

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO

JUAN HERBERT CHUCTAYA HUMARI

**Estudo do espectro Laplaciano na categorização
de imagens**

Ribeirão Preto
2016

JUAN HERBERT CHUCTAYA HUMARI

**Estudo do espectro Laplaciano na categorização
de imagens**

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo como parte das exigências para a obtenção do título de Mestre em Ciências.

Área de Concentração:
Física Aplicada à Medicina e Biologia.

Orientador:
Alexandre Souto Martinez.

Versão corrigida
Versão original disponível na FFCLRP-USP

Ribeirão Preto
2016

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

Juan Herbert Chuctaya Humari,

Estudo do espectro Laplaciano na categorização de imagens /
Juan Herbert Chuctaya Humari; orientador: Alexandre Souto
Martinez. - - Ribeirão Preto, 2016.

59 f. : il.

Dissertação (Mestrado) - - Faculdade de Filosofia, Ciências e
Letras de Ribeirão Preto, Universidade de São Paulo, 2016.

Inclui Bibliografia.

1. Análise de forma. 2. Processamento de imagens. 3. Teoria
de grafos. 4. Teoria espectral de grafos. 5. Espectro do
Laplaciano.

Nome: JUAN HERBERT CHUCTAYA HUMARI,

Título: Estudo do espectro Laplaciano na categorização de imagens

Dissertação apresentada à Faculdade de Filosofia,
Ciências e Letras de Ribeirão Preto da Universi-
dade de São Paulo como parte das exigências para
a obtenção do título de Mestre em Ciências.

Aprovado em: ____/____/____.

Banca Examinadora

Prof(a). Dr(a). : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof(a). Dr(a). : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof(a). Dr(a). : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Ao meus pais, Juan e Serafina. E em memoria de minha
querida amiga Natalia Biziak

Agradecimentos

Eu agradeço a meus pais Juan Chuctaya Sonco e minha mãe Serafina Humari e irmãos pelo apoio durante todos estes anos em minha formação como pessoa e profissional. Ao meu orientador Prof. Alexandre Souto Martinez pela oportunidade de trabalhar no Laboratorio de Modelagem de Sistemas Complexos. Também quero agradecer ao grupo de colaboração da IFSC-São Carlos, liderado pelo Prof. Odemir Martinez Bruno pela orientação em teorias de processamento de imagens fundamental para o trabalho.

A meus colegas do laboratório Gilberto Medeiros Nakamura, Fernando Meloni, Cristiano, Enock de Almeida Andrade Neto, Rayner Condori, Olavo Henrique Menin pela troca de conhecimentos e ajuda recebida.

Aos colegas do laboratório do Grupo de Computação Científica do IFSC, Marcos William da Silva Oliveira, Marina Jeaneth Machicao Justo pelas dicas fornecidas durante meu mestrado e o tempo que estive em São Carlos.

Aos amigos Tiago Alves, Carlitos Junior, Edwin Choquehuanca, Marina de Felipe, Jorge Vilchez, Tatiana Veloso, Priscila Rosse, Priscila Gomez, Daniele Maria, Daniel Gutierrez, Ana Gobbo, pela sua amizade e conselhos durante minha estadia em Ribeirão Preto e entre outros amigos que estou agradecido e foram de muita ajuda em instantes complicados.

Ao programa FAMB da FFCLRP, e aos professores e funcionários que direta e indiretamente colaboraram neste trabalho. Ao CNPq pela bolsa de mestrado.

Para conseguir grandes coisas, é necessário não apenas planejar, mas também acreditar; não apenas agir, mas também sonhar.

Anatole France

Resumo

JCHUCTAYA, N. D. **Estudo do espectro Laplaciano na categorização de imagens**. 2016. 59 f. Dissertação (Mestrado - Programa de Pós-Graduação em Física Aplicada à Medicina e Biologia) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2016.

Uma imagem engloba informação que precisa ser organizada para interpretar e compreender seu conteúdo. Existem diversas técnicas computacionais para extrair a principal informação de uma imagem e podem ser divididas em três áreas: análise de cor, textura e forma. Uma das principais delas é a análise de forma, por descrever características de objetos baseadas em seus pontos fronteira. Propomos um método de caracterização de imagens, por meio da análise de forma, baseada nas propriedades espectrais do laplaciano em grafos. O procedimento construiu grafos G baseados nos pontos fronteira do objeto, cujas conexões entre vértices são determinadas por limiares T_l . A partir dos grafos obtêm-se a matriz de adjacência A e a matriz de graus D , as quais definem a matriz Laplaciana $L = D - A$. A decomposição espectral da matriz Laplaciana (autovalores) é investigada para descrever características das imagens. Duas abordagens são consideradas: a) Análise do vetor característico baseado em limiares e a histogramas, considera dois parâmetros o intervalo de classes IC_l e o limiar T_l ; b) Análise do vetor característico baseado em vários limiares para autovalores fixos; os quais representam o segundo e último autovalor da matriz L . As técnicas foram testada em três coleções de imagens: sintéticas (Genéricas), parasitas intestinais (SADPI) e folhas de plantas (CNShape), cada uma destas com suas próprias características e desafios. Na avaliação dos resultados, empregamos o modelo de classificação *support vector machine* (SVM), o qual avalia nossas abordagens, determinando o índice de separação das categorias. A primeira abordagem obteve um acerto de 90% com a coleção de imagens Genéricas, 88% na coleção SADPI, e 72% na coleção CNShape. Na segunda abordagem, obtêm-se uma taxa de acerto de 97% com a coleção de imagens Genéricas; 83% para SADPI e 86% no CNShape. Os resultados mostram que a classificação de imagens a partir do espectro do Laplaciano, consegue categorizá-las satisfatoriamente.

Palavras-chave: 1. Análise de forma. 2. Processamento de imagens. 3. Teoria de grafos. 4. Teoria espectral de grafos. 5. Espectro do Laplaciano.

Abstract

JCHUCTAYA, N. D. **Study of the Laplacian spectrum in the categorization of images**. 2016. 59 f. Dissertation (M.Sc. - Postgraduate program in Physics Applied to Medicine and Biology) - Faculty of Philosophy, Sciences and Letters, University of São Paulo, Ribeirão Preto, 2016.

An image includes information that needs to be organized to interpret and understand its contents. There are several computational techniques to extract the main information of images and are divided into three areas: color, texture and shape analysis. One of the main of them is shape analysis, since it describes objects getting main features based on reference points, usually border points.

This dissertation proposes a shape analysis method based on the spectral properties of the Laplacian in graphs to represent images. The procedure builds G graphs based on object border points, whose connections between vertices are determined by thresholds T_i . From graphs G we obtain the adjacency matrix A and matrix degrees D , which define the Laplacian matrix $L = D - A$. Thus, spectral decomposition of the Laplacian matrix (eigenvalues) is investigated to describe image features. Two approaches are considered: a) Analysis of feature vector based on thresholds and histograms, it considers two parameters, classes range IC_i and threshold T_i ; b) Analysis of feature vector based on multiple linear for fixed eigenvalues, which represents the second and final eigenvalue matrix L . The techniques were tested in three image datasets: synthetic (Generic), human intestinal parasites (SADPI) and plant leaves (CNShape), each of these with its own features and challenges. Afterwards to evaluate our results, we used the classification model Support Vector Machine (SVM) to evaluate our approaches, determining the percentage of separation of categories. The first approach achieved 90% of precision with the Generic image dataset, 88% in SADPI dataset, and 72% in CNShape dataset. In the second approach, it obtains 97% of precision with the Generic image dataset, 83% for SADPI and 86 % in CNShape respectively. The results show that the classification of images from the Laplacian spectrum can categorize them satisfactorily.

Key-words: 1. Shape analysis. 2. Fourier transform. 3. Image processing. 4. Theory spectral graph. 5. Laplacian spectrum

Lista de Figuras

2.1	Uma visão geral das técnicas tradicionais de extração de características pela forma [1].	10
2.2	Onde a linha vermelha é o contorno do objeto e a elipse (<i>Linha preta</i>), que é o ajuste elíptico à forma.	11
2.3	A função e constitui-se pela soma das quatro funções $g(x) = f_1(x) + f_2(x) + f_3(x) + f_4(x)$ (Ref. [2]).	15
2.4	Grafo de incidência a partir de um grafo não orientado.	21
2.5	Imagem modelada em um grafo utilizando um raio $\sqrt{(2)}$ [3].	23
2.6	Imagem segmentada baseada em grafos, com informação de cores orientado para segmentação da imagem [4].	23
2.7	Detecção de características geométricas usando modelagem de grafos.	24
2.8	Grafo simples não orientado.	25
2.9	Possíveis hiperplanos de separação e hiperplano ótimo.	31
2.10	Transformação problema não linearmente separável em um problema linearmente separável	32
2.11	A figura apresenta a sequência do processo de classificação supervisionada utilizado neste trabalho. O conjunto de dados é separado em dois subconjuntos (conjunto de treino e conjunto de teste), sendo o primeiro utilizado na geração do modelo de classificação (treinamento) e o segundo para provas de teste.	34
3.1	Estrutura geral dos procedimentos desenvolvidos.	38
3.2	(a) Imagem binária, (b) Seguimento de contornos e (c) Contorno paramétrico obtido.	40
3.3	O conjunto S de elementos da borda do objeto característico.	41

3.4	Desenho do grafo $G(V, E)$ a partir da informação da fronteira do objeto.	41
3.5	A ilustração mostra o Grafo $E = W$ e o Grafo normalizado $E' = W'$.	43
3.6	Dado que T_l refere-se ao limiar, na ilustração mostra-se um vértice em verde e uma circunferência azul ao redor que define região de vizinhança definida por T_l , todo o que esteja inserido naquela região terá uma ligação com o vértice verde.	43
3.7	Evolução dinâmica do grafo gerado pelos limiares a) $T_l = 0.035$, b) $T_l = 0.065$, c) $T_l = 0.095$	44
3.8	As curvas representam os graus de cada um dos vértices do grafo, onde a linha preta corresponde à imagem original e a linha vermelha a imagem escalada em menor resolução. a) Sem normalizar e (b) Normalizado. ([5])	45
3.9	Exemplo de decompor do grafo regular, a partir de um limiar $T_l = 0.51$.	46
3.10	Exemplo do cálculo dos autovalores do espectro Laplaciano a partir da matriz Laplaciana de uma imagem do conjunto ShapeCN. (a) Contorno da imagem; (b) Espectro Laplaciano com um limiar $T_{0.75}$	48
3.11	Representação da histograma de um espectro do grafo $G(V, E)$ obtido de um limiar $T_l = 0.2$	50
3.12	Vetor característico $\beta = [\lambda_2(G_{t_1}), \lambda_n(G_{t_1}), \dots, \lambda_2(G_{t_2}), \lambda_n(G_{t_2})]$	51
3.13	Validação cruzada 10 pastas.	52
4.1	Coleção de imagens de formas genéricas.	54
4.2	Coleção de imagens SADPI.	55
4.3	Coleção de imagens ShapeCN.	56
4.4	IC_l vs T_l na coleção de imagens genéricas	57
4.5	Gráfico IC_l vs T_l , respeito da taxa de acerto.	58
4.6	Gráfico IC_l vs T_l , na coleção de imagens CNShape.	60
4.7	Segundo e último autovalor aplicado as imagens SADPI	63
4.8	Resultados do último autovalor na coleção de imagens SADPI.	64
4.9	Cálculos do segundo autovalor e ultimo autovalor nas imagens de folhas. 65	65

Lista de Tabelas

2.1	Representação de um grafo não orientado e sua matriz de adjacência.	20
2.2	Sumário dos principais Kernels nas SVMs	32
3.1	Representação de um grafo não orientado e sua matriz de adjacência A .	46
3.2	Representação de um grafo não orientado e sua matriz de graus D .	47
4.1	A tabela resume as melhores taxas de classificação obtidas na figura 4.4 por intervalo de classe na coleção de imagens genéricas.	57
4.2	Resumo dos melhores resultados obtidos por intervalo de classe IC_l vs T_l na coleção de imagens SADPI	59
4.3	Resumo dos melhores resultados CNShape.	60
4.4	Na tabela mostra-se as taxa de acerto, obtidas na classificação das categorias da coleção de imagens Genéricas.	66
4.5	Na tabela mostra-se a taxa de acerto na classificação de categorias da coleção de imagens SADPI. Onde observa-se uma maior taxa de acerto de 83% com 76 características.	67
4.6	Na tabela mostra-se os resultados da análise do segundo e último autovalor obtidos na evolução do grafo baseado nos limiares com a coleção de imagens CNShape.	68
4.7	Número de características por descritor	68
4.8	Na tabela mostra-se a taxa de acerto, obtida com a coleção de imagens Genéricas.	69
4.9	Na tabela mostra-se os resultados dos descritores utilizando a coleção de imagens SADPI.	69
4.10	Na tabela mostra-se os resultados da análise dos descritores com a coleção de imagens CNShape.	70

A.1	Análise de Vetor característico baseados em limiares e histogramas (Genérico). Parte 1	76
A.2	Análise de Vetor característico baseados em limiares e histogramas (Genérico). Parte 2	77
A.3	Análise de Vetor característico baseados em limiares e histogramas (SADPI). Parte 1	78
A.4	Análise de Vetor característico baseados em limiares e histogramas (SADPI). Parte 2	79
A.5	Análise de Vetor característico baseados em limiares e histogramas (CNShape). Parte 1	80
A.6	Análise de Vetor característico baseados em limiares e histogramas (CNShape). Parte 2	81
A.7	Matriz de confusão da melhor diferenciação entre as categorias, mediante o análise de vetor característico baseados em limiares e histogramas(banco de imagens Genérico) com $IC_l = 10$, $T_l = 0.475$ e taxa de acerto de 91%.(11 imagens por cada classe)	82
A.8	Análise de vetor característico baseados em limiares e histogramas com o banco de imagens (SADPI)($IC_l = 30$, $T_l = 0.575$ e 88 % de acerto)	82
A.9	Análise de vetor característico com limiar $T_l = 0.575$ e intervalo de classe $IC_l = 30$ obtendo um 72 % (CNShape).	83
A.10	Análise de vetor característico com limiar $T_l = 0.575$ e intervalo de classe $IC_l = 30$ obtendo um 72 % (CNShape).	84
A.11	Matriz de confusão obtida com limiares ($T_{inicial} = 0.025$, $T_{incremento} = 0.2$, $T_{final} = 0.95$), no qual se obteve uma taxa de acerto de 97 % no banco de dados Genérico	86
A.12	Matriz de confusão obtida com limiares ($T_{inicial} = 0.025$, $T_{incremento} = 0.025$, $T_{final} = 0.95$), no qual se obteve uma taxa de acerto de 83 %, com o banco SADPI	86
A.13	Matriz de confusão obtida ($T_0 = 0.025$, $T_{inc} = 0.025$, $T_{final} = 0.95$), taxa de acerto de 86% (CNShape)	87

A.14 Matriz de confusão obtida ($T_0 = 0.025, T_{inc} = 0.025, T_{final} = 0.95$), taxa de acerto de 86% (CNShape)	88
---	----

Sumário

Lista de Figuras	xv
Lista de Tabelas	xix
1 Introdução	1
2 Fundamentos teóricos	7
2.1 Análise de imagens	7
2.2 Métodos de extração de características ligadas à forma	11
2.2.1 Medição da elipticidade	11
2.2.2 Descritores de Fourier	14
2.2.3 Curvatura	16
2.2.4 Momentos de Hu	16
2.3 O uso de grafos na descrição de imagens	18
2.4 Teoria espectral de grafos	24
2.4.1 Espectro do Laplaciano	25
2.4.2 Autovalores e autovetores	28
2.5 Reconhecimento de padrões	29
2.5.1 Support vector machine	29
2.5.2 Validação cruzada (cross-validation)	33
2.5.3 Precisão e revocação	34
3 Metodologia desenvolvida	37
3.1 Pré-processamento da imagem	37
3.2 Construção dos grafos	40
3.3 Análise espectral dos grafos	46

3.4	Assinatura ou características	48
3.4.1	Vetor característico baseado em limiares e a histogramas . . .	49
3.4.2	Vetor característico baseado em vários limiares para autova- lores fixos	49
3.5	Máquina de aprendizagem	51
4	Experimentos e Resultados	53
4.1	Coleções de imagens usados como objetos de estudo	53
4.2	Análise do vetor característico baseados em limiares e a histogramas .	55
4.3	Análise do vetor característico baseado em vários limiares para auto- valores fixos	62
4.4	Comparação com outros descritores	65
5	Conclusões e Discussão	71
	Apêndice A - Anexos	75
A.1	Análise do vetor característico baseados em limiares e a histogramas .	75
A.2	Análise do vetor característico baseado em vários limiares para auto- valores fixos	85
	Referências	89

Introdução

Atualmente, imagens digitais desempenham um papel fundamental como fonte de informação em diversas áreas do conhecimento. Por exemplo na Biologia, imagens digitais são utilizadas para o estudo detalhado de estruturas microscópicas da fauna e flora, ou para detectar variações e deformidades anatômicas em organismos maiores [6]. Na Medicina, imagens digitais obtidas em ambiente hospitalar ajudam na visualização de padrões, anomalias e outros tipos de informação, auxiliam no diagnóstico precoce de diferentes doenças, como células e nódulos cancerígenos [7]. Basicamente, elas são analisadas por especialistas altamente qualificados, que identificam e interpretam os padrões observados baseando-se em conhecimentos previamente estudados.

Embora a qualidade dos resultados obtidos pelos especialistas seja satisfatória na análise das imagens, aspectos como a demasiada carga laboral e falta de especialistas em hospitais afastados das grandes capitais, implicam em obter um melhor aproveitamento desta tecnologia [7–10]. De fato, o treinamento de especialistas para analisar propriedades clínicas ou biológicas nos hospitais, envolve altos custos e as habilidades técnicas adquiridas tendem a ser restritas a um pequeno conjunto de situações. Isto tem estimulado o desenvolvimento de técnicas computacionais para análise de imagens com diferentes objetivos, tanto específicos como

mais gerais. Tais algoritmos visam reconhecer e comparar padrões em imagens, e as informações extraídas são comparadas com modelos previamente definidos. Em linhas gerais, as técnicas computacionais são capazes de extrair propriedades numéricas das imagens, as quais podem descrever padrões e detalhes representativos de cada imagem, formando um vetor característico *n dimensional*. Essa representação alternativa de imagens, algumas vezes também permite a recuperação das imagens originais e pode ser armazenada em uma base de dados para análises adicionais.

Em uma interpretação mais ampla, as características das imagens podem guardar informações do conteúdo textual, ou visual tais como cor, textura e a forma. Dentre esses atributos, a forma é considerada importante, porque nós utilizamos esta como aspecto primário para o reconhecimento de padrões [11]. Por esse motivo, muitas técnicas computacionais que analisam padrões em imagens são específicas para a detecção da forma do objeto representado, e devem permitir a correta classificação da imagem avaliada em uma categoria específica. Para este fim, são necessárias informações de imagens de referência, representadas por um objeto ou conjunto de objetos que compartilham a mesma forma daquele representado nas imagens de estudo. [1, 12, 13].

Segundo Pavlids [14], a análise de formas por algoritmos pode ser dividida em duas linhas de estudos principais: a) descritores de forma baseados em região, que usam como informação os descritores de momento obtidos a partir da imagem do objeto, tais como os momentos geométricos [15, 16], momentos de Zernike [17], momentos de Legendre [17], momentos de Tchebichef [18]; e (b) descritores de forma baseados em fronteira, as quais determinam as regiões de borda [19], tais como os descritores de Fourier [20], curvatura [19], descritores *Wavelet* [21], entre outros [14, 22]. Embora as técnicas de reconhecimento de formas em imagens tenham

avançado significativamente nos últimos anos [23, 24], o uso de algoritmos computacionais para a análise e interpretação de imagens ainda demanda aprimoramento técnico em muitas áreas. Isso deve-se ao fato de que os métodos computacionais devem obedecer diferentes condições para assegurar sua robustez, como por exemplo, a invariância à transformação geométrica do objeto (mudanças na escala, rotação, e obliquidade) [12, 25]. Como regra, o desenvolvimento de novas técnicas prima sempre por obter bons resultados de identificação e classificação de padrões, usando o menor esforço computacional possível.

Neste contexto, as imagens de ressonância, ultrassom e tomografia computacional, apresentam casos e ferramentas em que os profissionais de saúde por meio de procedimentos computacionais, permitem a detecção de padrões e anomalias em estruturas internas, do organismo humano. Aqui a observação da forma é importante para a interpretação dos resultados das imagens [22, 26–28]. Na biologia e análises clínicas, as técnicas computacionais são aplicadas em imagens de micrografias, buscando a identificação de micro-organismos patogênicos e parasitas [29], ou por busca de anomalias no padrão celular de diferentes órgãos [30] para determinar doenças.

Portanto, fica claro que as técnicas computacionais apresentam um grande potencial de aplicação prática, e que a forma talvez seja um dos aspectos mais interessantes em Medicina e Biologia. O objetivo geral desta dissertação é a análise e reconhecimento de padrões em imagens baseado no atributo de forma, através da matriz de adjacência e matriz de graus obtidas deste grafo, conseqüentemente pode-se caracterizar a forma do objeto. A abordagem utilizada é baseada no estudo da fronteira do objeto de interesse. O método que propomos utiliza a teoria do espectro Laplaciano para detecção de padrões e permite a correta classificação

das imagens [31]. Basicamente, utiliza-se a sequência de pixels obtidos da borda do objeto, como base para representar um grafo. Em seguida define-se a matriz Laplaciana desse grafo e com base na análise do espectro dos autovalores dessa matriz, definimos o descritor baseado nesses valores, sendo a assinatura representativa em uma imagem dada.

Para avaliar se nossas abordagens conseguem distinguir as imagens entre as diferentes categorias, utilizamos três coleções de imagens as quais dividem-se em distintos níveis de complexidade desde um básico até um nível complexo. A primeira coleção de imagens geradas ou sintéticas, possuem formas genéricas bem separáveis, sendo nossa primeira frente de avaliação. A segunda coleção de imagens, tem uma complexidade maior e maior e representam espécies de parasitas humanos chamadas helmintos, esta coleção caracteriza-se pela geometria oval dos parasitas. A terceira coleção de imagens de folhas de plantas, tem uma complexidade maior e casos de similaridade entre as categorias que extraíndo informações específicas é possível distinguir e categorizar as imagens. A avaliação da eficácia de nosso método consiste em medir a taxa de acerto na categorização, baseada nas características da imagem a partir do espectro Laplaciano. Os resultados obtidos pelo método proposto serão comparados com outros métodos tradicionais.

Esta dissertação está organizada da seguinte maneira: No Capítulo 2, é apresentado o contexto teórico fundamental para o desenvolvimento da metodologia proposta. Seguidamente, no Capítulo 3, é apresentada a metodologia desenvolvida e no Capítulo 4 apresentamos os resultados. Finalmente no Capítulo 5, se apresentam as conclusões e possíveis trabalhos futuros.

Como observação ao leitor, no Capítulo 2 faz-se uma simples revisão teórica das técnicas computacionais usadas no presente estudo. O objetivo desse capítulo é

apenas introduzir alguns conceitos aos leitores menos familiarizados com o assunto e, portanto, para os leitores que dominam esses conceitos, sugerimos o reinício da leitura no Capítulo 3.

Fundamentos teóricos

Neste capítulo, inicialmente definimos os conceitos básicos de processamento de imagens, e em seguida, estabelecemos uma revisão da teoria de grafos. Ao final, conceitos sobre o espectro Laplaciano, autovalores e autovetores são apresentados.

2.1 Análise de imagens

O processamento numérico de imagens é a manipulação de uma imagem por métodos computacionais, de modo que, tanto os dados de entrada, como os de saída, são imagens. A alteração minimiza ou maximiza as propriedades da imagem sujeito ao padrão de interesse.

Do ponto de vista computacional, uma imagem digital $I(m, n)$ é o resultado de um processo de discretização (*sampling*) de uma função contínua, sendo armazenada em computadores como um arranjo bidimensional, uma matriz de dimensões N_x e N_y , onde cada elemento da imagem (m, n) (com $m = 1 \dots N_x$ e $n = 1 \dots N_y$), corresponde a um pixel com intensidade igual a $I(m, n)$. O valor da intensidade de pixel pode ser representada em distintos espaços de cor [32]. Geralmente o espaço de cor é o *RGB* ou tons de cinza. O espaço *RGB* usa três valores para representar

um pixel (vermelho, verde e azul), cada um destes valores varia entre 0 até 255. Os tons de cinza só usam um valor para representar um pixel e o valor deste varia entre 0 até 255. Em imagens originalmente coloridas, a conversão prévia para tons de cinza se faz necessária, para utilização dos métodos propostos nessa dissertação.

A identificação de padrões em imagens digitais usa a intensidade dos pixels para encontrar regiões de interesse, as quais podem ser representadas em uma sequência numérica denominada *vetor de características*. A obtenção do vetor de características têm como objetivo reduzir a quantidade da informação da imagem, simplificando assim a aplicação dos algoritmos de classificação de padrões. Essa otimização é importante porque extrai só propriedades importantes da imagem inteira.

Filtrar e identificar as características mais relevantes das imagens são etapas importantes para codificar e gerar uma representação numérica adequada. Por exemplo, se tivermos uma imagem 64×64 pixels, o número total de pixels é 4096, o que apresenta muita informação relevante e também irrelevante. Contudo, se soubermos previamente que uma região de 12×12 pixels é suficiente para análise, teríamos apenas 144 pixels, o que significaria um conjunto de dados expressivamente menor a processar.

Os algoritmos que descrevem esta informação determinam medidas quantitativas e correlação dos pixels. Tipicamente, um descritor reduz a dimensionalidade de uma imagem de $\mathfrak{R}^{N_x \times N_y}$, onde N_x e N_y são a largura e altura na imagem (em pixels), em um vetor de informações \mathfrak{R}^n , de tamanho n que representa as características extraídas, e que depende do domínio e propriedades da imagem que deseja-se caracterizar [33].

Existem diversos algoritmos computacionais para extrair informação das

imagens denominados descritores de características [33,34]. Sonka et al. [35] definem a extração de características como a combinação das etapas de descrição de imagens e seleção de características. Embora o conjunto de técnicas conhecidas é baseado em cor, textura e forma [36,37], nesta dissertação tratamos apenas das técnicas para obtenção de descritores de forma. Os descritores de forma são amplamente utilizados para descrever, ou representar, objetos em imagens. A forma é uma propriedade importante para reconhecer um objeto no mundo real [38]. Assim, a forma representa um atributo de relevância no momento de caracterizar elementos representativos da imagem. Encontrar a forma do objeto de interesse na imagem significa isolar o objeto do meio onde está inserido. Por isso, a determinação da forma depende de uma etapa prévia de processamento, chamada segmentação, a qual divide a imagem em regiões homogêneas, a partir das quais pode-se extrair as fronteiras das regiões ou objetos de interesse. Descritores de forma baseiam seus princípios em duas abordagens, de contorno e de região da forma, como exemplificado na Figura 2.1.

Os descritores baseados em contorno expressam as propriedades das formas pelo seu esboço, que constitui a cadeia de pixels da fronteira de imagem considerado [14]. Estes descritores seguem a especificação do MPEG-7 [39], que estabelece propriedades importantes deste abordagem. Entre as principais propriedades do MPEG, o descritor deve ser eficiente em aplicações em que uma alta variabilidade nas formas é esperada, o que significa deformações do objeto e robustez ao ruído presente no contorno.

Os descritores de similaridade, baseados numa dada região utilizam a informação da quantidade dos pixels contidos dentro da fronteira objeto. As técnicas baseadas em momentos são as principais representantes desse tipo de descritor.

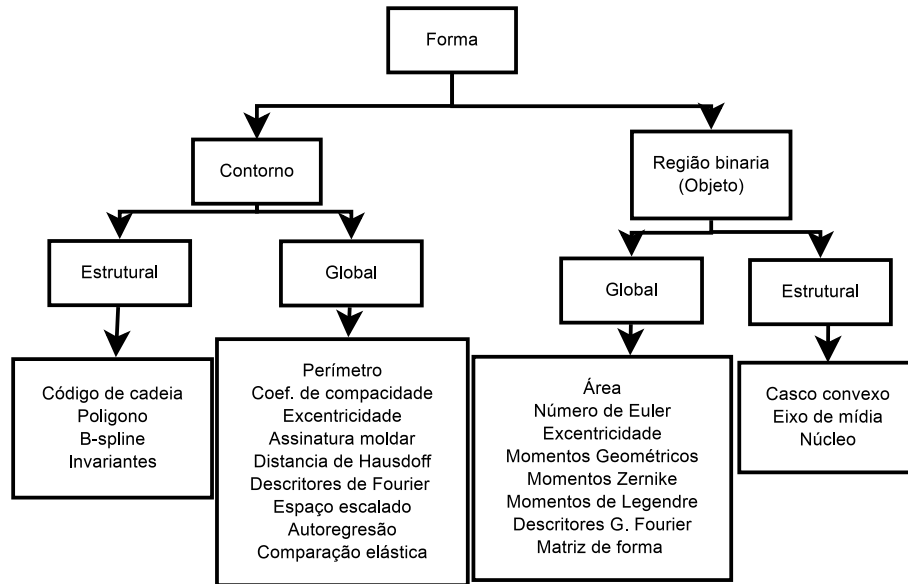


Figura 2.1 – Uma visão geral das técnicas tradicionais de extração de características pela forma [1].

Dentre as características que este tipo de descritor deve possuir, destacamos aquela que especifica que este descritor deve fornecer uma maneira compacta e eficiente de descrever propriedades de múltiplas regiões disjuntas simultaneamente, além de robustez a ruídos [35].

Recentemente, uma nova tendência de descritores de características baseados em grafos [40–42], foi proposta e têm sido utilizada com sucesso na descrição e processamento de imagens digitais. Esse tipo de abordagem tem ganhado importância devido sua capacidade de disponibilizar informações de naturezas distintas, de forma compacta e robusta. Embora o uso de grafos na análise de padrões em imagens apresente enorme potencial, o tópico ainda carece de estudos que detalhem melhor quais os procedimentos que proporcionam os melhores resultados na identificação de padrões.

2.2 Métodos de extração de características ligadas à forma

Nesta secção, apresentamos alguns metodos tradicionais utilizados como descritores de forma em imagens.

2.2.1 Medição da elipticidade

A medida da elipticidade é importante em sistemas de visão computacional [43]. Medidas tradicionais como convexidade, orientação e alongação são popularmente utilizadas, assim como medidas de primitivas geométricas, elipticidade e ajuste elíptico são motivadas por problemas de processamento de imagens do mundo real [44]. A medição de elipticidade consiste na comparação de duas formas, a borda da região de interesse da imagem digital e o ajuste elíptico formado a partir dos pontos da fronteira da região, a diferença entre os dois representa um valor característico [45].

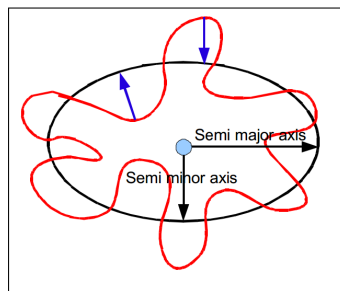


Figura 2.2 – Onde a linha vermelha é o contorno do objeto e a elipse (*Linha preta*), que é o ajuste elíptico à forma.

Existem diversos métodos de ajuste elíptico descritos [43, 46]. Destes um dos mais conhecidos e utilizados é o método *direct least square fitting of ellipse* [43]. A principal característica deste método é sua eficiência computacional e robustez, com bons resultados em situações que tenham ruído e é invariante à orientação [43].

O método baseia-se na equação Elíptica 2.1,

$$F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (2.1)$$

a qual pode ser expressada em sua forma vetorial,

$$F_a = \vec{x} \cdot \vec{a} = 0 \quad (2.2)$$

onde $\vec{a} = [a \ b \ c \ d \ e \ f]^T$ é o vetor dos coeficientes da equação, e $\vec{x} = [x^2 \ xy \ y^2 \ x \ y \ 1]$ é o vetor das coordenadas de pontos. Ambos vetores são relativos à seção cônica. Considerando que temos um conjunto de coordenadas da fronteira do objeto de interesse:

$$T = (x_i, y_i) : i \dots N \quad (2.3)$$

O objetivo desta técnica é minimizar a soma dos quadrados da distância da Eq. 2.1 aos pixels que representam as bordas (Eq. 2.3). Em outras palavras, admitimos $F(a, p_i)$ como a distância algébrica a partir da coordenada pixel $P_i = (x_i, y_j)$ a uma curva cônica expressa por Eq. 2.2 o problema da minimização não linear é demonstrado na Ref. [47]:

$$\min_a \left(\sum_{i=1}^N F(a, p_i) \right) = \min_a \left(\sum_{i=1}^N F(a \cdot p_i)^2 \right) \quad (2.4)$$

Na Ref. [43] é demonstrado que o problema pode ser solucionado a partir de determinadas restrições, resultando em uma única solução exata:

$$\left\{ \begin{array}{l} \min \|D \cdot a\|^2 \\ a^T \cdot C \cdot a = 1 \end{array} \right\} \quad (2.5)$$

tal que,

$$D = \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ x_N^2 & x_N y_N & y_N^2 & x_N & y_N & 1 \end{pmatrix} \quad (2.6)$$

e

$$C = \begin{pmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.7)$$

O extrator de medição de elipticidade baseia-se em 7 características descritas como segue :

- Centro da direção do eixo x.
- Centro da direção do eixo y.
- Abcissa menor
- Abcissa maior
- Ângulo de orientação
- Excentricidade

- Area

Representam as características geralmente empregadas pelo extrator.

2.2.2 Descritores de Fourier

A transformada de Fourier é uma técnica de ampla aplicação no processamento digital de sinais e imagens (sinais bidimensionais). Esta técnica permite representar um determinado sinal periódico por meio da soma de senos e cossenos de diferentes frequências, cada um multiplicado por um coeficiente próprio (séries de Fourier)(Fig. 3.7). A mudança de domínio do tempo (ou espaço) para o domínio de frequência (temporal ou espacial) pode facilitar o processamento das informações, pois permite que um sinal seja estudado de acordo com seus diferentes grupos de frequências, cada qual relacionada com as características específicas do sinal [2].

A transformada de Fourier de uma função discreta $f(x)$ é definida por,

$$F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{-j2\pi \frac{ux}{M}} \quad (2.8)$$

e sua inversa

$$F(x) = \sum_{u=0}^{M-1} F(u) e^{j2\pi \frac{ux}{M}} \quad (2.9)$$

onde $u, x = 0, 1, \dots, M - 1$ e $j = \sqrt{-1}$ tem que é o número imaginário puro. No caso de imagens (estruturas 2D), a transformada de Fourier e sua inversa são dadas respectivamente por:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi \left(\frac{ux}{M} + \frac{vy}{N} \right)} \quad (2.10)$$

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi \left(\frac{ux}{M} + \frac{vy}{N} \right)} \quad (2.11)$$

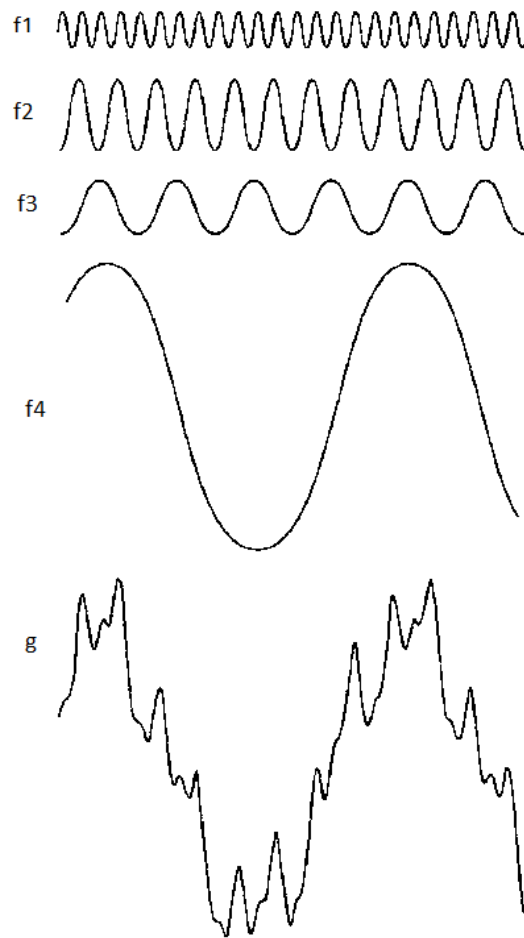


Figura 2.3 – A função e constitui-se pela soma das quatro funções $g(x) = f_1(x) + f_2(x) + f_3(x) + f_4(x)$ (Ref. [2]).

A transformada discreta de Fourier apresenta a desvantagem de deslocar os coeficientes de frequências mais baixas para as extremidades do espectro. Assim, é comum realizar uma operação de deslocamento, que consiste em mover a origem da transformada de Fourier para o centro das coordenadas de frequência $(\frac{M}{2}, \frac{N}{2})$. Normalmente coeficientes de baixa frequência estão associados à porção do espectro que descreve as informações mais importantes sobre o comportamento do sinal, enquanto as variações de alta frequência são associadas com variações bruscas tal como em bordas e ruído.

2.2.3 Curvatura

A curvatura é uma propriedade diretamente relacionada à característica da forma. O descritor de curvatura é uma curva associada às regiões côncavas e convexas do contorno do objeto [25]. Pontos de máximo e mínimo locais dessa curva correspondem às mudanças de direção no contorno, a curvatura trata as coordenadas paramétricas $C(t) = (x(u), y(u))$, em uma curvatura $k(t)$ que é definida em termos de derivada. Esta técnica utiliza a convolução de uma função gaussiana $g(t, \sigma)$ no sinal, reduzindo assim os efeitos provenientes do ruído ou das informações de alta frequência, antes do cálculo da curvatura:

$$k(s, \sigma) = \frac{\dot{X}(s, \sigma)\ddot{Y}(s, \sigma) + \ddot{X}(s, \sigma)\dot{Y}(s, \sigma)}{(\dot{X}(s, \sigma)^2 + \dot{Y}(s, \sigma)^2)^{\frac{3}{2}}} \quad (2.12)$$

onde $\dot{X}(s, \sigma) = x(s) * \dot{g}(s, \sigma)$, $\ddot{X}(s, \sigma) = x(s) * \ddot{g}(s, \sigma)$, $\dot{Y}(s, \sigma) = y(s) * \dot{g}(s, \sigma)$, $\ddot{Y}(s, \sigma) = y(s) * \ddot{g}(s, \sigma)$, $*$ é o operador de convolução, e $g(s, \sigma)$ é uma função gaussiana de desvio padrão σ , enquanto que $\dot{g}(s, \sigma)$ e $\ddot{g}(s, \sigma)$ são a primeira e segunda derivada de $g(s, \sigma)$ respectivamente.

Para este estudo utilizamos três características para extrair propriedades importantes na curvatura, como valor médio, desvio padrão e entropia.

2.2.4 Momentos de Hu

A técnica de momentos, também chamados momentos estatísticos, é um dos métodos tradicionais utilizados para extração de características de uma imagem. Estes momentos e as funções derivadas deles se caracterizam por valores numéricos calculados a partir da imagem binária previamente segmentada e que descrevem a distribuição espacial dos pontos contidos na imagem ou em uma região.

Os momentos mais utilizados numa imagem são os momentos regulares que

são definidos como,

$$m_{pq} = \sum_1^{nx} \sum_1^{ny} x^p y^q f(x, y) \quad (2.13)$$

onde m_{pq} é o momento de ordem $(p + q)$ da função intensidade $f(x, y)$. As variáveis nx e ny representam respectivamente a largura e a altura da imagem digital. Uma imagem binária terá valores da função $f(x, y)$ iguais a 0 ou 1.

Assim a partir dos momentos regulares encontram-se medidas importantes sobre os objetos de interesse e que são úteis na identificação de diferentes formas. Por exemplo, os momentos regulares de ordem 0 e 1 são usados para o cálculo do baricentro ou centro de massa do objeto. Assim temos a operação

$$x_c = \frac{m_{10}}{m_{00}}; y_c = \frac{m_{01}}{m_{00}} \quad (2.14)$$

que representa este cálculo. Com a informação obtida dos baricentros obtemos o que chamamos de momentos centrais que são definidos para imagens digitais pela fórmula:

$$u_{pq} = \sum_1^{nx} \sum_1^{ny} (x - x_c)^p (y - y_c)^q f(x, y) \quad (2.15)$$

Também podemos representar os momentos centrais normalizados representados por n_{pq} e definidos pela seguinte fórmula:

$$n_{pq} = \frac{u_{pq}}{m_{00}^{\frac{(p+q)}{2} + 1}} \quad (2.16)$$

Após a combinação de momentos que são invariantes a escala, rotação e translação. Estes momentos são 7 equações chamados momentos de Hu, que são as

características descritas da seguinte forma,

$$\begin{aligned}
I_1 &= n_{20} + n_{02} \\
I_2 &= (n_{20} - n_{02})^2 + (2n_{11})^2 \\
I_3 &= (n_{30} - 3n_{12})^2 + (3n_{21} - n_{03})^2 \\
I_4 &= (n_{30} + n_{12})^2 + (n_{21} - n_{03})^2 \\
I_5 &= (n_{30} - 3n_{12})(n_{30} + n_{12})[(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] \\
I_6 &= (n_{20} - n_{02})[(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] + 4n_{11}(n_{30} + n_{12})(n_{21} + n_{03}) \\
I_7 &= (3n_{21} - n_{03})(n_{30} + n_{12})[(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2] + (n_{30} - 3n_{12})(n_{21} + n_{03}) \\
&\quad [3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2]
\end{aligned} \tag{2.17}$$

2.3 O uso de grafos na descrição de imagens

Os Grafos são estruturas que podem representar imagens através das regiões que compõem e seus relacionamentos. Assim prévios cálculos de processamento de imagens obtêm os padrões de interesse, e a partir deste determinam-se as regiões. As regiões agrupam conjuntos de pixels com as mesmas características na imagem [48, 49]. Neste trabalho, são demonstradas técnicas que utilizam a teoria dos grafos na recuperação de informações por conteúdo. Por este motivo, serão apresentadas algumas definições básicas.

O *grafo*, é uma estrutura de dados descrita como um conjunto de *vértices* (ou *nós*) conectados por meio de *arestas* (ou *arcos*). Grafos podem ser utilizados para representar uma infinidade de modelos, com a vantagem de que suas propriedades e métricas podem ser utilizadas para interpretar o mundo real. Um exemplo dessa abordagem, são as propriedades estruturais dos grafos para o modelagem de

redes sociais, computacionais, transporte, internet, estrutura de moléculas, projeto de circuitos, dentre outros [50].

Definição 1 Um grafo é definido por $G = (V, E)$ sendo uma estrutura composta por um conjunto não vazio de vértices V , onde $V = \{v_i \mid i = 1 \dots N\}$, e por um conjunto de arestas (ou arcos) $E = \{(v_i, v_j) \in V \times V\}$, de pares não ordenados de elementos distintos de V .

Definição 2 Uma aresta conecta dois vértices $v_1, v_2 \in E$ e estes são considerados adjacentes. Se todos os vértices de $G(V, E)$ são adjacentes entre si, então G é chamado completo ou regular.

Definição 3 O valor da ligação entre dois vértices é definido por um valor $w(v_i, v_j)$, que representa o peso da aresta que conecta dois vértices e este valor obedecer uma métrica ou critério estabelecido.

Definição 4 Um grafo é dito direcional, ou dígrafo, quando é necessário estabelecer um sentido (orientação) para as arestas. O sentido da aresta é indicado através de uma seta.

Definição 5 Um grafo não orientado é um grafo simétrico, de maneira que para cada ligação $E(v_i, v_j)$ existe uma ligação $E(v_j, v_i)$. Um grafo finito o número total de vértices é o dobro do número de arestas.

Por último, revisamos a adjacência, que indica quando dois vértices estão ligados entre si (então eles são adjacentes, ou vizinhos [51]).

Definição 6 Uma das formas mais utilizadas para representar grafos é via a **Matriz de adjacência**. Seja $A = [a_{ij}]$ uma matriz $N \times N$, onde N é o número de vértices de um grafo $G(V, E)$, que representa o grafo [51], sendo construída da

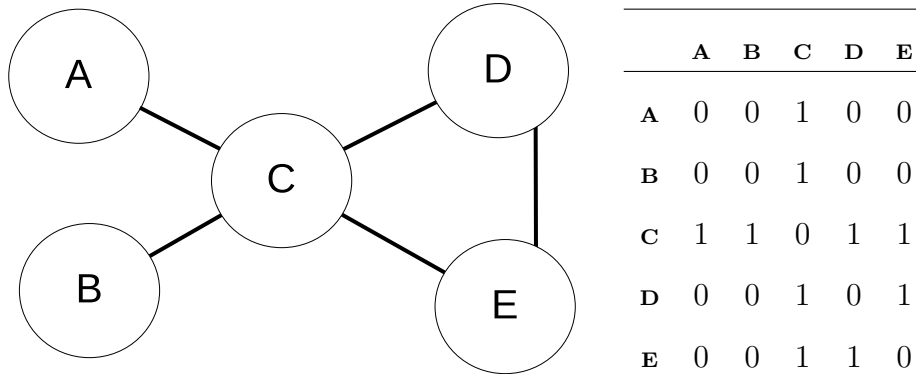


Tabela 2.1 – Representação de um grafo não orientado e sua matriz de adjacência.

seguinte forma:

$$A(i, j) = \begin{cases} 1 & \text{if } (i, j) \in E \\ 0 & \text{outro caso} \end{cases} \quad (2.18)$$

$A(i, j)$ é dita matriz de adjacência do grafo $G = (V, E)$. Um exemplo da matriz de adjacência para um grafo não orientado é ilustrado na Figura 3.1,

Definição 7 Outra definição importante para grafos não orientados é o grau do vértice definido como $D(v)$, o qual representa o número total de ligações do nó. O grau é definido, que nele da idéia de grafo orientado.

$$D(v_i) = \sum_{j \in V} A(i, j) \quad (2.19)$$

Definição 8 A matriz de incidência de um grafo, fornece uma descrição das relações de incidência das arestas nos vértices do grafo. Esta é uma representação computacional de um grafo através de uma matriz bidimensional, em que temos duas dimensões vértices n e a arestas m . Ela guarda informações sobre como os vértices se relacionam com cada aresta (isto é, informações sobre a incidência de

um vértice em uma aresta). Para representar um grafo sem pesos nas arestas e não direcionado, basta que as entradas da matriz de incidência Q sigam esta definição,

$$q_{ij} = \begin{cases} 1 & \text{Se } e_j \text{ uma aresta incidente no vértice } v_i \\ 0 & \text{caso contrário} \end{cases} \quad (2.20)$$

assim para este grafo 2.4, temos a matriz Q de incidência.

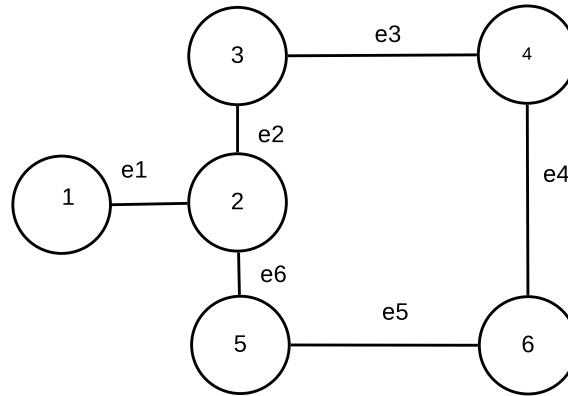


Figura 2.4 – Grafo de incidência a partir de um grafo não orientado.

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \quad (2.21)$$

A representação utilizada através dos grafos depende das informações que são extraídas das imagens. Nas aplicações desenvolvidas em imagens, os vértices contém informação para identificar as regiões relevantes da imagem obtidas pela

agrupação ou segmentação (processamento da imagem) e as arestas são os relacionamentos entre as regiões.

Representação de Imagens Através de Grafo

Os grafos são estruturas simples que podem representar imagens complexas através da informação de suas regiões e seus relacionamentos. As regiões consistem em qualquer conjunto de pixels da imagem que possuem uma mesma característica na percepção dos usuários e na aparência da imagem. Podemos apresentar esta informação em grafos relacionais que possuem atributos e a sua aplicabilidade na representação baseadas no conteúdo das imagens. Desta forma, um atributo é uma tupla (t, a) que consiste de um tipo (t) e um valor (a) associado. Os atributos utilizados para caracterizar as informações da imagem, dependem da necessidade e do objetivo da aplicação. Assim um atributo da imagem que pode ser representado pelo vértice do grafo que poderia ser:

- **Nível de cinza:** Cada pixel é mapeado em um vértice do conjunto V . As arestas são definidas pela distância Euclidiana. Dado dois vértices v_i e v_j , relacionados aos pixels p_i e p_j . Os vértices são conectados se a distância espacial entre os pixels p_i e p_j é menor que um dado raio r e o peso é definido pela diferença entre a intensidade dos pixels (Fig. 2.5).
- **Cor:** A informação de cor nas imagens, utilizam grafos para determinar as regiões na imagem (Fig. 2.6).
- **Textura:** A formação do grafo a partir da informação de textura baseia-se a partir das análises previamente feitas na imagem de entrada, por técnicas de extração de características locais (Transformada de Fourier, fractais, etc) . Assim os vértices e arestas serão fornecidos por estas transformações prévias.

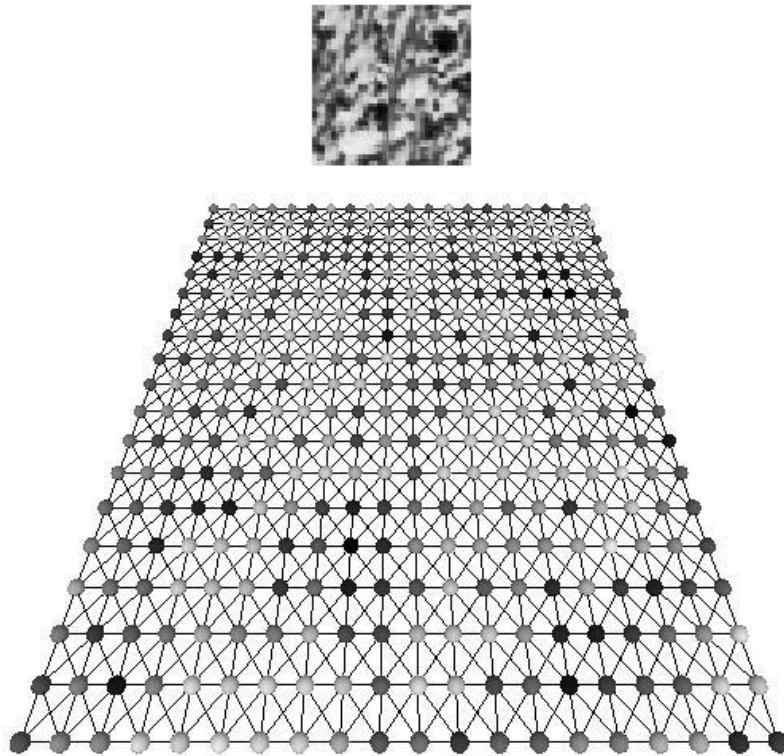


Figura 2.5 – Imagem modelada em um grafo utilizando um raio $\sqrt{2}$ [3].

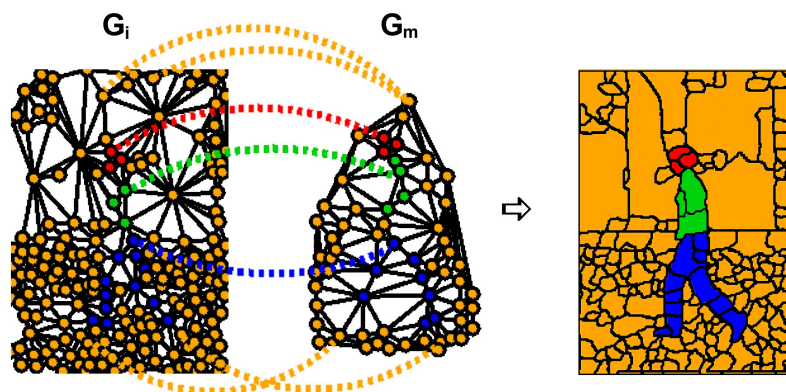


Figura 2.6 – Imagem segmentada baseada em grafos, com informação de cores orientado para segmentação da imagem [4].

- Características geométricas:

Baseiam-se na extração de propriedades do grafo. Critérios e medidas como número de triangulações, vizinhanças, distância entre os vértices de um grafo são importantes para procedimentos para aplicações de reconhecimentos de

rostros, impressões digitais [52,53]. Onde a localização dos pontos de interesse é fundamental para formar o grafo característico (Fig. 2.7) .

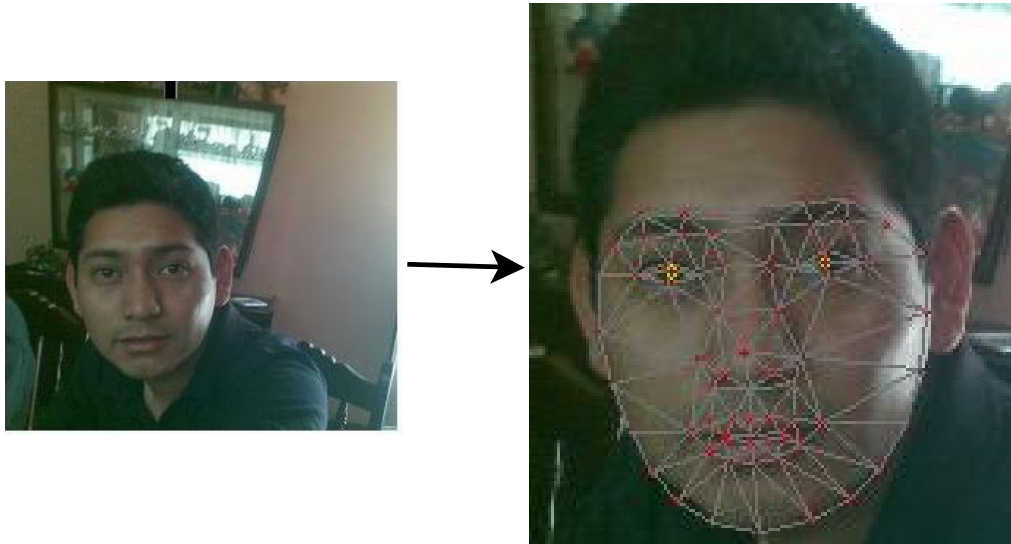


Figura 2.7 – Detecção de características geométricas usando modelagem de grafos.

Entre outras informações que podem ser modeladas usando grafos podemos citar a informação dos momentos, forma.

As arestas geralmente podem ser apresentadas com informações de:

- Distância,
- Orientação.

Esta informação obtida dos grafos deve estar sujeita a condições de invariância em relação as Transformações sofridas pelas imagens, como: rotação, escala e translação. É importante abordar esse problema no processo de formação do grafo para avaliar sua robustez.

2.4 Teoria espectral de grafos

A teoria espectral de grafos (TEG) foi inicialmente desenvolvida por Huc-
kel em 1931 no seu trabalho na área de química [54], em que representou através de

um grafo a estrutura da molécula dos hidrocarbonetos insaturados. Esta teoria vem atraindo interesse dos estudos desde a década de 80. em virtude da sua aplicação em diversas áreas, como na Química, na Matemática, na Engenharia e na Ciência da Computação, conforme Ref. [55]. A TEG é uma parte da matemática discreta que estuda as propriedades de um grafo a partir das informações fornecidas pelo espectro da matriz associada a este grafo, por exemplo a matriz de adjacência, a Laplaciana [56,57], ou seja dos autovalores destes. Nesta secção veremos o conceito da matriz laplaciana e alguns aspectos associados a ela.

2.4.1 Espectro do Laplaciano

Seja $G = (V, E)$ um grafo simples, não orientado, com n vértices e m arestas (Fig. 2.8).

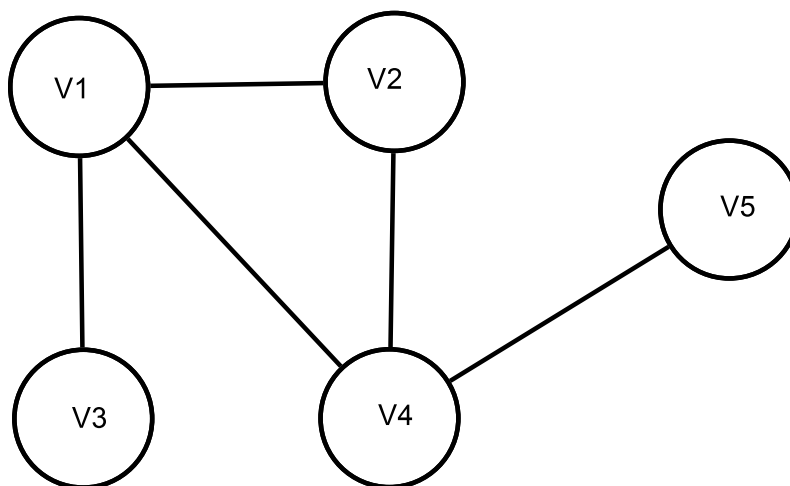


Figura 2.8 – Grafo simples não orientado.

Na matriz $n \times n$ as entradas são iguais a 1, se u e v são adjacentes, e 0 em caso contrário. Onde u e $v \in V$, é denominada *matriz de adjacência* A de G . Seja D a matriz diagonal dos vértices do grafo G (ou seja a matriz D tal que $D_{ii} = d(v_i)$), e seja A matriz de adjacência definimos a matriz L é chamada *matriz Laplaciana*

ou Laplaciano do grafo G (Eq. 2.22).

$$L_{ij} = (D - A)_{ij} = \begin{cases} -A_{ij} & \text{se } i \neq j \\ D_{ij} & \text{se } i = j \end{cases} \quad (2.22)$$

Por exemplo para o grafo da Figura 2.8 temos,

$$D = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.23)$$

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.24)$$

Logo,

$$L = \begin{bmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 2 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & -1 & 0 & 3 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \quad (2.25)$$

O **espectro do laplaciano** de um grafo G , denotado por $S(G)$, é a matriz linha cujos elementos são os autovalores de $L(G)$ ordenados de formas não crescente. Assim, $p_G(\lambda) = \det(\lambda I - L(G))$, onde λ representam as raízes do polinômio e dito ser um autovalor. Como o grafo tem n vértices, então ele possui n autovalores, sendo o maior deles o **raio espectral** de G , denominado índice do grafo.

$$S(G) = (\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n) \quad (2.26)$$

onde

$$\lambda_1 \geq \lambda_2 \geq \dots \lambda_n$$

são os autovalores de $L(G)$.

Assim, para o grafo da Fig.2.8, temos como espectro do laplaciano:

$$S(G) = (0, 0.69, 1.38, 3.61, 4.30) \quad (2.27)$$

O espectro do laplaciano é o multiconjunto de raízes do polinômio Laplaciano, com suas respectivas multiplicidades. O espectro do Laplaciano esta bem

definidos, independente da ordenação que damos aos vértices. A matriz laplaciana é também simétrica, e portanto possui todos seus autovalores reais, esta propriedade é provada pelo seguinte fato.

Definição 9 *Seja G um grafo. A matriz β de incidência com respeito a uma orientação dada é aquela cujas entradas são*

$$\beta_{i,j} = \begin{cases} +1 & \text{Se } v_i \text{ é o vértice onde chega } e_j \\ -1 & \text{Se } v_i \text{ é o vértice de onde parte } e_j \\ 0 & \text{nos outros casos.} \end{cases} \quad (2.28)$$

Não é difícil provar que $L = \beta\beta^T$. Segue daí que o Laplaciano L é uma matriz semifendida positiva que tem, portanto, todos os seus autovalores maiores ou iguais a zero.

Segue da proposição anterior que 0 é o menor autovalor da matriz laplaciana de qualquer grafo. Sejam $\mu_1 \geq \mu_2 \geq \dots \mu_n$, os autovalores da matriz laplaciana L de um grafo G . Então:

- a) $\mu_n = 0$ com autovetor associado $1 = [1, 1, \dots, 1]^T$;
- b) G é conexo se, e somente se, $\mu_{n-1} > 0$;
- c) Se G é regular de grau k então cada $\mu_i = k - \lambda_{n-i}$, onde λ_i é autovalor da matriz de adjacência A de G .

2.4.2 Autovalores e autovetores

Seja A uma matriz quadrada de ordem n e de componentes reais. Se diz que o escalar λ pertencente ao campo k é um autovalor, valor próprio ou valor

característico da matriz, se existe um vetor $x \in k^n$, diferente ao vetor nulo que satisfaz $Ax = \lambda x$.

Ao vetor x se chama autovetor, vetor próprio ou vetor característico da matriz A associado com o autovalor λ . Se define o auto-espaço ao conjunto que inclui todos os autovetores da matriz A associados ao autovalor λ e ao vetor nulo.

$$V_\lambda(A) = \{x \in k^n : x \neq \Theta \wedge Ax = \lambda x\} \cup \Theta$$

Algumas informações derivadas das funções do autovalor e do autovetor são as seguintes:

- a) Em primeiro lugar a equação matricial $Ax = \lambda x$ é não linear, já que envolve o produto dos valores desconhecidos λ e X . Porém se λ fosse conhecido, então estaríamos na presença de um sistema de equações lineares.
- b) A equação que define os autovalores e autovetores da matriz A escrita da seguinte forma:

$$(A - \lambda I)X = \Theta$$

- c) Todo autovetor esta associado a um e somente um autovalor.

2.5 Reconhecimento de padrões

Nesta seção são apresentados os métodos utilizados para a avaliação de resultados desenvolvidos e sua comparação com outras metodologias.

2.5.1 Support vector machine

As máquinas de vetores suporte (*SVM*) constituem uma técnica baseada na Teoria de Aprendizado Estatístico [58], que vem recebendo grande atenção nos

últimos anos [59, 60]. Os resultados obtidos desta técnica são comparáveis aos obtidos por outros métodos de aprendizagem, como as redes neurais artificiais [61]. Aplicações feitas com este classificador são envolvidas em aplicações de detecção de face, impressões digitais, categorização de textos e em aplicações em bioinformática [60, 62].

Basicamente esta técnica identifica a qual grupo (Y) (variável dependente), uma determinada observação (X) (variáveis independentes) têm a probabilidade de pertencer em um determinado grupo Y , e quais variáveis características separa eficientemente uma população dos diferentes grupos. Esta técnica apresenta as seguintes propriedades que é importante saber:

- i. Boa capacidade de generalização: A capacidade de generalização de um classificador consiste na medida de sua eficiência na classificação de dados que não pertencem ao conjunto utilizado durante fase de treinamento. Na geração de preditores por SVMs, portanto, é evitado o *overfitting*, situação na qual o preditor se torna especializado no conjunto de treinamento, obtendo um baixo desempenho quando é confrontado com novos padrões. Contudo, minimizar o erro na fase de treinamento não implica, necessariamente, que o erro na fase teste será pequeno. Isso deve-se ao fato de que existe outro aspecto importante, chamado de subajuste (*underfitting*), em que a complexidade da função obtida é inferior à necessidade do problema. Os classificadores gerados por uma SVM, em geral, conseguem bons resultados de generalização.
- ii. Robustez em grandes dimensões: Diferente de outros classificadores, as SVMs são robustas diante de objetos de grandes dimensões como, por exemplo, as imagens digitais.

Um SVM consiste na construção de hiperplanos, em que os dados a serem classificados são espacialmente alocados. Por exemplo, na Figura 2.9. Encontra-se

um conjunto de classificadores lineares que separam dados de duas classes distintas, porém apenas um deles maximiza a margem de separação. Uma restrição importante, fornecida pela teoria de aprendizagem estatística, relaciona o risco esperado da função ao seu risco empírico, e a um termo de capacidade. Esse limite, apresentado 2.9,

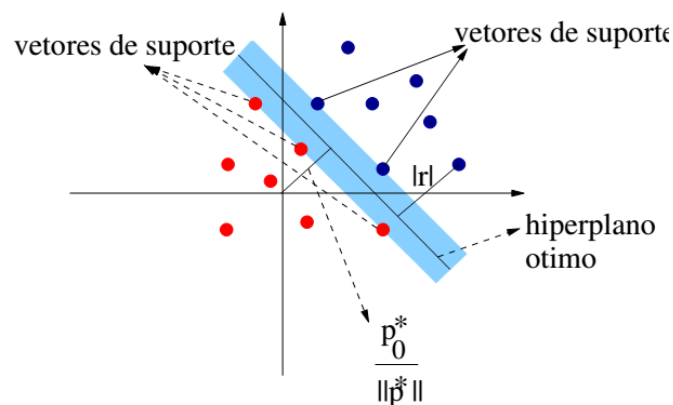


Figura 2.9 – Possíveis hiperplanos de separação e hiperplano ótimo.

é garantido com probabilidade $1 - \Theta$, em que $\Theta \in [0; 1]$.

$$R(f) \leq R_{\text{emp}}(f) + \sqrt{\frac{h \left(\ln \left(\frac{2n}{h} \right) + 1 \right) - \ln \left(\frac{\Theta}{4} \right)}{n}} \quad (2.29)$$

A funções *Kernel* têm a finalidade de projetar os vetores de características de entrada num espaço de características de alta dimensão para classificação de problemas que se encontram em espaços não-linearmente separáveis. Isso é feito, a medida que se aumenta o espaço da dimensão do problema, aumenta também a probabilidade desse problema se tornar linearmente separável em relação a um espaço de baixa dimensão. Para obter bons resultados de categorização de padrões requer-se um elevado conjunto de treinamento com um elevado número de instâncias [63] (Gonçalves 2010).

A figura 2.10 mostra a transformação de um domínio linearmente separável,

em um problema linearmente separável através do aumento da dimensão, em que é feito o mapeamento da função Kernel $F(x)$.

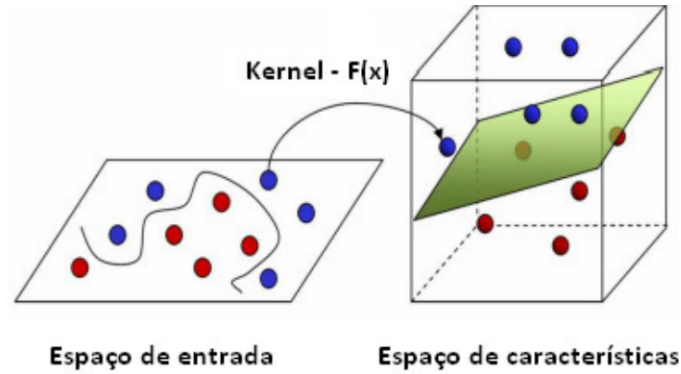


Figura 2.10 – Transformação problema não linearmente separável em um problema linearmente separável

Uma determinada função pode ser definida como *Kernel* se a matriz K é positivamente definida (autovalores maiores que zero), onde um Kernel é uma função que recebe dois pontos x_i e x_j do espaço de entradas e calcula o produto escalar $\Phi(x_i) \cdot \Phi(x_j)$ no espaço de características, como descrito em 2.30:

$$K = K_{ij} = K(x_i, x_j) \quad (2.30)$$

A utilidade dos *Kernels* se baseia na simplicidade de cálculo e na capacidade de representar espaços muito abstratos. Alguns dos *kernels* mais utilizados são os polinomiais, os Gaussianos ou RBF (*Radial basic Function*) e os Sigmoides apresentados na tabela 2.2.

Tipo de Kernel	Função $K(x_i, x_j)$	Comentários
Polinomial	$(x_i^T \cdot x_j + 1)^p$	A potencia p é definida pelo usuário
Gaussiano	$\exp\left(-\frac{1}{2\sigma^2}\ x_i x_j\ ^2\right)$	A amplitude σ^2 é especificada pelo usuário
Sigmoideo	$\tanh(\beta_0 x_i \cdot x_j + \beta_1)$	Utilizado para valores β_0 e β_1

Tabela 2.2 – Sumário dos principais Kernels nas SVMs

Segundo [64] nos Kernels em relação à tabela 2.2.

- i. No caso do Kernel polinomial a complexidade é crescente à medida que o expoente p aumenta.
- ii. O kernel gaussiano diferente do Kernel linear, é possível resolver problemas, originalmente não-linearmente separáveis em um mapeamento para espaços de maiores dimensões. Os parâmetros σ (ômega) buscam um melhor resultado de aprendizado. No kernel gaussiano de funções radiais e os seus respectivos centros são definidos pelos vetores de suporte obtidos.
- iii. O Kernel sigmoidal têm um comportamento de rede neural tipo *Perceptron* multicamadas. Neste caso, o número de neurônios da camada intermediária e os vetores de bias associados aos mesmos também são determinados pelos vetores de suporte.

2.5.2 Validação cruzada (cross-validation)

A validação cruzada é uma técnica bastante utilizada para determinar como os resultados de uma análise serão generalizados para conjuntos de dados independentes. Basicamente, essa técnica particiona um conjunto de observações em subconjuntos complementares, sendo a análise realizada sobre um conjunto (chamado treinamento) e a validação dessa análise sobre outro conjunto (chamado de conjunto de teste), cujos resultados serão amostrados numa matriz de confusão, uma explicação gráfica se amostra na figura 2.11.

Durante os trabalhos realizados optou-se pela utilização da validação cruzada *K-Folds*, em que o conjunto de treinamento original é dividido em K subconjuntos. Destes K subconjuntos, um subconjunto é retido para ser utilizado na validação do modelo e os $K - 1$ subconjuntos restantes são utilizados no treinamento. O processo de validação cruzada é, então repetido K vezes, de modo que cada um dos K

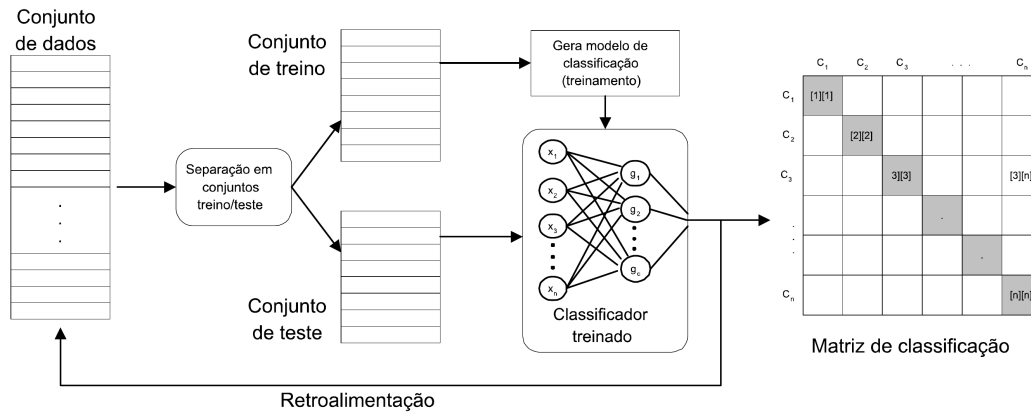


Figura 2.11 – A figura apresenta a sequência do processo de classificação supervisionada utilizado neste trabalho. O conjunto de dados é separado em dois subconjuntos (conjunto de treino e conjunto de teste), sendo o primeiro utilizado na geração do modelo de classificação (treinamento) e o segundo para provas de teste.

subconjuntos sejam utilizados exatamente uma vez como dado de teste de validação do modelo. O resultado final desse processo é o desempenho médio do classificador nos K testes. Então o objetivo é repetir os testes múltiplas vezes para aumentar a confiabilidade da estimativa de precisão do classificador.

2.5.3 Precisão e revocação

Em aplicações que envolvem a recuperação de imagens a partir do seu conteúdo (CBIR do inglês Content-based Image Retrieval) é muito importante validar e definir o real desempenho de um método. Isso permite que o mesmo seja facilmente comparado com outros métodos da literatura [65]. Um dos instrumentos mais utilizados para avaliar a eficácia dos sistemas de busca são os gráficos de precisão e revocação (*precisão vs recall*). A seguir são apresentados detalhes da construção de tais gráficos. *Revocação* indica a proporção de imagens recuperadas do banco de imagens por uma consulta 2.31. Precisão, por sua vez, indica a proporção de

imagens recuperadas que são relevantes para a consulta 2.32.

$$\text{Revocação} = \frac{R_{eR}}{R_e} \quad (2.31)$$

$$\text{Precisão} = \frac{R_{eR}}{I_R} \quad (2.32)$$

Onde R_e = total de itens relevantes existentes, R_{eR} = total de itens recuperados, e I_R = total de itens recuperados. A distância euclidiana aplicada na comparação de vetores característicos é usado para observar a semelhança das imagens.

Contextualizando as imagens digitais ao cálculo dos autovalores e autovetores, temos: (1) os objetos são representados em imagens digitais; (2) as formas dos objetos podem ser convertidas em grafos, (3) as matrizes de adjacência apresentam propriedades algébricas. (4) Os autovalores do espectro do Laplaciano podem ser considerados como um vetor identidade do objeto originalmente representado na imagem. (5) Esse vetor pode então ser comparado com outros vetores, calculados para outras imagens, permitindo que um sistema classificação possa ser processado.

Metodologia desenvolvida

Neste capítulo são apresentados os procedimentos e ferramentas utilizadas no desenvolvimento das abordagens propostas. Descreveremos os algoritmos de processamento de imagens utilizados na extração de informação do objeto de interesse, sua representação através do grafo, o cálculo do espectro em suas duas abordagens propostas e a avaliação em sua caracterização. A figura (Fig. 3.1) esquematiza-se um diagrama com o fluxo dos passos.

3.1 Pré-processamento da imagem

A rotina de processamento geralmente consiste em uma série de operações nas intensidades dos pixels da imagem. Os procedimentos objetivam detectar as informações referentes à forma, através da delimitação do contorno do objeto de interesse e seu posterior isolamento. Assim, consideramos uma imagem $I(x, y)$ sujeita a uma função $f(x, y)$ de processamento, um filtro, que produz uma imagem de saída diferente da original, que contém apenas uma quantidade reduzida de informação S ,

$$I(x, y) \rightarrow f(x, y) = S. \quad (3.1)$$

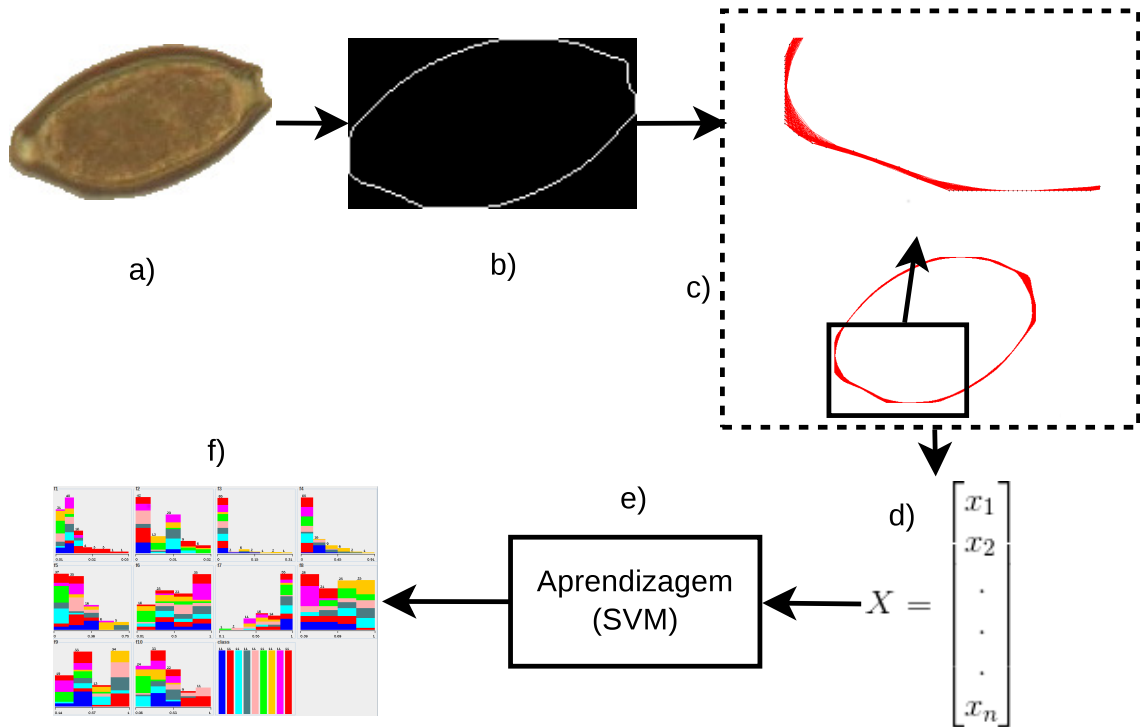


Figura 3.1 – A Ilustração mostra os passos seguidos pelo presente estudo, desde a obtenção dos contornos do objeto mediante técnicas de processamento de imagens até o cálculo dos vetores característicos e o processo de classificação. Temos: a) imagem de entrada; (b) borda do objeto da imagem processada; (c) representação do grafo baseados nos pontos da borda; (d) extração de características; (e) treinamento e (f) classificação das categorias.

Assim, a imagem $I(x, y)$ pode ser representada em distintos modelos de cores [32], sendo os mais populares o RGB (Red, Green, Blue), CMY (Cyan, Magenta, Yellow) e HSI (Hue, Saturation, Intensity). A imagem $I(x, y)$ neste trabalho esta representada pelo RGB ou tons de cinza. O RGB consiste de três camadas para sua apresentação, assim,

$$I_{\vec{x}} = \langle R_{\vec{x}}, G_{\vec{x}}, B_{\vec{x}} \rangle$$

onde *Red* (vermelho), *Green* (verde) e *Blue* (azul) são chamados camadas ou canais e \vec{x} é um vetor de coordenadas x, y de um ponto na imagem $I(x, y)$. Em nosso trabalho precisamos que a imagem colorida esteja em níveis de cinza que contém apenas uma camada de informação. O procedimento de transformação segue as normas NTSC

(International Telecommunication Union, 1998) [2] que estabeleceu a formula para calcular o valor de intensidade do nível de cinza a partir dos componentes *RGB*:

$$I(x, y) = T[I(\vec{x})] = 0.299R_{\vec{x}} + 0.58G_{\vec{x}} + 0.114B_{\vec{x}}. \quad (3.2)$$

Uma vez que a imagem foi padronizada em tons de cinza, a etapa seguinte é reduzir a informação dela a uma representação binária, o que facilita e simplifica o tratamento e a análise da imagem. A limiarização (*thresholding*) é uma técnica simples e bem conhecida para a segmentação de uma imagem. Esta técnica transforma uma imagem formatada em tons cinza, em uma imagem binária (preto e branco), processo também conhecido como binarização. Na forma mais simples, a limiarização é uma operação que atribui o valor de 0 ou 1 a cada pixel da imagem, de acordo com o parâmetro de comparação de um valor global de limiarização L . Assim a limiarização pode ser expressa como:

$$B(i, j) = \begin{cases} 0 & \text{Se } I(i, j) < L \\ 1 & \text{Se } I(i, j) > L \end{cases} \quad (3.3)$$

No presente estudo, o valor L foi calculado pelo algoritmo *Otsu* (Vease Ref. [66]). O resultado da binarização é utilizado para detectar o objeto de interesse na imagem, sobre o qual será realizada a detecção do contorno. Este procedimento é realizado através do algoritmo de seguimento de contorno (*Contour following*) [22] que consiste na extração paramétrica do contorno em uma imagem binária, $S = (x, y)$. O algoritmo inicia por definir um ponto de partida, o qual faz parte do contorno externo do objeto, assumindo que os pixels do objeto são pretos (*valor = 0*) e os pixels do fundo da imagem são brancos (*valor = 1*). Através de sucessivas chamadas recursivas, o algoritmo localiza o pixel seguinte do contorno, por isso o

termo seguimento de contorno (Fig. 3.2).

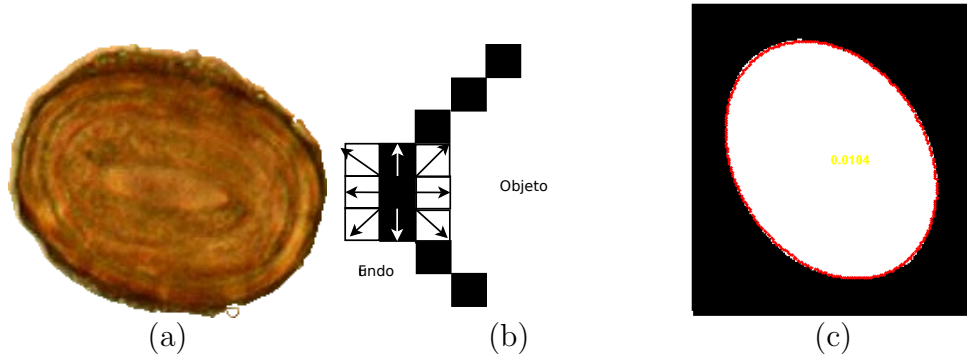


Figura 3.2 – (a) Imagem binária, (b) Seguimento de contornos e (c) Contorno paramétrico obtido.

3.2 Construção dos grafos

No presente estudo, a construção do grafo é determinado pelos pontos fronteira do objeto e esta baseado no trabalho de André Ricardo [5]. O grafo contém informação importante de conectividade em sua estrutura e este enfoque para caracterizar a forma é uma proposta interessante. Segundo o modelagem de André [5], considera um conjunto S o qual contém as coordenadas dos pixels i fronteira do objeto. Este é calculado pelo algoritmo de seguimento de bordas $S = [s_1, s_2, s_3, \dots, s_N]$, onde $s_i = [x_i, y_i]$ apresentam as posições (Fig. 3.3).

A partir da representação do contorno S será desenhado o grafo $G(V, E)$. Onde S representa os vértices do grafo, assim $V = \{S_1, S_2, \dots, S_n\}$ (Fig. 3.4). As arestas E do grafo $G(V, E)$, serão definidas pela distância euclidiana $d(v_i, v_j)$ entre os elementos que pertencem ao conjunto V .

$$d(v_i, v_j) = \sqrt{(v_i - v_j)^2 + (v_i - v_j)^2} \quad (3.4)$$

Inicialmente forma-se um grafo totalmente ligado, isto significa que a rede tem um comportamento regular, todos os vértices estão ligados com seus vizinhos e

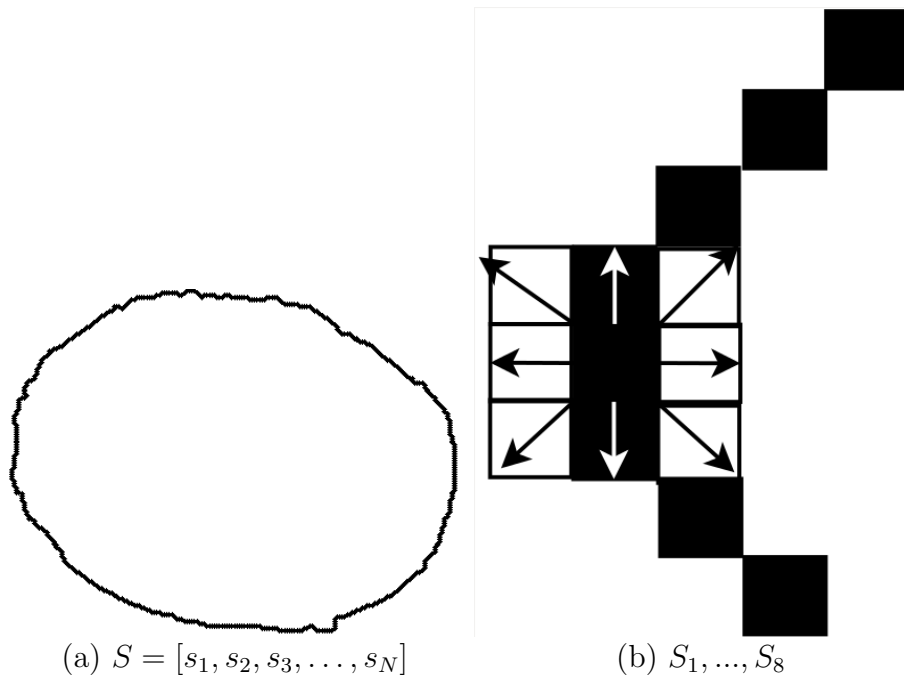


Figura 3.3 – O conjunto S de elementos da borda do objeto característico.

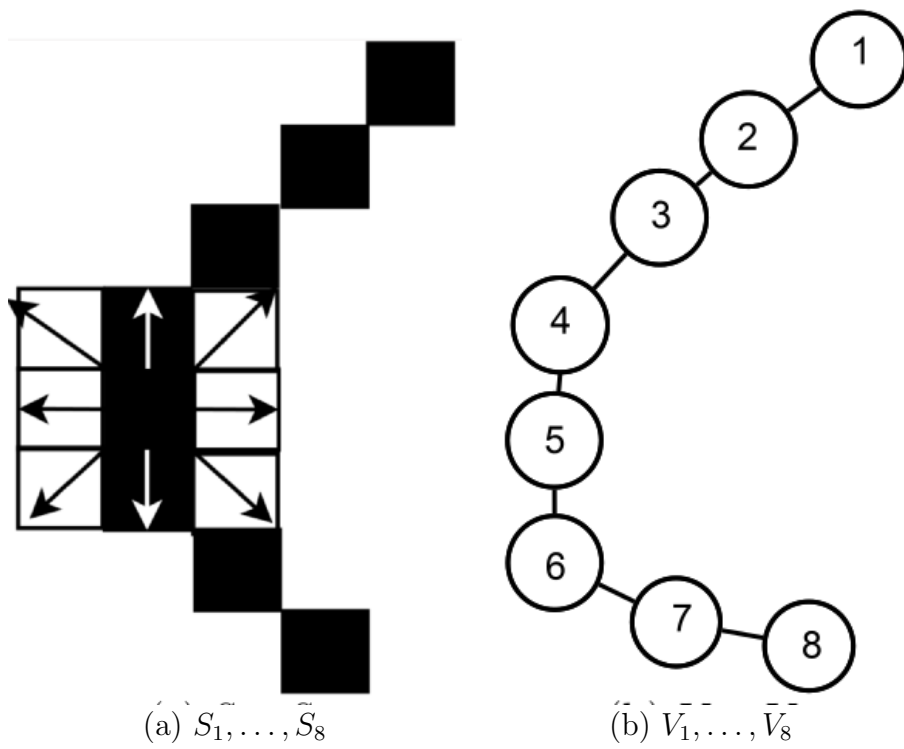


Figura 3.4 – Desenho do grafo $G(V, E)$ a partir da informação da fronteira do objeto.

têm o mesmo número de conexões. As distâncias entre os vértices foram armazenadas em uma matriz que denominamos de pesos com tamanho $N \times N$, definida como

$$W_{ij} = W([w_i, w_j]) = d(v_i, v_j), \quad (3.5)$$

onde N é o número de vértices.

A matriz de pesos W seguidamente é normalizada, dividendo os elementos da matriz entre o maior valor da matriz $\max W_{i,j}$. Assim criamos uma nova matriz W' com os valores normalizados no intervalo de $[0-1]$. Dessa forma também aborda-se o problema relacionado as escalas nas imagens (ou seja imagens capturadas com distinta proximidade).

$$W' \rightarrow \frac{W}{\max W_{ij}} \in W. \quad (3.6)$$

A normalização de W' em um intervalo de $[0, 1]$, assegura propriedades como invariantes a escala e rotação (Fig. 3.5). Contudo, essa configuração de ter todos os vértices ligados não permite analisar as características relevantes da forma do objeto. Assim é necessário decompor o grafo regular, em grafos que possuam distintos graus de conectividade, cujas ligações entre vértices sejam limitadas por uma regra.

A conectividade de cada vértice V do grafo $G(V, E)$ será definido pelos limiares T_l . Os limiares selecionam um grupo de arestas do grafo regular, cujos valores são menores que o limiar T_l . Este determina a regra pela qual os vértices possuam uma conectividade ou ligação (Fig. 3.6).

As ligações entre os vértices basean-se nos cálculos realizados na matriz de pesos normalizada W' , o qual define os valores de distância ou arestas E do grafo $G(V, E)$ (Fig. 3.9). Em consequência o valor do limiar deve ser definido entre valores maiores a 0 e menores a 1 e determina novas topologias, efeito de decompor o grafo

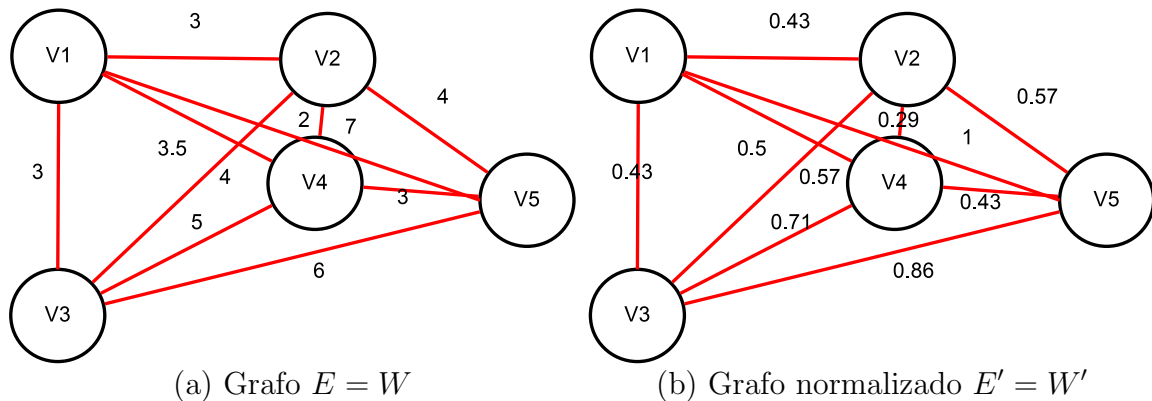


Figura 3.5 – A ilustração mostra em: a) Um grafo cujas arestas são determinadas pela distância euclidiana entre dois vértices, cujos resultados são armazenadas na matriz de pesos W , (b) O grafo mostrado têm as distancias normalizadas do grafo e armazenadas na matriz de pesos W' , com valores no intervalo de 0 – 1.

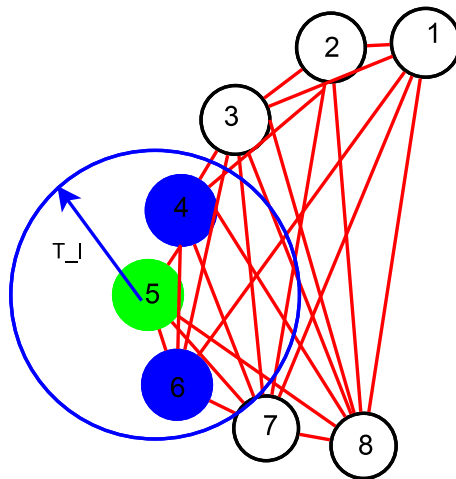
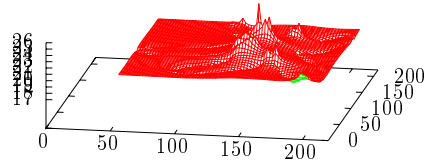


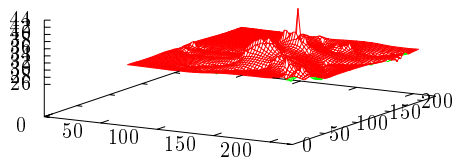
Figura 3.6 – Dado que T_l refere-se ao limiar, na ilustração mostra-se um vértice em verde e uma circunferência azul ao redor que define região de vizinhança definida por T_l , todo o que esteja inserido naquela região terá uma ligação com o vértice verde.

regular.

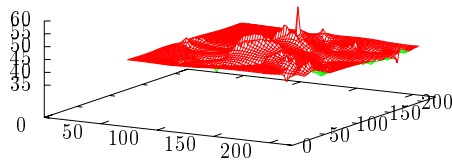
Se consideramos duas imagens em escalas diferentes, a normalização de W por $\max W_{ij}$ fixa as distâncias entre dois vértices, até um valor máximo a 1. Para guardar as características de invariância na escala é necessário saber que duas imagens em escalas diferentes vão ter diferentes números de pontos em seu contorno. Assim se temos $S = V$, dois contornos similares $S_A = [s_{A1}, s_{A2}, \dots, s_{AN}]$ e



a)



b)



c)

Figura 3.7 – Evolução dinâmica do grafo gerado pelos limiares a) $T_l = 0.035$, b) $T_l = 0.065$, c) $T_l = 0.095$.

$S_B = [s_{B1}, s_{B2}, \dots, s_{BM}]$, com $(N \neq M)$, produz em grafos diferentes ($G_A \neq G_B$), com número distinto de vértices. O cálculo do grau $deg(v_i)$ (Cap. 2, Eq. 2.19) é diretamente afetado pelo número de vértices do grafo. A solução é normalizar o grau $deg(v_i)$ ao tamanho do grafo modelado. A Figura 3.8 apresenta esta transformação por cada $deg(v_i)$, e a figura 3.7 mostra a evolução do grafo com respeito aos limiares T_l .

Contudo, não há a priori um limiar ótimo. Assim, a caracterização da forma pode ser obtida por diferentes valores T_l de uma transformação δ . Esta operação é representada como $A = \delta_{T_l}$, é aplicada em cada um dos elementos da matriz de

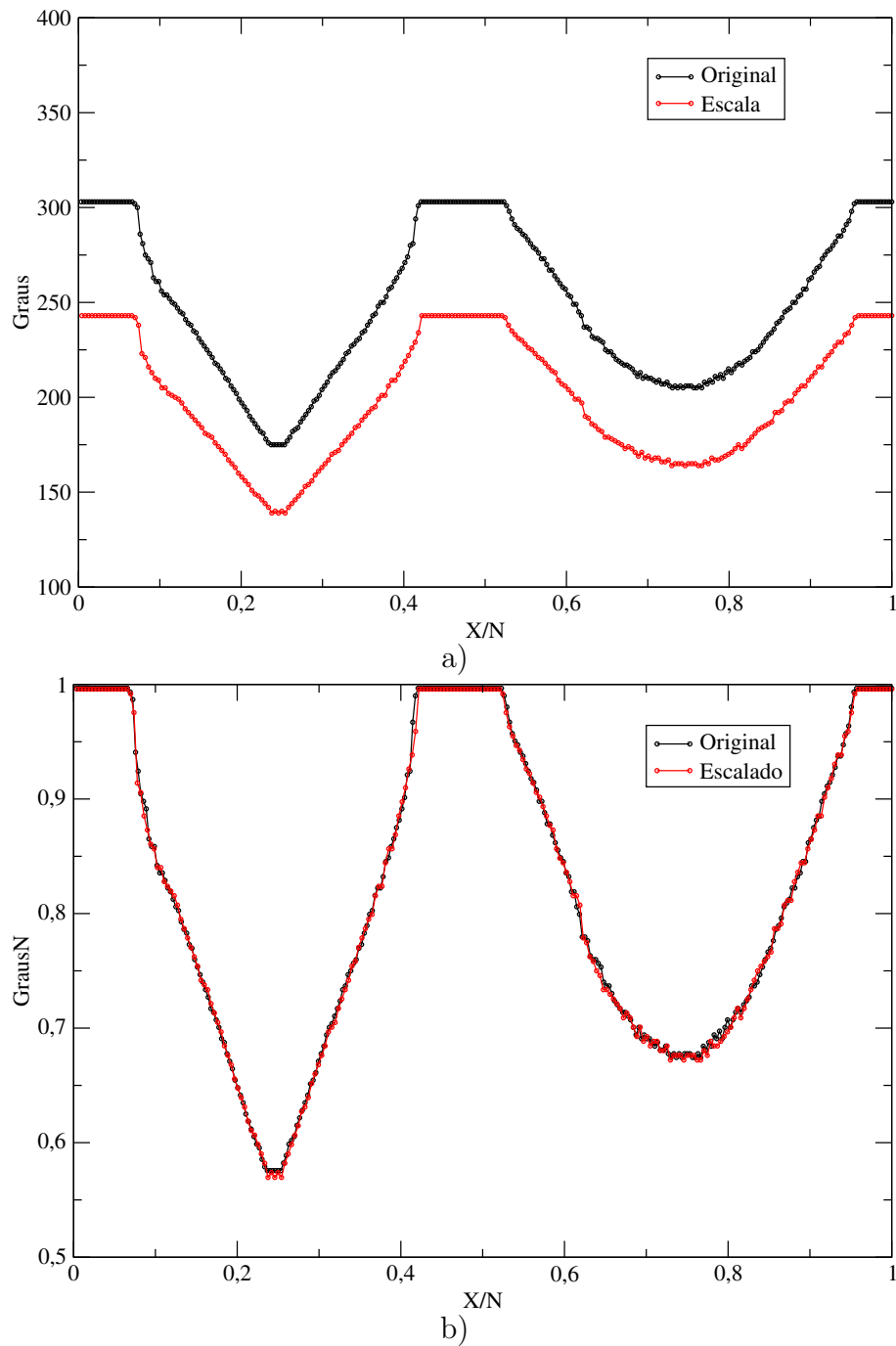


Figura 3.8 – As curvas representam os graus de cada um dos vértices do grafo, onde a linha preta corresponde à imagem original e a linha vermelha a imagem escalada em menor resolução. a) Sem normalizar e (b) Normalizado. ([5])

pesos normalizados W' , conduzindo à formação da matriz de adjacência. Assim :

$$A_{T_l} = \delta_{T_l}(W) = \forall w \in W \begin{cases} a_{ij} = 0 & \text{if } w_{ij} \geq T_l \\ a_{ij} = 1 & \text{if } w_{ij} < T_l \end{cases} \quad (3.7)$$

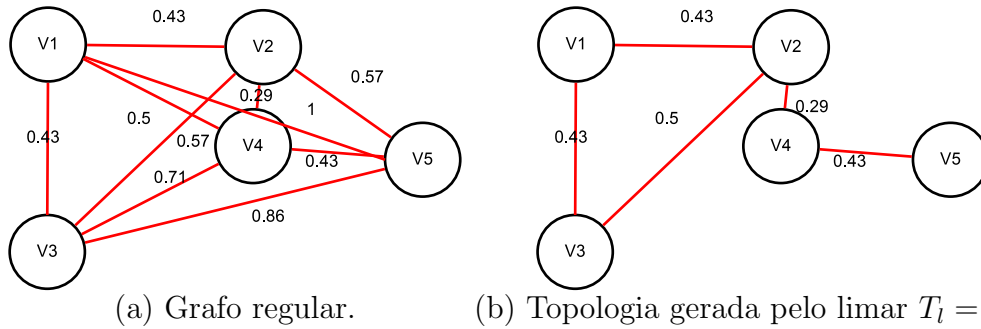


Figura 3.9 – Exemplo de decompor do grafo regular, a partir de um limiar $T_l = 0.51$.

3.3 Análise espectral dos grafos

Neste ponto, temos como objeto um grafo $G(V, E)$ construído e normalizado segundo o proposto na secção anterior (3.2), bem com sua respectiva matriz de adjacência. A estrutura topológica gerada pelos limiares pode ser capturada de forma numérica pela análise do espectro Laplaciano. Para o cálculo, precisamos de dois ingredientes: a matriz de adjacência (Fig. 3.1) e a matriz diagonal de grau D (Fig. 3.2).

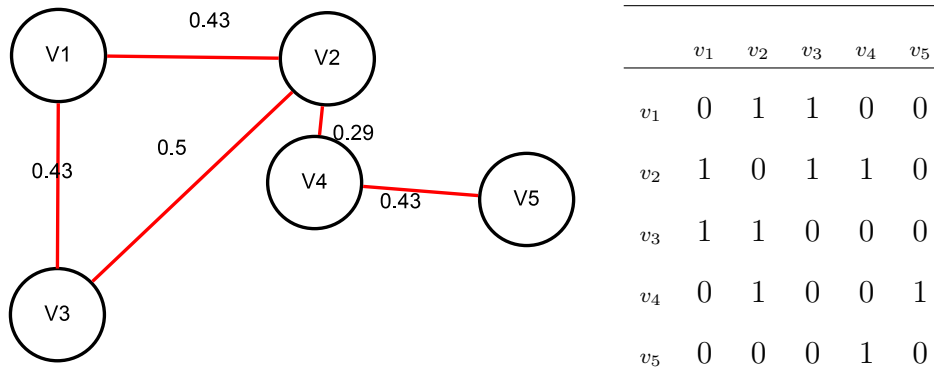


Tabela 3.1 – Representação de um grafo não orientado e sua matriz de adjacência A .

Assim, temos que a matriz Laplaciana,

$$L(G) = D(G) - A(G) \quad (3.8)$$

onde L é uma matriz simétrica semi positiva definida, e o espectro Laplaciano con-

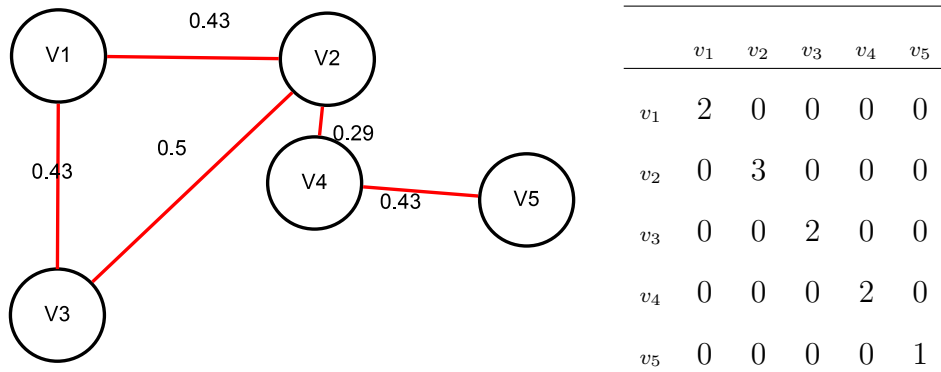


Tabela 3.2 – Representação de um grafo não orientado e sua matriz de graus D .

siste nas raízes obtidas a partir do cálculo do polinômio característico,

$$P_L = \det(\lambda I - L(G)) \quad (3.9)$$

da matriz $L(G)$.

Considerando os grafos gerados as raízes do polinômio característico P_μ foram calculadas pela decomposição dos valores singulares ("Singular values decomposition" ou SVD), onde L pode ser composto como $L = X \Delta Y^T$. Dessa maneira, obteve-se os autovalores

$$0 = \lambda_0 \leq \lambda_1 \leq \dots \leq 1$$

e auto vetores

$$u_0, u_1, \dots, u_{n-1}$$

.

A Figura 3.10 apresenta um exemplo da aplicação do espectro do Laplaciano nas informações oriundas de uma imagem do conjunto ShapeCN.

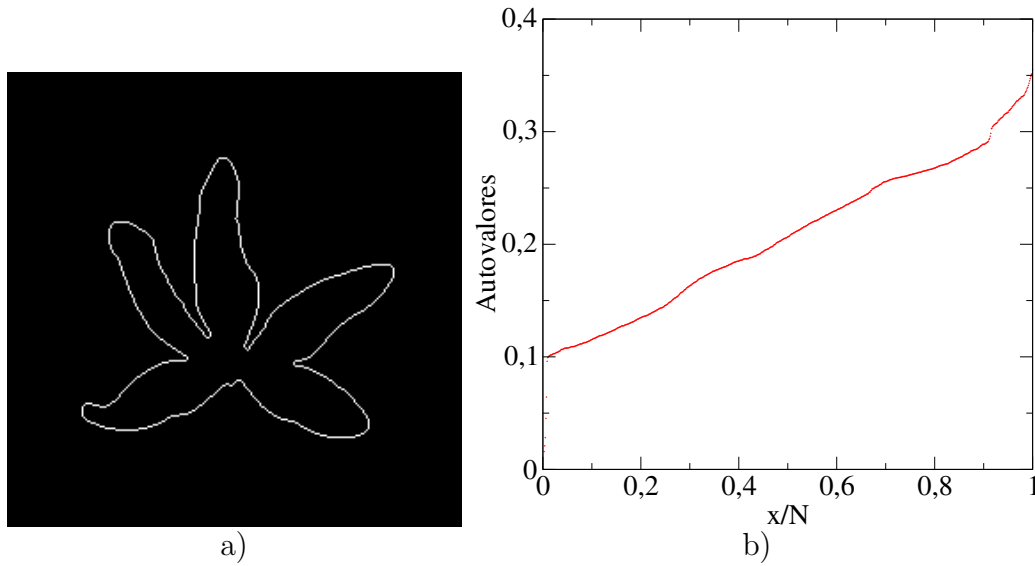


Figura 3.10 – Exemplo do cálculo dos autovalores do espectro Laplaciano a partir da matriz Laplaciana de uma imagem do conjunto ShapeCN. (a) Contorno da imagem; (b) Espectro Laplaciano com um limiar $T_{0.75}$.

3.4 Assinatura ou características

Após o cálculo dos autovalores da matriz Laplaciana $\lambda(L)$, dois procedimentos distintos foram feitos para obtenção de um vetor característico da imagem, para que então pudessem ser analisadas e classificadas. A primeira abordagem utilizou um conjunto de diferentes limiares, normalizando o comprimento dos vetores resultantes através de um histograma com as frequências dos autovalores, em que variaram os intervalos de classe utilizados na construção do histogramas (variação de 2 parâmetros). Na segunda abordagem, também utilizou-se um conjunto de diferentes limiares, contudo adotou-se como informação sempre o segundo e último autovalores (apenas 1 parâmetro). A seguir, apresentamos mais detalhes de cada abordagem.

3.4.1 Vetor característico baseado em limiares e a histogramas

Dado um determinado limiar T_l , construiu-se um grafo E , e sua respectiva matriz de adjacência, então calculou-se os autovalores do espectro Laplaciano. Como resultado, obteve-se uma representação de autovalores $\lambda_1 \geq \lambda_2, \dots \geq \lambda_n$. Dado que diferentes imagens produzem tamanhos de matrizes distintos, estes precisam de uma normalização prévia, uma etapa não trivial do problema.

Assim, adotou-se como estratégia de normalização do vetor característico pela análise de frequências de autovalores, segundo histogramas. Por meio destes histogramas, obteve-se uma nova representação característica da imagem através de $\beta = H(\lambda_1, \dots, \lambda_n)$. Determinando-se previamente um número n de intervalos de classes, o tamanho dos vetores de diferentes imagens puderam então ser padronizados. Contudo, espera-se que cada imagem possua um intervalo de classes ótimo, para o qual a perda de informações é mínima. Por isso, o tamanho dos intervalos de classes também foi considerado como um parâmetro avaliado dentro do problema. Portanto, nesta abordagem analisamos os autovalores obtidos com diferentes limiares T_l e diferentes intervalos de classes IC_l , quando processadas as frequências dos autovalores (Fig.3.11).

3.4.2 Vetor característico baseado em vários limiares para autovalores fixos

Temos aqui as mesmas informações resultantes da construção dos grafos e cálculo do espectro Laplacianos já mencionadas no item anterior. Embora esta abordagem tenha utilizado diferentes limiares T_l , como na anterior, aqui são tomados apenas o segundo e último autovalores de cada conjunto dos mesmos para formar o vetor de informação. Como resultado, o tamanho dos vetores foi o mesmo para

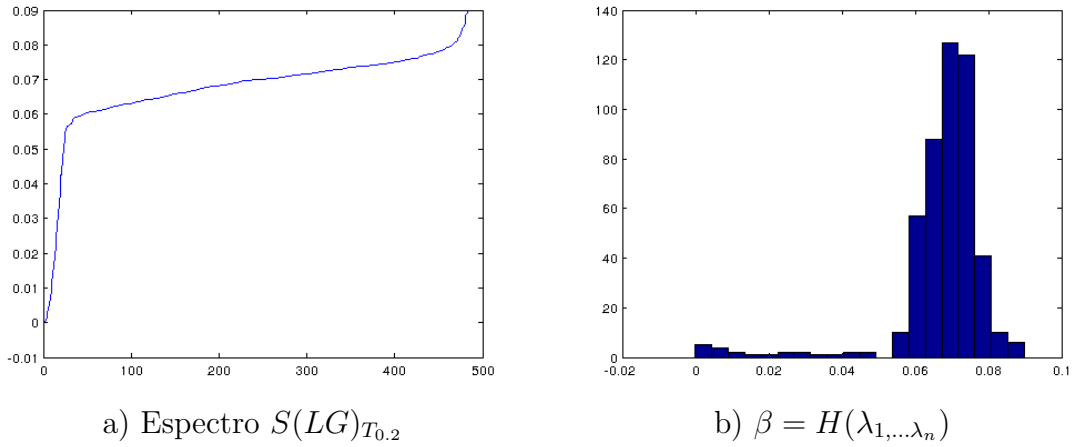


Figura 3.11 – Representação da histograma de um espectro do grafo $G(V, E)$ obtido de um limiar $T_l = 0.2$

todas as imagens e a comparação foi possível de forma mais direta, sem normalização prévia. Se comparado com a abordagem anterior (limiares e histogramas), o tempo de processamento é muito menor.

O segundo autovalor é citado como um bom descritor para segmentação da imagem e da topologia do grafo [67]. Por outro lado, testes preliminares sugeriram que o último autovalor também é capaz de descrever importantes características pertinentes à topologia do grafo. Assim, utilizou-se as informações de ambos autovalores como vetor caracterizador das imagens, visando um melhor custo benefício entre qualidade da classificação e tempo de processamento. Dado que o número de autovalores foi constante, o vetor característico,

$$\beta = [\lambda_2(G_{t_1}), \lambda_n(G_{t_1}), \dots, \lambda_2(G_{t_M}), \lambda_n(G_{t_M})], \quad (3.10)$$

onde $t_i \in T_l$, não demanda subsequente normalização (Fig. 3.12) .

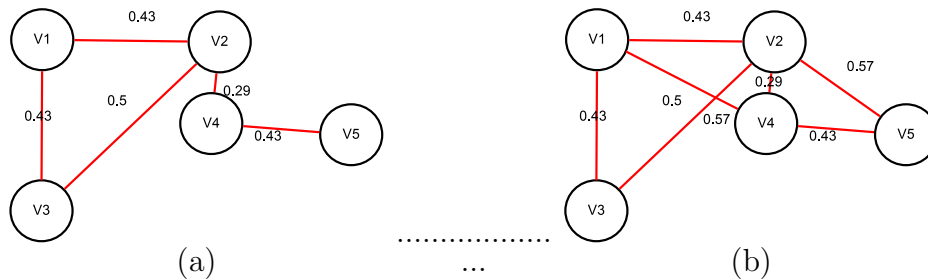


Figura 3.12 – Vetor característico $\beta = [\lambda_2(G_{t_1}), \lambda_n(G_{t_1}), \dots, \lambda_2(G_{t_2}), \lambda_n(G_{t_2})]$

3.5 Máquina de aprendizagem

Os resultados obtidos pelo presente estudo foram avaliados por uma máquina de aprendizagem. Considerando os sistemas de classificação previamente estabelecidos para as coleções de imagens **Genéricas**, **ShapeCN** e **SADPI**, o propósito foi avaliar a taxa de acerto em função dos vetores característicos obtidos a partir das duas abordagens acima detalhadas (3.4). Utilizou-se para tal o classificador SVM (Máquina de vetores de Suporte) da livreria WEKA [68], usando os parâmetros de ajuste padrão do classificador. A classificação foi avaliada mediante a validação cruzada de 10 pastas (10-Fold Cross-validation) 2.5.2. O processo consiste em dividir o conjunto da coleção de imagens em 10 partes, destas 9 partes são utilizadas para o treinamento e 1 para o teste (Fig. 3.13). O processo é repetido 10 vezes, de forma que cada parte seja usada uma vez como conjunto de testes. Ao final, a correção total é calculada pela média dos resultados obtidos em cada etapa, obtendo-se assim uma estimativa da qualidade do modelo de conhecimento gerado.

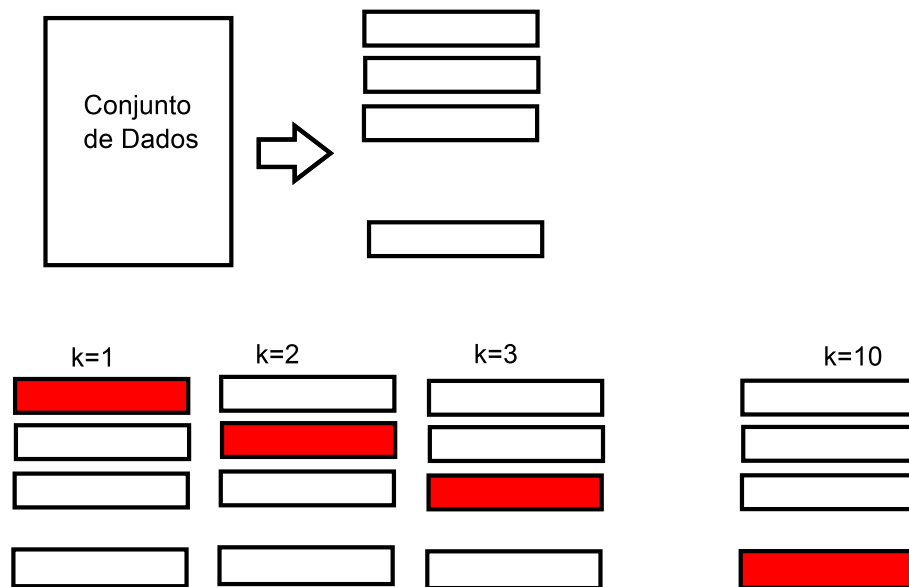


Figura 3.13 – Validação cruzada 10 pastas.

Experimentos e Resultados

4.1 Coleções de imagens usados como objetos de estudo

O presente estudo considerou três coleções de imagens digitais para avaliar nossas abordagens. Cada coleção caracteriza-se por seu nível de dificuldade. Assim temos uma coleção de imagens sintéticas que apresentam figuras genéricas [24], uma coleção de imagens de parasitologia SADPI (Sistema automático de diagnóstico de parasitas intestinais) [69], e uma coleção de imagens de folhas de plantas (**ShapeCN**, disponível na referência [70]).

A coleção de imagens sintéticas contém um total de 99 imagens, que caracterizam-se por representar figuras genéricas. O conjunto de imagens está classificado em 9 categorias. Cada categoria contém 11 imagens e possui figuras que representam coisas e animais [24]. Olhar a ilustração 4.1. Esta primeira coleção tem um nível básico de complexidade, o qual permite avaliar em primeira instância se as abordagens conseguem separar as imagens em suas categorias.

A coleção de imagens de Parasitas (SADPI) tem como foco as imagens de microscopia dos ovos de vermes helmintos, entes infecciosos presentes em quase



Figura 4.1 – Coleção de imagens de formas genéricas.

todas as regiões do planeta, e que afetam tanto seres humanos como outros animais. Esta coleção de imagens é formada por 656 amostras as quais estão divididas em 8 espécies (classes) [71] (Olhar figura 4.2). Além de um número maior de imagens por classe, as imagens do SADPI foram adquiridas de maneira sistemática e padronizada, com escalas e planos bem definidos. O grupo de pesquisa em biologia da Universidade Nacional de San Agustín pertencente ao projeto do Sistema automático para detecção de parasitas intestinais sob a condução do professor Cesar Beltran, disponibilizou as imagens para testarmos nossas abordagens. Por outro lado, todos as classes compartilham uma silhueta em comum (forma predominantemente elíptica), o que pode dificultar o processo de classificação das imagens em suas categorias quando apenas a forma do objeto é considerada.

O conjunto ShapeCN é composto por 600 imagens de folhas distribuídas em 30 categorias destas. Cada categoria contém 20 imagens, cuja resolução é 256×256 , em formato binário, na qual cada imagem possui a informação do contorno da folha [5] (Ver Fig.4.3). A classificação de folhas vegetais pode ser uma tarefa difícil. Considerando o conjunto de ShapeCN, como exemplo, nota-se similaridade considerável entre objetos de diferentes classes, variabilidade nas formas entre os elementos de uma mesma classe. Além disso, a obtenção das imagens não pode ser facilmente

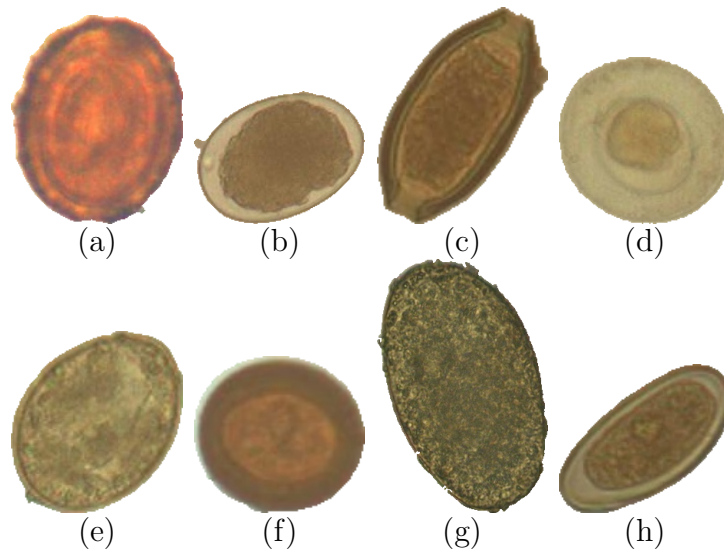


Figura 4.2 – Exemplificando as imagens contidas na coleção SADPI. Cada imagem representa uma categoria; a) *Ascaris lumbricoides*, b) *Uncinárias*, c) *Trichuris Trichura*, d) *Hymenopelis nana*, e) *Dyphillobotrium pacificum*, f) *Taenia solium*, g) *Fasciola hepática*, h) *Enterobius vermicularis*.

padronizada, e o conjunto de imagens, via de regra, apresenta variações de escala e orientação. A ideia de utilizar esta coleção de imagens especialmente problemática como objeto de estudo é avaliar o rendimento dos métodos propostos pelo presente estudo quando tais condições estão presentes.

4.2 Análise do vetor característico baseados em limiares e a histogramas

A análise desta primeira abordagem está baseada em dois parâmetros: limiares T_l e intervalos de classes dos histogramas IC_l . Estes produzem 722 vetores característicos diferentes por imagem e aplica-se em cada coleção de imagens. O número de vetores resultou das 38×19 possíveis combinações entre T_l e IC_l , os testes consideraram valores de T_l dentro do intervalo de 0.025 e 0.95, com incremento de

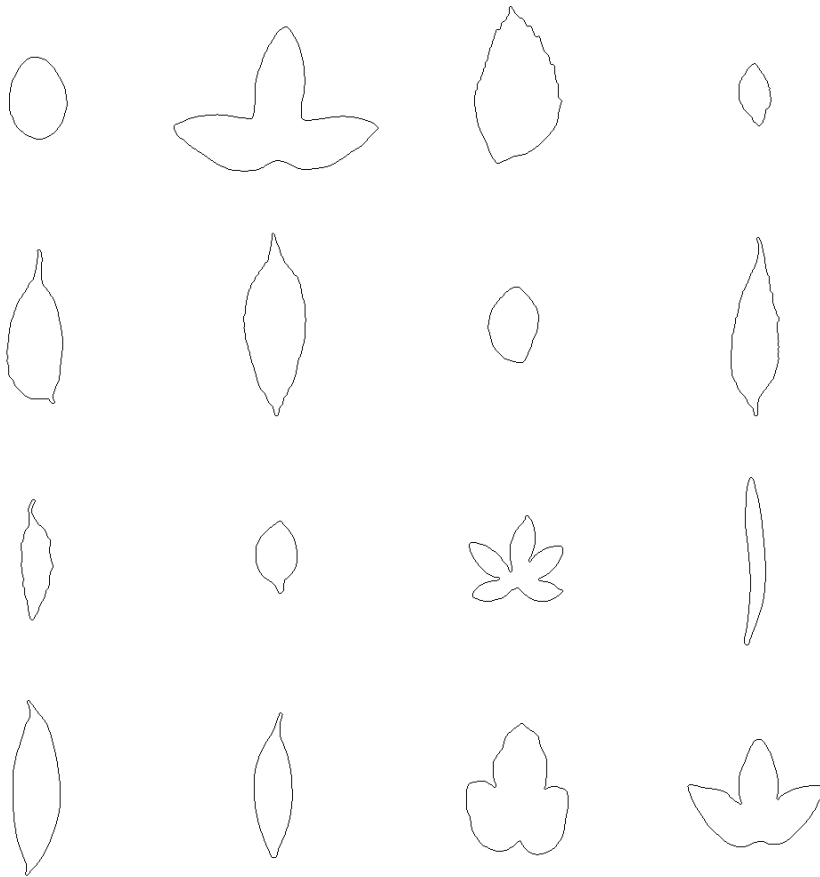


Figura 4.3 – A ilustração mostra um exemplo das distintas categorias da coleção de imagens de folhas de plantas (ShapeCN). Como podemos observar existem categorias que são semelhantes.

0.025, e valores de IC_l dentro do intervalo 10 e 100, com um incremento de 5.

Primeiramente utilizamos a coleção de imagens de formas genéricas. As seguintes observações foram ditas a partir da figura (Fig. 4.4), da qual podemos descrever que nos valores $T_l = 0.7-0.75$, $IC_l = 15-25$ e $T_l = 0.3-0.48$, $IC_l = 20-85$ as taxas de acerto são superiores a 80% (Para maior informação dos valores olhar a Tabela Resumo A.1). Das taxas de acerto observou-se que com os parâmetros $IC_l = 10$, $T_l = 0.475$ uma alta taxa de acerto de 91% foi possível (Ver a tabela resumo Tab. 4.1 e A.1).

Os resultados obtidos na coleção de imagens dos parasitas (SADPI), consi-

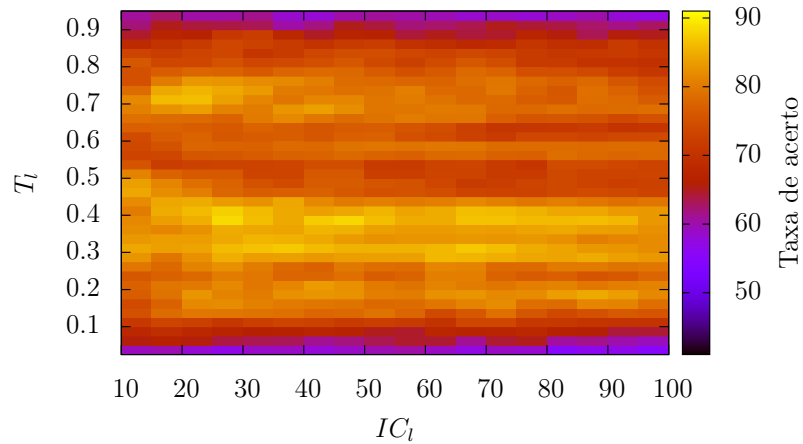


Figura 4.4 – Representação gráfica de duas dimensões para ilustrar dados com 3 variáveis, através da escala de cores. O gráfico apresenta um quadro IC_l vs T_l da coleção de imagens genéricas, onde as intensidades de cores representam as taxa de acerto obtidas pela maquina de aprendizagem SVM e está definido na parte direita. Observa-se, duas regiões que se destacam das demais, nas faixas amarelas as taxas de acerto foram acima do 80%.

Tabela 4.1 – A tabela resume as melhores taxas de classificação obtidas na figura 4.4 por intervalo de classe na coleção de imagens genéricas.

IC_l	T_l	%	IC_l	T_l	%
10	0.475	91	60	0.300	87
15	0.400	87	65	0.300	89
20	0.725	88	70	0.375	90
25	0.400	88	75	0.400	88
30	0.400	91	80	0.375	89
35	0.325	88	85	0.375	88
40	0.375	89	90	0.375	88
45	0.400	89	95	0.400	85
50	0.375	89	100	0.275	85
55	0.400	86			

derando as classes pré-determinadas e a taxa de acerto produzida pela máquina de aprendizagem SVM, sugerem que, embora existam inúmeras configurações de valores de T_l e IC_l com resultados de acerto próximos ou acima de 80%, há uma região bastante característica no espaço de parâmetros que concentrou as melhores taxas de acerto (Fig. 4.5). Essa região apresentou entre os valores limiares T_l de 0.75 e 0.85, e valores de IC_l entre 25 e 75 elevadas taxas de acerto. Dentro dessa região característica do espaço de parâmetros, também foi encontrada a configuração de parâmetros que produziu o máximo acerto (88%), cujos valores foram respectivamente $T_l = 0.775$ e $IC_l = 30$ (Tab. 4.2). (Para analisar os resultados em geral deste banco ver a tabelas A.3 e A.4).

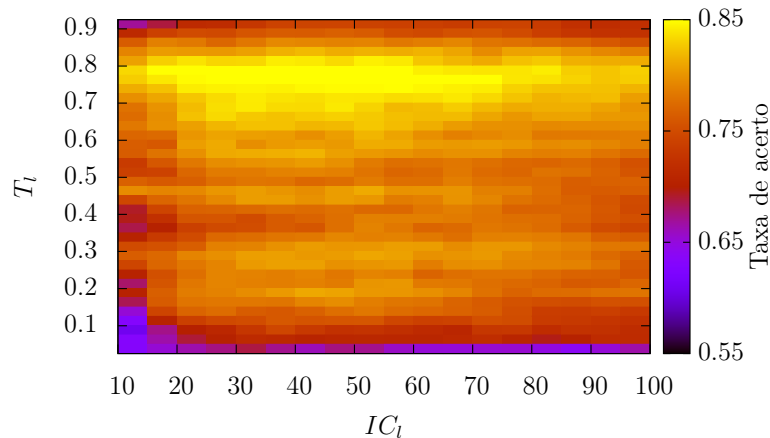


Figura 4.5 – A gráfico apresenta um quadro IC_l vs T_l respeito da taxa de acerto obtida pela maquina de aprendizagem SVM, aplicada na coleção de imagens SADPI. Este apresenta diferentes intensidades de cor, onde as regiões amarelas definem as maiores taxas de sucesso. Assim observamos que nos valores para IC_l no intervalo de 25 – 75, e nos limiares 0.75 – 0.85, obtêm-se os melhores resultados.

Considerando a coleção de imagens ShapeCN e as taxas de acerto pelo SVM, detectou-se alguns contrastes em relação ao observado em outras coleções de imagens SADPI e genéricas. Primeiro, de forma geral, os testes com as imagens de

Tabela 4.2 – A tabela descreve o resumo dos melhores resultados obtidos por intervalo de classe IC_l vs T_l na coleção de imagens SADPI. Assim podemos observar que a melhor taxa de acerto foi obtida com os seguintes parâmetros: Intervalo de classe $IC_l = 30$, limiar $T_l = 0.775$, obtendo-se uma taxa de acerto de 88 %

IC_l	T_l	%	TC_l	T_l	%
10	0.725	84	60	0.775	86
15	0.775	86	65	0.775	85
20	0.800	87	70	0.775	88
25	0.775	86	75	0.775	85
30	0.775	88	80	0.775	85
35	0.775	88	85	0.775	85
40	0.775	88	90	0.700	83
45	0.775	87	95	0.775	83
50	0.775	88	100	0.775	84
55	0.775	88			

ShapeCN apresentaram taxas menores de acerto em relação aos testes feitos com as coleções de imagens SADPI e genéricas (Fig. 4.6). O segundo contraste diz respeito à relação entre o espaço de parâmetros com as melhores taxas de acerto. Os testes com ShapeCN sugeriram duas regiões dentro do espaço de parâmetros cujos acertos produziram resultados acima de 65%: primeira região com $0.1 < T_l < 0.25$ e $25 < IC_l < 65$; e a segunda região com $0.4 < T_l < 0.6$ e $10 < IC_l < 65$, a qual também apresentou a máxima taxa de acerto (72%), para $T_l = 0.57$ e $IC_l = 15$ (Tab. 4.3)). (Para olhar os resultados ver as Tabelas A.5 e A.6).

Apesar das diferenças nas taxas de acerto e espaço de parâmetros obtidas para as três coleções de imagens analisadas, Genérico, ShapeCN e SADPI, a interpretação detalhada dos resultados permite encontrar alguns aspectos em comum. Primeiro, considerando as melhores taxas de acerto para cada banco de imagens, o

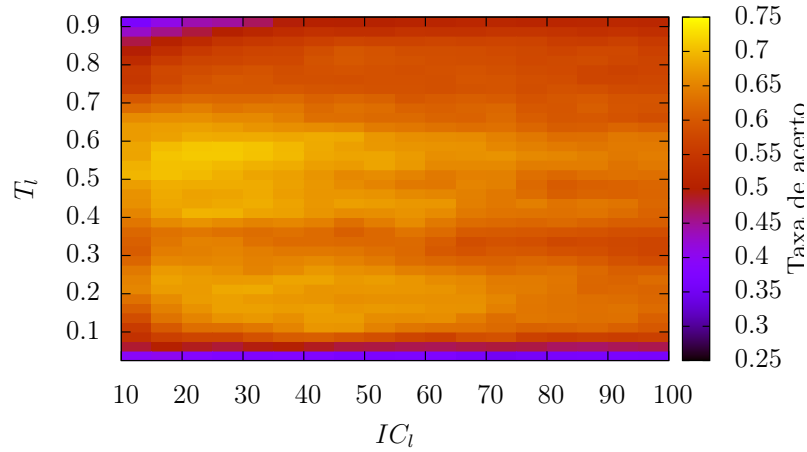


Figura 4.6 – A figura apresenta um quadro IC_l vs T_l , onde podemos observar as seguintes características. Existe mais de uma região onde as taxas de acerto são melhor apresentadas. Assim nos limiares T_l de 0.1 – 0.25 e 0.4 – 0.7 e intervalos de classe IC_l de 10 até 90 aprecia-se melhor diferenciação.

Tabela 4.3 – A tabela descreve o resumo dos melhores resultados obtidos por intervalo de classe IC_l a partir da coleção de imagens CNShape.

IC_l	T_l	%	IC_l	T_l	%
10	0.525	69	60	0.550	67
15	0.575	72	65	0.600	69
20	0.575	72	70	0.550	68
25	0.575	72	75	0.550	68
30	0.575	71	80	0.550	67
35	0.600	72	85	0.550	67
40	0.175	70	90	0.425	66
45	0.600	70	95	0.600	65
50	0.550	71	100	0.550	66
55	0.550	70			

parâmetro T_l sempre foi mais restritivo que o parâmetro IC_l . Além disso, há o fato de, os testes com SADPI apresentarem apenas uma região ótima de acertos bem definida. Porém no ShapeCN, encontrar duas regiões características, parece não ser um resultado casual. A afirmação está baseada no fato de que SADPI apresenta apenas imagens com padrões elipsoidais, enquanto o banco ShapeCN apresenta duas formas predominantes bastante distintas, folhas alongadas (peninérvias, curvinérvias ou paralelinérvias) e folhas palmadas (palminérveas). Portanto, é possível que haja uma configuração ótima de limiares e/ou intervalos de classes para cada forma geométrica, como é indicado pelos resultados com ShapeCN. Ainda, as formas de ShapeCN e SADPI são diferentes, e os limiares ótimos T_l foram diferentes em ambos os casos.

Tecnicamente, a relação entre o limiar T_l e a forma é suportada pela sua correlação com a densidade de conectividade (relação de vizinhança entre os pixels), que é expressa posteriormente pelos autovalores do espectro do laplaciano do grafo. Portanto, os resultados sugerem que diferentes formas geométricas tendem a apresentar diferentes limiares ótimos para descrever as características globais. Os resultados sugerem que não seria possível utilizar apenas um único limiar para classificar diferentes formas geométricas, dado que o número de formas tenderia a prejudicar a taxa de acerto de forma sistemática, embora o tema não tenha sido explicitamente testado. Por outro lado, os resultados aqui obtidos sugerem que o uso de limiares ótimos e espectro laplaciano podem ser uma alternativa viável para a classificação de imagens, cujos objetos possuem formas muito similares. Consequentemente, o método aqui apresentado pode ser considerado um método complementar a outros métodos já propostos para a classificação de formas, como uma técnica complementar, onde a distinção de características finas é o objetivo. Como uma ideia para futuros estudos, o uso sistematizado de imagens sintéticas poderia elucidar melhor como diferentes limiares estão correlacionados à diferentes formas geométricas.

4.3 Análise do vetor característico baseado em vários limiares para autovalores fixos

Há diferença da primeira abordagem, na qual utilizamos os autovalores obtidos por um limiar e estes distribuídos nos intervalos de classe. Nesta segunda abordagem como foi explicado na seção 3.4.2, analisa-se o segundo e último autovalor em vários limiares T_l . Os limiares modificam a topologia do grafo, podendo-se assim observar mudanças na quantidade de ligações na medida que o valor do limiar aumenta. Analisar a influência destas modificações no segundo e último autovalor é o objetivo desta seção de resultados. Na ilustração 4.7 apresentamos os valores do segundo com limiares $T_l = 0.3, 0.5, 0.7$ e 0.9 , realizado na coleção de imagens de parasitas (SADPI).

Na ilustração 4.8 apresentamos os valores do último autovalor com os limiares $T_l = 0.3, 0.5, 0.7$ e 0.9 , realizado na coleção de imagens de parasitas (SADPI), mostrando comportamentos similares com o segundo autovalor.

Observou-se independentemente o efeito do segundo e último autovalor aplicado nas coleções de imagens de parasitas e folhas para exemplificar. Nas seguintes tabelas mostramos os resultados obtidos, empregando distintas quantidades de limiares, dos quais por cada limiar se obtém duas características (segundo e último autovalor) armazenadas em um vetor característico. Os experimentos foram realizados com um intervalo fixo de limiares ($T_{ini} = 0.025$ até $T_f = 0.95$), onde o intervalo é fracionado por meio de passos ou incrementos T_{inc} que determinam o tamanho do vetor característico.

Na tabela 4.4 mostramos as distintas combinações de limiares aplicados à coleção de imagens Genérica. Observou-se que com um incremento de $T_{inc} = 0.2$ sendo 10 as características, obteve-se uma taxa de acerto de 97 %. Conseguindo-se com poucas características distinguir eficientemente as imagens em suas respectivas

4.3 - Análise do vetor característico baseado em vários limiares para autovalores fixos⁶³

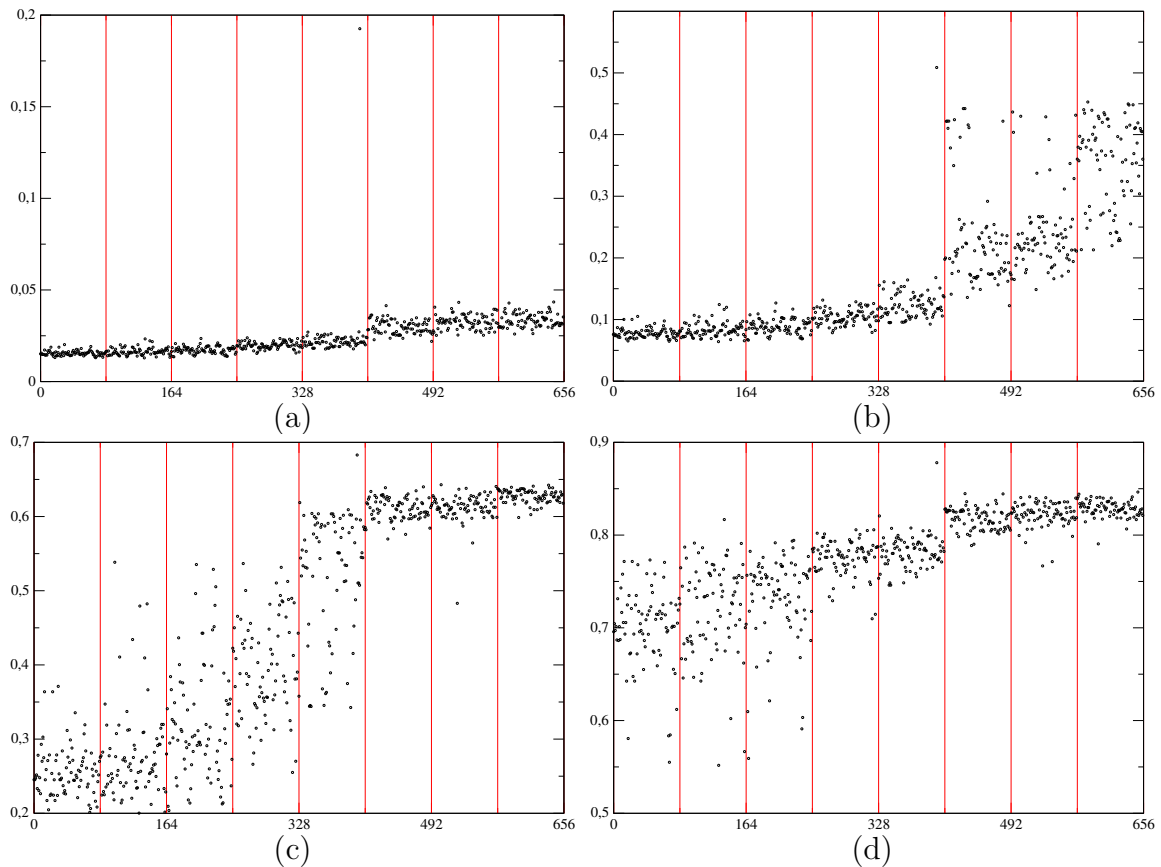


Figura 4.7 – Na ilustração, as linhas verticais separam as distintas 8 categorias de parasitas, cada bloco contém 82 imagens que são apresentadas por um ponto que representa o cálculo do segundo autovalor. As categorias foram ordenadas pelo valor médio obtido a partir do cálculo do segundo autovalor. Deste ordenamento se observou que as 3 categorias ao lado direito do gráfico possuem uma proximidade à forma de elipse, as 2 categorias da metade aproximam-se a formas elípticas e circulares e as últimas categorias uma similaridade a forma circular. Este mesmo comportamento não muda nos 4 gráficos, assim o valor limiar está aumentando.

categorias (Mais detalhes são apresentados no apêndice A.11).

Na tabela 4.5 observamos que uma alta taxa de acertos pode ser alcançada quando uma quantidade maior de limiares é utilizada.

Observamos na tabela 4.6, que um incremento de 10% na taxa de acerto é alcançado quando se utiliza 12 características em relação ao passo anterior, que usava 10 características. Assim também de 10 até 14 características incrementa em 5 % e

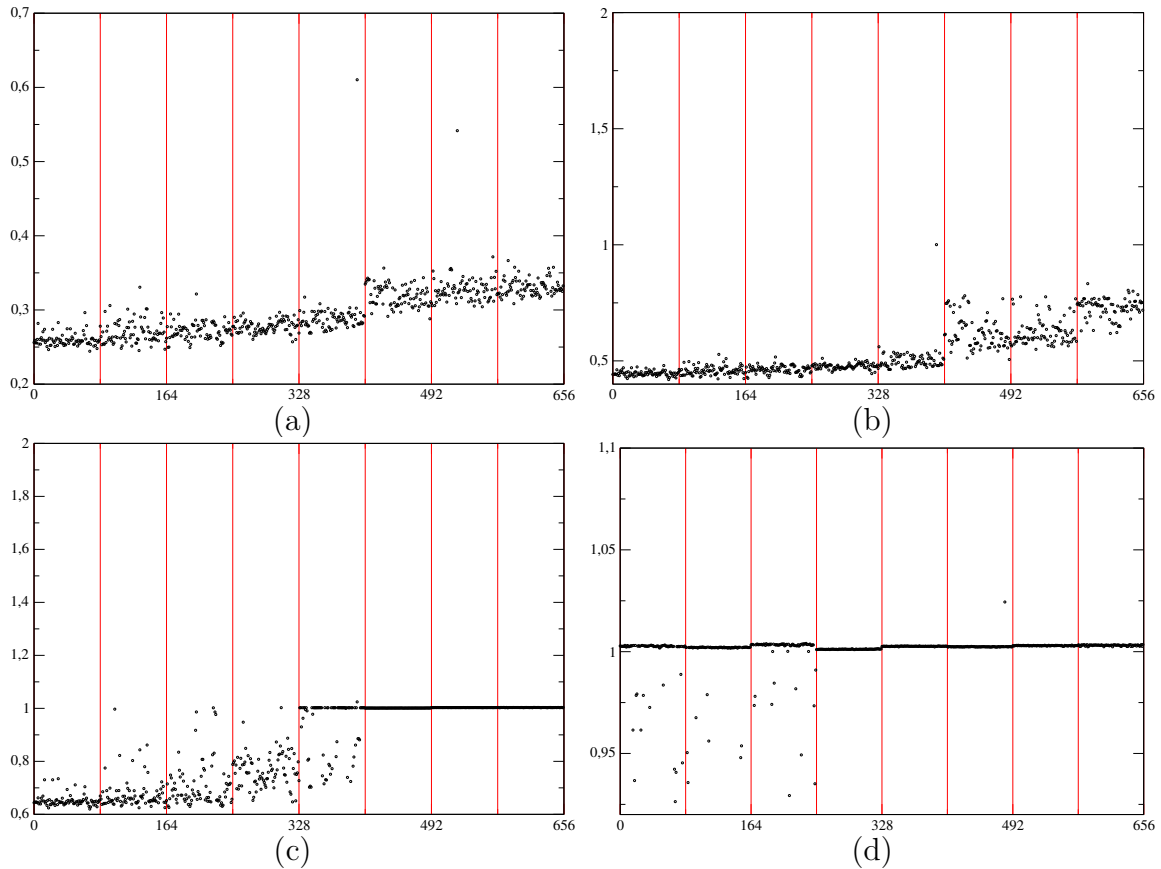


Figura 4.8 – As ilustrações mostram os mesmos procedimentos de cálculo realizados na figura 4.8, neste caso utilizamos o último autovalor que obtivemos nos gráficos com os limiares (a) $T_l = 0.3$, (b) $T_l = 0.5$, (c) $T_l = 0.7$ e (d) $T_l = 0.9$. Onde observou-se resultados similares com o segundo autovalor. As categorias estão divididas por linhas verticais, e seu ordenamento mostra na parte direita espécies próximas às elipses, na metade as espécies tem uma proximidade na forma entre a elipse e a circunferência e na parte esquerda as espécies com maior similaridade a uma circunferência.

seguidamente as porcentagens aumentam em pequenos passos até chegar a um valor máximo de 86 % com 76 características. Assim, podemos concluir que o segundo e último autovalores possuem propriedades importantes, capazes de representar os padrões que definem as formas dos objetos originais, e que o uso de diferentes limiares permite bons resultados na classificação de imagens em classes predefinidas.

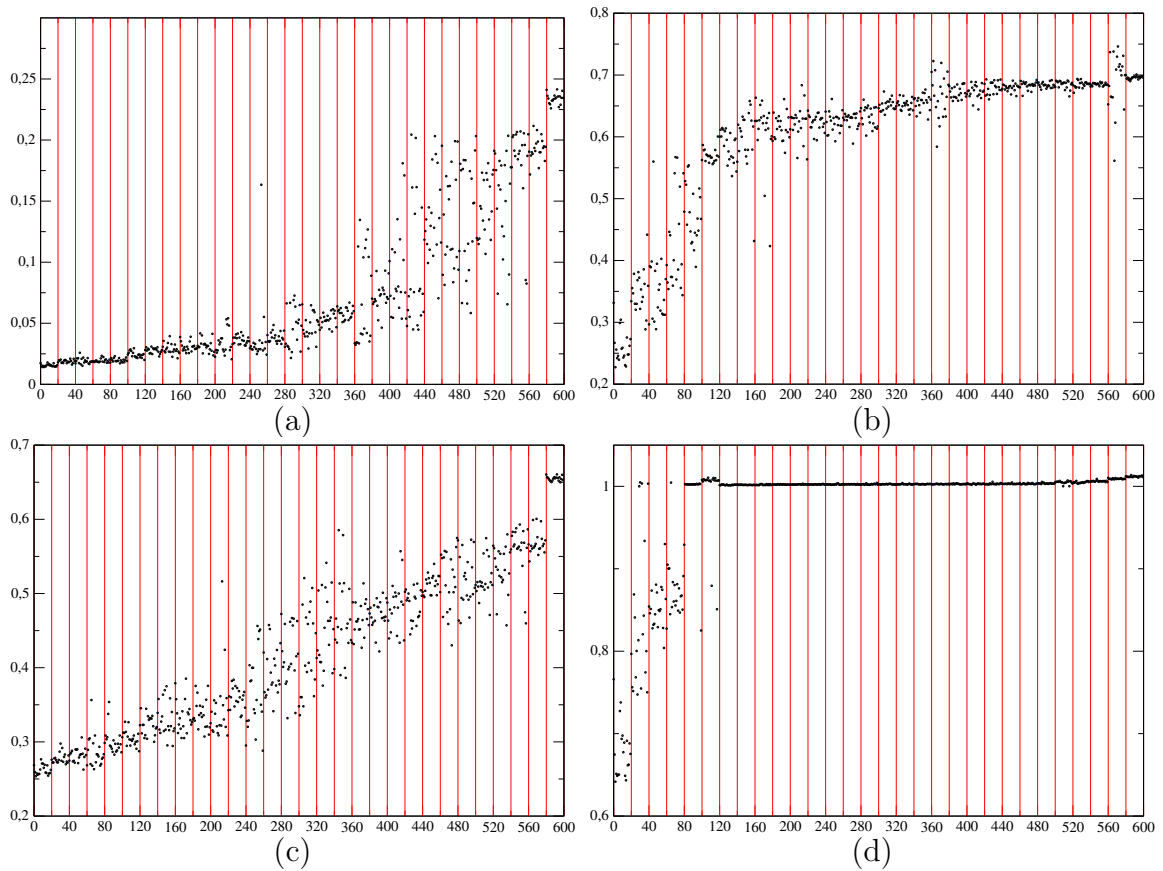


Figura 4.9 – Os quadros da primeira fila apresentam os cálculos do segundo autovalor e na segunda fila do ultimo autovalor. Os limiares utilizados são (a) $T_l = 0.3$, (b) $T_l = 0.7$, (c) $T_l = 0.3$ e (d) $T_l = 0.7$. Cada ponto representa uma imagem, as linhas vermelhas representam a separação das categorias e estas estão ordenadas pelo valor meio dos resultados obtidos por imagem. Nos gráficos mostra-se da direita à esquerda primeiramente os valores que apresentam uma tendência na forma elíptica e no lado esquerdo as folhas que conservam uma forma com proximidade à circunferência. Cada limiar extrai características da topologia do grafo que permite captar informações locais que são importantes para caracterizar os padrões.

4.4 Comparação com outros descritores

Para avaliar a eficácia dos procedimentos propostos pelo presente estudo, utilizamos algumas outras técnicas já consagradas pela literatura, tais como descritores de forma como redes complexas do trabalho de Andre [5], Momentos de Hu [72], medição de elipticidade (*FitElliptic*), curvatura e transformada de fourier

Tabela 4.4 – Na tabela mostra-se as taxa de acerto, obtidas na classificação das categorias da coleção de imagens Genéricas.

T_{ini}	T_{inc}	T_f	Número de características	Taxa de acerto (%)
0.025	0.025	0.95	76	98
0.025	0.050	0.95	38	98
0.025	0.075	0.95	26	97
0.025	0.1	0.95	20	97
0.025	0.125	0.95	16	97
0.025	0.15	0.95	14	97
0.025	0.175	0.95	12	97
0.025	0.2	0.95	10	97

(técnicas descritas no capítulo 2). Na tabela 4.7, apresentamos a quantidade de características utilizadas por descritor.

Estes descritores são aplicados às distintas coleções de imagens (Genérico, SADPI e CNShape), utilizados nesta pesquisa na qual compara-se as taxas de acerto obtidas pela máquina de aprendizagem SVM na classificação das categorias. O SVM utiliza os parâmetros ideais fornecidos pela biblioteca WEKA, e um parâmetro de complexidade $C = 25$. Os resultados são apresentados nas tabelas 4.8, 4.9, 4.10.

Na tabela 4.8, observamos que com as coleções de imagens Genéricas, as abordagens propostas e a de redes complexas obtiveram resultados acima de 90 %.

Enquanto que na coleção de imagens de parasitas SADPI 4.9, suas propriedades de medidas elípticas *FitEllipse*, redes complexas, e as abordagens aqui propostas obtiveram boas taxas de acerto, acima de 80 %.

Tabela 4.5 – Na tabela mostra-se a taxa de acerto na classificação de categorias da coleção de imagens SADPI. Onde observa-se uma maior taxa de acerto de 83% com 76 características.

T_{ini}	T_{inc}	T_f	Número de características	Taxa de acerto (%)
0.025	0.025	0.95	76	83
0.025	0.050	0.95	38	78
0.025	0.075	0.95	26	76
0.025	0.1	0.95	20	75
0.025	0.125	0.95	16	72
0.025	0.15	0.95	14	76
0.025	0.175	0.95	12	72
0.025	0.2	0.95	10	69

Na coleção de imagens das folhas CNShape 4.10, observa-se um conjunto mais complexo e técnicas como redes complexas e vetor característico baseado em vários limiares para autovalores fixos (abordagem 2) obtiveram resultados acima de 80 %. As técnicas como redes complexas, e o uso de segundo e último autovalores apresentaram resultados destacados, em vista de que estes dois procedimentos conseguem captar melhor as características gerais e também as características locais das imagens (Detalhes são apresentados nas tabelas do apêndice A.11 A.12, A.13, A.14) apresentando a capacidade de distinção dos padrões nas diferentes coleções de imagens.

Tabela 4.6 – Na tabela mostra-se os resultados da análise do segundo e último autovalor obtidos na evolução do grafo baseado nos limiares com a coleção de imagens CNShape.

T_{ini}	T_{inc}	T_f	Número de características	Taxa de acerto (%)
0.025	0.025	0.95	76	86.33
0.025	0.050	0.95	38	84.00
0.025	0.075	0.95	26	84.00
0.025	0.1	0.95	20	83.00
0.025	0.125	0.95	16	82.83
0.025	0.15	0.95	14	81.17
0.025	0.175	0.95	12	78.17
0.025	0.2	0.95	10	68.67

Tabela 4.7 – Número de características por descritor

Descritor	Numero de características
Fourier	32
<i>FitElliptic</i>	7
Curvature	3
Redes complexas	76
Hu	7
Abordagem 1	30
Abordagem 2	76

Tabela 4.8 – Na tabela mostra-se a taxa de acerto, obtida com a coleção de imagens Genéricas.

Tecnica	Numero de formas correctas	Taxa de acerto %
Fourier	87	88
<i>FitElliptic</i>	66	67
Curvature	33	33
Redes complexas	97	98
Hu	7	62
Abordagem 1	90	91
Abordagem 2	97	98

Tabela 4.9 – Na tabela mostra-se os resultados dos descritores utilizando a coleção de imagens SADPI.

Tecnica	Número de formas corretas	%
Fourier	513	78
<i>FitElliptic</i>	591	90
Curvature	276	42
Redes complexas	572	87
Hu	391	59.6
Abordagem 1	576	88
Abordagem 2	545	83

Tabela 4.10 – Na tabela mostra-se os resultados da análise dos descritores com a coleção de imagens CNShape.

Técnica	Número de formas correctas	%
Fourier	440	73
<i>FitElliptic</i>	334	40
Curvatura	64	11
Redes complexas	513	85
Hu	147	25
Abordagem 1	434	72
Abordagem 2	518	86

Conclusões e Discussão

O ponto crucial de nossa pesquisa é a análise de formas baseadas nos autovalores do espectro do laplaciano aplicada em imagens para a distinção delas dentre suas categorias. Nessa dissertação, trabalhou-se com coleções de imagens genéricas sintéticas, parasitas de humanos e folhas. Nossa proposta consistiu em duas abordagens: *Análise de vetor característico baseados em limiares e a histogramas* (a qual caracteriza a forma pela análise de um só limiar e seu espectro associado) e a *Análise da vetor característico baseado em vários limiares para autovalores fixos*.

Na abordagem *Análise de vetor característico baseados em limiares e a histogramas* obteve-se resultados promissórios nas coleções de imagens sintéticas e parasitas com resultados acima de 85% na categorização dos padrões, porém nas imagens de folhas não desenvolveram o mesmo comportamento pela complexidade dos padrões de folha.

Na segunda abordagem *Análise da vetor característico baseado em vários limiares para autovalores fixos*, conseguiu-se resultados aceitáveis para os três bancos de imagens. Nela, são utilizados como elementos de caracterização do segundo e do último autovalor da matriz laplaciana. Eles carregam propriedades importantes do espectro do grafo calculadas para cada limiar. Assim, observamos que este último

procedimento obtêm informações *globais e locais* que o método precisa para caracterizar padrões complexos como folhas e extrair aquelas peculiaridades locais. Nas coleções de imagens consideradas obteve-se valores acima de 80%, o que permite e motiva melhorar posteriormente esta análise.

Outro aspecto importante observado nesta segunda abordagem é a importância do segundo autovalor para caracterizar um padrão. Na coleção de imagens observou-se que este distingue entre formas circulares e elípticas, e desempenha um papel importante no grau de conectividade dos padrões de seus grafos gerados pelos limiares.

Observamos que das duas abordagens o segundo autovalor é estável na representação das imagens, conseguindo mediante o segundo e último autovalor capturar as características globais e locais da forma. A primeira abordagem apresenta bons resultados em imagens que não requerem características locais.

Neste trabalho, observou-se a importância da topologia de grafos mediante a análise da matriz laplaciana, pois ela introduz complexidade como mecanismo de separação entre as categorias de imagens. Alguns padrões são mais complexos que outros, o que torna necessário o desenvolvimento de novos atributos ou informações para melhorar a distinção entre as categorias. Nossos resultados sugerem que esse objetivo será alcançado por meio de estudos mais aprofundados dos autovalores, autovetores e dos limiares que caracterizam a matriz laplaciana. Isso permite uma nova frente de pesquisas em análise de imagens, textos e redes sociais.

A construção do grafo feita pelos limiares, captura informação importante nos espectros. Um procedimento futuro consiste em otimizar a formação do grafo para caracterizar eficientemente e em menos tempo os pontos principais a serem considerados do padrão. Durante a avaliação obtivemos resultados promissores com

a coleção de imagens de parasitas, obtendo um valor limiar adequado para sua caracterização, diminuindo o número de limiares e custos de cálculo.

O desenvolvimento de processos que minimizem a quantidade de limiares (elementos do vetor característico) necessários para caracterização de padrões, sem perder precisão, reduz o custo computacional, ampliando as possíveis aplicações do método proposto nesta dissertação. Com a obtenção de autovalores mais relevantes e com melhor eficiência os autovalores mais relevantes e com melhor eficiência. Um ponto importante será a implementação do método aplicado em texturas, pois carregam mais informações de conteúdo, aumentando a complexidade da topologia do grafo correspondente. Isto está de acordo com o delineamento da análise do espectro laplaciano. Os limiares e cálculo do espectro pode ser combinados com outros elementos de caracterização que podem melhorar a observação destes resultados sendo um tema de pesquisa no futuro.

Anexos

A.1 Análise do vetor característico baseados em limiares e a histogramas

Tabela A.1 – Análise de Vetor característico baseados em limiares e histogramas (Genérico). Parte 1

IC_l	T_l																		
	0.025	0.050	0.075	0.100	0.125	0.150	0.175	0.200	0.225	0.250	0.275	0.300	0.325	0.350	0.375	0.400	0.425	0.450	0.475
10	53	65	65	68	73	73	76	79	70	76	87	84	82	83	83	75	77	82	91
15	57	64	70	69	76	79	70	83	78	81	76	85	87	81	82	87	83	86	79
20	57	63	70	69	80	80	81	84	74	77	79	79	85	82	81	85	85	84	80
25	53	62	69	65	77	83	83	84	79	83	86	86	87	81	87	88	87	84	75
30	55	60	66	69	74	85	77	81	83	77	84	85	87	84	88	91	85	78	75
35	60	62	67	64	83	76	86	78	78	82	79	87	88	84	82	85	84	84	71
40	59	57	68	68	79	79	80	86	84	68	84	86	88	82	89	83	87	83	71
45	62	56	65	69	76	78	82	84	80	80	78	85	86	83	89	89	77	80	76
50	62	57	71	63	79	80	78	87	79	78	85	87	84	81	89	86	86	78	72
55	58	64	63	65	83	76	79	79	76	75	78	86	85	81	85	86	85	75	76
60	57	64	69	63	77	76	84	80	83	77	85	87	84	82	86	83	83	78	73
65	59	58	68	71	71	73	86	80	82	81	83	89	84	82	84	87	87	75	73
70	53	61	59	71	82	79	87	79	81	79	82	89	87	80	90	86	86	79	67
75	60	64	75	65	77	79	83	78	81	70	83	87	83	82	83	88	85	74	73
80	53	59	64	69	75	73	81	86	78	78	83	85	85	83	89	85	84	77	69
85	54	57	64	70	72	76	85	83	78	75	85	84	83	79	88	83	86	76	74
90	54	62	64	67	74	79	86	85	74	76	85	86	80	81	88	87	82	76	72
95	57	56	60	67	74	79	81	83	81	75	82	83	83	82	84	85	80	79	69
100	52	58	65	67	74	76	80	83	78	78	85	83	79	84	80	84	84	76	69

Tabela A.2 – Análise de Vetor característico baseados em limiares e histogramas (Genérico). Parte 2

IC_l	T_l																		
	0.500	0.525	0.550	0.575	0.600	0.625	0.650	0.675	0.700	0.725	0.750	0.775	0.800	0.825	0.850	0.875	0.900	0.925	0.950
10	81	84	75	76	72	76	78	80	80	77	72	73	75	76	69	62	59	63	57
15	85	75	69	74	74	76	79	80	84	86	80	76	77	77	72	72	67	61	64
20	82	69	76	76	78	79	76	87	87	88	85	81	74	67	73	66	68	59	
25	80	75	72	80	78	76	77	87	87	84	86	81	76	78	76	64	60	56	
30	81	70	76	77	76	75	78	82	82	83	86	79	75	73	68	71	62	65	
35	75	68	79	75	82	74	84	81	81	80	85	78	74	73	70	64	58	62	
40	81	70	77	79	78	74	76	84	84	74	83	81	79	71	74	64	57	59	
45	79	76	76	79	77	75	79	87	87	81	82	77	77	73	77	66	62	57	
50	80	71	76	82	80	77	76	82	82	76	78	78	81	74	75	68	66	58	
55	75	64	77	76	81	72	83	78	78	83	81	77	71	71	79	63	59	62	
60	78	73	73	83	78	70	78	80	80	81	78	76	77	72	73	66	64	60	
65	71	72	75	82	77	73	78	78	78	79	76	81	72	77	71	68	64	57	
70	77	73	74	83	77	65	75	77	77	79	79	80	81	75	73	62	64	51	
75	72	66	74	79	76	70	74	76	76	82	80	80	70	73	73	68	66	54	
80	80	69	75	81	79	70	71	79	79	71	78	75	75	70	72	71	65	53	
85	73	74	80	80	75	71	74	79	79	77	79	78	72	70	73	63	65	58	
90	74	72	72	81	78	70	73	80	80	80	81	75	72	68	75	67	60	57	
95	75	72	79	80	76	72	67	79	79	75	75	73	69	73	72	74	63	63	52
100	79	68	74	80	78	71	74	81	77	76	76	74	69	69	70	65	60	67	49

Tabela A.3 – Análise de Vetor característico baseados em limiares e histogramas (SADPI). Parte 1

IC_i	T_i																		
	0.025	0.050	0.075	0.100	0.125	0.150	0.175	0.200	0.225	0.250	0.275	0.300	0.325	0.350	0.375	0.400	0.425	0.450	0.475
10	60	65	58	56	61	56	70	62	63	71	73	80	66	67	69	69	63	82	78
15	63	65	64	73	71	73	75	75	72	76	77	75	79	71	68	72	73	80	78
20	65	68	69	78	75	79	79	77	79	75	80	76	72	72	78	74	80	80	78
25	65	68	71	75	80	80	80	78	79	78	80	80	78	78	75	80	75	83	77
30	66	71	71	78	77	80	80	79	81	79	80	79	78	74	77	76	83	83	78
35	66	71	72	75	75	80	80	80	76	80	82	80	78	80	77	80	83	83	78
40	64	67	74	77	77	81	81	81	78	80	80	80	81	74	77	77	82	82	78
45	63	71	70	76	77	83	80	80	77	81	82	80	81	77	75	79	83	84	79
50	66	70	73	77	80	82	82	79	78	77	80	80	79	79	78	78	78	84	79
55	63	70	72	78	78	78	78	82	81	80	83	80	81	76	82	76	82	82	80
60	61	70	72	78	77	79	79	81	76	78	79	80	81	78	77	80	80	80	78
65	60	70	71	77	77	78	78	81	76	78	79	80	80	74	79	79	78	82	80
70	59	72	72	74	79	79	79	80	77	80	81	79	78	78	79	77	80	80	78
75	61	70	74	76	76	76	76	77	77	78	80	81	82	77	79	77	79	79	79
80	61	70	69	74	75	74	76	78	77	81	78	79	79	77	78	77	79	78	79
85	60	68	71	73	73	73	77	78	75	78	80	80	81	78	77	79	77	78	75
90	60	68	74	72	76	72	77	77	75	81	78	77	79	76	75	77	76	77	76
95	62	71	73	70	73	74	79	79	75	77	77	78	79	73	76	76	76	75	77
100	61	71	71	73	73	73	76	79	76	76	75	77	79	74	74	75	72	76	77

Tabela A.4 – Análise de Vetor característico baseados em limiares e histogramas (SADPI). Parte 2

IC_l	T_l																		
	0.500	0.525	0.550	0.575	0.600	0.625	0.650	0.675	0.700	0.725	0.750	0.775	0.800	0.825	0.850	0.875	0.900	0.925	0.950
10	77	73	71	81	78	79	76	78	76	84	82	83	82	78	79	71	69	67	59
15	75	71	78	75	72	80	80	78	80	79	81	86	83	80	80	76	69	63	59
20	74	77	78	81	77	81	79	83	79	85	84	86	86	82	80	80	71	73	64
25	78	80	79	82	83	80	83	84	82	84	84	86	85	82	78	77	72	70	68
30	79	78	82	82	78	83	84	84	84	85	86	88	83	83	79	76	74	70	65
35	78	79	80	81	78	84	85	84	85	83	84	88	85	82	79	79	74	73	69
40	78	80	81	84	79	81	84	82	84	83	87	88	87	83	82	77	75	74	68
45	77	81	80	82	78	81	84	84	84	84	85	87	84	82	80	78	75	73	68
50	78	79	79	81	80	83	84	84	84	85	86	88	85	82	80	78	76	73	68
55	77	78	78	83	81	83	82	82	82	85	84	88	87	83	81	77	74	73	70
60	76	78	77	82	79	81	82	82	81	86	85	86	85	81	80	75	75	74	70
65	77	79	78	81	80	80	83	81	84	83	84	85	82	82	83	77	73	73	69
70	80	78	78	82	78	78	83	83	83	85	84	88	82	81	77	79	72	75	69
75	75	77	78	81	81	81	82	80	82	83	84	85	82	80	79	76	75	74	68
80	79	78	77	81	79	79	80	81	82	82	83	85	85	83	81	75	74	74	70
85	79	77	77	81	78	79	80	82	80	82	83	85	83	80	79	75	72	74	71
90	77	79	77	81	79	78	83	80	83	83	82	83	81	82	79	75	70	73	68
95	77	75	76	78	78	79	80	80	80	82	82	83	83	80	77	73	72	73	69
100	76	75	76	80	78	80	81	80	79	80	82	84	82	81	77	74	71	73	68

Tabela A.5 – Análise de Vetor característico baseados em limiares e histogramas (CNShape). Parte 1

IC_l	T_l																		
	0.025	0.050	0.075	0.100	0.125	0.150	0.175	0.200	0.225	0.250	0.275	0.300	0.325	0.350	0.375	0.400	0.425	0.450	0.475
10	33	43	52	55	54	57	60	63	62	61	61	54	59	60	63	64	63	64	65
15	37	46	54	56	59	65	62	67	69	68	67	66	64	62	63	66	68	70	68
20	33	46	56	61	66	64	68	67	67	68	66	63	65	64	64	70	67	69	68
25	32	43	52	61	64	67	68	68	68	67	67	65	66	65	62	70	71	67	68
30	33	43	56	61	66	70	67	69	69	65	67	64	66	62	63	68	68	69	69
35	31	44	56	64	65	68	66	67	70	66	63	66	63	62	63	65	71	68	69
40	34	43	57	66	67	66	70	65	69	69	66	66	62	63	64	69	66	67	67
45	31	43	54	67	69	69	67	66	67	68	66	65	64	61	61	65	67	68	67
50	30	43	53	66	66	66	66	67	68	66	66	63	64	61	61	65	67	69	65
55	31	43	53	65	64	69	68	65	69	68	64	64	62	62	65	67	65	69	64
60	32	42	52	61	66	66	66	67	67	64	63	60	63	62	65	67	67	67	63
65	33	42	53	61	65	68	67	66	67	66	64	61	56	61	62	63	66	67	65
70	29	40	56	65	61	64	67	65	65	64	60	61	59	58	61	64	63	63	65
75	31	40	56	60	64	63	66	64	64	63	62	63	59	62	61	63	65	66	63
80	30	42	53	58	64	65	62	64	65	64	63	61	58	57	61	62	64	64	58
85	31	42	51	62	62	63	64	64	61	64	63	60	58	58	63	63	64	62	60
90	32	41	53	61	64	63	63	62	61	61	61	59	59	58	59	62	66	65	60
95	31	43	52	59	61	64	63	63	63	62	60	58	57	57	61	62	62	62	59
100	30	41	54	61	60	59	63	59	63	64	59	57	58	57	59	61	61	63	63

Tabela A.6 – Análise de Vetor característico baseados em limiares e histogramas (CNShape). Parte 2

IC_l	T_l																		
	0.500	0.525	0.550	0.575	0.600	0.625	0.650	0.675	0.700	0.725	0.750	0.775	0.800	0.825	0.850	0.875	0.900	0.925	0.950
10	68	69	67	64	69	66	68	64	62	55	53	55	53	49	47	45	38	31	29
15	70	71	70	72	68	66	69	63	62	61	55	57	56	52	52	48	41	37	34
20	70	71	70	72	69	68	66	66	64	60	60	58	60	56	54	49	45	41	37
25	69	70	71	72	71	68	69	64	65	60	60	59	57	57	56	52	47	43	38
30	69	68	71	71	70	67	67	64	62	61	59	62	58	58	56	54	48	41	40
35	68	68	71	70	72	67	66	66	65	61	59	59	59	57	58	53	52	47	41
40	68	68	69	70	68	67	65	62	62	61	59	59	58	57	57	55	52	48	42
45	66	65	69	67	70	70	63	63	62	59	59	59	61	61	60	55	53	48	43
50	66	65	71	66	68	66	64	62	65	61	59	58	60	59	61	55	51	49	45
55	66	67	70	66	68	65	62	62	62	58	60	60	59	61	57	56	53	50	44
60	63	65	67	67	67	65	63	63	60	61	59	58	59	59	58	58	50	51	46
65	65	64	67	64	69	64	61	59	63	59	58	61	59	59	59	55	53	51	46
70	67	63	68	66	67	64	62	63	62	61	60	60	58	58	57	57	52	52	45
75	65	64	68	64	65	64	62	62	61	58	59	59	59	59	58	57	53	50	45
80	64	60	67	64	66	61	61	59	58	60	58	60	56	58	56	57	52	50	46
85	60	65	67	64	63	62	59	61	61	61	59	58	59	59	55	55	53	51	48
90	64	62	64	64	63	63	59	61	60	60	57	56	56	59	58	55	52	51	47
95	63	63	64	64	65	65	60	58	60	58	58	58	57	59	56	57	53	52	47
100	63	63	66	64	64	61	59	59	61	59	57	57	57	56	55	54	50	52	50

Tabela A.7 – Matriz de confusão da melhor diferenciação entre as categorias, mediante o análise de vetor característico baseados em limiares e histogramas(banco de imagens Genérico) com $IC_l = 10$, $T_l = 0.475$ e taxa de acerto de 91%.(11 imagens por cada classe)

	G1	G2	G3	G4	G5	G6	G7	G8	G9
G1	11	0	0	0	0	0	0	0	0
G2	0	11	0	0	0	0	0	0	0
G3	0	0	9	0	0	2	0	0	0
G4	0	0	0	10	1	0	0	0	0
G5	1	1	0	1	8	0	0	0	0
G6	0	0	2	0	0	9	0	0	0
G7	1	0	0	0	0	0	10	0	0
G8	0	0	0	0	0	0	0	11	0
G9	0	0	0	0	0	0	0	0	11

Tabela A.8 – Análise de vetor característico baseados em limiares e histogramas com o banco de imagens (SADPI)($IC_l = 30$, $T_l = 0.575$ e 88 % de acerto)

	S1	S2	S3	S4	S5	D6	S7	S8
S1	72	2	0	8	0	0	0	0
S2	0	74	2	0	5	0	0	1
S3	0	0	64	0	0	1	0	17
S4	8	0	0	69	2	3	0	0
S5	1	6	0	4	71	0	0	0
S6	0	3	0	7	1	71	0	0
S7	0	0	0	0	0	0	82	0
S8	0	0	7	0	0	0	0	75

A.2 Análise do vetor característico baseado em vários limiares para autovalores fixos

Tabela A.11 – Matriz de confusão obtida com limiares ($T_{inicial} = 0.025, T_{incremento} = 0.2, T_{final} = 0.95$), no qual se obteve uma taxa de acerto de 97 % no banco de dados Genérico

	G1	G2	G3	G4	G5	G6	G7	G8	G9
G1	11	0	0	0	0	0	0	0	0
G2	0	11	0	0	0	0	0	0	0
G3	0	0	11	0	0	0	0	0	0
G4	1	0	0	8	1	0	1	0	0
G5	0	0	0	0	11	0	0	0	0
G6	0	0	0	0	0	11	0	0	0
G7	0	0	0	0	0	0	11	0	0
G8	0	0	0	0	0	0	0	11	0
G9	0	0	0	0	0	0	0	0	11

Tabela A.12 – Matriz de confusão obtida com limiares ($T_{inicial} = 0.025, T_{incremento} = 0.025, T_{final} = 0.95$), no qual se obteve uma taxa de acerto de 83 %, com o banco SADPI

	S1	S2	S3	S4	S5	S6	S7	S8
S1	51	5	0	18	7	1	0	0
S2	8	71	2	0	1	0	0	0
S3	0	0	75	0	0	0	0	7
S4	10	0	0	63	4	5	0	0
S5	6	2	0	3	69	2	0	0
S6	2	2	0	15	5	58	0	0
S7	0	1	0	0	0	0	80	1
S8	0	0	4	0	0	0	0	78

Referências¹

- [1] Zhang D, Lu G. Review of shape representation and description techniques. *Pattern Recognition*. 2004;37(1):1–19. xv, 2, 10
- [2] Gonzalez RC, Woods RE. *Digital image processing 3rd edition*. Prentice Hall; 2007. xv, 14, 15, 39
- [3] Gonçalves WN. *Caminhadas determinísticas em redes complexas aplicadas em visão computacional*. Universidade de São Paulo;. xv, 23
- [4] Noma A, Graciano AB, Cesar Jr RM, Consularo LA, Bloch I. Interactive image segmentation by matching attributed relational graphs. *Pattern Recognition*. 2012;45(3):1159–1179. xv, 23
- [5] Backes AR, Casanova D, Bruno OM. A complex network-based approach for boundary shape analysis. *Pattern Recognition*. 2009;42(1):54–67. xvi, 40, 45, 54, 65
- [6] Castañón CA, Fraga JS, Fernandez S, Gruber A, Costa LdF. Biological shape characterization for automatic image recognition and diagnosis of protozoan parasites of the genus *Eimeria*. *Pattern Recognition*. 2007;40(7):1899–1910. 1

¹De acordo com o estilo LMSC (baseado no estilo IEEE Transactions - Electrical and Electronics Engineers).

-
- [7] He L, Long LR, Antani S, Thoma GR. Histology image analysis for carcinoma detection and grading. *Computer Methods and Programs in Biomedicine*. 2012;107(3):538–556. 1
- [8] Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: A review. *Biomedical Engineering, IEEE Reviews in*. 2009;2:147–171. 1
- [9] Wertheimer-Hatch L, Hatch I GF, Davis, et al. Tumors of the oral cavity and pharynx. *World Journal of Surgery*. 2000;24(4):395–400. 1
- [10] Eid RA, Landini G. The complexity of the oral mucosa: a review of the use of fractal geometry. *Journal of Control Engineering and Applied Informatics*. 2010;12(1):10–14. 1
- [11] Schomaker L, de Leau E, Vuurpijl L. Using pen-based outlines for object-based annotation and image-based queries. In: *Visual Information and Information Systems*. Springer; 1999. p. 585–592. 2
- [12] Alajlan N, El Rube I, Kamel MS, Freeman G. Shape retrieval using triangle-area representation and dynamic space warping. *Pattern Recognition*. 2007;40(7):1911–1920. 2, 3
- [13] Reljin B, Milosević Z, Stojić T, Reljin I. Computer aided system for segmentation and visualization of microcalcifications in digital mammograms. *Folia Histochemica et Cytobiologica*. 2010;47(3):525–524. 2
- [14] Loncaric S. A survey of shape analysis techniques. *Pattern Recognition*. 1998;31(8):983–1001. 2, 9
- [15] Hu MK. Visual pattern recognition by moment invariants. *Information Theory, IRE Transactions on*. 1962;8(2):179–187. 2

-
- [16] Flusser J. On the independence of rotation moment invariants. *Pattern Recognition*. 2000;33(9):1405–1410. 2
- [17] Teague MR. Image analysis via the general theory of moments*. *JOSA*. 1980;70(8):920–930. 2
- [18] Mukundan R, Ong S, Lee PA. Image analysis by Tchebichef moments. *Image Processing, IEEE Transactions on*. 2001;10(9):1357–1364. 2
- [19] Mokhtarian F, Bober M. Curvature scale space representation: theory, applications, and MPEG-7 standardization. vol. 25. Springer Science & Business Media; 2013. 2
- [20] Zhao Y, Belkasim S. Multiresolution Fourier descriptors for multiresolution shape analysis. *Signal Processing Letters, IEEE*. 2012;19(10):692–695. 2
- [21] Chuang GC, Kuo CJ. Wavelet descriptor of planar curves: theory and applications. *Image Processing, IEEE Transactions on*. 1996;5(1):56–70. 2
- [22] da Fontoura Costa L, Cesar Jr RM. Shape analysis and classification: theory and practice. CRC Press; 2010. 2, 3, 39
- [23] Torres RdS, Falcão AX, Costa LdF. A graph-based approach for multiscale shape analysis. *Pattern Recognition*. 2004;37(6):1163–1174. 3
- [24] Sebastian TB, Klein PN, Kimia BB. Recognition of shapes by editing their shock graphs. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 2004;26(5):550–571. 3, 53
- [25] El-ghazal A, Basir O, Belkasim S. Invariant curvature-based Fourier shape descriptors. *Journal of Visual Communication and Image Representation*. 2012;23(4):622–633. 3, 16

-
- [26] McInerney T, Terzopoulos D. Deformable models in medical image analysis: a survey. *Medical image analysis*. 1996;1(2):91–108. 3
- [27] Ye X, Lin X, Dehmeshki J, Slabaugh G, Beddoe G. Shape-based computer-aided detection of lung nodules in thoracic CT images. *Biomedical Engineering, IEEE Transactions on*. 2009;56(7):1810–1820. 3
- [28] Chalana V, Kim Y. A methodology for evaluation of boundary detection algorithms on medical images. *Medical Imaging, IEEE Transactions on*. 1997;16(5):642–652. 3
- [29] Castañon CAB. Análise e reconhecimento digital de formas biológicas para o diagnóstico automático de parasitas do gênero *Eimeria*. Universidade de São Paulo; 2006. 3
- [30] Vromen J, McCane B. Red blood cell segmentation from SEM images. In: *Image and Vision Computing New Zealand, 2009. IVCNZ'09. 24th International Conference*. IEEE; 2009. p. 44–49. 3
- [31] Carcassoni M, Hancock ER. Spectral correspondence for point pattern matching. *Pattern Recognition*. 2003;36(1):193–204. 4
- [32] Manjunath BS, Ohm JR, Vasudevan VV, Yamada A. Color and texture descriptors. *Circuits and Systems for Video Technology, IEEE Transactions on*. 2001;11(6):703–715. 7, 38
- [33] Guyon I. *Feature extraction: foundations and applications*. vol. 207. Springer Science & Business Media; 2006. 8, 9
- [34] Weinland D, Ronfard R, Boyer E. A survey of vision-based methods for action representation, segmentation and recognition. *Computer Vision and Image Understanding*. 2011;115(2):224–241. 9

-
- [35] Sonka M, Hlavac V, Boyle R. Image processing, analysis, and machine vision. Cengage Learning; 2014. 9, 10
- [36] Yampolskiy RV. Feature Extraction Methods for Character Recognition. 2004;. 9
- [37] Devijver PA, Kittler J. Pattern recognition: A statistical approach. vol. 761. Prentice-Hall London; 1982. 9
- [38] Yang M, Kpalma K, Ronsin J. A survey of shape feature extraction techniques. Pattern Recognition. 2008;p. 43–90. 9
- [39] Sikora T. The MPEG-7 visual standard for content description-an overview. Circuits and Systems for Video Technology, IEEE Transactions on. 2001;11(6):696–702. 9
- [40] Jorge-Hernandez F, Chimeno G. Graph theory for feature extraction and classification: A migraine pathology case study. Bio-medical materials and engineering. 2014;24(6):2979–2986. 10
- [41] Landini G, Othman IE. Architectural analysis of oral cancer, dysplastic, and normal epithelia. Cytometry Part A. 2004;61(1):45–55. 10
- [42] Landini G, Othman I. Estimation of tissue layer level by sequential morphological reconstruction. Journal of microscopy. 2003;209(2):118–125. 10
- [43] Fitzgibbon AW, Pilu M, Fisher RB. Direct least-squares fitting of ellipses. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1999;21:476–480. 11, 12
- [44] Stojmenovic M, Nayak A. Direct ellipse fitting and measuring based on shape boundaries. In: Advances in Image and Video Technology. Springer; 2007. p. 221–235. 11

-
- [45] Wong Y, Lin S, Ren T, Kwok N. A survey on ellipse detection methods. In: Industrial Electronics (ISIE), 2012 IEEE International Symposium on. IEEE; 2012. p. 1105–1110. 11
- [46] Greggio N, Manfredi L, Laschi C, Dario P, Carrozza MC. Robotcub implementation of real-time least-square fitting of ellipses. In: Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on. IEEE; 2008. p. 174–181. 11
- [47] Bonin-Font F, Ortiz A, Oliver G. Visual navigation for mobile robots: A survey. Journal of intelligent and robotic systems. 2008;53(3):263–296. 12
- [48] Harel J, Koch C, Perona P. Graph-based visual saliency. In: Advances in neural information processing systems; 2006. p. 545–552. 18
- [49] Cheddad A, Mohamad D, Manaf AA. Exploiting Voronoi diagram properties in face segmentation and feature extraction. Pattern recognition. 2008;41(12):3842–3859. 18
- [50] Foggia P, Percannella G, Vento M. Graph matching and learning in pattern recognition in the last 10 years. International Journal of Pattern Recognition and Artificial Intelligence. 2014;28(01):1450001. 19
- [51] Netto POB. Grafos: teoria, modelos, algoritmos. Edgard Blücher; 2003. 19
- [52] Zhu X, Ramanan D. Face detection, pose estimation, and landmark localization in the wild. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE; 2012. p. 2879–2886. 24
- [53] Milborrow S, Nicolls F. Active Shape Models with SIFT Descriptors and MARS. In: VISAPP (2); 2014. p. 380–387. 24

-
- [54] Hückel E. Quantentheoretische beiträge zum benzolproblem. *Zeitschrift für Physik A Hadrons and Nuclei*. 1931;70(3):204–286. 24
- [55] Abreu. Teoria Espectral dos Grafos: um Híbrido entre a Álgebra Linear e a Matemática Discreta e Combinatória com Origens na Química Quântica doi: 10.5540/tema. 2005.06. 01.0001. *Trends in Applied and Computational Mathematics*. 2005;6(1):1–10. 25
- [56] Hogben L. Spectral graph theory and the inverse eigenvalue problem of a graph. *Electronic Journal of Linear Algebra*. 2005;14(1):3. 25
- [57] Abreu N, Del-Vecchio R, Vinagre C, Stevanovic D. Introdução à teoria espectral de grafos com aplicações. *Notas em Matemática Aplicada*. 2007;27. 25
- [58] Vapnik V. *The nature of statistical learning theory*. Springer Science & Business Media; 2013. 29
- [59] Cristianini N, Shawe-Taylor J. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press; 2000. 30
- [60] Hearst M, Scholkopf B, Dumais S, Osuna E, Platt J. Trends and controversies—support vector machines. *IEEE Intelligent Systems*. 1998;13(4):18–28. 30
- [61] Haykin S, Network N. *A comprehensive foundation. Neural Networks*. 2004;2(2004). 30
- [62] Zien A, Rätsch G, Mika S, Schölkopf B, Lengauer T, Müller KR. Engineering support vector machine kernels that recognize translation initiation sites. *Bioinformatics*. 2000;16(9):799–807. 30
- [63] Gonçalves AR. *Maquina de Vetores Suporte*. Acesso em;21. 31

-
- [64] Tsuda K. Support vector classifier with asymmetric kernel functions. In: in European Symposium on Artificial Neural Networks (ESANN. Citeseer; 1999. . 33
- [65] Ribeiro MX. Suporte a sistemas de auxílio ao diagnóstico e de recuperação de imagens por conteúdo usando mineração de regras de associação. Universidade de São Paulo; 2008. 34
- [66] Otsu N. A threshold selection method from gray-level histograms. *Automatica*. 1975;11(285-296):23–27. 39
- [67] Shi J, Malik J. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 2000;22(8):888–905. 50
- [68] Weka Data mining Software in Java;. Accessed: 2015-11-27. <http://www.cs.waikato.ac.nz/ml/weka/>. 51
- [69] Universidade de São Paulo. www.pibapunsa.com; 2 de fevereiro de 2013. 53
- [70] Universidade de São Paulo. <http://scg.ifsc.usp.br/dataset/ShapeCN.php>; 22 de Dezembro de 2015. 53
- [71] Humpire G, Chuctaya J, Cáceres J, Bruno O, Beltrán C. Sistema Automático de Diagnóstico de Parasitos Helmintos. VII WORKshop de Visão Computacional WVC 2011;. 54
- [72] Mercimek M, Gulez K, Mumcu TV. Real object recognition using moment invariants. *Sadhana*. 2005;30(6):765–775. 65