1925), tendo este dividido em: “retinitis centralis punctuate”, “haemorrhagic form”, “retinal infarction” e “haemorrhagic glaucoma” [23]. No início do século XX, o escocês Arthur James Ballantyne (1876-1954) mostrou pela primeira vez o papel das alterações na parede capilar no desenvolvimento da retinopatia diabética, assim como a presença de exsudatos moles [23].

A DM é uma desordem no metabolismo da glicose com complicações graves, incluindo complicações como disfunções microvasculares (retinopatia, nefropatia e neuropatia) e macrovasculares, doença cerebrovascular, doença coronária isquêmica e doença arterial periférica. O avanço da retinopatia progride de forma lenta e gradual, sendo a maior causa de perda de visão em pacientes com diabetes [15]. A Tabela 2.1 mostra as principais

complicações oftalmológicas causadas pela DM, sendo a retinopatia a mais importante e comum entre elas [3].

Tabela 2.1. Principais complicações oftalmológicas causada pela diabetes mellitus.
Fonte: [3].

Local	Lesão/Alteração
Retina	Retinopatia diabética
	Maculopatia
	Oclusão da veia central da retina
Nervo Óptico	Papilite diabética
	Glaucoma
	Neuropatia óptica isquêmica não-arterial
Cristalino	Cataratas
	Alterações refractivas
Íris	Rubeosis iridis
	Aumento da transiluminação da íris
Córnea	Diminuição da sensibilidade corneana
	Síndrome da erosão corneana recorrente
	Diminuição da contagem de células endoteliais

As imagens de retina são utilizadas por oftalmologistas para o diagnóstico de várias doenças, tais como retinopatia diabética e maculopatia. A detecção de vasos sanguíneos é um primeiro passo importante para o desenvolvimento de sistemas de diagnóstico assistido por computador para análise de imagens de fundo de olho. Um dos métodos usados para obtenção dessas imagens é o exame ocular do fundo do olho, feito com o aparelho chamado oftalmoscópio. Este método pode ser direto, quando se obtém uma imagem ampliada quinze vezes maior, mas com restrito campo de visão, ou indireto, quando se usa uma imagem com ampliação menor, porém com visualização mais ampla da retina. Outro exame feito é a biomicroscopia da retina sob midríase medicamentosa (dilatação) em que se utiliza um microscópio com grande aumento possibilitando a visualização de detalhes do olho, que vão da córnea ao nervo óptico. A documentação fotográfica, chamada de retinografia, usando o aparelho retinógrafo mostrado na Figura 2.2 também é importante para a detecção e avaliação da progressão da doença e dos resultados do tratamento [8] e [9].

A retina é composta por dez camadas: membrana limitante interna, camada de fibras nervosas, camada de células ganglionares, plexiforme interna, nuclear interna, plexiforme externa, fotorreceptores (nuclear externa, segmento interno e o externo) e epitélio pigmentado da retina [26]. A camada dos fotorreceptores é composta de células chamadas cones e bastonetes. Essas células estão próximas à superfície externa da retina e a luz, para atingi-la, deve atravessar toda a cavidade vítrea e a retina interna. Após a fotorrecepção, o sinal é conduzido para as células bipolares (camada nuclear interna), que

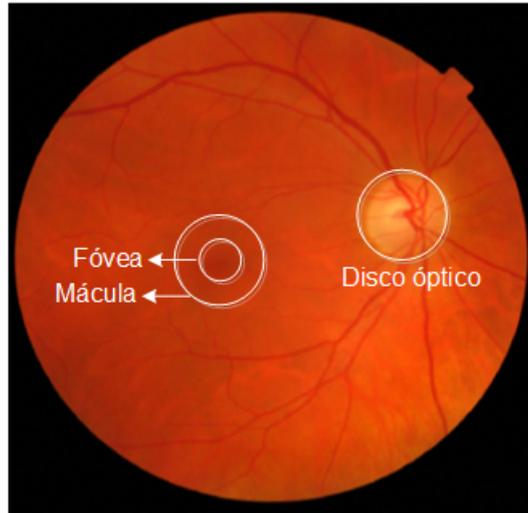


Figura 2.1. Principais partes anatômicas da imagem de retina: disco óptico, mácula e fóvea. Adaptado de [24].

transmitem os sinais para a camada de células ganglionares, cujos axônios se agrupam na superfície interna da retina para formar o nervo óptico [26].

A retinopatia é classificada em duas fases. Na fase inicial, chamada de retinopatia diabética não-proliferativa (RDNP), as modificações encontradas são: microaneurismas, hemorragias, exsudatos duros, manchas algodinosas e edema macular. Os microaneurismas são normalmente o primeiro sinal clínico detectável de RD [3].

Os microaneurismas podem ser visualizados no exame de fundoscopia como pequenos pontos vermelhos localizados nas camadas intermediárias da retina. A sua formação ocorre pelo enfraquecimento focal da parede vascular, sendo que normalmente aumentam em número e tamanho com a progressão da doença.

Com as rupturas dos microaneurismas ocorrem as hemorragias diferenciadas em relação a sua profundidade: hemorragias profundas são normalmente redondas ou ovais e hemorragias superficiais adquirem a forma de chama [27].

Outra alteração decorrente desta fase é a formação de exsudatos duros. Esses consistem em depósitos de lipídeos nas zonas adjacentes a anomalias vasculares, podendo formar um padrão espiralado em volta do vaso envolvido. Os exsudatos duros têm maior significado clínico quando se localizam na mácula uma vez que podem diminuir a acuidade visual do paciente [3].

Na segunda e avançada fase da retinopatia, chamada de retinopatia diabética proliferativa (RDP), ocorre o agravamento da doença com o aparecimento da neovascularização no disco óptico e/ou nos grandes vasos da retina. Esse estágio é uma ameaça para a



Figura 2.2. Retinógrafo responsável pelo exame de fundo de retina. Fonte: [25].

visão e quando não tratado pode levar à formação de hemorragias vítreas. Olhos com RDP também são susceptíveis a neovascularização de íris e glaucoma neovascular. As principais causas de perda de visão na RDP são: descolamento da retina, hemorragia vítrea e glaucoma neovascular [27]. A Fig. 2.3 mostra exemplos de imagens para as duas fases da retinopatia diabética.

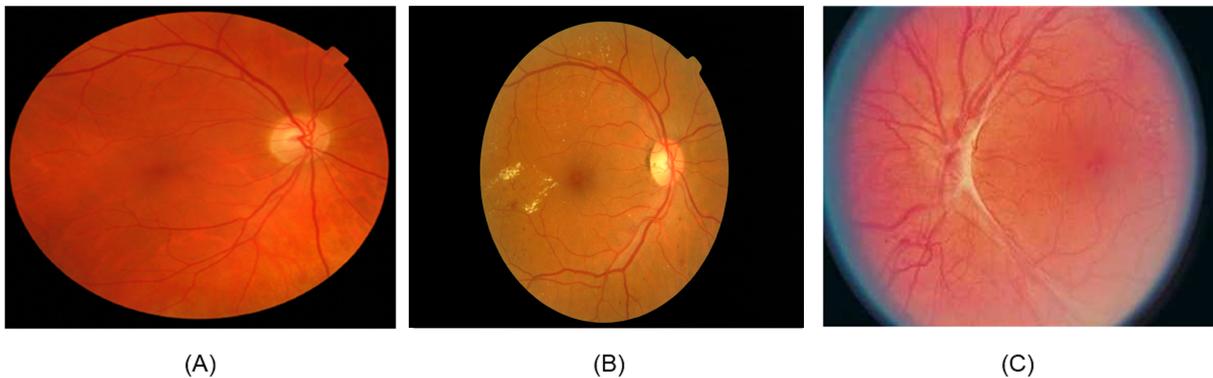


Figura 2.3. Fases Retinopatia Diabética: (A) retina saudável, (B) retina com retinopatia não-proliferativa, (C) retina com retinopatia proliferativa. Fonte: [24].

As imagens digitais da retina podem revelar características patológicas relacionadas com a retinopatia diabética, tais como aneurismas, hemorragias, exsudatos e edema macular. A detecção automática e análise quantitativa auxiliam na verificação de patologias relacionadas à RD [28].

2.2 EXTRAÇÃO DE CARACTERÍSTICAS

Uma das formas de classificação de imagens médicas é realizada por meio de extração de características. Nas imagens de retina, as características mais relevantes são: disco óptico, vasos sanguíneos, mácula e fóvea, como pode ser visto na Figura 2.4. Além disso, em imagens de pacientes com retinopatia a extração deve detectar os microaneurismas, hemorragias, exsudatos duros e moles e neovascularização. A Figura 2.5 apresenta um exemplo de detecção de exsudatos duros em uma imagem de retina.

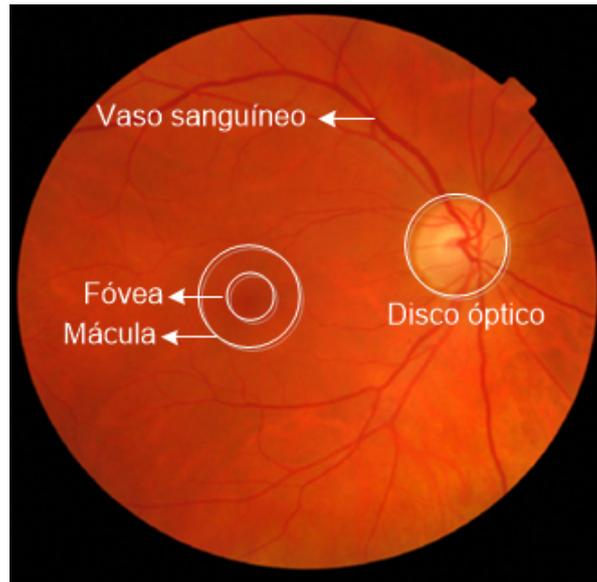


Figura 2.4. Principais características da imagem de retina: disco óptico, mácula, fóvea e vasos sanguíneos. Adaptado de [24].

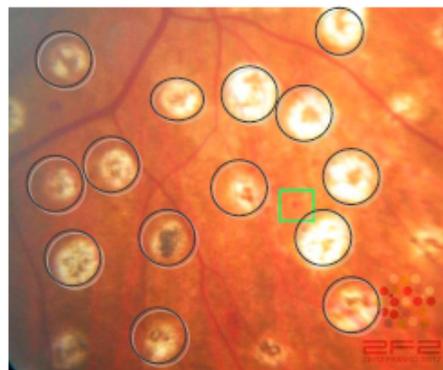


Figura 2.5. Detecção manual de exsudatos duros em imagens retinianas. Fonte: [27].

A extração pode ser feita manualmente, o que requer tempo, habilidade e esforço, ou por meio de algoritmos automáticos. Em trabalhos como [29] e [30] foram desenvolvidos algoritmos para ajudar na tarefa de discriminar vasos sanguíneos em imagens de retina. Esses trabalhos mostram que o processamento de imagens da retina é desafiador. Em primeiro lugar, as imagens da retina possuem baixo contraste, de modo que os vasos não

são facilmente discriminados de seu fundo. Em segundo, a detecção de vasos sanguíneos capilares é difícil em comparação com a detecção de vasos de espessura maior. Esse problema pode levar a um excesso de segmentação, resultando em falsa detecção de vasos. Finalmente, a presença de estruturas patológicas devido a doenças pode degradar o desempenho de algoritmos [29].

Outra forma de extração de características automática é detectar o disco óptico (DO). Rangayyan *et al.* [31] detectam essa estrutura anatômica da retina sem a necessidade de uma discriminação preliminar dos vasos sanguíneos, o que diminui a complexidade do processo. O DO aparece no lado esquerdo ou direito da imagem em uma área aproximadamente circular geralmente com um sexto da largura do diâmetro da imagem. Além disso, é uma região mais brilhante que a área circundante e converge a rede de vasos sanguíneos. Em uma imagem de uma retina saudável, todas as propriedades mencionadas (forma, cor, tamanho e convergência) contribuem para a identificação do DO. Esse trabalho se baseou na transformada de Hough para detecção da DO e obteve taxa de acerto de 92.5% usando os operadores Sobel e de 80% usando método Canny [31].

2.3 CLASSIFICADORES

2.3.1 Máquinas de Vetores de Suporte

As Máquinas de Vetores de Suporte (SVM, do inglês *Support Vector Machine*) constituem numa técnica fundamentada na Teoria de Aprendizado Estatístico visando a proposição de técnicas de aprendizado de máquina que buscam a maximização da capacidade de generalização e a minimização do risco estrutural. A maximização da capacidade de generalização em técnicas de aprendizado de máquina é a capacidade da máquina na classificação eficiente perante o conjunto de treinamento, e a minimização do risco estrutural é a probabilidade de classificação errônea de padrões ainda não apresentados à máquina [32].

Assume-se inicialmente que os dados do domínio em que o aprendizado está ocorrendo são gerados de forma independente e identicamente distribuída de acordo com uma distribuição de probabilidade $P(x, y)$, que descreve a relação entre os dados e os seus rótulos. O erro esperado para dados de teste é

$$R(f) = \int c(f(x), y) dP(x, y), \quad (2.1)$$

em que $c(f(x), y)$ é uma função de custo relacionado à previsão $f(x)$ quando a saída

desejada é y [33].

A classificação de dados é uma tarefa básica em aprendizado de máquina. Parte do princípio que cada ponto dado pertence a uma das duas classes e o objetivo é decidir em qual classe um novo ponto de dados será inserido. Seja T um conjunto de treinamento com n dados $x_i \in X$ e seus respectivos rótulos $y_i \in Y$, em que X constitui o espaço dos dados e $Y = \{-1, +1\}$. O conjunto T é linearmente separável se é possível separar os dados das classes $+1$ e -1 por um hiperplano [21].

Um hiperplano é um plano com uma dimensão a menos que o espaço, ou seja, é um subespaço com uma dimensão a menos que o espaço ambiente (espaço de configuração de ambiente). Existem muitos hiperplanos que podem classificar esses dados, conforme ilustrado na Figura 2.6.

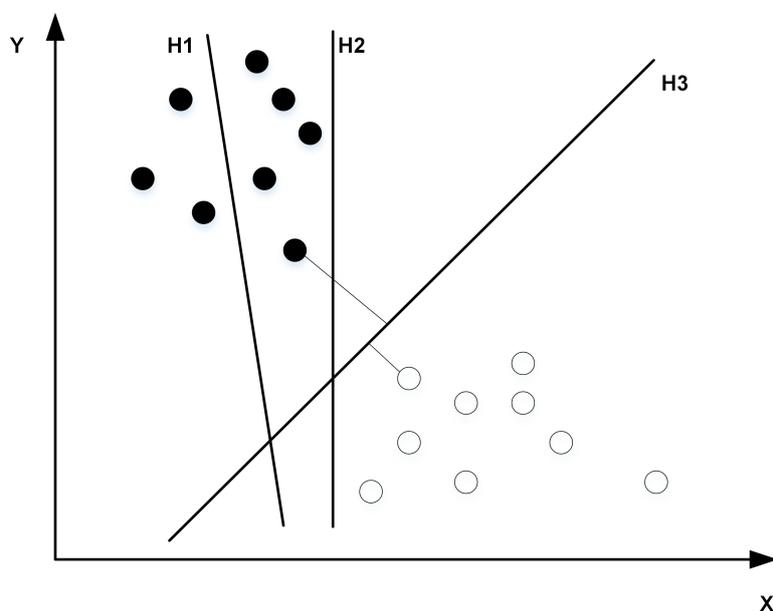


Figura 2.6. Ilustração da separação de classes SVM. Adaptado de: [34].

Assumindo-se que o conjunto de treinamento é linearmente separável, o melhor hiperplano é o que representa a separação mais distante, ou margem, entre as duas classes. Assim, opta-se pelo hiperplano no qual a distância entre os dados mais próximos é maximizada. Se tal hiperplano existe, ele é conhecido como hiperplano de máxima margem, e o classificador linear que o define é conhecido como classificador de margem máxima [35]. O hiperplano de máxima margem é definido como:

$$\langle w \cdot x \rangle + b = 0, \quad (2.2)$$

sendo \mathbf{x} e \mathbf{b} , o vetor peso e o *bias*, respectivamente. O vetor peso w define uma

direção perpendicular ao hiperplano, como mostra a Figura 2.7, e com a variação do *bias* o hiperplano é movido paralelamente a ele mesmo.

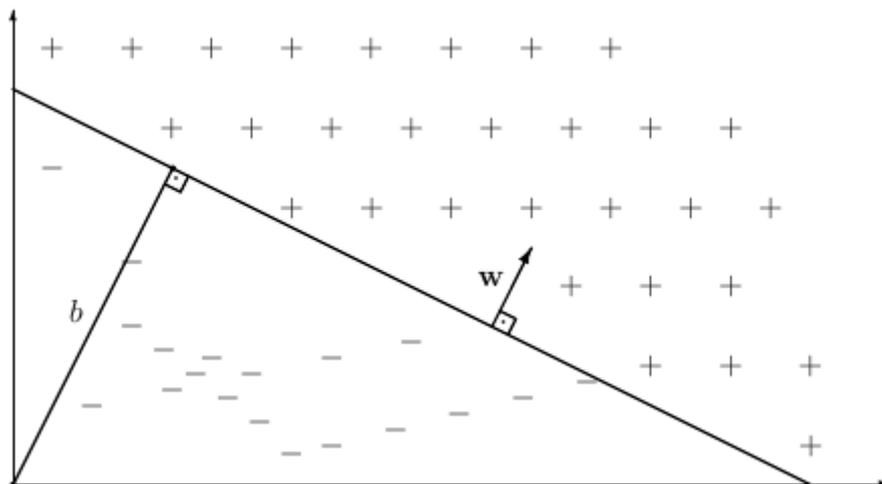


Figura 2.7. Interpretação geométrica de w e b sobre um hiperplano. Fonte: [34].

Considerando as restrições

$$\begin{aligned} \langle w \cdot x_1 \rangle + b &\geq +1, \text{ para } y_i = +1 \text{ e} \\ \langle w \cdot x_1 \rangle + b &\leq -1, \text{ para } y_i = -1, \end{aligned} \quad (2.3)$$

os classificadores lineares que separam um conjunto de treinamento possuem margem positiva. Ou seja, esta restrição afirma que não há nenhum dado entre $\langle w \cdot x \rangle + b = 0$ e $\langle w \cdot x \rangle + b = 1$.

As restrições podem ser combinadas em

$$y_i(\langle w \cdot x_i \rangle + b) \geq 1, \quad i = 1, 2, \dots, n. \quad (2.4)$$

O problema inicial do SVM é como separar em classes determinados padrões existentes no sistema pela comparação das entradas oferecidas pelo usuário (novos dados). Observando a Figura 2.8, pode-se entender melhor essa questão. Na Figura 2.8 observam-se dois padrões, os 'x' e os '*'. O SVM é capaz de identificar o que ele chama de vetores de suporte (*support vectors*), que são as linhas mais próximas das informações contidas no treinamento do sistema [36].

O hiperplano de margem máxima é dado pela minimização da norma $\|w\|$, considerando a restrição da Eq. 2.4. Formalmente pode-se escrever

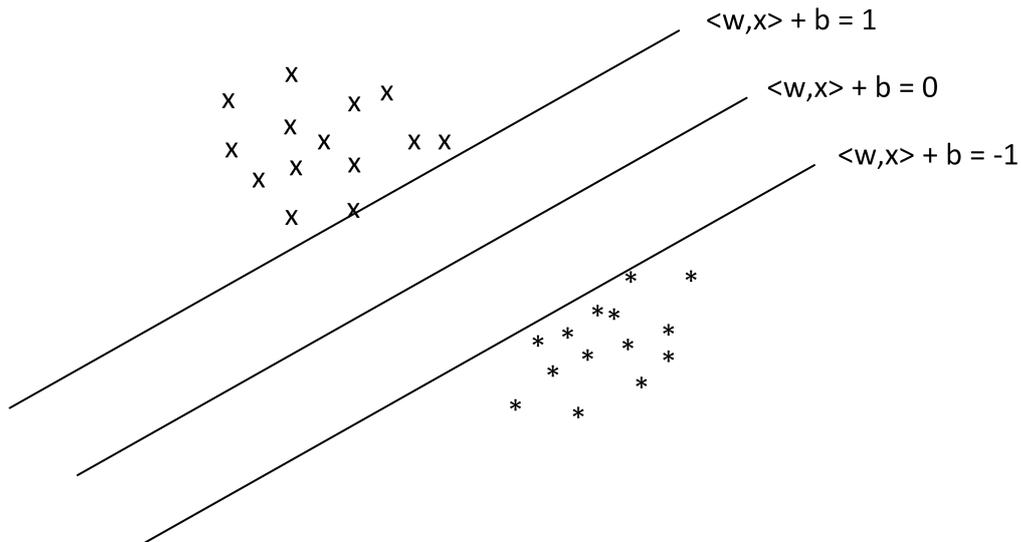


Figura 2.8. Ilustração de uma distribuição dos dados em duas classes perfeitamente separadas. Fonte: [36].

Problema P1

$$\text{Minimizar } \|w\| \quad (2.5)$$

$$\text{Sujeito a } y_i(\langle w \cdot x_i \rangle + b) \geq 1, \text{ para } i = 1, 2, \dots, n.$$

O problema P1 trata-se de um problema clássico de otimização. Ele pode ser resolvido com o método de multiplicadores de *Lagrange*. Utilizando a teoria dos multiplicadores de *Lagrange*, pode-se representar o problema 2.5 como:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i (y_i (\langle w \cdot x_i \rangle + b) - 1), \quad (2.6)$$

em que α_i são os multiplicadores de *Lagrange*. O problema passa a ser a minimização de $L(w, b, \alpha)$, em relação a w e b e a maximização dos α_i . O método dos multiplicadores encontra os pontos de máxima margem igualando as derivadas parciais a zero. Sendo assim, os pontos de margem máxima da Eq. 2.6 são obtidos por meio da resolução das igualdades:

$$\frac{\delta L}{\delta b} = 0 \quad (2.7)$$

e

$$\frac{\delta L}{\delta w} = 0, \quad (2.8)$$

sendo

$$\frac{\delta L}{\delta w} = \left(\frac{\delta L}{\delta w_1}, \frac{\delta L}{\delta w_2}, \dots, \frac{\delta L}{\delta w_n} \right). \quad (2.9)$$

A partir das Equações 2.7 e 2.8 obtém-se

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad (2.10)$$

e

$$w = \sum_{i=1}^n \alpha_i y_i x_i. \quad (2.11)$$

Retomando a equação de um hiperplano como sendo

$$f(x) = \langle w \cdot x \rangle + b = 0, \quad (2.12)$$

em que $w \cdot x$ é o produto escalar entre os vetores w e x , $w \in X$ é um vetor normal ao hiperplano descrito e $\frac{b}{\|w\|}$ corresponde à distância do hiperplano em relação à origem, com $b \in \mathfrak{R}$.

O plano descrito em 2.12 permite dividir o espaço dos dados X em duas distâncias iguais: $w \cdot x + b > 0$ e $w \cdot x + b < 0$. Uma função sinal $g(x) = \text{sgn}(f(x))$ pode então ser empregada para obter as classificações [37], como

$$g(x) = \text{sgn}(f(x)) = +1, \text{ se } w \cdot x + b > 0 \quad (2.13)$$

e

$$g(x) = \text{sgn}(f(x)) = -1, \text{ se } w \cdot x + b < 0. \quad (2.14)$$

Seja x_1 um ponto no hiperplano $H_1 : w \cdot x + b = +1$ e x_2 um ponto no hiperplano $H_2 : w \cdot x + b = -1$, conforme ilustrado na Figura 2.9. Projetando $x_1 - x_2$ na direção de w , perpendicular ao hiperplano separador $w \cdot x + b = 0$, é possível obter a distância entre os hiperplanos H_1 e H_2 como [38].

$$(x_1 - x_2) \left(\frac{w}{\|w\|} \cdot \frac{(x_1 - x_2)}{\|x_1 - x_2\|} \right). \quad (2.15)$$

Portanto, tem-se $w \cdot x_1 + b = +1$ e $w \cdot x_2 + b = -1$. A diferença entre essas equações fornece $w \cdot (x_1 - x_2) = 2$. Substituindo esse resultado na Equação 2.15, tem-se que a

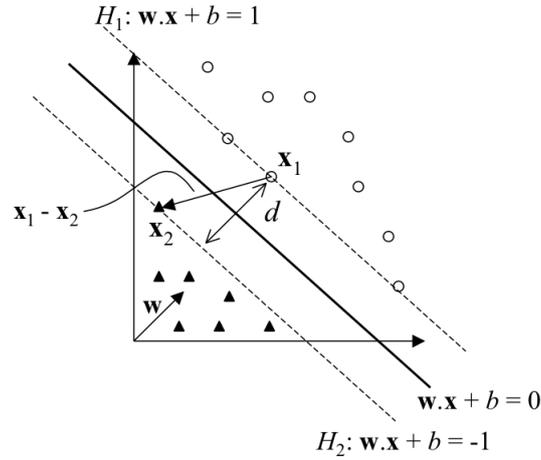


Figura 2.9. Cálculo da distância d entre os hiperplanos H_1 e H_2 . Fonte: [33].

distância entre os dois hiperplanos mais afastados é dada por [39]:

$$\frac{2}{\|w\|} = \frac{2}{\sqrt{w^T w}}. \quad (2.16)$$

Maximizar essa distância equivale a minimizar o inverso da relação à direita, ou seja,

$$\operatorname{argmax} \left(\frac{2}{\|w\|} \right) = \operatorname{argmin} \left(\frac{w^T w}{2} \right). \quad (2.17)$$

O problema 2.17 corresponde à minimização da norma l_2 de w , já que minimizar a norma ao quadrado é o mesmo que minimizar a própria função, dadas as restrições impostas pelo sistema:

$$\frac{w^T w}{2} = \frac{1}{2} \|w\|_2^2 \quad (2.18)$$

e

$$\operatorname{argmin} \left(\frac{w^T w}{2} \right) = \operatorname{argmin} \left(\frac{1}{2} \|w\|_2^2 \right), \quad (2.19)$$

em que a norma $\|w\|_2$ é definida por

$$\|w\|_2 = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}. \quad (2.20)$$

2.3.2 Classificadores do tipo Comitês

Diante de problemas complexos, são comuns as tomadas de decisões através de comitês, em que os membros apresentam suas opiniões, com base em suas experiências e conhecimentos sobre o domínio, e assim contribuem em sua resolução. O mesmo princípio, tão usualmente aplicado no cotidiano, vem sendo amplamente estudado e aplicado para problemas de aprendizado supervisionado. Segundo este princípio, pode-se obter um resultado para um problema de classificação através de comitês de classificadores, denominadas Comitês. Ele é um paradigma de aprendizado em que um grupo finito de propostas alternativas é utilizado para a solução de um dado problema [40].

O método em Comitês, também conhecido como *Ensemble*, tenta construir um conjunto de aprendizes e uma combinação deles. A aprendizagem do tipo Comitês pode ser chamada também aprendizagem baseada em comitês (*committee-based learning*) ou sistema classificador de múltipla aprendizagem (*multiple classifier learning systems*) [16]. A Figura 2.10 mostra uma arquitetura comum com Comitês, no qual o resultado final é uma combinação dos métodos de aprendizado. Os métodos podem ser diversos tipos de algoritmos de aprendizagem como, por exemplo, rede neural ou decisão em árvore. Se os métodos mostrados na Figura 2.10 foram do mesmo tipo então o método do tipo Comitês é chamado de Comitês homogêneos, porém se utilizar diversos tipos diferentes de métodos tem-se Comitês heterogêneos [16].

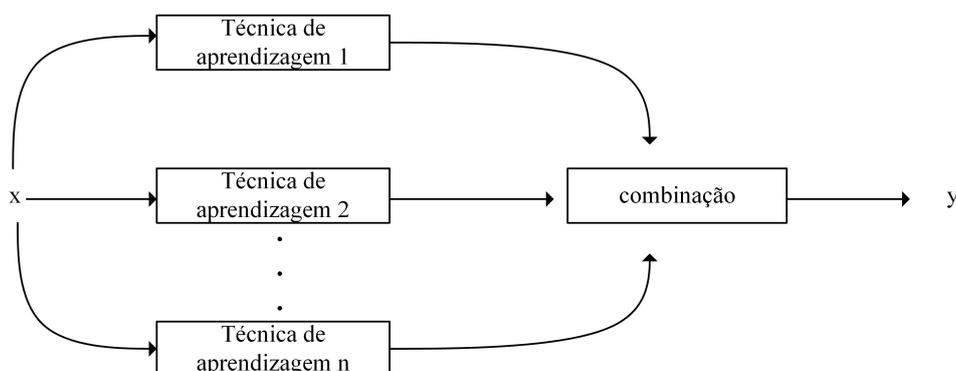


Figura 2.10. Arquitetura comum do tipo Comitês. Adaptado de: [16].

O uso da abordagem com Comitês tem sido bastante explorado na última década, por se tratar de uma técnica capaz de aumentar a capacidade de generalização de soluções baseadas em aprendizado de máquina. No entanto, para que os Comitês sejam capazes de promover melhorias de desempenho, os seus componentes devem apresentar bons desempenhos individuais e, ao mesmo tempo, devem ter comportamentos diversos entre si.

O método *Bagging* (*Bootstrap Aggregating*), proposto por Breiman (1996), é um algo-

ritmo para construção de classificadores agregados. O método consiste em gerar subconjuntos de exemplos através de sorteio simples com reposição, sobre o conjunto de dados de treinamento original. Cada subconjunto amostrado é utilizado para a construção de um novo classificador. A classificação final é realizada por um sistema de votação, em que usualmente se atribui para uma nova instância a classe com maior número de votos entre os classificadores. Uma vez que a classe prevista resulta da combinação das decisões individuais dos classificadores, demonstra-se que as previsões se tornam mais confiáveis à medida que se têm mais votos. Dessa forma, o método *bagging* aposta na expansão da quantidade de classificadores a fim de se reduzir a variância do conjunto, especialmente na presença de dados com ruídos.

O fato de combinar múltiplos classificadores ajuda quando eles apresentam diferenças entre si e quando cada um lida bem com uma parte do conjunto de dados, complementando-se, ao invés de duplicar um ao outro. O método *boosting*, proposto em 1996 por Freund e Schapire [41], explora esse conceito selecionando explicitamente modelos que se complementem. Embora se baseie em reamostragem dos dados do conjunto de treinamento, não adota a reamostragem uniforme. O processo de criação dos subconjuntos de treinamento os condiciona a serem dependentes dos desempenhos individuais de cada classificador.

O AdaBoost é uma implementação do *boosting*, de Freund e Schapire (1996). Foi projetado especificamente para problemas de classificação e pode ser aplicado combinado com qualquer algoritmo de aprendizagem de classificação. Ele trabalha apenas com problemas de classes nominais. O algoritmo inicia atribuindo pesos iguais a todas as instâncias nos dados de treinamento. Em seguida, executa o algoritmo de aprendizagem para gerar um classificador para estes dados e redistribui os pesos para cada instância de acordo com a saída deste classificador. Os pesos de instâncias corretamente classificadas são diminuídos e os pesos dos casos mal classificados são aumentados. Em todas as iterações subsequentes, um classificador é construído para os dados reponderados, concentrando-se, portanto, em classificar corretamente as instâncias erroneamente classificadas pelos classificadores anteriores.

Entre 2001 e 2004, Viola e Jones [42], [43] e [44] propuseram a detecção geral de objetos pela combinação AdaBoost com a arquitetura em cascata. Viola e Jones [44] reportaram que, em uma máquina de 466MHz, a detecção de face gastou apenas 0.067 segundos em uma imagem de 384x288; o que era quase quinze vezes mais rápido do que o estado-da-arte para reconhecimento facial na época. Foi considerado um dos maiores avanços em reconhecimento facial durante a última década.

Huang et al. [45] desenvolveram uma arquitetura do tipo Comitês para reconheci-

mento facial independente da pose da face, rotacionando as imagens das faces. A ideia básica foi utilizando uma combinação de redes neurais. Em contraste com as técnicas que requerem informações de pose, este método não necessita da informação da pose para o reconhecimento de face. *Huang et al.* [45] reportaram que a técnica superou as técnicas convencionais facilitadas pela estimação da pose da face inicial.

2.3.2.1 Algoritmo AdaBoost

O AdaBoost (do inglês, *Adaptive Boosting*) é uma das abordagens usadas no classificador do tipo Comitês para selecionar as características e treinar o classificador. O AdaBoost é o modelo mais utilizado dos algoritmos do *boosting*. Nesta seção é explicado o funcionamento do AdaBoost, bem como da unidade básica que compõe o classificador, denominada classificador fraco (do inglês, *weak learner*).

Um classificador fraco é uma estrutura simples que contém um vetor de características f , um limiar e uma paridade. Durante o treinamento do classificador fraco deve ser encontrado um limiar que melhor separe o valor de uma *feature* de exemplos definidos como positivos dos negativos. A paridade é utilizada para indicar se os objetos positivos devem ser classificados abaixo ou acima do limiar [46]. O classificador fraco, em que, p_j é a paridade, θ_j é o limiar e $f_j(x)$ é o valor da *feature* é

$$h_j(x) = \begin{cases} 1, & \text{se } p_j f_j(x) < p_j \theta_j; \\ 0, & \text{caso contrário.} \end{cases} \quad (2.21)$$

Caso os exemplos positivos sejam classificados abaixo do limiar, a polaridade recebe “1”, caso contrário recebe 0.

O AdaBoost combina um conjunto de classificadores fracos para fazer um classificador forte. Um classificador forte é composto de um conjunto de classificadores fracos, associados a pesos que classificam de forma precisa dois conjuntos de dados pré-rotulados, na qual as características com pesos maiores são mais significativas para a classificação de exemplos definidos como parte de um certo conjunto. A equação

$$H(x) = \alpha_1 h_1 + \alpha_2 h_2 + \dots + \alpha_n h_n(x) \quad (2.22)$$

cria um classificador forte por meio de um algoritmo de *Boosting*, em que $H(x)$ é o classificador forte, α_i é o peso associado ao classificador h_i . Quando se tem uma base de dados de entrada, a função do AdaBoost é encontrar o conjunto de características que

formarão o classificador forte e proverão a melhor classificação do conjunto de entradas, seguindo o algoritmo 1.

Algoritmo 1: Pseudocódigo do AdaBoost. Fonte:[16]

```

1  Obtém N exemplos de imagens  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , em que  $x$  corresponde
   à matriz de pixels de uma imagem, e  $y = 0$  para exemplos negativos e  $y = 1$  para
   exemplos positivos.
2  Inicializa os pesos  $\omega_{1,i} = \frac{1}{2m}$  para  $y_i = 0$  e  $\omega_{1,i} = \frac{1}{2l}$  para  $y_i = 1$ , em que  $m$  é o total
   de exemplos negativos e  $l$  o total de positivos.
3  início
4  |   para  $t = 1, \dots, T$  faça faça
5  |   |   Normaliza os pesos:  $\omega_{t,i} \leftarrow \frac{\omega_{t,i}}{\sum_{j=1}^N \omega_{t,j}}$ , em que  $\omega_t$  é uma distribuição de
   |   |   probabilidade.
6  |   |   para cada feature  $j$  faça faça
7  |   |   |   treine um classificador  $h_j$  restrito ao uso de uma única feature.
8  |   |   |   avalie o erro de acordo com:  $\omega_j, \epsilon_j = \sum_i \omega_i |h_j(x_i) - y_i|$ .
9  |   |   fim
10 |   |   escolha o classificador  $h_t$  com o menor erro  $\epsilon_t$ 
11 |   |   atualize os pesos:  $\omega_{t+1,i} = \omega_{t,i} \beta_t^{1-e_i}$ , em que  $e_i = 0$ , se o exemplo  $x_i$  for
   |   |   classificado corretamente,  $e_i = 1$  caso contrário, e  $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$ .
12 |   fim
13 |   O classificador forte é definido por:  $H(x) = 1$ , se  $\sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t$  e
   |    $H(x) = 0$ , caso contrário. Sendo que  $\alpha_t = \log \frac{1}{\beta_t}$ .
14 fim

```

Para sistemas extensos a tarefa descrita anteriormente demanda um tempo muito grande. Para resolver esse problema, *Viola e Jones* desenvolveram em [47] uma estrutura chamada de árvore de decisão, chamada também de cascata. A ideia principal é que os primeiros estágios da cascata sejam mais genéricos, e ao longo da cascata os classificadores fiquem mais específicos. Deste modo, a avaliação dos exemplos é acelerada, uma vez que grande parte dos exemplos negativos são eliminados nos primeiros estágios da cascata [47].

A estrutura de uma árvore de decisão pode ser observada na Fig. 2.11. O nó raiz são os dados de entrada do sistema, os nós de decisão dividem um atributo e geram ramificações, e por fim os nós folha que apresentam as informações de classificação do algoritmo.

Nesse modelo, cada nó representa um teste e cada percurso na árvore, da raiz até a folha, corresponde a uma regra de classificação. O diferencial desse modelo é fato de uma decisão complexa poder ser decomposta em decisões elementares. Entretanto, o sistema é considerado instável, pois pequenas perturbações do conjunto de treinamento podem provocar grandes alterações na classificação.

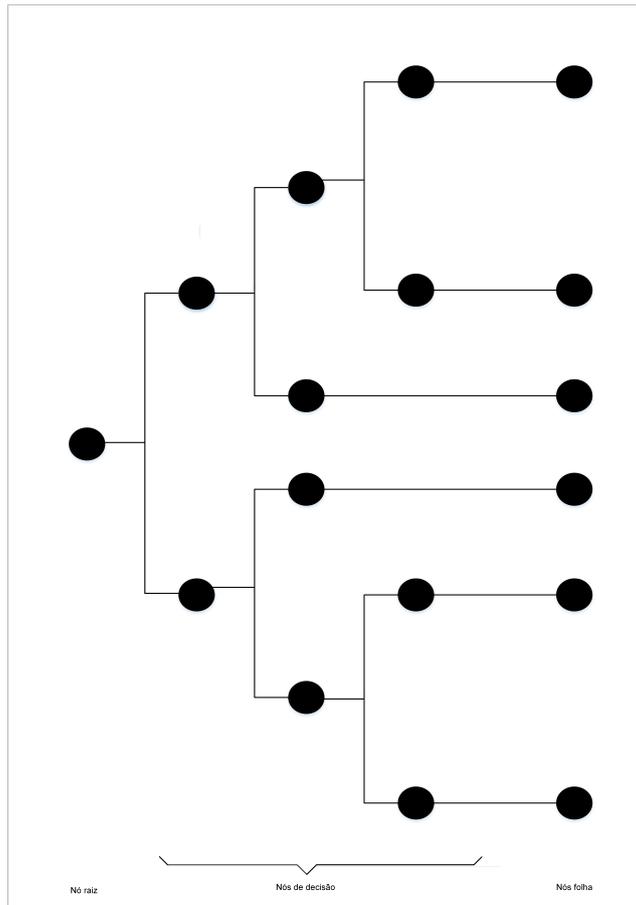


Figura 2.11. Modelo de uma árvore de decisão. Nesse modelo uma decisão complexa (nó raiz) pode ser decomposta em decisões elementares (nós folha). Adaptado de: [48].

Variações do AdaBoost

Após o desenvolvimento do AdaBoost original, muitas variações e extensões deste algoritmo foram propostas e desenvolvidas como alternativas para fornecerem melhores resultados para os mais diversos problemas.

Os problemas binários, ou seja, em que existem somente duas classes para que os exemplos sejam classificados, foram os primeiros a serem pesquisados. Diversas variações para resolver esses problemas foram desenvolvidas, como o Discrete AdaBoost [41], Real Adaboost [49], Gentle AdaBoost [50] e o *Modest AdaBoost* [51].

Em muitas tarefas, porém, o exemplo pode ser classificado em várias classes em vez de apenas duas. Esses problemas, denominados multi-classes, são enfrentados por outras variações do AdaBoost. Pode-se destacar: AdaBoost.MH e AdaBoost.MR, além do AdaBoost.M1 e AdaBoost.M2, que serão brevemente introduzidos.

O *AdaBoost.M1* é considerado o mais simples dentre as versões do *AdaBoost*, e foi sugerido por Yoav Freund e Robert E. Schapire em 1996 [41]. Os autores provaram que caso o erro do classificador fraco seja somente ligeiramente maior que $1/2$, o erro de treinamento da hipótese final $H(x)$ diminui de forma exponencial. O *AdaBoost.M1* segue praticamente o mesmo algoritmo mostrado no Algoritmo 1, com a diferença de que os algoritmos base podem ser classificadores de multi-classes ao invés de serem binários. Este foi o modelo adotado para a realização dos experimentos dessa pesquisa.

O AdaBoost.M2 é similar ao AdaBoost.M1, sendo também capaz de resolver problemas de multi-classes, e foi desenvolvido pelos mesmos cientistas no mesmo ano (1996). A diferença encontra-se na forma como trabalham as multi-classes. No AdaBoost.M1 o peso do classificador fraco é função do erro calculado em relação a todos os exemplos da base. Por outro lado, no AdaBoost.M2 o peso do classificador fraco é em função do erro calculado somente em relação aos exemplos classificados incorretamente. Dessa forma, o objetivo não é focar apenas nos exemplos difíceis de serem classificados como ocorre no AdaBoost.M1, mas sim nas classes utilizadas de forma incorreta durante a classificação e que portanto são difíceis de serem discriminadas [41].

2.3.3 Algoritmo de Aprendizagem por Limiarização Suave

O algoritmo de aprendizagem por limiarização suave (LAST, do inglês *Learning Algorithm for Soft-Thresholding*) foi desenvolvido por Fawzi et al. [19], em 2014. A classificação linear é um método computacionalmente eficiente usado para classificar amostras de teste para encontrar um separador linear entre duas classes.

O LAST está baseado na forma de interagir a extração de características e o classificador simultaneamente, em um único sistema de forma que o classificador esteja atualizado com novas informações relevantes. O sistema implementado treina um dicionário a partir de um banco de imagens, que extrai as características, e um classificador que usa esse dicionário simultaneamente. Dessa forma, é possível extrair um conjunto de características para a classificação desejada junto com o treinamento do classificador.

Antes da explicação detalhada de como funciona o algoritmo do LAST, se faz necessário uma explicação breve do que são a limiarização suave (*Soft-Thresholding*) e classificação linear. Para qualquer $\alpha \in \mathbb{R}$, define-se a limiarização suave $h_\alpha : \mathbb{R} \rightarrow \mathbb{R}$ como

$$h_\alpha(z) = \max(0, z - \alpha). \quad (2.23)$$

Assim, h_α pode ser estendido em vetores \mathbf{z} aplicando a transformada linear de coordenadas independentes. Esta transformação pode ser aplicada por $z = D^T x$ que representa os coeficientes das características no sinal \mathbf{x} . As características mais importantes de \mathbf{x} são usadas na classificação. Nesse caso, a classificação leva em conta, resumidamente, os seguintes procedimentos:

1. Extração de características: para $D \in \mathbb{R}^{n \times N}$ e $\alpha \in \mathbb{R}$ dado um ponto de teste $x \in \mathbb{R}^n$, calcula-se $h_\alpha(D^T x)$.
2. Classificação linear: para $w \in \mathbb{R}^N$. Se $w^T h_\alpha(D^T x)$ é positivo, então \mathbf{x} é da classe 1. Se negativo, \mathbf{x} é da classe -1.

A vantagem do LAST [19] é a capacidade de aprender ao mesmo tempo o dicionário de esparsificação (\mathbf{D}) e o classificador linear (\mathbf{w}). Para os sinais do conjunto de treinamento (X, y) , em que $X = [x_i]_{i=1}^m$ é o conjunto de sinais com cada sinal $x_i \in \mathbb{R}^n$ e $y = [y_i]_{i=1}^m$ são as respectivas classes de \mathbf{x} , com cada classe $y_i \in \{-1, 1\}$, o dicionário esparsificante \mathbf{D} , que contém d átomos, e o hiperplano classificador \mathbf{w} são estimados usando a otimização supervisionada

$$\arg \min_{\mathbf{D}, \mathbf{w}} \sum_{i=1}^m L(y_i \mathbf{w}^T h_\alpha(\mathbf{D}^T x_i)) + \frac{v}{2} \|\mathbf{w}\|_2^2, \quad (2.24)$$

em que, L é a função de perda $L(x) = \max(0, 1 - x)$ e v é o parâmetro de regularização que previne o classificador (\mathbf{w}) de se sobreajustar aos dados do conjunto de treinamento. Quando se necessita realizar o teste, a saída do classificador para um sinal \mathbf{x} normalizado

é a classe retornada pelo classificador; sendo $+1$ se $\mathbf{w}^T \max(0, \mathbf{D}^T x - \alpha) > 0$ ou -1 , caso isso não ocorra. Essa classificação pode ser visualizada na Figura 2.12, onde o sinal em tempo de teste \mathbf{x} normalizado (norma l_2 igual a 1) é primeiramente transformado de forma linear por meio da matriz D , resultando no vetor de características $f = D^T x$. Em seguida, o vetor de características f é transformado de forma não-linear por meio de limiarização suave, resultando no vetor de características esparsas $f_\alpha = \max(0, f - \alpha)$. Finalmente, esse vetor de características esparsas f_α é classificado pelo vetor hiperplano w , resultando no valor escalar $c = w^T f_\alpha$. A classe resultado do classificador será 1 se $c > 0$ e -1 , caso contrário [52].

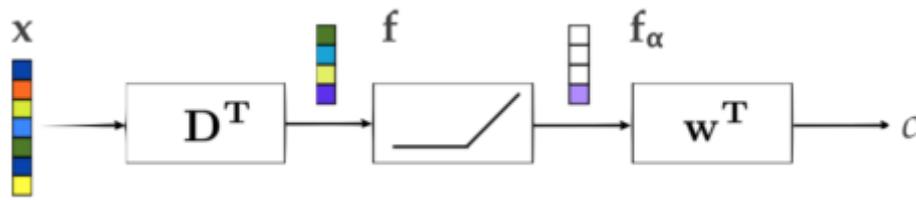


Figura 2.12. Classificação em tempo de teste realizados pelo LAST. Adaptado de [19].

Os sinais de treinamento devem ser normalizados para ter norma l_2 igual a 1 antes do treinamento. Isso é necessário, para evitar instabilidade numérica nos cálculos realizados durante a fase de aprendizagem. Além disso, como \mathbf{D} e \mathbf{w} são treinados com sinais normalizados, os sinais em tempo de teste também deve ser normalizados.

O algoritmo 2 mostra os procedimentos para implementar o classificador LAST em tempo de execução.

2.4 ANÁLISE DE COMPONENTES PRINCIPAIS

A Análise de Componentes Principais (PCA, do inglês *Principal Component Analysis*) é uma técnica que visa a analisar os dados com a intenção de reduzir, eliminar sobreposições e escolher formas mais representativas de dados a partir de combinações lineares das variáveis originais. Ela é bastante útil quando se deseja identificar relações entre as informações extraídas de dados e entre vetores de características com muitas dimensões. Um dos principais objetivos dessa técnica é a diminuição de dimensionalidade com o máximo aproveitamento de informações dos dados originais. Matematicamente, isso representa uma transformação linear dos dados para um novo conjunto de variáveis dos dados, chamados de componentes principais (CPs), que não estão correlacionados [53]. Ela inclui cálculos de autovalores e autovetores da matriz de covariância. Os autovetores são ortogonais entre si para formar o sistema de vetores e são classificados de forma decrescente na ordem de seus respectivos autovalores. O autovetor correspondente ao maior

Algoritmo 2: Algoritmo LAST

```
1 início LAST
2 // Carrega D, w, X
3 D: dicionário D
4 X: imagem
5 w: vetor classificador

6 // Extração de características
7 aux ← D[i][k] * X[k][j]

8 // Aplica limiar: max(0, d * x - α)
9 fα ← aux - α
10 if fα < 0 then
11     fα ← 0

12 // Aplica a classificação
13 c ← w * fα
14 if c ≥ 0 then
15     labels ← 1
16 else
17     labels ← -1
```

autovalor será o primeiro componente principal, o autovetor correspondente ao segundo maior autovalor será o segundo componente principal, e assim sucessivamente.

A meta é ter p variáveis X_1, X_2, \dots, X_p , como na Tabela 2.2, e encontrar combinações dessas para produzir índices Z_1, Z_2, \dots, Z_p , que sejam não correlacionados na ordem de sua importância, e que descrevam a variação de dados. A ordem é tal que $Var(Z_1) \geq Var(Z_2) \geq \dots \geq Var(Z_p)$, em que $Var(Z_1)$ denota a variância de (Z_1) , nos quais os índices Z são os componentes principais.

Tabela 2.2. Exemplo de dados para Análise de Componentes Principais.

<i>Caso</i>	X_1	X_2	\dots	X_p
1	a_{11}	a_{21}	\dots	a_{1p}
2	a_{21}	a_{22}	\dots	a_{2p}
\vdots	\vdots	\vdots	\dots	\vdots
n	a_{n1}	a_{n2}	\dots	a_{np}

Dadas as variáveis como indicado na Tabela 2.2, o primeiro componente principal dado por

$$Z_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p, \quad (2.25)$$

é a combinação linear das variáveis X_1, X_2, \dots, X_p e varia tanto quanto possível. A variância de Z_1 , $Var(Z_1)$, tem uma variância maior que qualquer variância entre os X_s . Além disso, precisa obedecer a restrição da equação

$$a_{11}^2 + a_{12}^2 + \dots + a_{1p}^2 = 1. \quad (2.26)$$

Seguindo a mesma ideia, o segundo componente principal será

$$Z_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p, \quad (2.27)$$

com variância de Z_2 , $Var(Z_2)$, sendo a maior possível e com a restrição da equação

$$a_{21}^2 + a_{22}^2 + \dots + a_{2p}^2 = 1. \quad (2.28)$$

Os componentes posteriores são definidos da mesma maneira.

Outro fato importante é de que Z_1, Z_2, \dots, Z_p tenham correlação zero para os dados. Se existem p variáveis, então existirão no máximo p componentes principais. Portanto, a análise de componentes principais se resume em encontrar os autovalores de uma matriz de covariâncias amostral

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \dots & c_{pp} \end{pmatrix}, \quad (2.29)$$

em que o elemento c_{ii} na diagonal é a variância de X_i e os termos fora da diagonal c_{ij} é a covariância entre as variáveis $X_i X_j$.

As variâncias dos componentes principais são os autovalores da matriz \mathbf{C} , sendo que alguns podem ser zero. Assumindo que os autovalores estão ordenados como:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0, \quad (2.30)$$

então λ_i corresponde ao i -ésimo componente principal sendo:

$$Z_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p. \quad (2.31)$$

Outra propriedade importante do PCA é que a soma dos autovalores é igual à soma dos elementos da diagonal (traço) da matriz \mathbf{C} dado por:

$$\lambda_1 + \lambda_2 + \cdots + \lambda_p = c_{11} + c_{22} + \cdots + c_{pp}. \quad (2.32)$$

Isso acontece porque c_{ii} é a variância de X_i e λ_i é a variância de Z_i , o que significa que a soma das variâncias dos CPs é igual à soma das variâncias das variáveis originais.

Uma das formas de se aplicarem os princípios do PCA é utilizando a Decomposição em Valores Singulares (SVD, do inglês *Singular Value Decomposition*). Esse método consiste em fatorar uma matriz qualquer em um produto $U\Sigma V^T$ [54]. Foi demonstrado pela primeira vez por Beltrami e Jordan na década de 1870, para matrizes quadradas reais. Mais tarde, em 1915, Autonne demonstrou a decomposição SVD para matrizes quadradas complexas. Em 1936, Carl Eckart e Gale Young demonstraram para matrizes retangulares gerais [55]. Em relação aos principais métodos de cálculo, o primeiro deles, baseado nas rotações de Givens, surgiu com o matemático Ervand G. Kogbetliantz em 1958 [56]. Em 1965 foi publicado por Golub e William, uma forma alternativa baseada nas reflexões Householder [57]. Finalmente, em 1970 foi publicado por Gene Howard Golub e Christian Reinsch uma atualização do algoritmo de Golub/Kahan que é o método mais utilizado atualmente [58].

Por definição, a SVD é uma técnica de fatoração de matrizes que consiste em representar qualquer matriz $A_{m \times n}$ na forma:

$$A_{m \times n} = U_{m \times m} \begin{pmatrix} \Sigma_{n \times n} \\ 0_{(m-n) \times n} \end{pmatrix} V_{n \times n}^T, \quad (2.33)$$

em que $U_{m \times m}$ e $V_{n \times n}^T$ são matrizes ortogonais, ou seja, $UU^T = I_m$ e $VV^T = I_n$ e $\Sigma_{n \times n}$ é uma matriz diagonal que contém os valores singulares σ_j em ordem decrescente $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_j \geq 0$ para $1 \leq j \leq \min(m, n)$. A decomposição em valores singulares de uma matriz A é dada pelo Teorema 2.4.1 [54].

Teorema 2.4.1 *Qualquer matriz $A_{m \times n}$ possui uma decomposição em valores singulares da forma $A = U\Sigma V^T$. Além disso, os seus valores singulares são sempre únicos. Porém, os vetores singulares à direita e os vetores singulares à esquerda somente serão únicos se a matriz A for quadrada e possuir autovalores distintos.*

Resumidamente, o Teorema 2.4.1 pode ser explicado da seguinte forma. Suponha que $\{v_1, v_2, \dots, v_n\}$ é uma base ortogonal para \mathbb{R}^n consistindo em autovetores de $A^T A$

ordenados de tal forma que os autovalores associados satisfaçam $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Considera-se que Av_1, Av_2, \dots, Av_r seja uma base ortogonal. Para normalizar cada Av_i e obter uma base ortogonal $\{u_1, u_2, \dots, u_r\}$, tem-se

$$u_i = \frac{1}{\|Av_i\|} Av_i = \frac{1}{\sigma_i} Av_i. \quad (2.34)$$

Portanto,

$$Av_i = \sigma_i u_i \quad (1 \leq i \leq r). \quad (2.35)$$

Agora estende-se $\{u_1, u_2, \dots, u_r\}$ a uma base ortogonal $\{u_1, u_2, \dots, u_m\}$ de \mathbb{R}^m e define-se $U = [u_1, u_2, \dots, u_m]$ e $V = [v_1, v_2, \dots, v_n]$. As matrizes U e V são ortogonais. Além disso, a partir da Eq. 2.35, tem-se a igualdade da equação

$$AV = [Av_1 \dots Av_r \ 0 \dots 0] = [\sigma_1 u_1 \dots \sigma_r u_r \ 0 \dots 0]. \quad (2.36)$$

Supondo uma matriz diagonal D cujos elementos diagonais são $\sigma_1, \sigma_2, \dots, \sigma_r$. Então $U\Sigma$ será

$$U\Sigma = [u_1 u_2 \dots u_m] \begin{pmatrix} \sigma_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \ddots & 0 & \dots & 0 \\ 0 & 0 & 0 & \sigma_r & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix} = [\sigma_1 u_1, \dots, \sigma_r u_r, 0, \dots, 0] = AV. \quad (2.37)$$

Como V é uma matriz ortogonal, $U\Sigma V^T = AVV^T = A$. As colunas das matrizes U e V são os vetores singulares esquerdos e os vetores singulares direitos de A , respectivamente [59].

2.5 TRANSFORMADA DE FOURIER

A Transformada de Fourier (FT, do inglês *Fourier Transform*), desenvolvida pelo matemático francês Jean Baptiste Joseph Fourier, que viveu de 1768 a 1830, é uma ferramenta

matemática fundamental na solução de problemas de processamento de imagens digitais. Esta transformada é uma das maneiras de realizar uma mudança de domínio espacial (x,y) para espectral, no qual um sinal fica caracterizado por suas componentes de frequência (u,v) . Na FT, não há perda de informação durante a mudança de domínios, apenas a informação visual da imagem está representada de uma outra forma, no domínio da frequência [60].

A FT de uma função contínua $f(x)$ de uma variável real x é definida como sendo

$$F(u) = \int_{-\infty}^{+\infty} f(x) \exp[-j2\pi ux] dx, j = \sqrt{-1}, \quad (2.38)$$

e a partir de $F(u)$, pode-se obter $f(x)$ por meio da transformada inversa de Fourier

$$f(x) = \int_{-\infty}^{+\infty} F(u) \exp[j2\pi ux] du. \quad (2.39)$$

As equações 2.38 e 2.39 são chamadas de par de transformada de Fourier e podem existir se forem absolutamente integráveis e se $f(x)$ em tempo for contínuo. A FT de uma $f(x)$ é uma função complexa e pode ser expressa pela soma de suas componentes real e imaginária, representadas por R e I , respectivamente, ou seja,

$$F(u) = R(u) + jI(u). \quad (2.40)$$

Para obter-se o espectro de Fourier (módulo) e o ângulo de fase da função $F(x)$, faz-se necessário expressar a equação 2.40 na forma exponencial

$$F(u) = |F(u)|e^{j\theta(u)}. \quad (2.41)$$

Com isso define-se o módulo de Fourier como

$$|F(u)| = [R^2(u) + I^2(u)]^{1/2} \quad (2.42)$$

e o ângulo de fase

$$\phi(u) = \tan^{-1}[I(u)/R(u)]. \quad (2.43)$$

A figura 2.13, mostra os espectro de Fourier de algumas imagens simples: um círculo, um segmento de reta vertical e quatro segmentos de reta. Um ponto no espaço, que foi representado em todo o domínio da frequência, pode conter informações sobre toda a imagem no domínio espacial, indicando quanto desta frequência há na imagem.

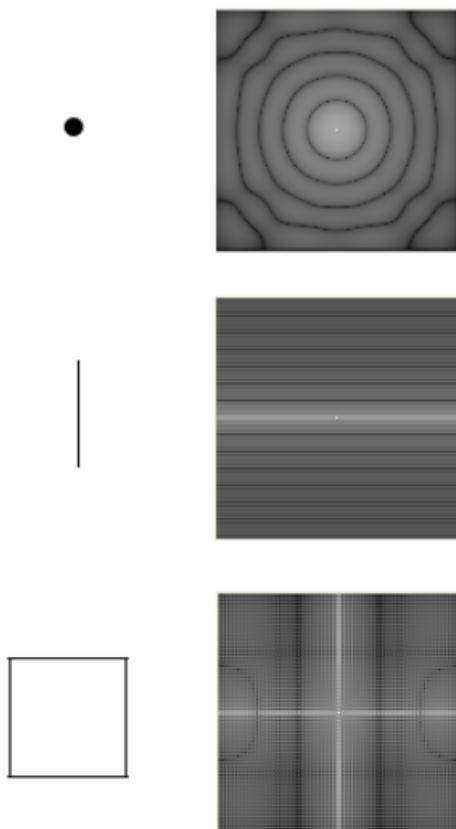


Figura 2.13. Algumas funções bidimensionais e seus espectros de Fourier.
Fonte: [60].

A variável u que está presente na transformada de Fourier é chamada de variável de frequência derivada do termo exponencial, $\exp[-j2\pi ux]$ e pela fórmula de Euler obtêm-se

$$\exp[-j2\pi ux] = \cos 2\pi ux - j \sin 2\pi ux. \quad (2.44)$$

Portanto, pode ser decomposta pelo somatório de senos e cossenos e a FT informa a distribuição de amplitudes, frequências e fases desses senos e cossenos [60].

A FT de uma função unidimensional (1D), em que a entrada é um vetor, pode ser estendida para uma função bidimensional (2D) $f(x,y)$, na qual a entrada é uma matriz. Em processamento de imagens é comum utilizar transformadas bidimensionais. E assim, tem-se o par de transformadas de Fourier

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \exp[-j2\pi(ux + vy)] dx dy \quad (2.45)$$

e a partir de $F(u, v)$, pode-se obter $f(x, y)$ através da transformada inversa de Fourier

$$F(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u, v) \exp[j2\pi(ux + vy)] dudv. \quad (2.46)$$

Assim como acontece na transformada unidimensional, a transformada de Fourier bidimensional de uma função $f(x, y)$ é uma função complexa e pode ser expressa pela soma de componentes real e imaginário, representados por R e I , respectivamente, ou seja,

$$F(u, v) = R(u, v) + jI(u, v). \quad (2.47)$$

Como ocorre no caso unidimensional, o espectro de Fourier e o ângulo fase podem ser representados, respectivamente, conforme as equações:

$$|F(u, v)| = [R^2(u, v) + I^2(u, v)]^{1/2} \quad (2.48)$$

e

$$\phi(u, v) = \tan^{-1}[I(u, v)/R(u, v)]. \quad (2.49)$$

A fase contém as informações essenciais sobre a estrutura da imagem, e o espectro de Fourier mostra o que a estrutura contém, mas sem nenhuma referência espacial [61].

A transformada de Fourier pode ser aproximada usando a Transformada Discreta de Fourier (DFT). Ela utiliza um número finito de pontos no domínio do tempo e define uma representação discreta do sinal no domínio da frequência. Uma função contínua pode assumir a seguinte sequência discreta: $\{f(x_0), f(x_0 + \Delta x), f(x_0 + 2\Delta x), \dots, f(x_0 + [N - 1]\Delta x)\}$, com um número finito de pontos.

Nesse caso, N são amostras separadas em intervalos de Δx unidades. Para usar x como uma variável discreta, defini-se: $f(x) = f(x_0 + n\Delta x)$, com n assumindo os valores discretos $(0, 1, 2, \dots, N - 1)$.

Com isso, pode-se denotar amostragens de N valores com intervalos uniformemente espaçados de uma função contínua através da sequência: $\{f(0), f(1), f(2), \dots, f(N - 1)\}$. Com base na transformada de Fourier de funções contínuas e a partir desta “discretização”

de x , pode-se definir o par de transformadas discretas de Fourier como sendo uma soma finita de exponenciais complexas [32]. Considerando o domínio da frequência também discreto e representado por $u = (0, \Delta u, 2\Delta u, \dots, (N-1)\Delta u)$, em que $\Delta = 1/N\Delta x$. A transformada discreta é obtida como sendo

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \exp[-j2\pi ux/N] \quad (2.50)$$

e sua inversa por

$$f(x) = \frac{1}{N} \sum_{u=0}^{N-1} F(u) \exp[j2\pi ux/N]. \quad (2.51)$$

Considerando x uma “discretização”, para $x = (0, 1, 2, \dots, N-1)$ e aplicando estas equações para uma função bidimensional, o par de transformadas discretas de Fourier passa a ser:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp \left[-j2\pi \left(\frac{ux}{M} + \frac{vy}{N} \right) \right] \quad (2.52)$$

para u e v , “discretizados” com $u = (0, 1, 2, \dots, M-1)$ e $v = (0, 1, 2, \dots, N-1)$ e a inversa

$$F(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \exp \left[j2\pi \left(\frac{ux}{M} + \frac{vy}{N} \right) \right] \quad (2.53)$$

para x e y assumindo valores discretos, isto é: $x = (0, 1, 2, \dots, M-1)$ e $y = (0, 1, 2, \dots, N-1)$. Em que

$$\Delta u = \frac{1}{M\Delta x} \quad (2.54)$$

e

$$\Delta v = \frac{1}{N\Delta y}. \quad (2.55)$$

Na prática, para diminuir o esforço computacional, em aplicações de processamento de imagens, as transformadas discretas de Fourier são calculadas utilizando o algoritmo da transformada rápida de Fourier (FFT, do inglês *Fast Fourier Transform*) que devido as suas características de implementação fazem com que a complexidade caia de N^2 para $N \log_2 N$ operações. Representando assim, uma significativa economia computacional, particularmente quando o valor de N é muito grande.

2.6 AVALIAÇÃO DE DESEMPENHO PARA CLASSIFICADORES

2.6.1 Métricas de desempenho

Para avaliar o desempenho de classificação, há algumas medidas de desempenho que podem ser usadas para comparar o comportamento dos classificadores como taxa de erro, acurácia, precisão, sensibilidade e Medida-F. Nessa seção é apresentada a definição de cada métrica.

A taxa de erro de um classificador h é

$$err(h) = \frac{1}{n} \sum_{i=1}^n ||y_i \neq h(x_i)||, \quad (2.56)$$

em que n é o número de exemplos, y_i é a classe verdadeira do exemplo i e $h(x_i)$ é a classe dada pelo classificador h para o exemplo i . Essa medida compara a classe verdadeira de cada exemplo com a classe atribuída pelo classificador h . Se as duas classes forem iguais, ou seja, se $y_i = h(x_i)$, então $||y_i \neq h(x_i)|| = 0$; caso contrário, $||y_i \neq h(x_i)|| = 1$.

A acurácia ou taxa de acerto é denotada por $acc(h)$ e corresponde ao complemento da taxa de erro

$$acc(h) = 1 - err(h). \quad (2.57)$$

As taxas de erro e acerto também podem ser obtidas por meio de uma matriz de confusão. Ela possui dimensão correspondente ao número de classes existentes em um determinado conjunto de exemplos. A sua diagonal principal corresponde ao número de acertos de cada classe e os elementos fora da diagonal principal correspondem ao número de erros. Na Tabela 2.3 é mostrado um exemplo de uma matriz de confusão referente a um conjunto de exemplos com duas classes geralmente denominadas como positiva e negativa.

Tabela 2.3. Exemplo de uma matriz de confusão.

	Predição Positiva	Predição Negativa
Classe Positiva	Verdadeiro Positivo(VP)	Falso Negativo(FN)
Classe Negativa	Falso Positivo(FP)	Verdadeiro Negativo(VN)

Nesta tabela, verdadeiros positivos corresponde ao exemplo que é positivo e foi clas-

sificado como positivo; falsos positivos são os exemplos negativos que foram classificados como positivos; verdadeiros negativos são os exemplos negativos e que foram classificados como negativos e falsos negativos são os exemplos positivos que foram classificados como negativos [62].

A partir da matriz de confusão, pode-se então obter-se a taxa de erro e a taxa de acerto:

$$err(h) = \frac{FN + FP}{VP + FN + FP + VN} \quad (2.58)$$

e

$$acc(h) = \frac{VP + VN}{VP + FN + FP + VN}, \quad (2.59)$$

respectivamente.

Outras medidas de desempenho que podem ser calculadas a partir da matriz de confusão são: precisão(P), sensibilidade(S) e medida-F. A precisão:

$$P = \frac{VP}{VP + FP}, \quad (2.60)$$

é a taxa com que todos os exemplos classificados como positivos são realmente positivos; nenhum exemplo negativo é incluído.

A sensibilidade:

$$S = \frac{VP}{VP + FN}, \quad (2.61)$$

é a taxa com que classifica como positivos todos os exemplos que são positivos; nenhum exemplo positivo é deixado de fora

A partir do cálculo da precisão e sensibilidade é possível encontrar mais uma métrica de desempenho, a medida-F, derivada em 1979 por *Van Rijsbergen* [63]. Medida-F é a média harmônica ponderada da precisão e sensibilidade dada por:

$$F_\beta = \frac{(1 + \beta) \times (P \times S)}{\beta \times P + S}, \quad (2.62)$$

em que

$$\beta = \frac{1 - \alpha}{\alpha}. \quad (2.63)$$

A Medida F_β mede a eficácia da recuperação em relação ao valor atribuído a β . Pesos comumente utilizados para β são: F_2 (sensibilidade é o dobro da precisão) e $F_{0.5}$ (precisão é o dobro da sensibilidade). A precisão tem peso maior para valores $\beta < 1$, enquanto que $\beta > 1$ favorece a sensibilidade. *Van Rijsbergen* baseou-se na medida de eficiência (*effectiveness*)

$$E = 1 - \left(\frac{1}{\left(\frac{\alpha}{P} + \frac{(1-\alpha)}{R} \right)} \right), \quad (2.64)$$

para mostrar que a relação entre a F_α e a medida de eficiência é: $F_\alpha = 1 - E$. Quando a precisão e a sensibilidade têm o mesmo peso ($\beta = 1$) tem-se a Medida-F:

$$\text{Medida F} = \frac{2 \times P \times S}{P + S}. \quad (2.65)$$

2.6.2 Plano ROC

Uma alternativa à avaliação utilizando medidas é o uso de gráficos e/ou diagramas. Uma das medidas que pode-se utilizar é o gráfico de sensibilidade (taxa de verdadeiros positivos) versus taxa de falsos positivos (1-especificidade), conhecido como curva das características de operação do receptor (ROC, do inglês *Receiver Operating Characteristics*). A especificidade do classificador é a razão entre os verdadeiros negativos e o total de exemplos negativos, ou seja,

$$\text{especificidade} = \frac{VN}{FP + VN}. \quad (2.66)$$

Alguns pontos no espaço ROC merecem destaque. O ponto (0,0) representa a estratégia de nunca classificar um exemplo como positivo. Modelos que correspondem a esse ponto não apresentam nenhum falso positivo, mas também não conseguem classificar nenhum verdadeiro positivo. A estratégia inversa, de sempre classificar um novo exemplo como positivo, é representada pelo ponto (100%, 100%). O ponto (0, 100%) representa o modelo perfeito com todos os exemplos positivos e negativos são corretamente classificados. O ponto (100%, 0) representa o modelo que sempre faz previsões erradas. Modelos próximos ao canto inferior esquerdo podem ser considerados "conservadores": eles fazem uma classificação positiva somente se têm grande segurança na classificação. Como consequência, eles cometem poucos erros falsos positivos, mas frequentemente têm baixas taxas de verdadeiros positivos. Modelos próximos ao canto superior direito podem

ser considerados “liberais”: eles predizem a classe positiva com maior frequência, de tal maneira que classificam a maioria dos exemplos positivos corretamente, mas, geralmente, com altas taxas de falsos positivos [64].

A linha diagonal ascendente $(0, 0) - (100\%, 100\%)$ representa um modelo de comportamento estocástico: cada ponto (p, p) pode ser obtido pela previsão da classe positiva com probabilidade p e da classe negativa com probabilidade $100\% - p$. Pontos pertencentes ao triângulo superior esquerdo a essa diagonal representam modelos que desempenham melhor que o aleatório e pontos pertencentes ao triângulo inferior direito representam modelos piores que o aleatório. A diagonal descendente $(0, 100\%) - (100\%, 0)$ representa classificadores que desempenham igualmente em ambas as classes. À esquerda dessa linha, estão os modelos que desempenham melhor para a classe negativa em detrimento da positiva e, à direita, os modelos que desempenham melhor para a classe positiva. O gráfico da Figura 2.14 representa o espaço ROC [64].

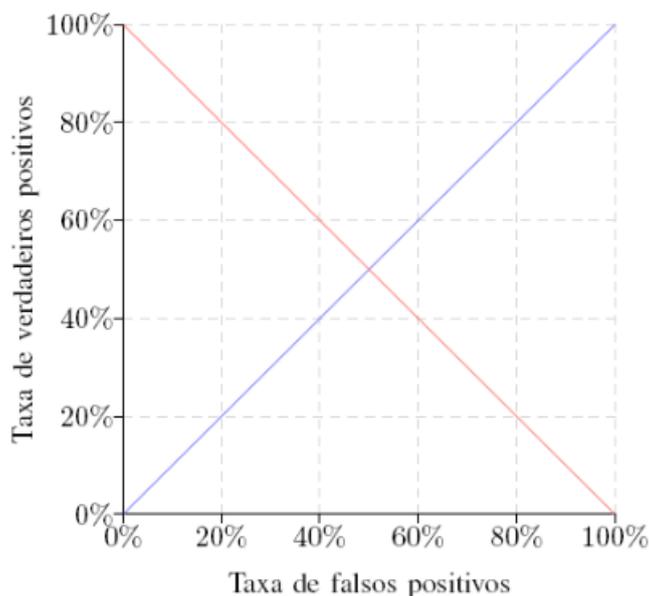


Figura 2.14. Espaço ROC. Adaptado de: [64].

Na Figura 2.15 é mostrado um gráfico ROC com 5 pontos arbitrários representando 5 modelos de classificação diferentes (A, B, C, D, E), para um conjunto de dados também arbitrário. Neste caso hipotético, A é o mais conservador (aceita poucos falsos negativos, mas consequentemente penaliza bastante o desempenho do verdadeiros positivos) e D é o mais liberal (não se importa muito em aceitar bastante falsos positivos, porém seu desempenho nos verdadeiros positivos é muito bom). Além disso, percebe-se que um ponto no espaço ROC é melhor que outro se e somente se ele está acima e à esquerda do outro ponto, porque tem uma maior taxa de verdadeiros positivos e uma menor taxa de falsos positivos.

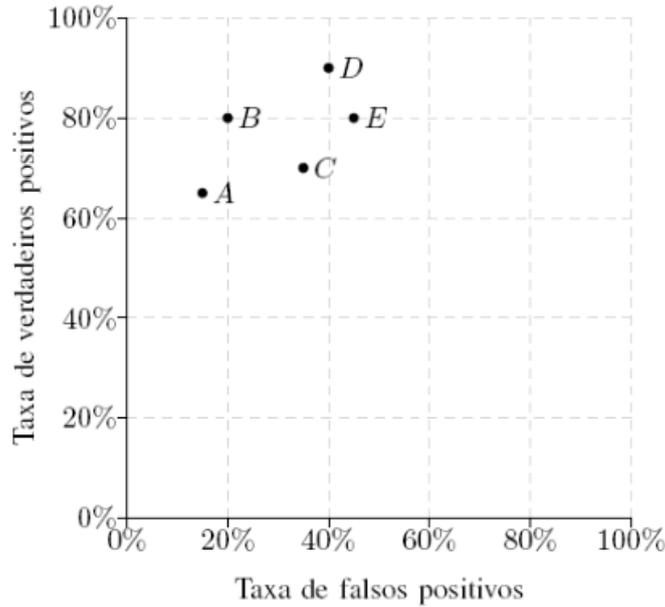


Figura 2.15. Diferentes pontos no espaço ROC. Adaptado de: [64].

2.6.3 Testes de hipótese

No desenvolvimento dos métodos da estatística moderna, as primeiras técnicas de inferência que apareceram foram as que faziam diversas hipóteses sobre a natureza da amostra da qual se extraíram os dados. Como os valores relacionados com a amostra são denominados “parâmetros”, tais técnicas estatísticas foram denominadas de paramétricas.

Os levantamentos feitos para realizar um teste de hipótese determina o grau de aceitação do comportamento dos dados. Formulada uma determinada hipótese particular é necessário coletar dados empíricos e com base nestes dados decide-se então sobre a validade ou não da hipótese. A decisão sobre a hipótese pode levar a rejeição, revisão ou aceitação da teoria que a originou.

Para se chegar a conclusão se uma determinada hipótese deverá ser aceita ou rejeitada, baseado em um particular conjunto de dados, é necessário dispor de um processo objetivo que permita decidir sobre a veracidade ou falsidade de tal hipótese.

Uma hipótese estatística é uma suposição ou afirmação que pode ou não ser verdadeira, relativa a uma ou mais populações. A veracidade ou falsidade de uma hipótese estatística nunca é conhecida com certeza.

Em estatística trabalha-se com dois tipos de hipótese: nula (H_0) e alternativa. A hipótese nula é a hipótese de igualdade. A rejeição da hipótese nula envolve a aceitação de outra hipótese denominada de alternativa. Esta hipótese é a definição operacional da

hipótese de pesquisa que se deseja comprovar. A natureza do estudo vai definir como deve ser formulada a hipótese alternativa. Por exemplo, se o teste é do tipo paramétrico, no qual o parâmetro a ser testado é representado por θ , então a hipótese nula seria: $H_0 : \theta = \theta_0$ e as hipóteses alternativas seriam:

- $H_1 : \theta = \theta_1$ (Hipótese alternativa simples) ou
- $H_1 : \theta \neq \theta_0; \theta > \theta_0$ ou $\theta < \theta_0$. (Hipóteses alternativas compostas).

No primeiro caso, $H_1 : \theta \neq \theta_0$, diz-se que o teste é bilateral(ou bicaudal), se $H_1 : \theta > \theta_0$, diz-se que o teste é unilateral (ou unicaudal) à direita e se $H_1 : \theta < \theta_0$, então, diz-se que o teste é unilateral (ou unicaudal) à esquerda.

Existem inúmeros testes estatísticos tanto paramétricos quanto não paramétricos. Alguns itens devem ser levados em conta na escolha da prova estatística para determinada situação. A maneira como a amostra foi obtida, a natureza dos dados da qual se extraiu a amostra e o tipo de mensuração ou escala empregado nas definições operacionais das variáveis envolvidas, isto é, o conjunto de valores numéricos e ainda o tamanho da amostra disponível.

Uma vez determinados a natureza da população e o método de amostragem ficará estabelecido o modelo estatístico. Associado a cada teste estatístico tem-se um modelo estatístico e condições de mensuração. O teste é válido sob as condições especificadas no modelo e pelo nível da escala de mensuração. Nem sempre é possível verificar se todas as condições do modelo foram satisfeitas e neste caso tem-se que admitir que estas condições foram satisfeitas. Estas condições do modelo estatístico são denominadas suposições ou hipóteses do teste. Qualquer decisão tomada por meio de um teste estatístico somente terá validade se as condições do modelo forem válidas. Quanto mais fracas forem as suposições do modelo mais gerais serão as conclusões. No entanto, as provas mais poderosas, isto é, as que apresentam maior probabilidade de rejeitar H_0 quando for falsa, são as que exigem as suposições mais fortes ou mais amplas.

Além dos conceitos já vistos, para o teste de hipóteses é necessário ainda definir os erros envolvidos e as regiões de rejeição e de aceitação. Para poder aceitar ou rejeitar uma hipótese H_0 e como consequência, rejeitar ou aceitar a hipótese alternativa H_1 , é necessário estabelecer uma regra de decisão, isto é, é necessário estabelecer para quais resultados vai-se rejeitar H_0 , ou seja, afirmar H_1 , e para quais resultados, vai-se aceitar H_0 . O conjunto de valores de uma amostra que levará a rejeição da hipótese nula é denominado de região crítica (RC) e a faixa restante de valores da amostra é denominada de região de aceitação (RA).

A probabilidade de que uma variável dentro dos dados assuma valores do conjunto RC é denominada de nível de significância do teste. O nível de significância do teste é, na realidade, a probabilidade de se rejeitar a hipótese nula, quando ela é verdadeira, sendo então a probabilidade de se cometer um erro. Como este é apenas um dos dois tipos de erro possível de ser cometido num teste de hipóteses, ele é denominado de erro do tipo I. O outro tipo de erro possível de ser cometido é aceitar H_0 quando ela é falsa e é denominado de erro do tipo II. Em resumo pode-se ter as seguintes situações em um teste de hipóteses:

Tabela 2.4. Possibilidades envolvidas em um teste de hipóteses.

		Realidade	
		H_0 verdadeira	H_0 falsa
Decisão	Rejeitar H_0	Erro Tipo I	Decisão Correta
	Aceitar H_0	Decisão Correta	Erro Tipo II

Os testes de hipóteses dividem-se em paramétricos e não paramétricos. Os paramétricos exigem que as amostras tenham uma distribuição normal, especialmente se tiverem uma dimensão inferior a 30. Os testes não paramétricos não precisam de requisitos tão fortes, como a normalidade. Para os testes paramétricos se tem os testes t-student e ANOVA e para os não paramétricos os testes Wilcoxon, McNemar, Friedman, entre outros. Para verificar se os dados seguem a Distribuição Normal e atendem à pressuposição de normalidade, podem-se utilizar alguns testes como: teste de assimetria, de curtose, de Qui-quadrado e de Lilliefors. Nessa pesquisa utilizou-se o teste de Lilliefors para o teste de normalidade, t-student para os testes paramétricos e Wilcoxon para os não paramétricos.

3 METODOLOGIA

Este capítulo detalha todos os métodos realizados para a classificação de imagens de retina. Para isso é necessário usar bancos de dados disponíveis, porém constatou-se, em testes preliminares, que a quantidade de imagens com exsudatos duros para a etapa de treinamento era baixa. Para resolver esse problema as imagens com exsudatos foram rotacionadas (como detalhado na Seção 3.1.2) com intuito de gerar mais exemplos para a classificação. A imagem de retina possui formas como os vasos sanguíneos, disco óptico, fóvea e mácula. Um obstáculo encontrado, comum na literatura, para a classificação de imagens com exsudato é similaridade que a região com disco óptico possui com os exsudatos, por isso é necessário detectar o local em que o disco óptico está para que não seja confundido com exsudato (detalhado na seção 3.1.1). Em um primeiro momento foi testada a possibilidade de classificar as imagens inteiras (1440 x 960 pixels), porém nos primeiros testes percebeu-se que os classificadores não suportariam classificar a imagem completa. Sendo assim, por meio de experimentos iniciais escolheu-se empiricamente trabalhar com janelas de tamanho 100 x 100 pixels em todos os métodos. Assim, optou-se, em cada etapa de treinamento ou classificação, por percorrer cada imagem em janelas de 100 x 100 pixels.

Todos os métodos desenvolvidos foram testados. Alguns métodos foram testados sistematicamente em centenas de vezes e outros foram testados uma única vez por conta do tempo elevado, aproximadamente uma semana, gasto para a realização de um teste (conforme está detalhado na Seção 3.2). Todos os resultados obtidos dos testes foram validados e comparados por meio de métricas de desempenho como acurácia, precisão, sensibilidade e Medida-F (Seção 2.6.1), calculadas a partir da matriz de confusão de cada teste, e técnicas de comparação da acurácia dos classificadores como testes de hipóteses 2.6.3 e a relação entre o verdadeiros positivos e falsos positivos, o plano ROC 2.6.2.

3.1 DESENVOLVIMENTO DOS ALGORITMOS DE CLASSIFICAÇÃO DAS IMAGENS DE RETINA

O desenvolvimento dos algoritmos de classificação seguiu algumas etapas descritas a seguir. Os treinamentos e validações foram feitos com imagens de um banco de dados disponibilizado por um grupo de pesquisa (conforme detalhado na Seção 3.2). As imagens do banco foram separadas em dois grupos: saudáveis (268) e retinopatia (320). As imagens identificadas como retinopatia eram de toda a retina, portanto foi necessário separar manualmente as regiões com exsudatos duros em janelas de 100 x 100 pixels, tamanho escolhido empiricamente. A quantidade de imagens com exsudatos duros utilizada foi 208, pois nem todas as imagens com retinopatia presentes no banco de dados possuem exsudatos. Com os dois grupos definidos, as imagens usadas no treinamento (60%, 161 imagens saudáveis e 125 imagens com exsudatos) e na validação (107 imagens saudáveis e 83 imagens com exsudatos) eram selecionadas aleatoriamente, a cada experimento realizado, para gerar novas combinações.

Ressalta-se que em cada experimento três módulos foram comuns: detecção do disco óptico, janelamento das imagens saudáveis e rotação de imagens com exsudatos. O módulo de detecção do disco óptico (DO) é necessário para que essa região não seja classificada posteriormente como exsudato duro, tendo em vista a semelhança entre essas estruturas. A Figura 3.1 ilustra a região do disco óptico e uma região com exsudatos duros (conforme detalhado na Seção 3.1.1). Esse problema é encontrado em algumas pesquisas de classificação de imagens de retina como [65], [66] e [67], nas quais mostram como é necessário detectar os discos ópticos antes de classificar.

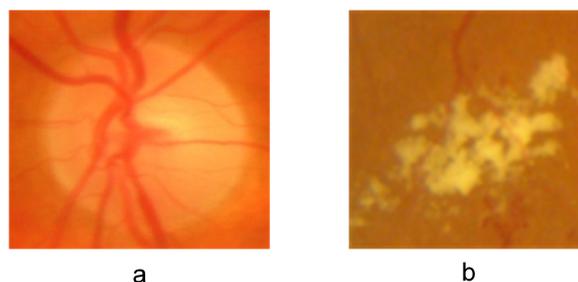


Figura 3.1. Semelhança entre o disco óptico e região com exsudatos duros. (a) disco óptico e (b) exsudatos duros.

Para a montagem das matrizes que eram usadas no treinamento e na validação, percorreu-se as imagens saudáveis em janelas com tamanho 100 x 100 pixels. Esse tamanho foi escolhido empiricamente para as janelas fossem do mesmo tamanho das imagens com exsudatos duros. O módulo para rotacionar as imagens com exsudatos, detalhado na Seção 3.1.2, serviu para resolver o problema de se ter poucos exemplos para treinar

em relação a quantidade de imagens saudáveis disponíveis.

Nos métodos que utilizaram o módulo com a técnica de PCA, cada imagem selecionada para treinamento, saudável e com exsudato, foi colocada como uma coluna numa matriz para montar a matriz de covariância 2.4. Portanto, como eram selecionadas 286 imagens de tamanho 100 x 100 pixels para treinamento, a matriz de covariância tinha dimensão 10000x286. A partir desta etapa, foram calculados os autovalores da matriz montada e ordenou-se os autovetores em ordem decrescente dos autovalores calculados. Os autovetores foram colocados em uma matriz contendo as novas funções de base representando as principais informações das imagens. A ideia central foi verificar como os classificadores se comportam a partir da seleção das componentes principais avaliadas. Dadas as funções de base determinadas pela PCA, optou-se, a partir de testes preliminares que levaram em conta o custo e a qualidade das imagens reconstruídas, por utilizar as 100 primeiras componentes principais.

Outra técnica abordada para redução de dimensionalidade foi associada à transformada de Fourier de cada imagem. A transformada de Fourier (FT) resulta em uma representação complexa que pode ser expressa em termos do módulo e fase, como visto na Seção 2.5. A informação visual na FT está representada de outra forma e não há perda de informação. O módulo da transformada de Fourier fornece a distribuição de energia na imagem e o ângulo da fase da FT fornece informações sobre as formas da imagem. Em algumas tarefas de aplicação a informação das formas é mais importante do que a distribuição de energia e outras ocorre o inverso. Nos experimentos que utilizaram módulo transformada de Fourier, a ideia foi verificar o comportamento do classificador LAST separando o espectro da fase e a fase do espectro da transformada de Fourier.

Os métodos explicados: detecção do disco óptico, janelamento das imagens saudáveis, rotação das imagens com exsudatos, PCA e transformada de Fourier foram utilizados nos algoritmos para investigar o comportamento dos classificadores SVM, Comitês e LAST. A técnica de PCA foi aplicada nos três classificadores. Além disso, classificou-se usando o LAST diretamente, com a fase ou espectro da FT e com a fase ou espectro da FT com PCA. Portanto, as classificações foram:

- SVM com PCA;
- do tipo Comitês com PCA;
- LAST com PCA;
- LAST com espectro da transformada de Fourier e PCA;
- LAST com fase da transformada de Fourier e PCA;

- LAST com fase da transformada de Fourier;
- LAST com espectro da transformada de Fourier;
- LAST.

A Figura 3.2 mostra o esquemático com todos os métodos implementados. No método 1 classificou-se as imagens usando o LAST com a fase ou espectro da transformada de Fourier com PCA. O método 2 foi o LAST com o espectro ou a fase da transformada de Fourier. No 3 classificou-se usando o SVM, Comitês e LAST com as componentes principais do PCA. Por fim, no método 4, a classificação foi realizada usando as imagens diretamente no LAST.

No método 1 classificou-se as imagens usando o LAST com a transformada de Fourier. As duas informações dadas pela FT (espectro e o ângulo da fase) foram usados em cada teste separadamente. Primeiro somente o espectro e depois somente a fase. Além disso, com o espectro ou a fase empregou-se também módulo de PCA (como explicado anteriormente) para treinar. Para a validação não se usou o módulo do cálculo de PCA, porque se pretende classificar a imagem real e direta. A diferença nesse método está no fato de se calcular a transformada de Fourier antes do cálculo do PCA no treinamento e nas imagens para validação.

O segundo método foi utilizado novamente a transformada de Fourier, mas sem o PCA, como ilustrado na Figura 3.2. Na primeira parte desse método usou-se a componente do espectro de Fourier antes da montagem das matrizes de treinamento e validação para serem utilizados na classificação diretamente no LAST. Depois repete-se todo o processo, porém usando a fase da transformada de Fourier. Nesse método utiliza-se, assim como nos outros, os módulos de rotacionar as imagens com exsudatos no treinamento e percorrer em janelas as imagens saudáveis.

Na etapa do terceiro método, o módulo para percorrer as imagens é usado para as imagens saudáveis, tanto para o treinamento quanto para a validação. O módulo para rotacionar as imagens foi usado para os exemplos com exsudatos duros somente no treinamento. As imagens resultantes desses módulos são usadas no módulo seguinte para o cálculo do PCA. Depois, o resultado dos principais componentes calculados servem para a montagem da matriz de treinamento. As imagens saudáveis selecionadas aleatoriamente para validação passam pelo mesmo processo de serem percorridas em janelas compondo a matriz de validação juntamente com as imagens em que estão os exsudatos. As matrizes de treinamento e validação montadas são usadas na classificação. O processo é o mesmo para classificação usando o SVM, Comitês e LAST.

O último método implementado foi a classificação direta, sem PCA e sem Fourier, com o LAST. Nessa etapa as imagens são separadas como nas etapas anteriores, percorre-se as imagens saudáveis em janelas e rotaciona-se as imagens com exsudatos e os resultados são usados para a montagem das matrizes de treinamento e validação, seguindo para a classificação no LAST.

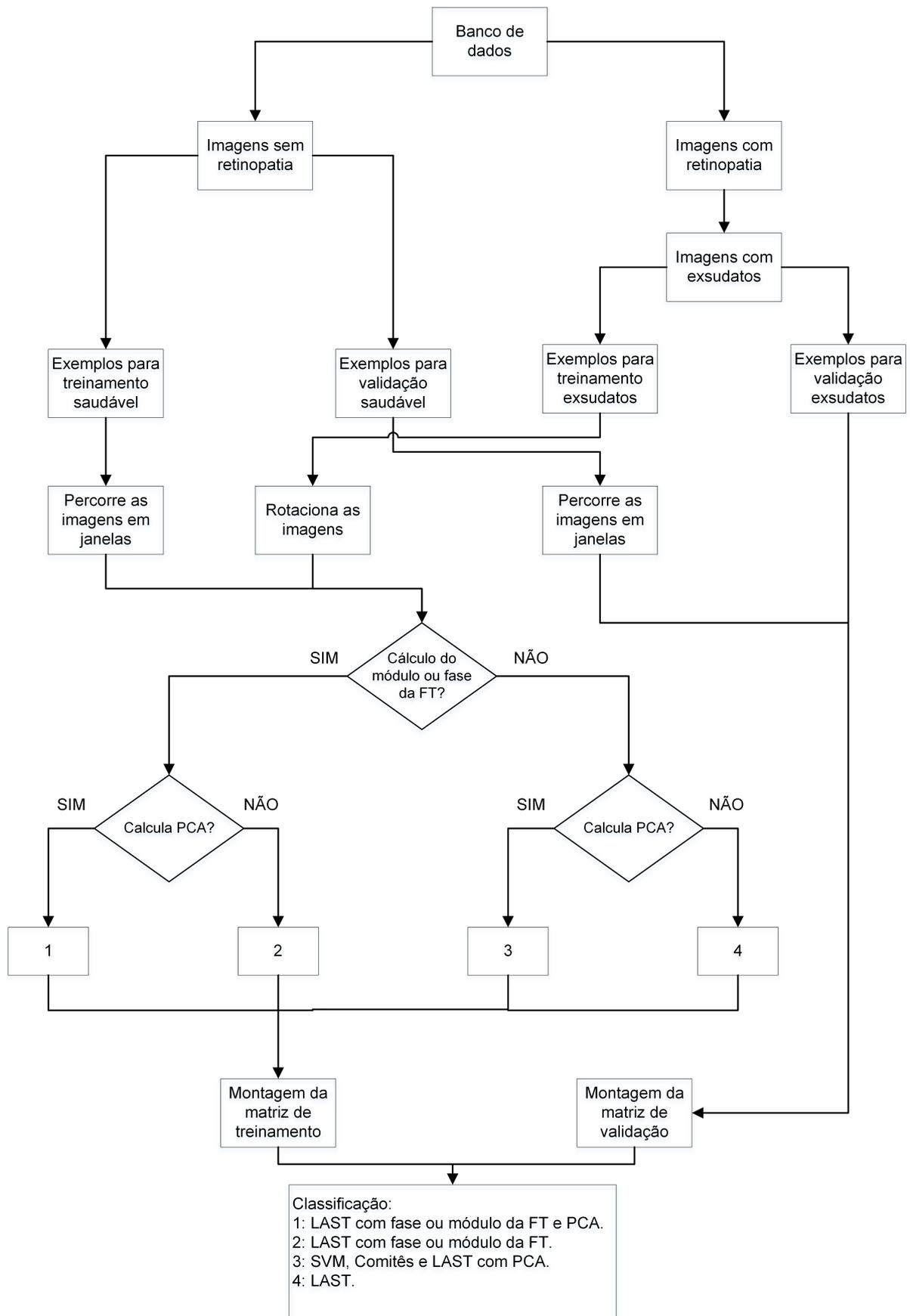


Figura 3.2. Diagrama esquemático com a metodologia implementada para o treinamento e classificação usando SVMs, Comitês e LAST.

3.1.1 Detecção do disco óptico

O disco óptico da retina diz respeito à porção do nervo óptico vista no fundo do olho, formado pelo encontro de todos os axônios das células ganglionares da retina, ou seja, o disco óptico é o fundo do olho. Na Figura 3.3 (A) mostra-se uma imagem de um olho esquerdo e na (B) de um olho direito. Um dos problemas frequentes em classificação de imagens de retina é a região do disco óptico, porque o classificador pode ser induzido a marcar esta região como sendo um exsudato duro. Esse problema é comum em classificadores de imagens de retina. *Rangayan* [30, 31], detecta o disco óptico e vasos sanguíneos da retina utilizando os filtros de Gabor. Outra forma que ele usou é a determinação de ângulos entre a fóvea e o centro do DO por meio da Transformada de Hough [68, 25] para determinar o grau da retinopatia.

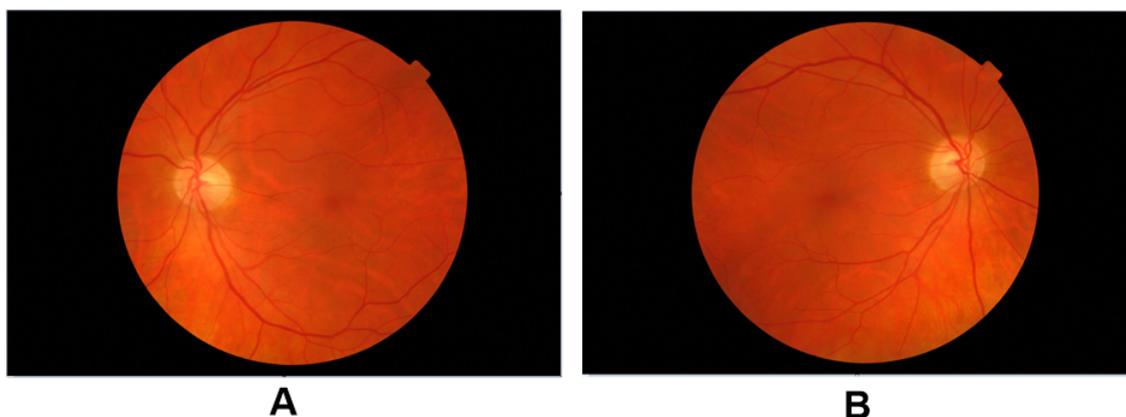


Figura 3.3. Exemplificação das posições diferentes do disco óptico. (A) Olho esquerdo e (B) Olho direito. Fonte: imagens extraídas do banco de dados MESSIDOR [24].

No sistema implementado, foi feita a detecção do disco óptico separadamente e classificando essa região como sendo saudável. A abordagem escolhida para detecção do disco óptico foi detectá-la manualmente. Um exemplo do DO marcado está na Figura 3.4. Como a imagem é RGB, a abordagem feita para encontrar o DO foi por meio da intensidade dos pixels na região vermelha, verde e azul, sendo que a intensidade na região do DO é diferente em comparação com as outras regiões da imagem. Após isso, através de histogramas foi possível determinar a posição do centro do disco óptico e demarcar uma região 160 x 160 pixels ao redor do centro como sendo o disco óptico. Em todas as imagens separa-se a janela da imagem que está presente o DO e a classifica como saudável. Este processo foi realizado para todos os classificadores, SVM, Comitês e LAST.

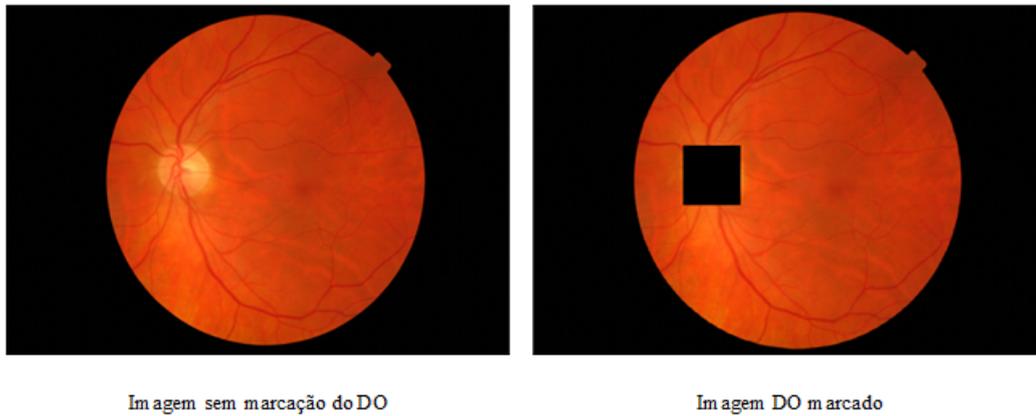


Figura 3.4. Exemplificação da marcação do disco óptico. Fonte: imagens extraídas do banco de dados MESSIDOR [24].

3.1.2 Rotação imagens com exsudato duro

As imagens de exsudatos duros foram obtidas manualmente, pela extração por meio de janelas (100 x 100 pixels) feita nas imagens dos bancos de dados disponíveis, Figura 3.5. Nesse processo foram geradas 208 imagens com exsudatos duros para serem utilizadas como a classe de retinopatia diabética nos classificadores. Estas imagens foram acrescentadas à matriz de treinamento do sistema.

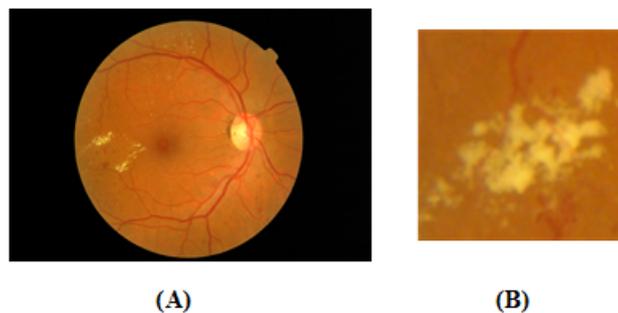


Figura 3.5. Exemplo da obtenção de imagem com exsudato duro. (A) Imagem da retina completa. (B) Janela separada para ser utilizada como a classe de exsudato duro.

A quantidade de imagens com exsudatos duros é muito inferior em comparação com a quantidade de imagens disponíveis classificadas como saudáveis. Uma das soluções encontrada para resolver esse problema é gerar mais exemplos de exsudatos rotacionando as próprias imagens. Para exemplificar o processo de rotação necessário, na Figura 3.6, mostra um exemplo de 25 rotações feitas com variação de, aproximadamente, 25 graus entre cada imagem. No caso do sistema, a quantidade necessária de rotações foi 100, ou seja, com variação de, aproximadamente, 3 graus entre cada imagem. Esta abordagem foi feita em todas as imagens para o sistema ter um melhor aprendizado, por conta da quantidade maior de imagens com exsudatos para imagem

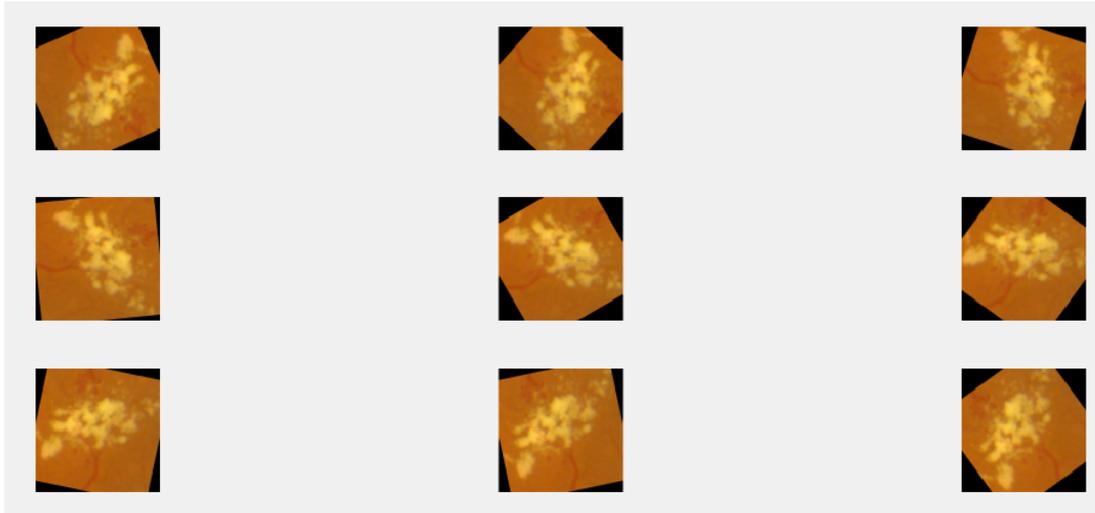


Figura 3.6. Exemplo de rotações em uma imagem com exsudato duro. Nesse caso, foram feitas 25 rotações com variação de, aproximadamente, 14 graus entre cada imagem.

3.2 EXPERIMENTOS DE AVALIAÇÃO DOS MÉTODOS IMPLEMENTADOS DE CLASSIFICAÇÃO DE IMAGENS DE RETINA

Para a implementação de testes experimentais, alguns bancos de dados de imagens de retina foram encontrados na literatura como: Messidor [24], e-optha [69], DRIVE [70], STARE [71], HRF [72], entre outros. Um dos fatores essenciais para a escolha do banco a ser utilizado foi a quantidade de imagens com RD. Assim, o banco de dados de imagens de retina escolhido foi o MESSIDOR [24]. Ele possui 1200 imagens de fundo de olho com três tamanhos diferentes: 1440 x 960, 2240 x 1488 e 2304 x 1536 pixels. O tamanho escolhido para o projeto foi 1440x960 pixels, pois é o tamanho que possui o maior número de imagens - 320 com retinopatia diabética e 268 sem retinopatia, Figura 3.7. Cabe ressaltar, que as imagens com exsudatos usadas nos experimentos foram retiradas manualmente das imagens do banco de dados selecionadas como portadores de retinopatia.

A parte experimental do projeto teve como objetivo colher resultados para posterior avaliação do desempenho do sistema. As classificações usando o SVM com PCA foram rodadas de forma sistemática 180 vezes na mesma máquina para validar a classificação. Os resultados como a SVM treinada e os resultados obtidos (acurácia, verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos) foram armazenadas em arquivos para serem avaliados na parte de validação do sistema.

O classificador do tipo Comitês possui diversas abordagens para classificação como: *Bag*, *AdaBoostM1*, *LogitBoost*, *GentleBoost*, *RobustBoost*, *LPBoost*, *RUSBoost* e *Subspace*. A escolha do melhor método para a pesquisa foi feita por meio de testes prelimi-

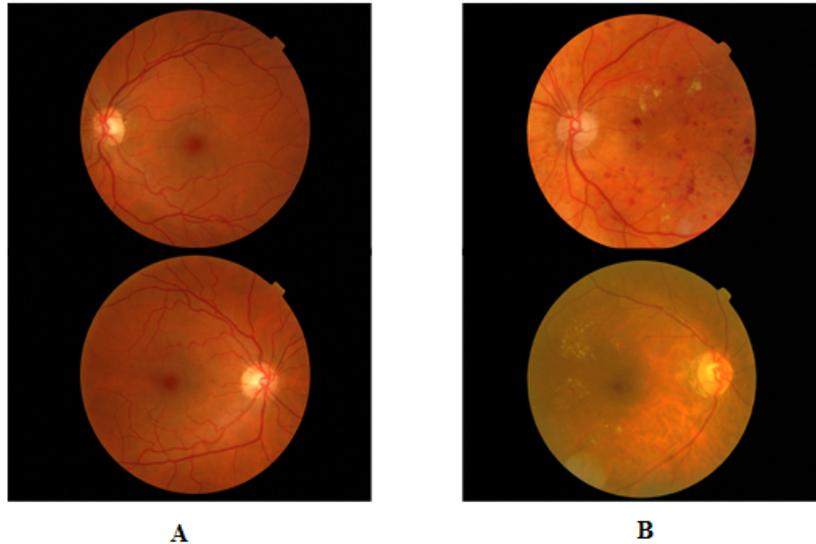


Figura 3.7. Exemplos de imagens do banco de dados MESSIDOR. (A) imagem sem a presença de sintomas de Retinopatia Diabética. (B) imagens com sintomas de Retinopatia Diabética. Fonte: [24].

nares, em que usou-se todas as abordagens individualmente. Os resultados obtidos com todos os métodos estão na Tabela 3.1. Como o *AdaBoostM1* foi a abordagem que mostrou melhor resultado com acurácia de 87.01% foi então escolhido para os experimentos futuros.

Tabela 3.1. Resultados de experimentos preliminares com os métodos do tipo Comitês.

Abordagem do tipo Comitês	Acurácia
Bag	79.87%
AdaBoostM1	87.01%
LogitBoost	85.06%
GentleBoost	85.54%
RobustBoost	85.26%
LPBoost	82.47%
RUSBoost	74.67%
Subspace	54.55%

Com a escolha da abordagem do classificador do tipo Comitês partiu-se para a etapa de sucessivos testes para a validação e comparação desse classificador com os demais. Foram realizados 200 testes sucessivos na mesma máquina armazenando todos os resultados em arquivos para que os resultados fossem comparados e validados posteriormente.

Outro experimento realizado sucessivamente foi usando o classificador LAST com o PCA. Nesse caso, foram feitos 164 repetições sistematicamente para obtenção de resul-

tados concretos da classificação para validação. Nesse caso, também os resultados foram armazenados em arquivos para que pudessem ser analisados posteriormente.

Um problema enfrentado na parte de experimentos foi o tempo gasto para rodar alguns programas que usam o classificador LAST, aproximadamente uma ou até duas semanas. Por esta razão, as classificações do LAST com PCA e o módulo e a fase da Transformada de Fourier, sem o PCA e com o módulo e a fase de Fourier e com a imagem diretamente (sem PCA e Fourier) foram rodados apenas uma vez. O tempo que seria gasto para rodar esses programas centenas de vezes era inviável.

3.3 METODOLOGIA DE ANÁLISE DOS RESULTADOS DE CLASSIFICAÇÃO – VALIDAÇÃO E COMPARAÇÃO DOS CLASSIFICADORES

Em aprendizagem de máquina, algumas métricas são usadas para validar o desempenho do classificador. Dentre tais medidas, algumas serão utilizadas como: taxa de erro, acurácia, precisão, sensibilidade e Medida-F. A partir desses cálculos, a definição da classificação com melhor desempenho foi definido pelo maior valor de Medida-F [2.6.1](#) Os métodos que foram testados sistematicamente serão comparados de duas formas: entre si e tomando um valor de acurácia como referência (0.8), por meio de testes de hipóteses. Outro método de comparação explicado nessa seção será a relação entre verdadeiros positivos e falsos positivos pelo método ROC (do inglês, *Receiver Operating Characteristics*). Todos esses métodos estatísticos servirão para comparar o desempenho entre os classificadores SVM, do tipo Comitês e LAST.

3.3.1 Métricas para avaliação de desempenho dos classificadores

As métricas de desempenho: acurácia, precisão, sensibilidade e Medida-F foram calculadas nos métodos de classificações feitas uma única vez (LAST, LAST com PCA, LAST com fase da transformada de Fourier, LAST com módulo da transformada de Fourier, LAST com fase da transformada de Fourier e PCA e LAST com módulo da transformada de Fourier e PCA). Por meio dessas medidas, poderá avaliar e comparar o desempenho do classificador LAST.

Os resultados dos testes realizados uma única vez, por conta do tempo de execução que levam, foram coletados e os seus valores de verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos colocados em matrizes de confusão como no modelo da Tabela [2.3](#). Na pesquisa os verdadeiros positivos correspondem ao exemplos de imagens

com exsudatos (positivos) e foram classificados como tendo exsudatos (positivos); os falsos positivos são os exemplos negativos (saudável) que foram classificados como positivos (possuindo exsudatos); verdadeiros negativos são as imagens saudáveis (negativos) e que foram classificados como imagens saudáveis (negativos) e falsos negativos são os exemplos positivos (imagens com exsudatos) que foram classificados como negativos (saudáveis).

A acurácia ou taxa de acerto de cada classificação é calculada pela Equação 2.59. Ela é medida pelo número de acertos (verdadeiros positivos e verdadeiros negativos) dividido pelo total (acertos e erros). A precisão, Eq. 2.60, é calculada pela razão entre os verdadeiros positivos pelo número total de classificados como positivos (verdadeiros positivos e falsos positivos). Na classificação de imagens de retina, a precisão é a taxa com que todas as imagens classificadas como tendo retinopatia realmente possuem exsudatos.

A sensibilidade, Equação 2.61, é definida pela razão entre verdadeiros positivos e a soma dos verdadeiros positivos e falsos negativos. Ela calcula a taxa com que o sistema classifica como positivos todos os exemplos que são positivos. Nos classificadores representa a taxa das imagens com exsudatos são classificadas com tendo retinopatia. A Medida-F 2.65 é a média ponderada de precisão e sensibilidade. Ela será usada como padrão para comparar os classificadores, ou seja, o classificador que tiver maior Medida-F será considerado como tendo o melhor desempenho de classificação com imagens de retina.

3.3.2 Testes de Hipóteses

Para comparar as acurácias das classificações feitas sistematicamente, foram realizados testes de hipóteses. Eles podem ser divididos em paramétricos e não paramétricos. Os paramétricos exigem que as amostras tenham distribuição normal. Os testes não paramétricos não precisam de requisitos tão fortes, como a normalidade. Antes de fazer o teste de hipótese, testa-se a normalidade. Para verificar se os dados seguem a distribuição normal e atendem à pressuposição de normalidade utilizou-se o teste de Lilliefors. Os testes paramétricos foram feitos usando o teste t-student e os não paramétricos o teste Wilcoxon.

Na metodologia de um teste de hipótese, determina-se uma hipótese nula, espera-se que seja rejeitada, com um certo nível de significância calculada. O nível de significância é a probabilidade de se rejeitar a hipótese nula e quanto maior for (mais próximo de 100%) maior a confiança de que a hipótese nula pode ser rejeitada.

Os experimentos em que foram feitos testes de hipóteses foram: SVM com PCA,

do tipo Comitês com PCA e LAST com PCA. O primeiro teste feito foi para definir se seguiam a distribuição normal ou não. Caso os dois fossem normais o teste seria paramétrico, caso contrário, o teste seria não paramétrico. Na comparação entre eles as hipóteses nulas definidas foram: Comitês tem acurácia melhor do que o LAST (o teste foi não paramétrico), SVM tem acurácia melhor do que LAST (teste foi não paramétrico) e a acurácia do SVM é melhor do que o classificador do tipo Comitês (teste foi paramétrico).

Outro teste de hipótese formulada foi de que as classificações realizadas sucessivamente obtiveram, como hipótese nula, desempenho menor do que uma acurácia de referência de 80%. Na comparação entre o LAST e o valor de referência o teste foi não paramétrico e entre Comitês e SVM os testes foram paramétricos. Essa comparação serviu para avaliar o desempenho dos classificadores desenvolvidos com os presentes na literatura [19, 67, 66].

3.3.3 Medida de desempenho no plano ROC

Para comparar todas as classificações foi usado o método ROC. No espaço ROC, como mostrado na Figura 2.15, plota-se taxa de falsos positivos, ou especificidade 2.66, pela taxa de verdadeiros positivos, ou sensibilidade 2.61, foi usado para comparar todas as classificações realizadas. Nos testes feitos uma vez (LAST, LAST com PCA, LAST com fase da transformada de Fourier, LAST com módulo da transformada de Fourier, LAST com fase da transformada de Fourier e PCA e LAST com módulo da transformada de Fourier e PCA) usaram-se as matrizes de confusão montadas para plotar os pontos no plano. Nos testes sistemáticos escolheu-se um ponto e a partir da matriz de confusão plotaram-se os pontos no plano junto com os outros pontos. A partir desse gráfico é possível avaliar o método de classificação que obteve melhor desempenho para imagens de retina.

A avaliação é dada pelo posicionamento do ponto no plano ROC, Figura 2.14. Os classificadores que tiveram ponto próximo ao canto inferior esquerdo podem ser considerados "conservadores", porque eles fazem uma classificação positiva somente se têm grande segurança na classificação. Como consequência, eles cometem poucos erros falsos positivos, mas frequentemente têm baixas taxas de verdadeiros positivos. Os classificadores que tiveram pontos próximos ao canto superior direito podem ser considerados "liberais", pois eles predizem a classe positiva com maior frequência, de tal maneira que classificam a maioria dos exemplos positivos corretamente, mas, geralmente, com altas taxas de falsos positivos. Portanto, os melhores classificadores serão aqueles que estiverem mais próximos do canto superior esquerdo do gráfico.

4 RESULTADOS E DISCUSSÃO

Os resultados obtidos nos experimentos são mostrados e avaliados de quatro formas nesse capítulo. A primeira será em forma de histogramas de desempenho obtidos nos testes sucessivos: SVM com PCA, Comitês com PCA e LAST com PCA. Na segunda parte, será por meio da comparação das medidas de desempenho nos métodos de classificação com único teste. Na terceira parte os testes sucessivos serão avaliados por meio de testes de hipóteses, tanto na comparação entre eles quanto comparando com um valor de referência. Por fim, todos os métodos usados na pesquisa são esboçados no plano ROC.

4.1 HISTOGRAMAS DE DESEMPENHO DOS MÉTODOS TESTADOS SISTEMATICAMENTE

O resultado das acurácias obtidas nos testes sistemáticos, SVM com PCA, Comitês com PCA e LAST com PCA, estão representadas nessa seção por meio de histogramas de frequência. Cabe ressaltar que a acurácia é taxa de acerto calculada como o número de acertos na validação dividido pelo número de acertos mais o número de erros na validação. Na Figura 4.1 está o histograma obtido nos experimentos usando o classificador SVM com o PCA com frequência total de 180. Na Figura 4.2 está o histograma obtido com o classificador do tipo Comitês com PCA com frequência total de 200. Por fim, na Figura 4.3 expõe o histograma do LAST com PCA com frequência total de 164.

Para verificar se os histogramas encontrados obedecem a curva de normalidade, foi feito previamente um teste de hipótese usando o método Lilliefors. Os resultados obtidos no tipo Comitês e SVM foram definidos como seguindo uma curva normal. Já para a classificação com o LAST essa hipótese foi rejeitada, não dando a curva a condição de normalidade. Na Figura 4.3 é possível perceber que os alguns resultados fugiram do padrão e ficaram entre 0.6 e 0.7.

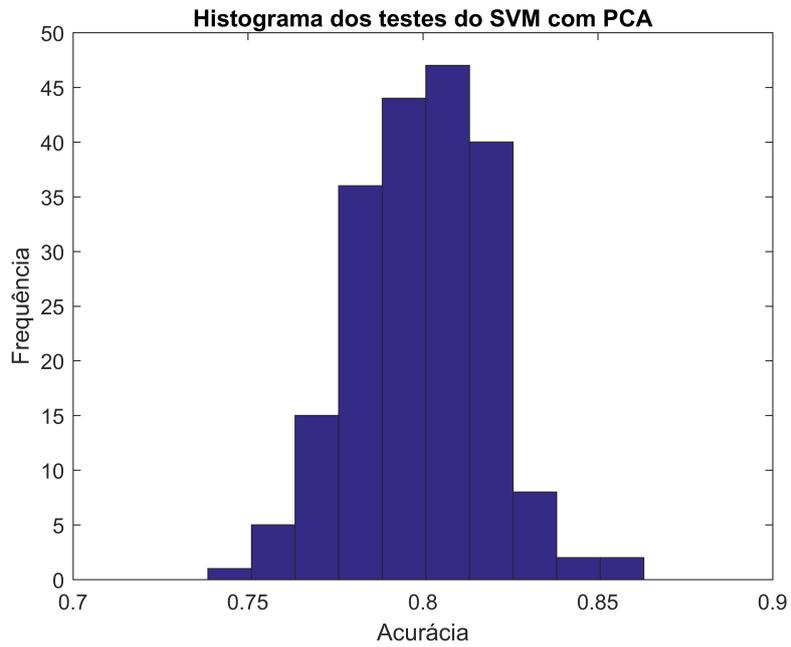


Figura 4.1. Histograma dos testes do classificador SVM com PCA.

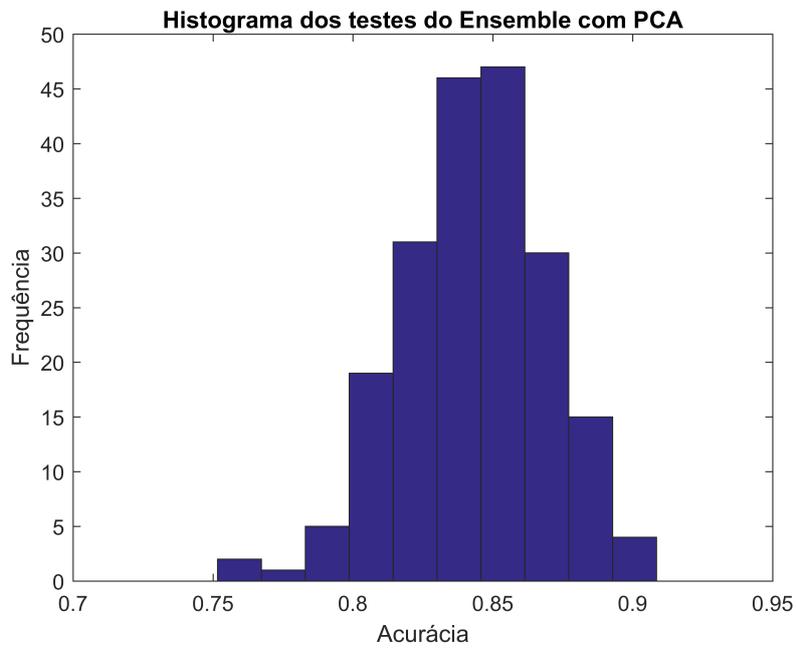


Figura 4.2. Histograma dos testes do classificador do tipo Comitês com PCA.

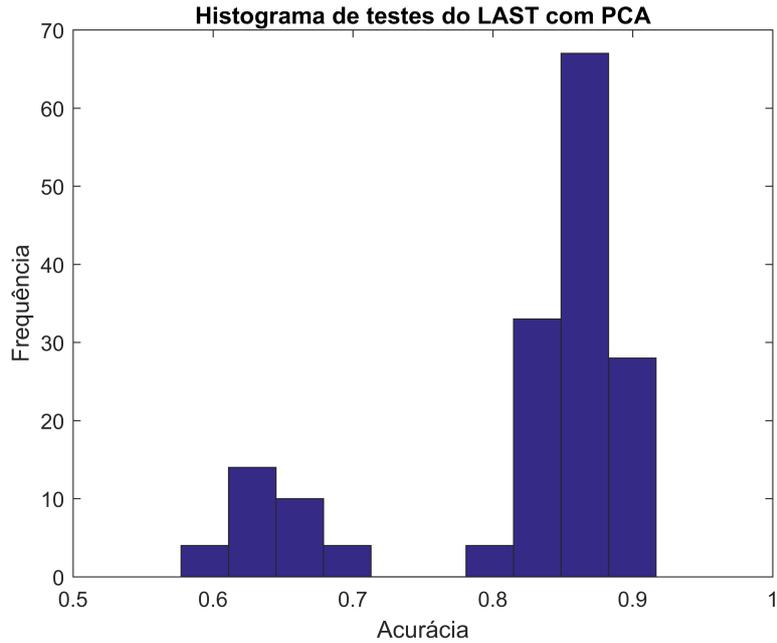


Figura 4.3. Histograma dos testes do classificador LAST com PCA.

4.2 MEDIDAS DE DESEMPENHO PARA OS TESTES REALIZADOS UMA VEZ

Para os métodos de classificação realizados uma vez, a comparação foi feita colocando-se os resultados obtidos na forma de matrizes de confusão como na Tabela 2.3, nas quais estão colocados os verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos encontrados, em cada caso. Na tabela 4.1 estão as matrizes de confusão para a classificação usando o LAST com a fase da transformada de Fourier das imagens, LAST com o módulo da transformada de Fourier, LAST com PCA e o módulo da transformada de Fourier, LAST com PCA e a fase da transformada de Fourier e LAST sem PCA e sem Fourier.

A comparação entre os métodos utilizados com o LAST foi feita por meio de métricas de desempenho como acurácia, precisão, sensibilidade e Medida-F, Tabela 4.2. Por meio das fórmulas de cada uma detalhadas na Seção 3.3.1 calculou-se a Medida-F de cada experimento. Tomando como medida padrão a Medida-F, o melhor resultado foi a classificação de imagens de retina usando o PCA com o módulo da transformada de Fourier no LAST com 0.8815. A partir desse resultado conclui-se que o uso do PCA foi importante para melhorar o desempenho e que a distribuição de energia nas imagens ajudou o classificador a separar as classes. Por outro lado, o que apresentou menor Medida-F foi a classificação com a fase da transformada de Fourier, mostrando que as informações sobre as bordas das estruturas que a imagem não ajudaram na classificação.

Tabela 4.1. Matrizes de confusão para os resultados da classificação com testes feitos uma vez.

	Predição Positiva	Predição Negativa
	LAST com fase de Fourier	
Classe Positiva	45	38
Classe Negativa	20	44
	LAST com módulo de Fourier	
Classe Positiva	40	41
Classe Negativa	0	66
	LAST com módulo de Fourier e PCA	
Classe Positiva	65	14
Classe Negativa	2	66
	LAST com fase de Fourier e PCA	
Classe Positiva	43	30
Classe Negativa	8	66
	LAST	
Classe Positiva	62	22
Classe Negativa	2	71

Tabela 4.2. Resultados de experimentos feitos com LAST e ferramentas como PCA e Transformada de Fourier.

	Acurácia	Precisão	Sensibilidade	Medida-F
LAST com fase de Fourier	0.6054	0.6923	0.5421	0.6065
LAST com módulo de Fourier	0.72	1	0.5422	0.7031
LAST com PCA e módulo de Fourier	0.8526	0.9710	0.8072	0.8815
LAST com PCA e fase de Fourier	0.74	0.8431	0.5310	0.6516
LAST	0.8526	0.9687	0.62	0.7560

4.3 TESTES DE HIPÓTESES PARA COMPARAÇÃO ENTRE OS TESTES SISTEMÁTICOS

Os testes de hipóteses foram feitos com os experimentos realizados sistematicamente (SVM com PCA, Comitês com PCA e LAST com PCA). A primeira parte comparou os três métodos entre si e na segunda foi comparar os três testes com um valor de referência. Antes de comparar fez-se o teste de normalidade (método Lilliefors) para determinar se os testes seriam paramétricos ou não paramétricos. Caso a hipótese de normalidade não fosse rejeitada para os dois casos comparados o teste seria paramétrico (t-student), caso contrário, o teste seria não paramétrico (Wilcoxon).

A Tabela 4.3 mostra os resultados obtidos nas comparações entre os classificadores. A primeira hipótese nula testada foi de que a acurácia obtida pelo classificador do tipo Comitês seria maior do que a obtida no LAST. Esse teste foi não paramétrico e a hipótese foi rejeitada com nível de significância de 98%. O segunda segunda hipótese nula testada foi de que a acurácia obtida no SVM foi maior do que LAST. Esse teste também foi não paramétrico e a hipótese foi rejeitada com nível de significância de 99.99%. No terceiro teste, a hipótese nula foi de que a acurácia no SVM foi maior do que no classificador do tipo Comitês. Ao contrário dos testes anteriores, esse teste foi paramétrico e a hipótese nula foi rejeitada com nível de significância de 99.99%. Esses testes mostraram, com nível de significância alto, que o LAST apresentou melhor desempenho na comparação do tipo Comitês e SVM. Em segundo lugar ficou o classificador do tipo Comitês e em terceiro a classificação com o SVM.

Tabela 4.3. Valores dos níveis de significância os testes de hipótese entre SVM com PCA, Comitês com PCA e LAST com PCA. O teste com hipótese nula de que a classificação com tipo Comitês era maior do que o LAST foi rejeitada com 98%. O teste com hipótese nula de que a classificação com SVM foi maior do que o LAST foi rejeitada com 99.99%. O teste com hipótese nula de que a classificação com SVM foi maior do que o tipo Comitês foi rejeitada com 99.99%.

	Comitês com PCA	SVM com PCA
LAST com PCA	98%	99.99%
Comitês com PCA		99.99%

Na segunda parte, os resultados obtidos foram comparados com um valor de referência para acurácia de 0.8. A primeira hipótese nula testada foi de que a acurácia usando o classificador LAST era menor do que 0.8. Esse teste realizado foi não paramétrico e a hipótese nula foi rejeitada com nível de significância de 91%. A segunda hipótese nula foi de que a acurácia usando o classificador do tipo Comitês era menor do que o valor de referência. Esse teste foi paramétrico e a hipótese nula foi rejeitada com nível de

significância de 99.99%. O terceiro teste foi testar se a acurácia usando o SVM foi menor do que o valor de referência. Esse teste também foi paramétrico e a hipótese nula foi rejeitada com nível de significância de 99%. Todos esses resultados estão presentes na Tabela 4.4.

Os resultados obtidos nessa avaliação mostram que os três classificadores funcionaram bem e que estão acima do esperado de 80% de acurácia. Cabe ressaltar que acredita-se que o nível de significância para o LAST foi menor por conta dos valores de acurácia na região entre 0.6 e 0.7 mostrados no histograma e discutido na Seção 4.1.

Tabela 4.4. Valores dos níveis de significância os testes de hipótese das médias das acurácias do SVM com PCA, do tipo Comitês com PCA e LAST com PCA com acurácia de referência de 0.8.

	Acurácia de referência 0.8
LAST com PCA	91%
do tipo Comitês com PCA	99.99%
SVM com PCA	99%

4.4 RESULTADOS DE DESEMPENHO DE TODOS OS TESTES NO PLANO ROC

Os desempenho mostrado no plano ROC exhibe a comparação entre todos os métodos testados na pesquisa. O gráfico da Figura 4.4 apresenta os pontos obtidos da relação entre a taxa de falsos positivos (1-especificidade) e a taxa de verdadeiros positivos (sensibilidade). Como detalhado na Seção 3.3.3 os pontos que estão na parte esquerda superior são os que possuem melhor desempenho.

Pelo gráfico é possível definir que a classificação do LAST com módulo de Fourier é o mais conservador na comparação, porque ele não aceita falsos positivos, mas prejudica bastante o desempenho dos verdadeiros positivos. O mais liberal é a classificação usando o LAST com a fase da transformada de Fourier, porque aceita bastante falsos negativos, porém tem bom desempenho nos verdadeiros positivos.

A maioria dos classificadores (SVM com PCA, Comitês com PCA, LAST com PCA, LAST e LAST com o módulo da transformada de Fourier e PCA) estão próximos do canto esquerdo superior do gráfico, o que mostra bom desempenho pela relação entre falsos positivos e verdadeiros positivos. Dentre eles, pode-se destacar o desempenho do SVM com PCA e o LAST com o módulo da transformada de Fourier e PCA como os melhores. Por outro lado, a classificação que mostrou pior desempenho foi com o LAST com a fase da transformada de Fourier, confirmando o que foi encontrado também na Seção 4.2.

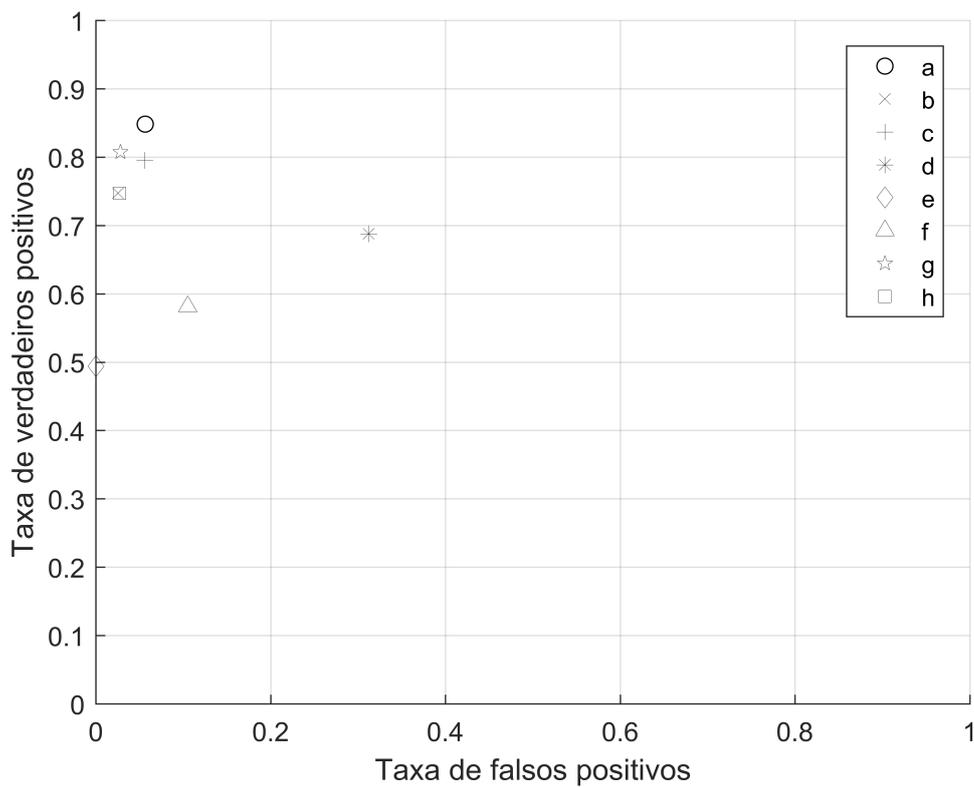


Figura 4.4. Métricas de desempenho no plano ROC para todos métodos testados. Os pontos no gráfico representam: (a)SVM com PCA, (b)Comitês com PCA, (c)LAST com PCA, (d)LAST com a fase da transformada de Fourier, (e)LAST com o módulo da transformada de Fourier, (f)LAST com a fase da transformada de Fourier e PCA, (g)LAST com o módulo da transformada de Fourier e PCA e (h)LAST.

5 CONCLUSÃO

A pesquisa desenvolveu e avaliou um sistema para classificação de imagens de retina como forma de auxílio a diagnóstico da retinopatia diabética. Como o diagnóstico por imagem é a forma de detecção da doença, o trabalho teve como finalidade comparar os métodos tradicionais de classificações de imagens médicas utilizando a extração de características com algoritmos que dispensam a extração prévia de características por meio de treinamento de dicionários. Avaliou-se também a utilização de algumas ferramentas que possam melhorar o desempenho da classificação como PCA e o uso da transformada discreta de Fourier para extração de características espectrais.

Os resultados mostraram, por meio de testes de hipóteses, realizados após avaliação sistemática que a classificação usando o método do LAST com PCA teve melhor desempenho nas comparações com o classificador do tipo Comitês com o PCA com nível de significância de 98% e com o SVM com PCA com nível de significância de 99.99%. Já na comparação entre Comitês com PCA e SVM com PCA, o método do tipo Comitês obteve resultado melhor com nível de significância de 99.99%. Outros testes de hipóteses mostraram que as três classificações possuem acurácia superior a 0.8 com nível de significância de 91% para o LAST, 99.99% para Comitês e 99% para o SVM. Em experimentos com outras variáveis, tendo a Medida-F como parâmetro obteve-se a seguinte ordem de desempenho: LAST com PCA e o módulo da Transformada de Fourier (0.88), LAST (0.75), LAST com módulo de Fourier (0.70), LAST com PCA e fase de Fourier (0.65) e LAST com fase de Fourier (0.60). A análise no plano *Receiver Operating Characteristic* mostrou que as classificações da imagens com SVM com PCA, Comitês com PCA, LAST com PCA, LAST e LAST com o módulo da transformada de Fourier e PCA possuem boa relação entre verdadeiros positivos e falsos negativos, o mesmo não acontece nos casos do LAST módulo da transformada de Fourier, LAST com fase da transformada de Fourier e PCA e no LAST com a fase da transformada de Fourier.

Conclui-se que no treinamento usando dicionários, sem a extração prévia de características, o LAST obteve melhor acurácia média (0.84) do que a acurácia comumente encontrada na literatura (0.8) e com as acurácias médias encontradas na pesquisa usando o classificador do tipo Comitês (0.81) e o SVM (0.79). A pesquisa mostrou que a classifi-

cação usando o LAST ajudaria no diagnóstico de retinopatia diabética, porque o sistema apresentou bom índice de verdadeiros positivos e baixo índice de falsos negativos. Em trabalho futuro, pretende-se combinar técnicas de aprendizagem profunda com as abordagens aqui empregadas, para avaliar o impacto sobre as métricas de desempenho, com base em bancos de imagens maiores a serem obtidos em parcerias de pesquisa.

Referências

- [1] J. H. Ma, J. J. Wang, e S. X. Zhang. The unfolded protein response and diabetic retinopathy. *Hindawi Publishing Corporation - Journal of Diabetes Research*, 2014.
- [2] Sang Jin Kim et al. Jonghwa Jin, Hophil Min. Development of diagnostic biomarkers for detecting diabetic retinopathy at early stages using quantitative proteomics. *Journal of Diabetes Research*, 2016(6571976):22 pages, 2016.
- [3] Vernon S. Negi, A. An overview of the eye in diabetes. *Journal of the Royal Society of Medicine*, 96(6):266–272, 2003.
- [4] K. Kirboga, A. V. Ozec, M. Kosker, A. Dursun, e M. I. Toker. The association between diabetic retinopathy and levels of ischemia-modified albumin, total thiol, total antioxidant capacity, and total oxidative stress in serum and aqueous humor. *Hindawi Publishing Corporation - Journal of Ophthalmology*, 2014.
- [5] International Diabetes Federation. *IDF Diabetes Atlas*. International Diabetes Federation, 6nd edition, 2014.
- [6] L. Rocha e D. Del Rio Selles. Censo 2014. *Conselho Brasileiro de Oftalmologia*, 2014.
- [7] M. Karaca, E. Coban, S. Ozdem, M. Unal, O. Salim, e O. Yucel. The association between endothelial dysfunction and hypertensive retinopathy in essential hypertension. *Medical Center Monitor*, 2014.
- [8] Donald S. Fong, Lloyd P. Aiello, Frederick L. Ferris, e Ronald Klein. Diabetic retinopathy. *Diabetes Care*, 27(10):2540–2553, 2004.
- [9] Sharon D. Solomon, Emily Chew, Elia J. Duh, Lucia Sobrin, Jennifer K. Sun, Brian L. VanderBeek, Charles C. Wykoff, e Thomas W. Gardner. Diabetic retinopathy: A position statement by the american diabetes association. *Diabetes Care*, 40(3):412–418, 2017.
- [10] Clara Sanchez, Roberto Hornero, Maria I. Lopez, Mateo Aboy, Jesus Poza, e Daniel. Abasolo. A novel automatic image processing algorithm for detection of hard exuda-

- tes based on retinal image analysis. *Medical Engineering and Physics*, 30:350–357, 2007.
- [11] Daniel Welfera, Jacob Scharcanskia, e Diane R Marinho. A coarse-to-fine strategy for automatically detecting exudates in color eye fundus images. *Computerized Medical Imaging and Graphics*, 34:228–235, 2009.
- [12] C. Jaya Kumaria e R. Maruthib. Detection of hard exudates in color fundus images of the human retina. *Elsevier*, 2011.
- [13] Doaa Youssef e Nahed H. Solouma. Accurate detection of blood vessels improves the detection of exudates in color fundus images. *Computer Methods and Programs in Biomedicine*, 108(3):1052–1061, 2012.
- [14] Nayomi Geethanjali Ranamuka e Ravinda Gayan N. Meegama. Detection of hard exudates from diabetic retinopathy images using fuzzy logic. *IET Image Processing*, 7:121–130(9), March 2013.
- [15] Mohamed Al-Shabrawey Ahmed M. Abu El-Asrar, Edoardo Midea e Ghulam Mohammad. New developments in the pathophysiology and management of diabetic retinopathy. *Journal of Diabetes Research*, 2013(424258):2 pages, 2013.
- [16] Zhi-Hua Zhou. *Ensemble Methods: Foundations and Algorithms*. Chapman & Hall/CRC, 1st edition, 2012.
- [17] Ian Goodfellow, Yoshua Bengio, e Aaron Courville. *Deep Learning*. MIT Press, 2016.
- [18] Guang-Bin Huang, Dian Hui Wang, e Yuan Lan. Extreme learning machines: a survey. *International Journal of Machine Learning and Cybernetics*, 2(2):107–122, 2011.
- [19] A. Fawzi, M. Davies, e P. Frossard. Dictionary learning for fast classification based on soft-thresholding. *International Journal of Computer Vision*, pages 1–16, 2014.
- [20] Xiaofeng Yang, Ning Wu, Guanghui Cheng, Zhengyang Zhou, David S. Yu, Jonathan J. Beitler, Walter J. Curran, e Tian Liu. Automated segmentation of the parotid gland based on atlas registration and machine learning: A longitudinal {MRI} study in head-and-neck radiation therapy. *International Journal of Radiation Oncology Biology Physics*, 90(5):1225 – 1233, 2014.
- [21] Bernhard Scholkopf e Alexander J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA, USA, 2001.

- [22] Thomas G. Dietterich. *Ensemble Methods in Machine Learning*, pages 1–15. Springer Berlin Heidelberg, Berlin, Heidelberg, 2000.
- [23] Angelou M.-Poulakou-Rebelakon E. Kalantzis, G. Diabetic retinopathy: An historical assesement. *Hormones*, 5(1):72–75, 2006.
- [24] Etienne Decencière, Xiwei Zhang, Guy Cazuguel, Bruno Lay, Bèatrice Cochener, Caroline Trone, Philippe Gain, Richard Ordonez, Pascale Massin, Ali Erginay, Bèatrice Charton, e Jean-Claude Klein. Feedback on a publicly distributed image database: The messidor database. *Image Analysis & Stereology*, 33(3):231–234, 2014.
- [25] Faraz Oloumi, Rangaraj M. Rangayyan, e Anna L. Ells. Computer-aided diagnosis of proliferative diabetic retinopathy via modeling of the major temporal arcade in retinal fundus images. *Journal of Digital Imaging*, 26(6):1124–1130, 2013.
- [26] A. Bosco, A. C. Lerario, e D. Soriano et al. Retinopata diabetica. *Arq Bras Endocrinol Melab*, 49(2):217–227, 2005.
- [27] Kalesnykiene V.-Kamarainen-J. Kauppi, T. Diaretdb1 diabetic retinopathy database and evaluation protocol. In *Proc of the 11th Conf. on Medical Image Understanding and Analysis*, 2009.
- [28] R. M. Rangayyan, F. Oloumi, F. Oloumi, P. Eshghzadeh-Zanjani, e F. J. Ayres. Detection of blood vessels in the retina using gabor filters. In *2007 Canadian Conference on Electrical and Computer Engineering*, pages 717–720, April 2007.
- [29] Akram Youssry, Ahmed El-Rafei, e Salwa Elramly. A quantum mechanics-based algorithm for vessel segmentation in retinal images. *Quantum Information Processing*, 15(6):2303–2323, 2016.
- [30] A. K. Dhara, R. M. Rangayyan, F. Oloumi, e S. Mukhopadhyay. Methods for the detection of blood vessels in retinal fundus images and reduction of false-positive pixels around the optic nerve head. In *E-Health and Bioengineering Conference (EHB), 2013*, pages 1–6, Nov 2013.
- [31] Xiaolu Zhu e R. M. Rangayyan. Detection of the optic disc in images of the retina using the hough transform. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3546–3549, Aug 2008.
- [32] S.S. HAYKIN e B. VAN VEEN. *Sinai's E Sistemas*. Bookman, 2001.
- [33] Ana Carolina Lorena e André CPLF de Carvalho. Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada*, 14(2):43–67, 2007.

- [34] N. Cristianini e J. Shawe-Taylor. *An Introduction to Support Vector Networks and other kernel-based learning methods*. Cambridge University Press, 2000.
- [35] Corinna Cortes e Vladimir Vapnik. Support-vector networks. *Mach. Learn.*, 20(3):273–297, September 1995.
- [36] A. R. G. Lima. Máquinas de vetores suporte na classificação de impressões digitais. Master’s thesis, Universidade Federal do Ceará, Fortaleza, Ceará, 2002.
- [37] Dustin Boswell. Introduction to support vector machines. 2002.
- [38] C. Campbell. *An introduction to kernel methods*. Springer Verlag, Berlin, 2000.
- [39] Jenny Balfer e Jurgen Bajorath. Visualization and Interpretation of Support Vector Machine Activity Predictions, May 2015.
- [40] L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.
- [41] Yoav Freund e Robert E. Schapire. Experiments with a new boosting algorithm, 1996.
- [42] P. Viola e M. Jones. Rapid object detection using a boosted cascade of simple features. pages 511–518, 2001.
- [43] Paul Viola e Michael Jones. Fast and robust classification using asymmetric adaboost and a detector cascade. In *Advances in Neural Information Processing System 14*, pages 1311–1318. MIT Press, 2002.
- [44] Paul Viola e Michael J. Jones. Robust real-time face detection. *Int. J. Comput. Vision*, 57(2):137–154, May 2004.
- [45] Fu Jie Huang, Zhihua Zhou, Hong-Jiang Zhang, e Tsuhan Chen. Pose invariant face recognition. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, pages 245–250, 2000.
- [46] Peter S. Carbonetto. Robust object detection using boosted learning, 2002.
- [47] Paul Viola e Michael Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57:137–154, 2004.
- [48] B.B. Chaves. Estudo do algoritmo adaboost de aprendizagem de máquina aplicado a sensores e sistemas embarcados. Master’s thesis, Universidade de São Paulo, 2012.
- [49] Robert E. Schapire e Yoram Singer. Improved boosting algorithms using confidence-rated predictions. *Machine Learning*, 37(3):297–336, 1999.

- [50] Jerome Friedman, Trevor Hastie, e Robert Tibshirani. Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, 28:2000, 1998.
- [51] Alexander Vezhnevets e Vladimir Vezhnevets. Modest adaboost' – teaching adaboost to generalize better. In *GraphiCon*, Novosibirsk Akademgorodok, Russia, 2005.
- [52] Emerson Lopes Machado. Redução de custo computacional em classificações baseadas em transformadas aprendidas. *Tese (Doutorado em Informática)-Universidade de Brasília*, 2015.
- [53] Jonathon Shlens. A tutorial on principal component analysis. *CoRR*, abs/1404.1100, 2014.
- [54] G. H. Golub e C. Reinsch. Singular value decomposition and least squares solutions. *Numerische Mathematik*, 14(5):403–420, 1970.
- [55] Carl Eckart e Gale Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936.
- [56] Magnus R. Hestenes. Inversion of matrices by biorthogonalization and related results. *Journal of the Society for Industrial and Applied Mathematics*, 6(1):51–90, 1958.
- [57] W. Kahan G. Golub. Calculating the singular values and pseudo-inverse of a matrix. *Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis*, 2(2):205–224, 1965.
- [58] Nobre G. G. Sandra. A decomposição em valores singulares e suas aplicações. *Universidade do Algarve - Faculdade de Ciências e Tecnologia*, 2007.
- [59] James W. Demmel. Matrix computations. *SIAM*, 32(4):690–691, 1990.
- [60] Alan Watt e Fabio Policarpo. *The Computer Image*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1998.
- [61] Bernd Jahne. *Digital Image Processing: Concepts, Algorithms, and Scientific Applications*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 4th edition, 1997.
- [62] T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.
- [63] C. J. Van Rijsbergen. *Information Retrieval*. Butterworth-Heinemann, Newton, MA, USA, 2nd edition, 1979.
- [64] R. C. Prati, Batista, e M. C. Monard. Curvas ROC para avaliação de classificadores. *Revista IEEE América Latina*, 2008.

- [65] Javeria Amin, Muhammad Sharif, e Mussarat Yasmin. A review on recent developments for detection of diabetic retinopathy. *Scientifica*, 2016, 2016.
- [66] Morium Akter e Mohammad Shorif. A review on automated diagnosis of diabetic retinopathy. *International Journal of Advanced Computer Technology*, 3(10):1161–1166, 2014.
- [67] Akara Sopharak, Bunyarit Uyyanonvara, e Sarah Barman. Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy c-means clustering, 2009.
- [68] F. Oloumi, R. M. Rangayyan, e A. L. Ells. Parabolic modeling of the major temporal arcade in retinal fundus images. *IEEE Transactions on Instrumentation and Measurement*, 61(7):1825–1838, July 2012.
- [69] E. Decenciere, G. Cazuguel, X. Zhang, G. Thibault, J.-C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Lay, e A. Chabouis. Teleophta: Machine learning and image processing methods for teleophthalmology. *{IRBM}*, 34(2):196 – 203, 2013. Special issue : {ANR} {TECSAN} : Technologies for Health and Autonomy.
- [70] M. Niemeijer, J.J. Staal, B. van Ginneken, M. Loog, e M.D. Abramoff. Comparative study of retinal vessel segmentation methods on a new publicly available database. In J. Michael Fitzpatrick e M. Sonka, editors, *SPIE Medical Imaging*, volume 5370, pages 648–656. SPIE, SPIE, 2004.
- [71] A. Hoover e M. Goldbaum. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE Transactions on Medical Imaging*, 22(8):951–958, Aug 2003.
- [72] J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, e E. Angelopoulou. Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. *IET Image Processing*, 7:373–383(10), June 2013.