



UNIVERSIDADE ESTADUAL PAULISTA  
“JÚLIO DE MESQUITA FILHO”  
Campus de São José do Rio Preto

Vitor Emanuel da Silva Rozeno

# **Caracterização de redes complexas utilizando Fisher Vector para análise de contornos**

São José do Rio Preto

2024

Vitor Emanuel da Silva Rozeno

## **Caracterização de redes complexas utilizando Fisher Vector para análise de contornos**

Monografia apresentada como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação, junto ao Departamento de Ciências de Computação e Estatística do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de São José do Rio Preto.

Universidade Estadual Paulista "Júlio de Mesquita Filho" - (UNESP)

Instituto de Biociências, Letras e Ciências Exatas - (IBILCE)

Departamento de Ciências de Computação e Estatística

Bacharelado em Ciência da Computação

Orientador: Prof. Dr. Lucas Correia Ribas

São José do Rio Preto

2024

R893c

Rozeno, Vitor Emanuel da Silva

Caracterização de redes complexas utilizando Fisher Vector para análise de contornos / Vitor Emanuel da Silva Rozeno. -- São José do Rio Preto, 2024

67 p. : il., tabs.

Trabalho de conclusão de curso (Bacharelado - Ciência da Computação) - Universidade Estadual Paulista (UNESP), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto

Orientador: Lucas Correia Ribas

1. Reconhecimento de padrões. 2. Análise de imagem. 3. Teoria dos grafos. 4. Ciência da computação. I. Título.

Vitor Emanuel da Silva Rozeno

## **Caracterização de redes complexas utilizando Fisher Vector para análise de contornos**

Monografia apresentada como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação, junto ao Departamento de Ciências de Computação e Estatística do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de São José do Rio Preto.

Banca Examinadora:

**Prof. Dr. Lucas Correia Ribas**

UNESP - Câmpus de São José do Rio Preto  
Orientador

**Prof. Dr. Rodrigo Capobianco Guido**

UNESP - Câmpus de São José do Rio Preto

**Prof. Dr. Diego Renan Bruno**

UNESP - Câmpus de São José do Rio Preto

São José do Rio Preto

30 de Outubro de 2024

Aos meus pais, Gerolino e Sandra, que sempre estiveram ao meu lado.

# Agradecimentos

Primeiramente, agradeço a Deus, que dá sentido à minha vida, proporcionando infinita graça, alegria e paz que vão além de qualquer entendimento humano. Ele me sustenta nos momentos mais difíceis e me lembra que, em Jesus, alcanço a maior vitória: a salvação.

Aos meus pais, não há palavras suficientes para expressar minha gratidão. Vocês me moldaram em quem sou hoje, me apoiaram e possibilitaram que eu chegasse até aqui. Ao meu pai, agradeço por todo o esforço ao longo da vida, buscando sempre nos prover uma vida tranquila. À minha mãe, só posso dizer que é um privilégio saber que, após um dia difícil, posso voltar para casa e encontrar alguém tão maravilhosa. Amo vocês.

Aos meus queridos avós, Nelson e Celina, sempre presentes em minha vida e ajudando de toda forma que podem, obrigado por tudo. Ao meu avô Alcides, por quem tenho grande amor, e à minha avó Glauceir, com quem infelizmente não poderei compartilhar essa conquista, mas que certamente se alegraria muito comigo, meu muito obrigado.

Aos meus tios, Dito e Soninha, pessoas queridas em quem sempre posso confiar, agradeço pelo apoio e por tantos momentos felizes em família, especialmente nos finais de ano que passamos juntos, sempre com muita alegria.

À minha prima Gabi, em quem sempre encontrei uma irmã. Sou grato por ter alguém tão especial como ela em minha vida, que sempre me motiva a continuar, seja por suas palavras ou por suas ações, que tanto me inspiram.

Ao Matheus Penha, meu melhor amigo desde os 11 anos de idade e presente em todas as etapas da minha vida desde então. Muito obrigado pela amizade e parceria; é muito valioso para mim ter com quem crescer junto e compartilhar esse processo.

Aos amigos que fiz durante essa jornada acadêmica, em especial ao Heitor Xavier e Lucas Shinji, que foram meus grandes parceiros nos estudos, apresentações e conversas despreocupadas, meu muito obrigado por tornarem os dias mais leves.

Aos meus queridos amigos da Aliança Bíblica Universitária (ABU), que entraram na minha vida no momento em que eu mais precisava, mostrando que unidos somos mais fortes, agradeço pelas amizades, estudos e orações.

Aos meus amigos da PIB Norte Rio Preto, minha igreja, sou profundamente grato. Em especial à Jéssica, à Bia e ao Ney, pessoas com quem sempre pude compartilhar pedidos de oração e verdadeiras amizades. Muito obrigado!

Por fim, agradeço imensamente ao meu professor e orientador Lucas Ribas, que me guiou e apoiou durante toda essa jornada, tornando essa tarefa desafiadora mais fácil de enfrentar.

*“As aventuras nunca acabam?  
Acho que não. Outra pessoa sempre tem de continuar a história.”  
(Bilbo, em O Senhor dos Anéis: A Sociedade do Anel - J.R.R. Tolkien)*

# Resumo

A área de análise de dados e reconhecimento de padrões tem se beneficiado significativamente do uso de redes complexas, que capturam as relações e interações entre os componentes de um sistema. Na visão computacional, redes complexas são utilizadas para representar elementos em imagens, possibilitando sua análise e classificação com base em suas propriedades estruturais. Este trabalho propõe o uso da codificação Fisher Vector para caracterizar redes complexas aplicadas à análise de contornos, abordando desafios tradicionais de reconhecimento de padrões, como robustez a ruído e variações de rotação e escala. A proposta surge da necessidade de métodos mais eficientes e robustos para essa caracterização. A aplicação dessa abordagem em diferentes bases de dados, como Aforo Otolith e Otolith, obteve acurácias de 75,56% e 96,03% utilizando o LDA, e em Portuguese Leaves e ETH-80, obteve acurácias de 79,40% e 93,63% com o SVM, demonstrando um desempenho notável em relação aos métodos tradicionais, comprovando a robustez e versatilidade da proposta. Dessa forma, a combinação de redes complexas com Fisher Vector aprimora a caracterização de contornos e oferece uma nova alternativa para tarefas de reconhecimento de padrões em visão computacional.

**Palavras-chave:** Redes Complexas, Fisher Vector, Análise de contornos, Reconhecimento de padrões.



# Abstract

*The field of data analysis and pattern recognition has greatly benefited from the use of complex networks, which capture the relationships and interactions between the components of a system. In computer vision, complex networks are used to represent elements in images, enabling their analysis and classification based on their structural properties. This work proposes the use of Fisher Vector encoding to characterize complex networks applied to shape contour analysis, addressing traditional pattern recognition challenges such as robustness to noise and variations in rotation and scale. The proposal arises from the need for more efficient and robust methods for this characterization. The application of this approach on different datasets, such as Aforo Otolith and Otolith, achieved accuracies of 75.56% and 96.03% using LDA, and on Portuguese Leaves and ETH-80, achieved accuracies of 79.40% and 93.63% with SVM, demonstrating remarkable performance compared to traditional methods, proving the robustness and versatility of the proposal. The combination of complex networks and Fisher Vector enhances shape contour characterization and offers a new alternative for pattern recognition tasks in computer vision.*

**Keywords:** *Complex Networks, Fisher Vector, Shape contour analysis, Pattern recognition.*

# Sumário

	<b>Lista de ilustrações</b>	<b>10</b>
	<b>Lista de tabelas</b>	<b>12</b>
<b>1</b>	<b>INTRODUÇÃO</b>	<b>14</b>
1.1	Considerações iniciais	14
1.2	Justificativa e Motivação	15
1.3	Objetivo	16
1.4	Organização da Monografia	16
<b>2</b>	<b>FUNDAMENTAÇÃO E REVISÃO DA LITERATURA</b>	<b>17</b>
2.1	Redes Complexas	17
2.1.1	Representação de redes complexas	17
2.1.2	Medidas de redes complexas	19
2.1.3	Modelos de redes complexas	21
2.2	Fisher Vector	22
2.2.1	Modelo de Mistura Gaussiana	23
2.2.2	Entendendo o Fisher Vector	25
2.2.3	Aplicando o Fisher Vector	26
2.3	Trabalhos relacionados	28
2.3.1	Redes Complexas	28
2.3.2	Fisher Vector	30
2.3.3	Análise de contornos	31
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>33</b>
3.1	Bases de dados de contornos	33
3.1.1	Kimia	33
3.1.2	ETH-80	34
3.1.3	Otolith	34
3.1.4	Aforo otolith	34
3.1.5	Portuguese Leaves	35
3.1.6	USP Leaves	36
3.2	Classificadores	37
3.2.1	Support Vector Machine	38
3.2.2	Linear Discriminant Analysis	38
3.3	Validação e métricas de desempenho	38

3.3.1	Validação cruzada <i>K-Fold</i> estratificada . . . . .	38
3.3.2	Matriz de confusão: . . . . .	39
3.3.3	Acurácia . . . . .	39
3.3.4	Desvio-padrão . . . . .	40
<b>3.4</b>	<b>Ferramentas . . . . .</b>	<b>40</b>
3.4.1	Python . . . . .	40
3.4.2	Bibliotecas Python . . . . .	40
<b>3.5</b>	<b>Métodos de classificação de contornos comparados . . . . .</b>	<b>41</b>
<b>4</b>	<b>ABORDAGEM PROPOSTA . . . . .</b>	<b>43</b>
<b>4.1</b>	<b>Visão Geral . . . . .</b>	<b>43</b>
<b>4.2</b>	<b>Modelagem do contorno como RC . . . . .</b>	<b>44</b>
<b>4.3</b>	<b>Evolução dinâmica da RC . . . . .</b>	<b>44</b>
4.3.1	Grau . . . . .	45
4.3.2	Força . . . . .	46
4.3.3	Coeficiente de <i>clustering</i> . . . . .	46
<b>4.4</b>	<b>Obtenção do vetor de características . . . . .</b>	<b>46</b>
<b>5</b>	<b>RESULTADOS . . . . .</b>	<b>49</b>
<b>5.1</b>	<b>Análise de parâmetros . . . . .</b>	<b>49</b>
5.1.1	Genéricas . . . . .	50
5.1.2	Otólitos . . . . .	52
5.1.3	Folhas . . . . .	52
<b>5.2</b>	<b>Discussão dos resultados . . . . .</b>	<b>55</b>
5.2.1	Genéricas . . . . .	55
5.2.2	Otólitos . . . . .	56
5.2.3	Folhas . . . . .	56
5.2.4	Invariância a escala e rotação . . . . .	57
5.2.5	Robustez a ruído . . . . .	57
5.2.6	Explicabilidade do vetor e custo computacional . . . . .	58
5.2.7	Desempenho geral . . . . .	58
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>59</b>
<b>6.1</b>	<b>Contribuição científica . . . . .</b>	<b>59</b>
<b>6.2</b>	<b>Trabalhos futuros . . . . .</b>	<b>59</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>61</b>

# Lista de ilustrações

Figura 1	– Representação de um grafo não-dirigido (a), dirigido (b) e ponderado não-dirigido (c) com $N = 6$ vértices e $E = 8$ arestas. No grafo direcionado, vértices adjacentes são conectados por setas que indicam a direção de cada conexão. No grafo ponderado, os valores $w_{i,j}$ em cada conexão representam os respectivos pesos que são graficamente representados pela espessura da linha . . . . .	18
Figura 2	– Grafo não-dirigido (a) e sua matriz de adjacência $A$ . . . . .	19
Figura 3	– Representação de uma rede pequeno mundo (a), aleatória (b) e livre de escala (c). . . . .	21
Figura 4	– Ilustração do processo de aplicação do Fisher Vector em imagens . . .	23
Figura 5	– Exemplo de uma mistura de três distribuições gaussianas . . . . .	24
Figura 6	– Exemplo das 9 classes da base de dados Kimia, sendo elas: (a) quadrúpede, (b) coelho, (c) homem, (d) ave, (e) mão, (f) jato, (g) raia, (h) ferramenta e (i) peixe. . . . .	33
Figura 7	– Quatro variações para as oito classes da base de dados ETH-80, contendo imagens de contornos de (a) maçãs, (b) carros, (c) vacas, (d) canecas, (e) cachorros, (f) cavalos, (g) peras e (h) tomates. . . . .	34
Figura 8	– Contornos de otólitos de quatorze espécies de peixes, retirados da base de dados Otolith. . . . .	35
Figura 9	– Contornos de otólitos de vinte espécies de peixes, retirados da base de dados Aforo. . . . .	35
Figura 10	– Contornos de folhas de 30 espécies diferentes, retirados da base de dados Portuguese Leaves. . . . .	36
Figura 11	– Contornos de folhas de 30 espécies diferentes, retirados da base de dados USP Leaves. . . . .	37
Figura 12	– Base USP Leaves: Exemplos de amostras modificadas. (a) rotacionadas, (b) dimensionadas e (c) com ruído. . . . .	37
Figura 13	– Ilustração da metodologia adotada na abordagem proposta. . . . .	43
Figura 14	– Evolução dinâmica da RC por um limiar $T_l$ e zoom aplicado na área: (a) $T_l = 0,1$ ; (b) $T_l = 0,15$ ; (c) $T_l = 0,2$ . . . . .	45
Figura 15	– Processo de obtenção do Fisher Vector que será o vetor de característica final para descrição da evolução da rede complexa modelando um contorno. . . . .	48

Figura 16 – Mapas de calor das acurácias (%) nas bases de dados (a) Kimia e (b) ETH-80 para diferentes valores de $K$ e $N$ , com descritores $\phi = [g, s, c]$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada. . . . .	51
Figura 17 – Mapas de calor das acurácias (%) nas bases de dados (a) Aforo Otolith e (b) Otolith para diferentes valores de $K$ e $N$ , com descritores $\phi = [g, s, c]$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada. . . . .	52
Figura 18 – Mapas de calor das acurácias (%) nas bases de dados (a) Portuguese Leaves e (b) USP Leaves para diferentes valores de $K$ e $N$ , com descritores $\phi = [g, s, c]$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada. . . . .	54

# Lista de tabelas

Tabela 1	– Recorte dos experimentos iniciais com diferentes combinações de parâmetros em todas as bases de dados, demonstrando a porcentagem de acurácia para os classificadores LDA e SVM. . . . .	50
Tabela 2	– Número de features e acurácias (%) nas bases de dados Kimia e ETH-80 para diferentes combinações $\phi$ de descritores grau, força e clustering, nos conjuntos os conjuntos $(N, K) = \{(60, 24), (70, 24), (70, 20)\}$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada (Destaque na combinação escolhida). . . . .	51
Tabela 3	– Número de features e acurácias (%) nas bases de dados Aforo Otolith e Otolith para diferentes combinações $\phi$ de descritores grau, força e clustering, nos conjuntos os conjuntos $(N, K) = \{(70, 8), (35, 24), (35, 14)\}$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada (Destaque na combinação escolhida). . . . .	53
Tabela 4	– Número de features e acurácias (%) nas bases de dados Portuguese Leaves e USP Leaves para diferentes combinações $\phi$ de descritores grau, força e clustering, nos conjuntos os conjuntos $(N, K) = \{(70, 24), (70, 20), (60, 20)\}$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada (Destaque na combinação escolhida). . . . .	54
Tabela 5	– Comparação de resultados para todas as bases de dados e métodos da literatura utilizando os classificadores LDA e SVM. . . . .	55
Tabela 6	– Comparação da acurácia da abordagem proposta com outros métodos na base de dados USP Leaves submetida a diferentes níveis de ruído. Acurácias alcançadas usando o classificador LDA. . . . .	57

# Lista de abreviaturas e siglas

RC	<i>Redes Complexas</i>
FV	<i>Fisher Vector</i>
GMM	<i>Gaussian Mixture Model</i>
BoVW	<i>Bag of Visual Words</i>
FK	<i>Fisher Kernel</i>
FIM	<i>Fisher Information Matrix</i>
SVM	<i>Support Vector Machine</i>
LDA	<i>Linear Discriminant Analysis</i>
ACC	<i>Acurácia</i>
DP	<i>Desvio-padrão</i>
VP	<i>Verdadeiro positivo</i>
VN	<i>Verdadeiro negativo</i>
FP	<i>Falso positivo</i>
FN	<i>Falso negativo</i>

# 1 Introdução

## 1.1 Considerações iniciais

Atualmente, na área de análise de dados e reconhecimento de padrões, as redes complexas se destacam como ferramentas de grande valia para a modelagem e análise de dados e sistemas complexos. Sua capacidade de capturar as relações, interações e padrões entre os componentes de um sistema as torna instrumentos eficazes em diversas áreas de estudo, como biologia (BARABASI; OLTVAI, 2004), sociologia (NEWMAN; PARK, 2003) e computação (CHEN et al., 2024; WU; JING; WANG, 2024; RIBAS, 2021), especialmente em visão computacional nos últimos tempos.

No âmbito da visão computacional, as redes complexas têm sido utilizadas para representar elementos em imagens de maneira que seja possível usar descritores da topologia para análise. O objetivo das redes complexas nesse contexto é diferenciar e classificar esses elementos de acordo com suas características inerentes. Tratam-se, portanto, de estudos que visam classificar, comparar e agrupar diferentes redes analisadas de forma holística (RIBAS, 2021). Tipos de características em imagens comumente modeladas como redes complexas são: texturas (RIBAS et al., 2020; SCABINI; GONÇALVES; CASTRO, 2015; SCABINI; RIBAS; BRUNO, 2020; RIBAS et al., 2022) e contornos de formas (MIRANDA; MACHICAO; BRUNO, 2018; RIBAS; NEIVA; BRUNO, 2019; SCABINI et al., 2017).

Na análise de redes complexas, pode-se adotar diferentes abordagens para mensurar e caracterizar a rede a fim de classificá-la. Comumente, utilizam-se, por exemplo, o grau médio, grau máximo e entropia da rede (BACKES; CASANOVA; BRUNO, 2009). Em outros estudos, o caminho geodésico entre vértices, grau hierárquico, coeficiente de *clustering*, correlação de grau, centralidade *betweenness* e evolução dinâmica da rede foram experimentados como medidas de caracterização para a rede modelada (MIRANDA; MACHICAO; BRUNO, 2018; RIBAS; NEIVA; BRUNO, 2019; RIBAS; BRUNO, 2024; SCABINI et al., 2017).

Sendo assim, percebe-se que nem sempre as medidas propostas são capazes de analisar toda a complexidade da topologia da rede. Portanto, o maior desafio nessa área tem sido encontrar métodos cada vez mais eficientes para caracterização de redes complexas, buscando explorar a totalidade de suas representações.

Nesse sentido, na última década, vários trabalhos utilizaram técnicas mais sofisticadas para agregar características locais. O Fisher Vector, uma dessas técnicas, é um método poderoso para computar representações ricas de imagens ou vídeos, fornecendo alto desempenho entre os algoritmos de extração de características (BACH et al., 2015). Ele combina abordagens generativas e discriminativas para gerar representações, codificando probabilidades e



estatísticas de ordem superior, resultando em uma representação robusta e de alto desempenho (PERRONNIN; DANCE, 2007; SÁNCHEZ et al., 2013).

Em estudos realizados, o Fisher Vector foi utilizado na codificação de descritores locais de imagens e vídeos para gerar o vetor de características de entrada ao classificador (CHANDRASEKHAR et al., 2016; GÓMEZ-SIRVENT et al., 2022; ASLAN et al., 2015). Os resultados dessas pesquisas demonstraram alta precisão de representação, robustez e menor custo computacional.

A aplicação escolhida neste trabalho foi a de análise de contornos de formas. Contornos em visão computacional são elementos importantes que podem ser usados para caracterização e classificação de objetos em diferentes tarefas e aplicações (MIRANDA; MACHICAO; BRUNO, 2018). Existem diversas aplicações de reconhecimento de padrões para análise de contornos, como na agricultura (PLOTZE et al., 2005), neurociência (GERIG et al., 2001) e em imagens médicas (TSAI et al., 2003), por exemplo. Na literatura, existem diversos trabalhos que exploram diferentes técnicas para extrair informações de contornos, como métodos estocásticos, sintáticos, aproximação e decomposição de borda, transformações elásticas e descritores de Fourier (LONCARIC, 1998; PURNOMO et al., 2024; YU et al., 2024; BROPHY et al., 2024). Técnicas baseadas em redes complexas se destacam devido a vantagens intrínsecas da modelagem e flexibilidade, como invariância a rotação, robustez a ruído, degradação e escala, por exemplo (BACKES; CASANOVA; BRUNO, 2009; RIBAS; BRUNO, 2024).

Portanto, a proposta deste trabalho foi empregar a codificação Fisher Vector para caracterização de redes complexas, com aplicação em análise de contornos, buscando capturar as relações e interações entre os vértices do contorno de forma mais abrangente e precisa, possibilitando um melhor desempenho na caracterização de redes complexas e, consequentemente, em tarefas de classificação de formas de contornos.

## 1.2 Justificativa e Motivação

Tendo em vista o desafio da caracterização de redes complexas, este trabalho explorou uma forma de caracterização mais robusta, que capturasse as características intrínsecas da topologia da rede, potencializando aplicações, como a análise de contornos. Diante desse desafio, propôs-se explorar o Fisher Vector como uma nova abordagem, dado que ele apresentou bons resultados em aplicações de visão computacional.

Estudos prévios que utilizam modelagem em redes complexas já indicaram a possibilidade de inovação e otimização nos métodos de caracterização da rede (RIBAS et al., 2022; MIRANDA; MACHICAO; BRUNO, 2018; RIBAS et al., 2020). Com a introdução do Fisher Vector, foi possível obter uma descrição mais detalhada, robusta e completa das propriedades e padrões presentes nas redes, resultando em um melhor desempenho na sua classificação.

Assim, a combinação de redes complexas com a codificação do Fisher Vector apresentou-se como uma estratégia promissora para avançar no estudo da caracterização de redes complexas e na aplicação em análise de contornos, que se beneficiou das vantagens de ambas as técnicas utilizadas em conjunto na abordagem proposta.

## 1.3 Objetivo

Este trabalho teve como objetivo investigar, desenvolver e testar um método aprimorado de caracterização de redes complexas usando Fisher Vector. A abordagem foi estudada para o problema de análise de contornos. Dessa forma, buscou-se alcançar alta precisão na classificação dos contornos ao utilizar o vetor de características extraído no processo.

Como objetivo final, a abordagem proposta agregou aos estudos de análise e caracterização de redes complexas e às aplicações do Fisher Vector. Além disso, contribuiu cientificamente para a pesquisa em análise de contornos, com foco específico na extração de características.

## 1.4 Organização da Monografia

A monografia está dividida em seis capítulos. O Capítulo 2 apresenta a fundamentação teórica necessária para compreender Redes Complexas e Fisher Vector, que são os principais conceitos utilizados no estudo e nos experimentos. Além disso, neste capítulo, também são abordados trabalhos correlatos, ou seja, pesquisas relacionadas ao reconhecimento de padrões com redes complexas, à caracterização de dados por meio da codificação do Fisher Vector e à análise de contornos. No Capítulo 3, são apresentados os materiais e métodos utilizados na configuração experimental deste trabalho. A descrição das etapas da metodologia proposta é detalhada no Capítulo 4. No Capítulo 5, os resultados obtidos nos experimentos são apresentados e analisados. Por fim, no Capítulo 6, são feitas as conclusões gerais a respeito desta monografia, bem como suas contribuições científicas e possíveis aperfeiçoamentos para pesquisas futuras.

## 2 Fundamentação e Revisão da Literatura

Neste capítulo, a teoria dos conceitos que envolvem a abordagem proposta neste trabalho é fundamentada, fornecendo uma visão geral das bases que compõem o método. Além disso, são apresentados trabalhos relacionados, a fim de documentar o atual estado da arte.

### 2.1 Redes Complexas

Uma rede complexa (RC) é uma maneira de representar e observar uma parte do mundo de maneira sistêmica. Quando se tem um ponto de vista isolado daquilo que se observa, é possível focar nas interações que ocorrem entre elementos próximos. (GARRIDO, 2011)

RCs possuem uma natureza verdadeiramente interdisciplinar (COSTA et al., 2007): Seguindo os preceitos da teoria dos grafos e indo além, ao utilizar da mecânica estatística para denotar suas conexões, construir modelos e analisar sua dinâmica e topologia. RCs, ou ciência de redes (em inglês, *network science*) (BARABÁSI; PÓSFAL, 2016), são populares pois com elas se torna possível representar, de maneira flexível e generalizada, qualquer estrutura natural, mesmo aquelas que mudam sua topologia ao longo do tempo (COSTA et al., 2007).

A análise de uma RC atende à diversas finalidades (RIBAS et al., 2022; WU; JING; WANG, 2024; RIBAS et al., 2020; RIBAS; NEIVA; BRUNO, 2019). Uma delas é classificar diferentes tipos de estruturas, como contornos, no caso deste trabalho. Isso ocorre porque RCs podem auxiliar no reconhecimento de padrões em grandes quantidades de dados, o que abre portas para novas descobertas.

Dessa forma, a ciência de redes muda a forma como estudamos sistemas complexos. Ao invés de focar em partes isoladas, é possível analisar as conexões e como elas geram propriedades novas. Essa visão holística é poderosa para criar novas soluções para problemas difíceis.

#### 2.1.1 Representação de redes complexas

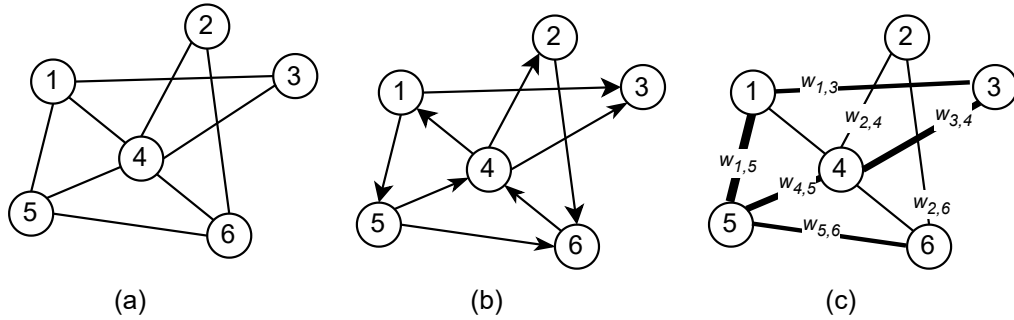
Formalmente, uma Rede Complexa (RC) é uma estrutura matemática que pode ser representada por um grafo  $G = (V, E)$ , onde:

- $V$  é o conjunto de vértices  $v_1, \dots, v_N$ .
- $E$  é o conjunto de arestas.

Cada aresta  $e_{v_i, v_j}$  indica uma conexão entre dois vértices  $v_i$  e  $v_j$ , e, portanto, o conjunto  $E = e_{v_i, v_j}, \dots$  consiste em todas as arestas que conectam os vértices em  $V$ .

Assim como a Figura 1 demonstra, é possível que a rede seja dirigida, o que significa que as arestas  $e_{v_i, v_j}$  têm uma direção de  $v_i$  para  $v_j$ , representando assim a transferência de informações ou influência de um vértice para outro no sistema modelado. Além disso, as arestas podem ter pesos  $w(e_{v_i, v_j})$  associados, representando características entre os vértices do sistema modelado, o que as torna redes ponderadas.

Figura 1 – Representação de um grafo não-dirigido (a), dirigido (b) e ponderado não-dirigido (c) com  $N = 6$  vértices e  $E = 8$  arestas. No grafo direcionado, vértices adjacentes são conectados por setas que indicam a direção de cada conexão. No grafo ponderado, os valores  $w_{i,j}$  em cada conexão representam os respectivos pesos que são graficamente representados pela espessura da linha



Fonte: Elaborado pelo autor

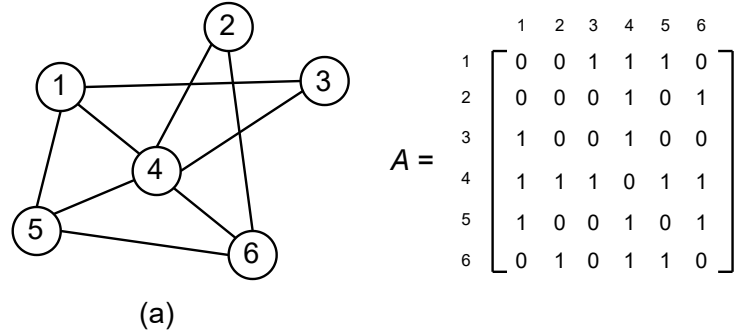
Uma RC pode ser formada por diversos componentes desconectados, que são conhecidos como subgrafos. Em certos contextos, o componente com o maior número de vértices interconectados é considerado o componente principal.

Do ponto de vista computacional, as RCs podem ser representados e armazenados de diferentes formas, incluindo:

- **Listas de adjacência:** cada vértice  $v_i$  da rede é associado a uma lista  $L_i$ , onde  $L_i$  contém os vértices aos quais  $v_i$  está conectado. Cada elemento da lista  $L_i$  representa uma aresta e pode conter informações adicionais, como o peso da aresta, em redes ponderadas. Por exemplo,  $L_i = [(v_j, w_{ij}), (v_k, w_{ik})]$  indica que  $v_i$  está conectado a  $v_j$  e  $v_k$  com os respectivos pesos associados.
- **Listas de arestas:** nessa representação, cada aresta da rede é representada como um par de vértices  $(i, j)$ , origem e destino, podendo ser acompanhado pelo peso da aresta em redes ponderadas. Essas informações são armazenadas em uma lista. Por exemplo,  $L = [(v_i, v_j, w_{ij}), (v_j, v_k, w_{jk}), \dots]$  indica as arestas de uma rede, em que a primeira tupla representa uma aresta entre os vértices  $v_i$  e  $v_j$ , com o peso associado  $w_{ij}$ .

- **Matrizes de adjacência:** em uma matriz de adjacência  $A$ , os vértices  $v_i$  e  $v_j$  estão conectados se a entrada  $A(i, j)$  contiver o valor 1 (ou o valor do peso da aresta), caso contrário,  $A(i, j) = 0$ , assim como demonstrado na Figura 2.

Figura 2 – Grafo não-dirigido (a) e sua matriz de adjacência  $A$ .



Fonte: Elaborado pelo autor

Em muitos estudos relacionados a RCs, o primeiro passo é definir como modelar o problema em questão como uma rede, o que implica na definição dos conjuntos de vértices e arestas (RIBAS, 2021). Uma vez que o grafo  $G$  tenha sido construído de forma adequada, uma variedade de medidas pode ser calculada para caracterizar sua estrutura.

### 2.1.2 Medidas de redes complexas

Múltiplas medidas podem ser empregadas para estudar e analisar a topologia e estrutura das RCs. Com essas medidas, é possível caracterizar, comparar e classificar as redes que modelam um determinado sistema ou objeto que está sendo o alvo de estudo (RIBAS, 2021). A seguir, são apresentadas algumas das medidas clássicas de RCs:

- **Grau e grau médio:** Com o grau de um vértice é possível estabelecer diversas relações, possibilitando derivar mais medidas e análises. O grau  $g_{v_i}$  de um vértice  $v_i$  é o número das arestas  $e_{v_i, v_j}$  que incidem em outros vértices:

$$g_{v_i} = \sum_{j=1}^N A(i, j). \quad (2.1)$$

O grau médio, por sua vez, é obtido ao calcular a média aritmética entre os graus de todos os vértices:

$$\langle g \rangle = \frac{1}{N} \sum_{i=1}^N g_{v_i}. \quad (2.2)$$

Ao trabalhar com redes dirigidas, mais medidas podem ser calculadas:

- **Grau de saída** ( $\bar{g}_{v_i}$ ): soma de arestas de saída do vértice  $v_i$ .
- **Grau de entrada** ( $\hat{g}_{v_i}$ ): soma de arestas de entrada no vértice  $v_i$ .
- **Força**: em redes ponderadas, a força  $s_{v_i}$  é dada pela soma dos pesos das arestas que incidem ao vértice  $v_i$ :

$$s_{v_i} = \sum_{j=1}^N w_{ij}. \quad (2.3)$$

Além disso, com redes dirigidas e ponderadas, pode-se calcular:

- **Força de saída** ( $\bar{s}_{v_i}$ ): soma dos pesos das arestas de saída de  $v_i$ .
- **Força de entrada** ( $\hat{s}_{v_i}$ ): soma dos pesos das arestas de entrada em  $v_i$ .
- **Distribuição do grau**: essa função descreve a probabilidade  $P(g)$  de um vértice específico ter um determinado número de conexões. Em outras palavras, ela indica a chance de um vértice  $v_i$  ter grau  $g$ .

A partir da distribuição do grau, pode-se medir a heterogeneidade da rede utilizando medidas derivadas, tais como:

- **Entropia**: em uma RC, a entropia pode ser usada para medir a incerteza ou a complexidade da distribuição dos vértices e arestas: Quanto mais uniformemente distribuídos os graus dos vértices, maior é a entropia.

$$g_h = \sum_i P(g_i) \log P(g_i). \quad (2.4)$$

- **Energia**: essa medida aumenta à medida que a distribuição dos graus dos vértices se torna mais concentrada em alguns valores específicos. Isso pode indicar uma estrutura mais organizada ou coesa na rede.

$$g_e = \sum_i P(g_i)^2. \quad (2.5)$$

- **Contraste**: indica a variação dos graus dos vértices na rede. Um contraste maior indica uma maior variação na distribuição dos graus dos vértices, enquanto um contraste menor indica uma distribuição mais homogênea.

$$g_c = \sum_i P(g_i) g_i^2. \quad (2.6)$$

Além das medidas clássicas, pode-se calcular e quantificar outras propriedades de uma RC. Uma medida importante é o coeficiente de aglomeração, que é uma métrica de transitividade.

O coeficiente de aglomeração mede o número de agrupamentos em uma RC, onde dois vértices com uma conexão em comum são prováveis de serem conectados. De forma genérica, essa medida de transitividade representa a presença de triângulos na rede (WATTS; STROGATZ, 1998). O coeficiente de aglomeração local  $c_i$  é calculado para cada vértice  $v_i$  da rede:

$$c_i = \frac{2e_i}{g_i(g_i - 1)}, \quad (2.7)$$

onde  $e_i$  representa as conexões entre vizinhos de  $v_i$  e  $g_i$  é o grau de  $v_i$ . Ao fim, obtém-se a fração de possíveis conexões entre os vizinhos de  $v_i$ .

Desse modo, o coeficiente de aglomeração da rede é dado pela média dos coeficientes locais  $c_i$  em todos os vértices  $N$  da RC:

$$C = \langle c \rangle = \frac{1}{N} \sum_{i \in N} c_i \quad (2.8)$$

Essa métrica é útil para medir padrões de texturas e contornos em imagens, uma vez que esses padrões se baseiam na interação entre um pixel e seus pixels vizinhos.

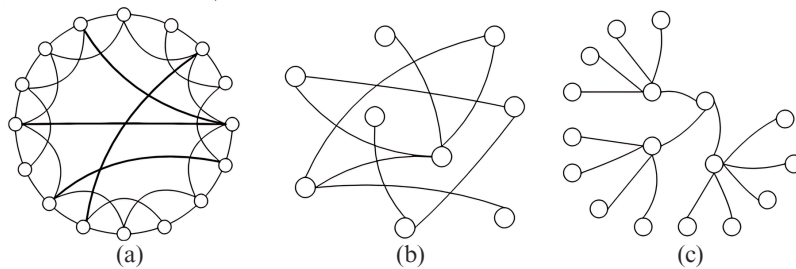
Finalmente, múltiplas outras medidas podem ser utilizadas na caracterização de uma RC, como medidas de conectividade, centralidade, detecção de comunidades e motifs, completamente descritas em Costa et al. (2007).

### 2.1.3 Modelos de redes complexas

RCs podem ser classificadas em diferentes modelos com base nas características estruturais e propriedades estatísticas quantificadas. Essa prática envolve rotular os modelos de acordo com os padrões identificados na rede.

Os principais modelos de RCs são: redes pequeno mundo, redes aleatórias e redes livre de escala, representadas na Figura 3.

Figura 3 – Representação de uma rede pequeno mundo (a), aleatória (b) e livre de escala (c).



Fonte: Adaptado de Huang, Sun e Lin (2005)

- **Redes pequeno mundo:** inicia-se com um grafo regular de  $n$  dimensões, onde cada vértice está conectado a  $z$  vizinhos. Em seguida, algumas arestas são aleatoriamente reconectadas com uma probabilidade  $p$ . Em uma variação do modelo original, são inseridas conexões de longo alcance, também conhecidas como "atalhos", entre pares de vértices escolhidos aleatoriamente. Essa variação resulta em uma rede que apresenta propriedades de pequeno mundo e de agrupamento, semelhantes às encontradas em redes sociais humanas (WATTS; STROGATZ, 1998).
- **Redes aleatórias:** uma rede desse modelo é criada adicionando-se conexões entre pares de vértices escolhidos aleatoriamente, seguindo uma distribuição de probabilidade de Poisson para a distribuição do grau dos vértices (ERDOS; RENYI, 1960). Embora capaz de exibir propriedades de pequeno mundo com a adição suficiente de conexões, essas redes geralmente não exibem agrupamento significativo, uma característica incomum em redes do mundo real (WANG; CHEN, 2003).
- **Redes livre de escala:** esse tipo de rede começa com um pequeno número de vértices  $m_0$ , aos quais novos vértices são adicionados e conectados a  $m$  vértices existentes de acordo com a probabilidade baseada no grau de cada vértice. Novos vértices são preferencialmente ligados a vértices existentes com graus altos, resultando em uma distribuição de graus em lei de potência, com alguns vértices altamente conectados, semelhante à estrutura de hiperlinks da World Wide Web (WANG; CHEN, 2003) (BARABÁSI; ALBERT, 1999).

## 2.2 Fisher Vector

O Fisher Vector (FV) é um método de caracterização usado principalmente para classificação de imagens. O processo de aplicação do FV se inicia com a extração de descritores locais da imagem para depois serem ajustados a um modelo probabilístico generativo, tipicamente um Modelo de Mistura de Gaussiana (GMM), resultando em um vocabulário visual probabilístico (PERRONNIN; DANCE, 2007). Em seguida, calcula-se o gradiente do logaritmo da probabilidade dos descritores em relação aos parâmetros do GMM. Esses gradientes são, então, concatenados e formam o FV, que representa a imagem para um classificador (SÁNCHEZ et al., 2013).

Esse método é uma alternativa à técnica de codificação *Bag of Visual Words* (BoVW) para classificação e recuperação de imagens. O BoVW, derivado do *Bag of Words* (BoW), é uma técnica que representa imagens como "sacos" de palavras visuais, desconsiderando a ordem ou a disposição espacial. O vocabulário é construído agrupando o conjunto de descritores locais da imagem pelo algoritmo *k-means*, tipicamente. Cada grupo formado representa uma palavra visual do vocabulário. A representação final de uma imagem é obtida contando quantas

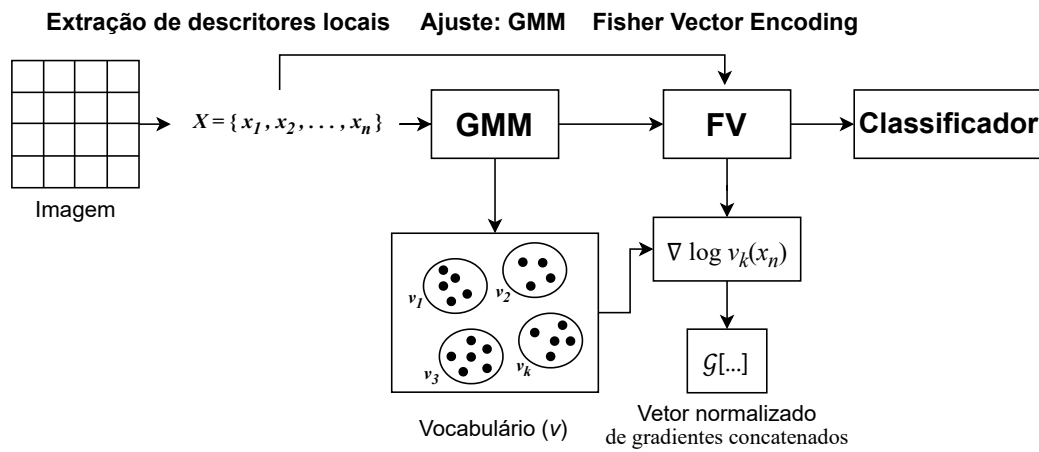


vezes cada palavra visual do vocabulário aparece na imagem, resultando em um vetor de características que representa a frequência das palavras visuais na imagem.

O FV se torna vantajoso pois, utilizando vocabulários compactos, pode gerar representações informativas de alta dimensionalidade sem abrir mão da generalização. Ele codifica não apenas a contagem probabilística de ocorrências, mas também estatísticas de ordem superior, resultando em uma representação mais eficiente, exigindo vocabulários menores para alcançar alto desempenho (PERRONNIN; DANCE, 2007; SÁNCHEZ et al., 2013).

Em seguida, será explicado mais detalhadamente o processo de aplicação do FV, ilustrado na Figura 4.

Figura 4 – Ilustração do processo de aplicação do Fisher Vector em imagens



Fonte: Elaborado pelo autor

### 2.2.1 Modelo de Mistura Gaussiana

Representado na Figura 5, um GMM assume que os dados observados são gerados por uma combinação linear de várias distribuições gaussianas (também denominadas distribuições normais), ou seja uma mistura de gaussianas (DEMPSTER; LAIRD; RUBIN, 1977).

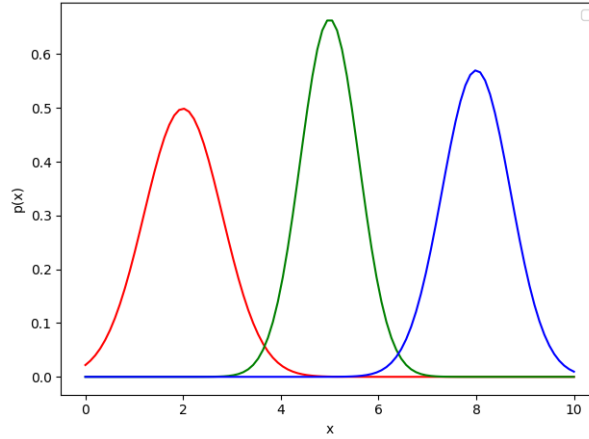
Formalmente, para um conjunto de dados  $X = \{x_1, x_2, \dots, x_n\}$ , onde  $x_n$  é um vetor de observação de  $d$  dimensões, a função de densidade de probabilidade de misturas de gaussianas de  $X$  é modelada como:

$$p(X|\Theta_k) = \sum_{k=1}^K \pi_k \phi(X|\theta_k), \quad (2.9)$$

onde:

- O vetor  $\Theta = (\pi_1, \dots, \pi_k, \theta_1, \dots, \theta_k)$  representa os parâmetros do modelo, para  $K$  componentes, isto é, o número de distribuições (ou *clusters*).

Figura 5 – Exemplo de uma mistura de três distribuições gaussianas



Fonte: Elaborado pelo autor

- O peso,  $\pi_k$ , é a probabilidade de escolha do componente  $k$ , tal que  $\sum_{k=1}^K \pi_k = 1$ .
- $\phi(X|\theta_k) = \phi(X|\mu_k, \Sigma_k)$  define a função de densidade de probabilidade de uma distribuição gaussiana, com média  $\mu_k$  e matriz de covariância  $\Sigma_k$ , definida por:

$$\phi(x_n|\mu_k, \Sigma_k) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)} \quad (2.10)$$

No estágio inicial do treinamento do modelo, é necessário decidir o número de componentes,  $K$ . Para realizar essa tarefa, pode-se adotar várias abordagens, incluindo: conhecimento do domínio e critérios estatísticos. Conhecimento do domínio baseia-se em utilizar conhecimento prévio dos dados ou do problema para determinar um número adequado. Critérios estatísticos, como o BIC (*Bayesian Information Criterion*) e o AIC (*Akaike Information Criterion*), são usados para comparar modelos, penalizando modelos mais complexos, e, ao final, auxiliando na tarefa de escolher o número adequado de componentes (SCHWARZ, 1978; AKAIKE, 1974).

Os parâmetros do GMM (médias, covariâncias e pesos) geralmente são inicializados aleatoriamente ou por métodos heurísticos, isto é investigações baseadas em aproximação progressiva.

Depois da inicialização dos parâmetros, se iniciam duas etapas iterativas, que compõem o algoritmo de Expectativa-Maximização (EM) (DEMPSTER; LAIRD; RUBIN, 1977), usado para estimar e atualizar os parâmetros do GMM:

- **Expectativa (E-step):** para cada ponto  $x_n$ , calcula-se a probabilidade  $z_{nk}$  do ponto pertencer ao componente  $k$ :

$$\gamma(z_{nk}) = \frac{\pi_k \phi(x_n | \theta_k)}{\sum_{j=1}^K \pi_j \phi(x_n | \theta_j)}. \quad (2.11)$$

O E-step calcula as probabilidades usando as estimativas atuais dos parâmetros do GMM. As probabilidades são comumente chamadas de "responsabilidades" (DEISENROTH; FAISAL; ONG, 2020) das distribuições gaussianas. Elas são representadas pelas variáveis  $z_{nk}$ , onde  $n$  é o índice do ponto de dados, e  $k$  é o índice da distribuição gaussiana. A responsabilidade mede o quão responsável pelo  $n$ -ésimo ponto de dados é a  $k$ -ésima distribuição gaussiana. O teorema de Bayes (1763) é utilizado nessa etapa, já que ele trata de probabilidades condicionais.

- **Maximização (M-step):** nessa etapa, as responsabilidades das distribuições gaussianas, obtidas no E-Step, são utilizadas para atualizar as estimativas dos parâmetros do GMM:

$$\pi_k' = \frac{1}{N} \sum_{n=1}^N \gamma(z_{nk}) \quad (2.12)$$

$$\mu_k' = \frac{\sum_{n=1}^N \gamma(z_{nk}) x_n}{\sum_{n=1}^N \gamma(z_{nk})} \quad (2.13)$$

$$\Sigma_k' = \frac{\sum_{n=1}^N \gamma(z_{nk}) (x_n - \mu_k') (x_n - \mu_k')^T}{\sum_{n=1}^N \gamma(z_{nk})} \quad (2.14)$$

O processo de EM é repetido até que haja convergência, isto é, quando a mudança nos parâmetros se torna pequena o suficiente ou até que a log-verossimilhança dos dados, ou seja, o ajuste do modelo, atinja um máximo.

Dessa forma, com os parâmetros, obtém-se o vocabulário visual probabilístico que será usado no processo de codificação do FV.

## 2.2.2 Entendendo o Fisher Vector

O FV é derivado do Fisher Kernel (FK), que mede a similaridade entre amostras  $X$  e  $Y$  usando o gradiente da verossimilhança dos dados em relação aos parâmetros (vocabulário) do modelo generativo (JAAKKOLA; HAUSSLER, 1998). É definido como:

$$K_{FK}(X, Y) = G_{\lambda}^{X'} F_{\lambda}^{-1} G_{\lambda}^Y, \quad (2.15)$$

onde  $G$  é a função de pontuação, ou *score function*, e  $F$  é a matriz de informação de Fisher, ou *Fisher Information Matrix* (FIM).

A *score function*  $G_\lambda^X$  é um vetor gradiente que descreve como os parâmetros  $\lambda$  do modelo  $u_\lambda$  devem ser ajustados para melhor se adequar aos dados  $X$ . Matematicamente, a *score function* é representada como:

$$G_\lambda^X = \nabla_\lambda \log u_\lambda(X) \quad (2.16)$$

A FIM  $F_\lambda$  é uma matriz quadrada de dimensão  $M \times M$ , onde  $M$  é o número de parâmetros no vetor  $\lambda$ . É definida como uma maneira de medir a quantidade de informação que uma variável  $X$  carrega sobre um parâmetro  $\lambda$  de uma distribuição que modela  $X$  (AMARI; NAGAOKA, 2000). Nesse caso, ela é a variância do *score*, ou seja, o valor esperado da informação observada:

$$F_\lambda = \mathbb{E}_{x \sim u_\lambda} [G_\lambda^X G_\lambda^{X'}] \quad (2.17)$$

Usando a decomposição de Cholesky (LÉVY, 1924) da inversa da FIM,  $F_\lambda^{-1} = L_\lambda' L_\lambda$ , o FK pode ser reescrito explicitamente como:

$$K_{FK}(X, Y) = \mathcal{G}_\lambda^{X'} \mathcal{G}_\lambda^Y, \quad (2.18)$$

onde

$$\mathcal{G}_\lambda^X = L_\lambda G_\lambda^X = L_\lambda \nabla_\lambda \log u_\lambda(X). \quad (2.19)$$

Esse gradiente normalizado, calculado multiplicando o gradiente pelo fator de normalização  $L_\lambda$ , é chamado Fisher Vector de  $X$ , possuindo a mesma dimensionalidade que o vetor gradiente original  $G_\lambda(X)$ . A diferença entre os dois se encontra na linearidade: diferentemente do gradiente original, o FV é linear, sendo vantajoso para a sua aplicação em classificadores lineares (SÁNCHEZ et al., 2013).

### 2.2.3 Aplicando o Fisher Vector

Primeiramente, extrai-se um conjunto de descritores locais da imagem, e, então, ajusta-se um GMM K-componente com parâmetros  $\lambda = \{\pi_k, \mu_k, \Sigma_k, k = 1, \dots, K\}$ :

$$u_\lambda(x) = \sum_{k=1}^K \pi_k u_k(x), \quad (2.20)$$

Como descrito por Sánchez et al. (2013), no processo, assume-se que as matrizes de covariância são diagonais, o que simplifica a modelagem e os cálculos. Nesse caso, a matriz de covariância  $\Sigma_k$  para o  $k$ -ésimo componente é representada por um vetor de variâncias  $\sigma_k^2$ , que contém as variâncias de cada dimensão da distribuição gaussiana.

Além disso, para garantir que os pesos  $\pi_k$  dos componentes gaussianos formem uma distribuição válida, isto é, somem a 1 e sejam não-negativos, adota-se o formalismo de soft-

max, conforme descrito por Krapac, Verbeek e Jurie (2011). A fórmula para os pesos é dada por:

$$\pi_k = \frac{\exp(\alpha_k)}{\sum_{j=1}^K \exp(\alpha_j)} \quad (2.21)$$

A re-parametrização usando  $\alpha_k$ , da função soft-max, garante que  $\sum_{k=1}^K \pi_k = 1$ .

Após essas alterações, e com o modelo ajustado, calcula-se o gradiente da verossimilhança dos dados em relação aos parâmetros  $\lambda = \{\alpha_k, \mu_k, \sigma_k\}$  do GMM. Os gradientes para os parâmetros de média, Eq. (2.22), desvio-padrão, Eq. (2.23), e peso, Eq. (2.24), da mistura são os seguintes:

$$\nabla_{\mu_k} \log u_\lambda(x_n) = \gamma(z_{nk}) \frac{(x_n - \mu_k)}{\sigma_k^2} \quad (2.22)$$

$$\nabla_{\sigma_k} \log u_\lambda(x_n) = \gamma(z_{nk}) \left[ \frac{(x_n - \mu_k)^2}{\sigma_k^3} - \frac{1}{\sigma_k} \right] \quad (2.23)$$

$$\nabla_{\alpha_k} \log u_\lambda(x_n) = \gamma(z_{nk}) - \pi_k \quad (2.24)$$

O FV de um conjunto de descritores  $X = \{x_n, n = 1, \dots, N\}$  é formado concatenando os gradientes em relação a todos os parâmetros em um vetor de dimensão  $(1 + 2D)K$ :

$$\mathcal{G}_\lambda^X = (\mathcal{G}_{\alpha_1}^X, \dots, \mathcal{G}_{\alpha_K}^X, \mathcal{G}_{\mu_1}^{X'}, \dots, \mathcal{G}_{\mu_K}^{X'}, \mathcal{G}_{\sigma_1}^{X'}, \dots, \mathcal{G}_{\sigma_K}^{X'})' \quad (2.25)$$

onde  $\mathcal{G}_{\mu_k}^X$ ,  $\mathcal{G}_{\sigma_k}^X$  e  $\mathcal{G}_{\alpha_k}^X$  são definidos por:

$$\mathcal{G}_{\mu_k}^X = \frac{1}{\sqrt{\pi_k}} \sum_{n=1}^N \gamma(z_{nk}) \frac{(x_t - \mu_k)}{\sigma_k}, \quad (2.26)$$

$$\mathcal{G}_{\sigma_k}^X = \frac{1}{\sqrt{\pi_k}} \sum_{n=1}^N \gamma(z_{nk}) \left[ \frac{(x_t - \mu_k)^2}{\sigma_k^2} - 1 \right], \quad (2.27)$$

$$\mathcal{G}_{\alpha_k}^X = \frac{1}{\sqrt{\pi_k}} \sum_{n=1}^N (\gamma(z_{nk}) - \pi_k), \quad (2.28)$$

em que  $\frac{1}{\sqrt{\pi_k}}$  é o fator  $L_\lambda$  de normalização para a FIM  $F_\lambda$  diagonal (SÁNCHEZ et al., 2013).

Para melhorar o desempenho, aplicam-se técnicas de normalização como a normalização de potência, Eq. (2.29), e a normalização  $l_2$ , Eq. (2.30), ao FV:

$$\mathcal{G}_\lambda^X = \text{sign}([\mathcal{G}_\lambda^X]_i) \sqrt{|[\mathcal{G}_\lambda^X]_i|}, \text{ para } i = 1, \dots, (1 + 2D)K. \quad (2.29)$$

$$\mathcal{G}_\lambda^X = \frac{\mathcal{G}_\lambda^X}{\sqrt{\mathcal{G}_\lambda^{X'} \mathcal{G}_\lambda^X}} \quad (2.30)$$

Finalmente, o FV normalizado é usado como vetor de características da imagem para classificação em um classificador linear, como uma *Support Vector Machine* (SVM).

## 2.3 Trabalhos relacionados

Nesta seção são apresentados alguns trabalhos que empregam os conceitos apresentados e utilizados neste trabalho.

### 2.3.1 Redes Complexas

Costa et al. (2007) possui um trabalho que descreve as medidas de RCs, além de, em outro trabalho, examinar o impacto e as aplicações das RCs em problemas do mundo real, destacando sua contribuição tanto para o desenvolvimento teórico quanto para a resolução de problemas práticos (COSTA et al., 2011). Após uma introdução aos conceitos e modelos principais, são analisadas as representações, análises e modelagens de uma variedade de fenômenos, abrangendo até 22 áreas distintas. Isso demonstra claramente a amplitude do campo das RCs e sua influência em diversas disciplinas.

Em um estudo focado na mecânica estatística da topologia e dinâmica de RCs, Albert e Barabási (2002) abordam a transição de modelagem desses sistemas como grafos aleatórios para o reconhecimento de princípios organizacionais robustos que regem a topologia e a evolução das redes reais. Além deste estudo, Barabási possui uma alta gama de trabalhos voltados para RCs (BARABÁSI; ALBERT, 1999; NEWMAN; BARABÁSI; WATTS, 2006; BARABÁSI; PÓSFAL, 2016).

O estudo de Ribas et al. (2022) apresenta uma nova abordagem para a detecção precoce da osteoartrite do joelho através da análise de imagens de raios-X. Cada pixel da imagem é modelado como um vértice e conexões são estabelecidas com base em distâncias euclidianas. Um conjunto de limiares é aplicado para revelar propriedades de textura específicas. A partir da função de densidade de probabilidade do grau do vértice, a média, variância e outras medidas estatísticas extraídas da rede resultante formam um vetor de características, que é então utilizado em experimentos de classificação. O método foi comparado a modelos de aprendizado de última geração (como AlexNet, VGG e ResNet) e descritores de textura tradicionais, mostrando-se competitivo e promissor para a detecção precoce da osteoartrite do joelho. Os resultados indicam que a abordagem baseada em RC não só consegue modelar

as características principais das imagens de raios-X, como também melhora a separação entre grupos controle e de pacientes com osteoartrite, atingindo taxas de acurácia superiores em comparação com métodos existentes.

Já em outra aplicação, a pesquisa apresentada por Ribas et al. (2020) introduz um método de análise de texturas, combinando RCs e redes neurais randomizadas (RNNs). No método proposto, a imagem de entrada é modelada como uma RC, cujas propriedades topológicas, juntamente com os pixels da imagem, são utilizadas para treinar redes neurais randomizadas. Essa abordagem visa criar uma assinatura que represente as características profundas da textura. Os resultados experimentais demonstraram uma precisão superior em comparação com vários métodos existentes na literatura, indicando que a combinação das RCs com as redes neurais pode ser uma abordagem promissora na análise de texturas.

Em outra perspectiva, Chen et al. (2024) apresenta uma abordagem para identificar vértices influentes em RCs utilizando o modelo Transformer (uma arquitetura de rede neural conhecida por sua abordagem de atenção, permitindo a captura de relações complexas entre elementos em uma sequência de dados, sem depender de estruturas sequenciais convencionais). Este método aborda as limitações dos algoritmos de aprendizado profundo atuais, que muitas vezes têm dificuldades em agregar informações de vizinhos em RCs. Incorporando informações específicas dos vértices e dos vizinhos na sequência de entrada, o modelo proposto identifica eficazmente os vértices influentes. A metodologia foi testada em 9 redes sintéticas e 12 redes reais, demonstrando melhorias significativas na precisão em comparação com métodos existentes, especialmente nas redes Netscience e Faa. Este avanço possui potenciais aplicações em diversas áreas, incluindo supressão de rumores, detecção de comunidades e detecção de surtos de doenças.

De outra maneira, demonstrando a interdisciplinaridade das aplicações de RCs, o trabalho de Wu, Jing e Wang (2024) aborda a questão da segurança da Internet das Coisas (IoT) a partir da perspectiva da observabilidade em RCs, explorando a capacidade de reconstruir os dados completos da rede a partir de informações parciais dos vértices. Os resultados do estudo sugerem que essa abordagem pode representar um primeiro passo promissor para lidar com questões de segurança na IoT, embora reconheçam a simplificação da estrutura da rede IoT, que pode não capturar todas as suas complexidades reais.

Já o estudo de Zhang et al. (2024) se foca em uma nova métrica para identificar a centralidade de intermediação em redes complexas, baseada na topologia da rede e na dinâmica dos fluxos de informação. Esta métrica, que combina aspectos estruturais e dinâmicos, foi testada em diversas redes, demonstrando ser mais eficaz do que as métricas tradicionais em identificar vértices chave para a conectividade e robustez da rede. Os resultados sugerem que a nova métrica pode oferecer uma compreensão mais profunda da influência dos vértices na propagação da informação e na estabilidade das redes complexas, com potencial para aplicações em áreas como redes sociais, biologia de sistemas e infraestruturas críticas.

### 2.3.2 Fisher Vector

Os estudos de Sánchez et al. (2013) sobre o FV apresentam contribuições significativas para a área de processamento e análise de imagens. Ao estender o método de BoVW, o FK combina abordagens generativas e discriminativas, oferecendo melhorias na precisão da classificação de imagens. Além disso, ao aplicar FV na recuperação de imagens em larga escala, os estudos demonstram que os FVs comprimidos são altamente eficazes, superando abordagens anteriores baseadas em BoVW comprimidos. Por fim, ao propor o FV como uma estratégia alternativa para a codificação de patches em imagens, os estudos mostram que essa abordagem oferece eficiência computacional, excelente desempenho mesmo com classificadores lineares simples e capacidade de compressão sem perda significativa de precisão (PERRONNIN et al., 2010; PERRONNIN; SÁNCHEZ; MENSINK, 2010).

Partindo para uma abordagem mista, o estudo de Chandrasekhar et al. (2016) propõe uma avaliação abrangente e sistemática dos descritores globais de imagem, FVs e redes neurais convolucionais (CNNs), para recuperação de instâncias de imagem. A pesquisa compara o desempenho desses descritores em múltiplos conjuntos de dados públicos e critérios diversos, demonstrando que nenhum descritor é consistentemente superior ao outro e que ganhos de desempenho geralmente são obtidos ao utilizar ambos em conjunto.

Já o trabalho de Gómez-Sirvent et al. (2022) examina a classificação de defeitos em wafers semicondutores a partir de imagens de microscopia eletrônica de varredura, com foco no uso de CNNs e vocabulários visuais como descritores. Ao comparar o desempenho das CNNs com a codificação do FV, observa-se que esta última é mais eficiente computacionalmente e produz melhores resultados de classificação. Os vocabulários visuais demonstram ser uma alternativa viável, alcançando uma precisão de até 93,13% na classificação dos defeitos. Apesar de apresentar disparidades na classificação entre diferentes classes de defeitos, o FV ainda se mostra útil e eficaz na identificação dos mesmos.

Em uma aplicação similar a deste trabalho, o artigo de Aslan et al. (2015) apresenta um sistema de detecção de quedas utilizando a caracterização de quedas baseada em contornos de imagens de vídeos, utilizando o classificador SVM. Características CSS (Curvature Scale Space) dos contornos são codificadas usando o FV. Testes no conjunto de dados SDUFall mostram uma precisão de detecção de quedas de até 88,83%, superando métodos anteriores baseados em Bag-of-Words. Esta abordagem oferece melhorias significativas em termos de precisão e eficiência computacional para a detecção automática de quedas em vídeos de profundidade.

Além dos estudos relacionados à aplicação do Fisher Vector (FV) em processamento e análise de imagens, a literatura também apresenta contribuições relevantes no uso dessa técnica para a detecção de distúrbios de voz. O trabalho de Yagnavajjula et al. (2022) investiga o potencial do FV na detecção automática de distúrbios de voz de origem neurológica (NVDs).



Neste estudo, os autores extraem características cepstrais, como Mel-Frequency Cepstral Coefficients (MFCCs) e Perceptual Linear Prediction Cepstral Coefficients (PLPCCs), tanto do sinal de voz acústico quanto do sinal de fonte glotal estimado por meio de filtragem inversa. Os descritores locais são então convertidos em descritores globais utilizando a codificação por FV. Os resultados indicam que o uso do FV leva a uma melhor identificação de pessoas com NVDs, em comparação com a codificação temporal convencional.

Outro estudo relevante é o de HM, Karjigi e Sreedevi (2023), que também explora o uso do FV para a detecção de distúrbios de voz, desta vez em um contexto de reconhecimento de fala. Neste trabalho, os autores extraem MFCCs e PLPCCs dos sinais de voz e os codificam usando o FV para obter descritores globais. Esses descritores são então utilizados para treinar um classificador SVM, que demonstra um desempenho superior na detecção de vozes saudáveis e patológicas em comparação com a codificação temporal.

### 2.3.3 Análise de contornos

No âmbito geral da análise de contornos, o estudo de Purnomo et al. (2024) apresenta uma análise detalhada do tamanho e forma de partículas de diferentes transportadores de oxigênio à base de ferro em diferentes graus de oxidação. Os autores utilizaram a técnica de análise de imagem dinâmica (DIA) para examinar a distribuição de tamanho, esfericidade e razão de aspecto das partículas. Já a pesquisa de Brophy et al. (2024) apresenta uma abordagem para classificação de dentes fósseis da família Bovidae usando análise de contornos elástica. Os autores derivam características de contornos e as utilizam para classificar dentes de bovídeos usando abordagens de aprendizado de máquina. Por fim, Yu et al. (2024) apresenta uma abordagem para identificar populações do peixe *Konosirus punctatus* nas águas costeiras da Coreia, utilizando análise do contorno dos otólitos e comparação de parâmetros de crescimento.

De maneira semelhante ao que será feito na primeira etapa deste trabalho, o estudo de Backes, Casanova e Bruno (2009) propõe um método para análise e classificação de contornos baseado em RCs. O método modela o contorno como uma RC, onde cada ponto do contorno é um vértice e as conexões entre os vértices são definidas pela distância euclidiana entre os pontos. Métricas da teoria de RCs, como grau médio, grau máximo e entropia, são então calculadas em diferentes limiares da RC para gerar um vetor de características que representa o contorno.

O método foi avaliado em três bases de dados: contornos genéricos com variações estruturais, peixes com rotações e escalas, e folhas. Os resultados mostraram que os descritores propostos baseados em grau da RC superam métodos tradicionais como descritores de Fourier, momentos de Zernike, descritores de curvatura e dimensão fractal multi-escala em termos de taxa de acerto na classificação. O método também apresentou propriedades desejáveis como invariância à rotação e escala, tolerância a ruído e robustez a degradação parcial do contorno.

Já a pesquisa apresentada por (RIBAS; NEIVA; BRUNO, 2019) apresenta a *Distance Transform Network* (DTN) para análise de contornos, combinando a robustez das RCs com a riqueza de informações da transformada de distância euclidiana. A técnica começa gerando um mapa de distância aplicando a transformada de distância euclidiana em cada contorno. Cada raio de dilatação é modelado como uma RC, e as medições de grau da evolução dinâmica da RC são usadas para caracterizar o contorno. Um vetor de características robusto é composto pelas características de diferentes raios de dilatação. Testes em sete bases de dados, incluindo contornos naturais como otólitos e folhas, demonstraram que a DTN é eficaz na classificação de contornos, mostrando taxas de sucesso superiores aos métodos existentes na literatura. As vantagens incluem robustez contra degradação e ruído, e tolerância a variações na escala e orientação dos contornos.

Avançando mais nos métodos, o trabalho de Ribas e Bruno (2024) apresenta uma nova abordagem para análise e classificação de contornos a partir de imagens, utilizando redes neurais randomizadas (RNNs) aplicadas em mapas de características topológicas derivados de uma estrutura de RC. A metodologia propõe modelar o contorno como uma RC, computando suas medidas topológicas através de uma estratégia de evolução dinâmica. As medidas topológicas resultantes são incorporadas em uma representação matricial chamada Mapa de Características Topológicas (TFM). A RNN é então usada para aprender representações a partir do TFM através de uma estratégia de janela deslizante. Os resultados experimentais mostram melhorias significativas no desempenho da classificação de contornos, incluindo a robustez ao ruído e invariância a transformações de escala e rotação. Esta técnica foi testada em dois conjuntos de dados genéricos e no reconhecimento de folhas de plantas, obtendo alto desempenho nessas tarefas.

## 3 Materiais e métodos

Este capítulo tem como objetivo descrever os principais materiais e métodos utilizados na etapa de configuração experimental, a fim de fornecer uma visão mais aprofundada a respeito dessa importante etapa. São apresentadas as bases de dados escolhidas, classificadores utilizados, validações e métricas de desempenho adotadas e ferramentas usadas no processo.

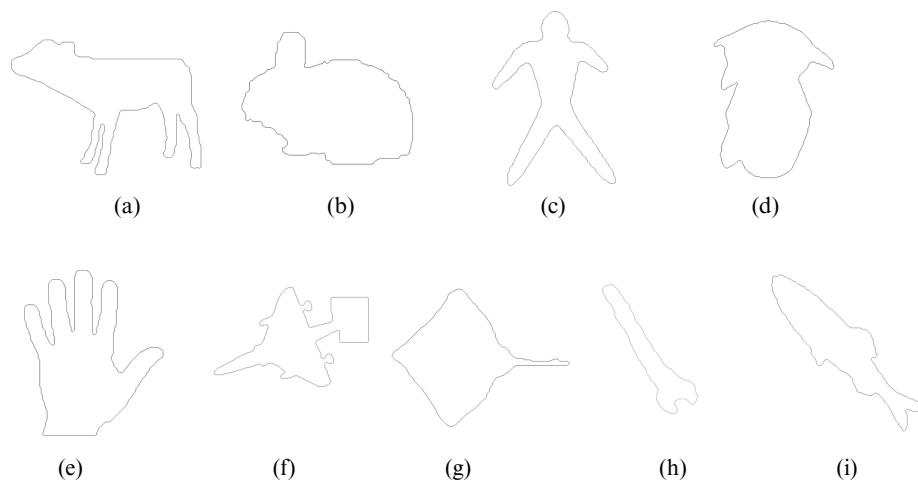
### 3.1 Bases de dados de contornos

A fim de avaliar a proposta deste trabalho, foram utilizadas seis bases de dados de imagens de contornos: *Kimia*, *ETH-80*, *Otolith*, *Aforo otolith*, *Portuguese Leaves* e *USP Leaves*.

#### 3.1.1 Kimia

Formada por 99 contornos genéricos divididos em 9 classes, exemplificadas na Figura 6. As imagens (PNG) foram obtidas por projeções de contornos 3D, o que gera certa degradação em algumas partes. Além disso, algumas amostras são rotacionadas, o que contribui para a complexidade da análise (SEBASTIAN; KLEIN; KIMIA, 2004).

Figura 6 – Exemplo das 9 classes da base de dados Kimia, sendo elas: (a) quadrúpede, (b) coelho, (c) homem, (d) ave, (e) mão, (f) jato, (g) raia, (h) ferramenta e (i) peixe.

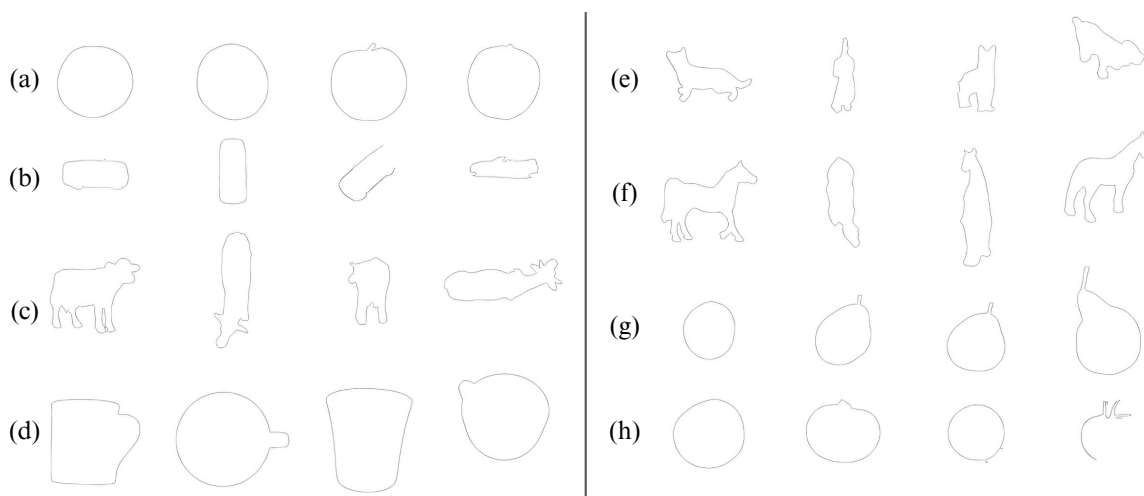


Fonte: Sebastian, Klein e Kimia (2004)

### 3.1.2 ETH-80

A base de dados ETH-80 contém 3280 imagens PNG separadas em 8 classes, podendo ser objetos feitos por humanos ou objetos naturais, ambos em diferentes estilos, inclusive representações em diferentes pontos de vista, como demonstrado na Figura 7. Essa organização permite que, para cada tipo de objeto, os contornos obtidos sejam bem diferentes, devido à variância de rotação adicionada (LEIBE; SCHIELE, 2003).

Figura 7 – Quatro variações para as oito classes da base de dados ETH-80, contendo imagens de contornos de (a) maçãs, (b) carros, (c) vacas, (d) canecas, (e) cachorros, (f) cavalos, (g) peras e (h) tomates.



Fonte: Leibe e Schiele (2003)

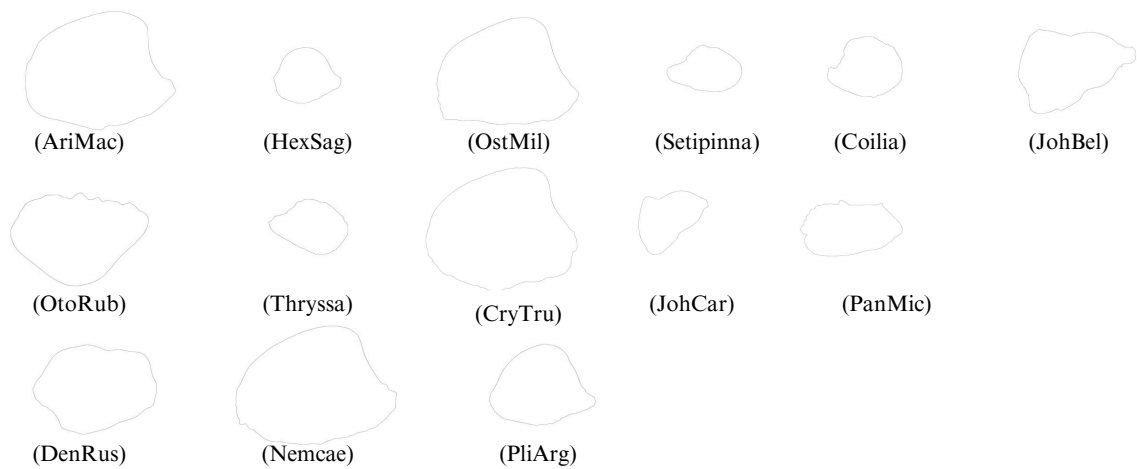
### 3.1.3 Otolith

Encontradas no labirinto do ouvido interno dos peixes, os otólitos são estruturas compostas por cristais de carbonato de cálcio. Essas formações desempenham um papel importante na taxonomia, sendo amplamente utilizadas para a identificação e classificação de diferentes espécies de peixes, devido às suas características morfológicas específicas que variam entre as espécies. Exemplificada na Figura 8, a base de dados contém imagens PNG de otólitos de 14 espécies pertencentes a três famílias: Engraulidae, Sciaenidae e Ariidae. As imagens foram capturadas utilizando um estereomicroscópio (Olympus DP25FW, com ampliação de 6,3x) e uma câmera digital, registrando diferentes posições das estruturas (SALIMI et al., 2015).

### 3.1.4 Aforo otolith

Oferecida pela AFORO (LOMBARTE et al., 2006), esta base de dados para reconhecimento de peixes é um catálogo online de imagens BMP de otólitos. Para o experimento,

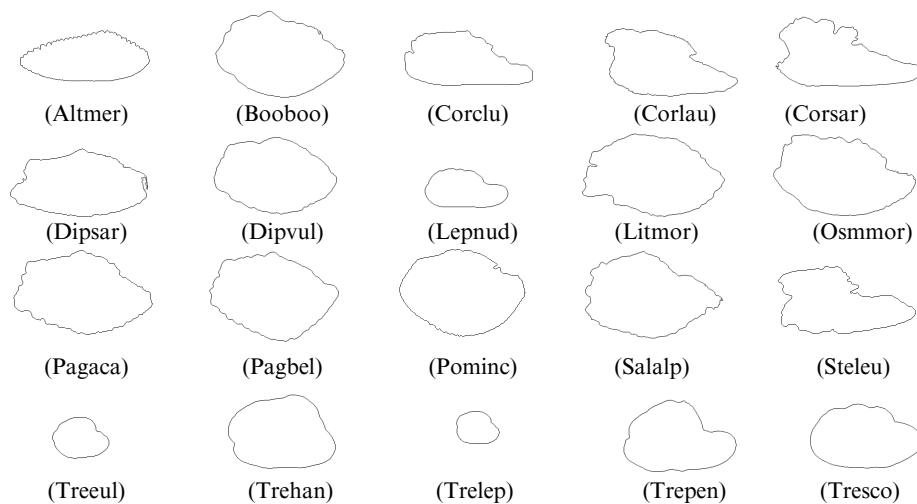
Figura 8 – Contornos de otólitos de quatorze espécies de peixes, retirados da base de dados Otolith.



Fonte: Salimi et al. (2015)

foram utilizadas 180 imagens de contornos de otólitos de 20 espécies diferentes, como demonstrado na Figura 9.

Figura 9 – Contornos de otólitos de vinte espécies de peixes, retirados da base de dados Aforo.



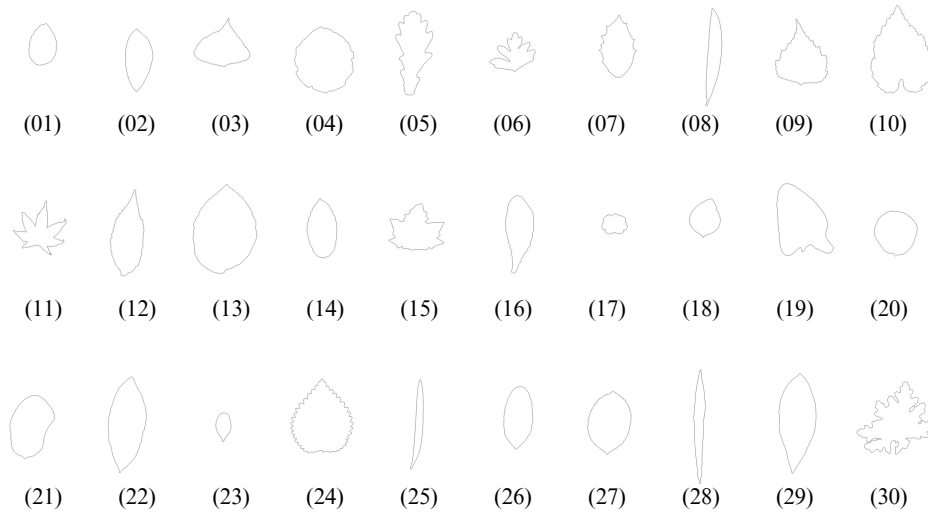
Fonte: Lombarte et al. (2006)

### 3.1.5 Portuguese Leaves

Esta base de dados, apresentada em Silva, Marcal e Silva (2013), possui, ao todo, 340 imagens BMP de folhas de 30 espécies diferentes, exemplificadas na Figura 10. A base de dados original contém imagens RGB e binárias. Para este experimento, foi utilizada a base

de dados modificada em Ribas, Neiva e Bruno (2019), onde os contornos foram extraídos utilizando o Matlab.

Figura 10 – Contornos de folhas de 30 espécies diferentes, retirados da base de dados Portuguese Leaves.



Fonte: Silva, Marcal e Silva (2013)

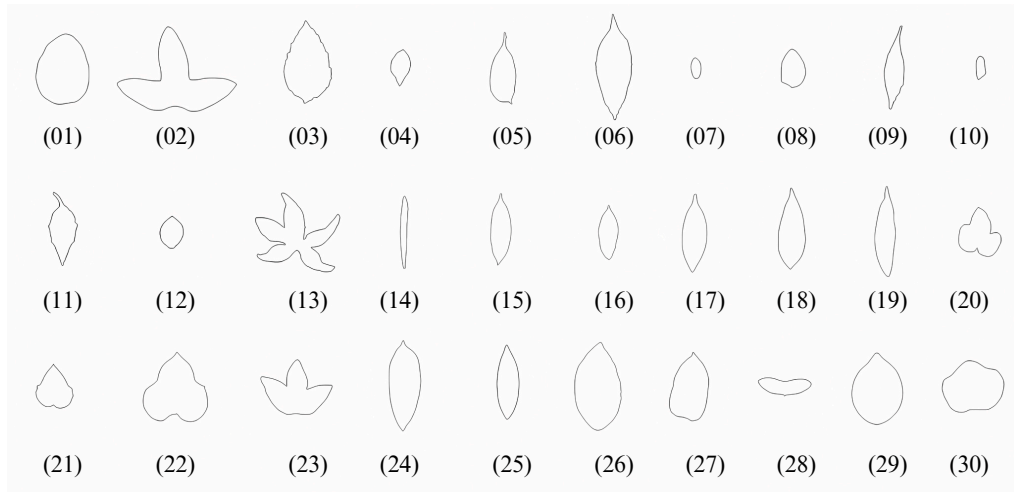
### 3.1.6 USP Leaves

Base de dados criada por pesquisadores da Universidade de São Paulo (BACKES; CASANOVA; BRUNO, 2009), contendo 600 imagens PNG de 30 espécies de folhas, demonstradas na Figura 11. A classificação dessa base de dados é uma tarefa desafiadora, já que diferentes espécies de plantas podem conter contornos de folhas parecidos. Ademais, sobreposições nos contornos podem acontecer, devido ao processo de digitalização das imagens.

Além disso, para que seja possível avaliar a robustez do método a variância de rotação, escala e ruído, esta base também possui conjuntos suplementares com as seguintes características:

- Rotação: Contém imagens rotacionadas a 7, 35, 132, 201 e 298 graus, gerando um novo conjunto com 3600 contornos.
- Escala: Conjunto com imagens em escalas de 125%, 150%, 175% e 200% em relação às imagens originais, resultando em uma base diferente, agora com as imagens originais e em diferentes escalas.
- Ruído: Neste conjunto, diferentes níveis de ruído uniforme foram adicionados às imagens, seguindo um intervalo  $[-l \dots l]$ , onde  $l$  é o nível de ruído. Serão adotados 4 níveis neste experimento, o que resulta em 80 novas imagens para cada classe.

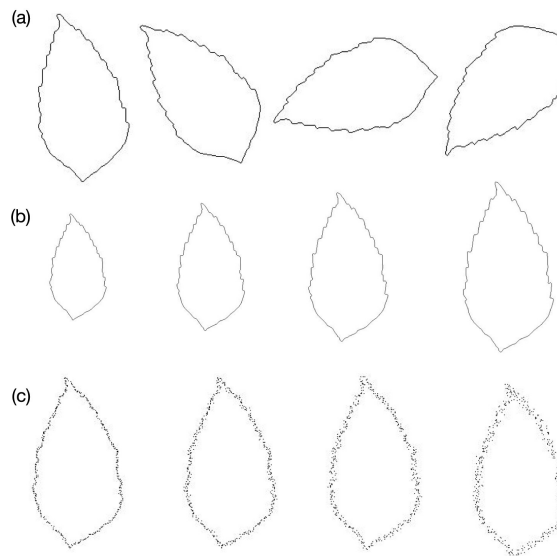
Figura 11 – Contornos de folhas de 30 espécies diferentes, retirados da base de dados USP Leaves.



Fonte: Backes, Casanova e Bruno (2009)

Exemplos de conjuntos para avaliação da robustez do método são apresentados na Figura 12.

Figura 12 – Base USP Leaves: Exemplos de amostras modificadas. (a) rotacionadas, (b) dimensionadas e (c) com ruído.



Fonte: Ribas e Bruno (2024), Backes, Casanova e Bruno (2009)

## 3.2 Classificadores

Para fins de teste e, depois, validação, o vetor de características obtido será utilizado em classificadores, a saber: Support Vector Machine (SVM) e Linear Discriminant Analysis (LDA)

### 3.2.1 Support Vector Machine

O SVM é um algoritmo de aprendizado supervisionado que é usado tanto para classificação quanto para regressão. O objetivo do SVM é encontrar um hiperplano no espaço de características que melhor separe as classes dos dados de treinamento (CORTES; VAPNIK, 1995; BURGESS, 1998).

Em termos simples, o SVM mapeia os dados de entrada para um espaço de características de alta dimensionalidade onde é mais fácil encontrar um hiperplano de separação. O SVM então determina o hiperplano ótimo que maximiza a margem entre as classes, permitindo uma boa generalização para dados não vistos.

### 3.2.2 Linear Discriminant Analysis

O LDA é um método de aprendizado supervisionado utilizado para a classificação de padrões que busca encontrar a melhor combinação linear das características, maximizando a separação entre as classes e minimizando a dispersão intra-classe. Em resumo, o LDA maximiza a separabilidade entre classes, permitindo uma melhor compreensão da distribuição dos dados e uma classificação mais precisa (FISHER, 1936; BALAKRISHNAMA; GANAPATHIRAJU, 1998).

## 3.3 Validação e métricas de desempenho

Com o objetivo de medir o desempenho da performance do método proposto, foi adotada a técnica de validação cruzada *k-fold* estratificada. Adicionalmente, algumas métricas foram utilizadas, como a acurácia e desvio-padrão.

### 3.3.1 Validação cruzada *K-Fold* estratificada

O validação cruzada *k-fold* é uma técnica utilizada para avaliar a capacidade de generalização de um modelo de aprendizado de máquina. Nesse método, o conjunto de dados é dividido aleatoriamente em  $K$  subconjuntos mutuamente exclusivos, chamados de *folds*. O modelo é treinado  $K$  vezes, sendo treinado em  $K - 1$  *folds* e testado na *fold* restante em cada iteração (KOHAVI, 1995).

A variação dessa técnica que será usada neste trabalho é a validação cruzada *k-fold* estratificada, que visa melhorar a representatividade das *folds* em relação ao conjunto de dados original. No método tradicional, os dados são divididos aleatoriamente em  $k$  *folds*, o que pode resultar em *folds* com distribuições de classes significativamente diferentes do conjunto de dados original, especialmente em casos de desbalanceamento de classes. O *k-fold* estratificado aborda esse problema garantindo que cada *fold* tenha aproximadamente a mesma proporção de cada classe presente no conjunto de dados completo. Isso é importante para



obter uma avaliação mais precisa do desempenho do modelo, pois evita vieses que podem ocorrer devido à má representação das classes em algumas *folds*.

### 3.3.2 Matriz de confusão:

Uma matriz de confusão é uma ferramenta que fornece a visualização do desempenho de um classificador, no âmbito do aprendizado de máquina. Essa matriz fornece uma representação clara das previsões feitas pelo modelo em comparação com os valores reais.

A matriz de confusão é estruturada como uma tabela que apresenta, para cada classe, as frequências dos seguintes resultados:

- **Verdadeiro positivo (VP):** Número de instâncias positivas corretamente classificadas pelo modelo. Ou seja, é o número de vezes que o modelo previu a classe positiva corretamente.
- **Falso positivo (FP):** Número de instâncias negativas incorretamente classificadas como positivas. Isso ocorre quando o modelo prediz a classe positiva para uma instância que é negativa.
- **Verdadeiro negativo (VN):** Número de instâncias negativas corretamente classificadas pelo modelo. Representa o número de vezes que o modelo previu a classe negativa corretamente.
- **Falso negativo (FN):** Número de instâncias positivas incorretamente classificadas como negativas. Isso ocorre quando o modelo prediz a classe negativa para uma instância que é positiva.

### 3.3.3 Acurácia

A acurácia (ACC), estimada do modelo em cada interação  $k$  é calculada como o número total de classificações corretas dividido pelo número total de instâncias no conjunto de dados. Esse valor é uma estimativa da capacidade do modelo de generalizar para novos dados.

$$ACC = \frac{\text{Número de classificações corretas}}{\text{Número total de classificações}} = \frac{VP + VN}{VP + VN + FP + FN} \quad (3.1)$$

A acurácia média do modelo, portanto, é calculada como:

$$\overline{ACC} = \frac{1}{k} \sum_{i=1}^k ACC_i \quad (3.2)$$

### 3.3.4 Desvio-padrão

O desvio padrão (DP),  $\sigma$ , fornece uma medida da variabilidade das acurácias entre as diferentes *folds*, indicando a estabilidade do modelo, e pode ser calculado como:

$$\sigma = \sqrt{\frac{1}{k-1} \sum_{i=1}^k (\text{ACC}_i - \overline{\text{ACC}})^2} \quad (3.3)$$

## 3.4 Ferramentas

Algumas ferramentas foram necessárias para a implementação da abordagem proposta. A linguagem de programação escolhida foi o Python, juntamente com algumas bibliotecas que se adequam à solução.

### 3.4.1 Python

A linguagem Python é interpretada, interativa e orientada a objetos. Por ser interpretada, executando o código linha a linha, torna-se fácil encontrar rapidamente os erros. Além disso, devido à sua tipagem dinâmica, a codificação é acelerada, pois não é necessário especificar os tipos das variáveis. Ademais, Python pode ser executado em várias plataformas, como Linux, Mac e Windows, e, por ser um software livre, pode ser compartilhado e modificado, resultando em uma linguagem constantemente aperfeiçoada (ROSSUM; DRAKE, 2009).

No entanto, um ponto negativo a ser considerado é que a execução de código Python pode ser mais lenta em comparação a outras, como C ou C++, que são compiladas. Apesar dessa desvantagem, a facilidade de uso, a legibilidade do código e a diversidade de bibliotecas disponíveis justificam a escolha por Python.

### 3.4.2 Bibliotecas Python

Python é conhecido por ter uma grande quantidade de bibliotecas, abrangendo os mais diversos tipos de aplicações. Estas foram as principais bibliotecas utilizadas na implementação deste trabalho:

- *NumPy*: Uma biblioteca essencial para computação científica, oferecendo suporte a arrays e matrizes multidimensionais, fornecendo uma ampla gama de funções otimizadas para operações rápidas sobre essas estruturas (HARRIS et al., 2020).
- *Scikit-learn*: O Scikit-learn é uma biblioteca open source de aprendizado de máquina que oferece suporte a algoritmos de aprendizado supervisionado e não supervisionado. Além disso, auxilia no pré-processamento de dados, seleção de modelos e avaliação de desempenho, por exemplo, facilitando a implementação de soluções de aprendizado

de máquina (PEDREGOSA et al., 2011). Esta biblioteca foi utilizada nas tarefas de classificação e validação deste trabalho.

- *scikit-image*: Voltado ao processamento de imagens e baseada no NumPy, o scikit-image oferece algoritmos para tarefas como filtragem, segmentação, transformação geométrica, restauração, análise de regiões e detecção de bordas (WALT et al., 2014). Neste trabalho, esta biblioteca foi importante para a implementação do Fisher Vector.
- *Joblib*: O Joblib é uma biblioteca voltada para a execução de tarefas que envolvem computação intensiva, especialmente em cenários de paralelização. Ela é comumente utilizado para melhorar o tempo de execução de tarefas de machine learning e processamento de dados (Joblib Development Team, 2024).
- *Pandas*: Uma biblioteca para manipulação e análise de dados, permitindo a realização de tarefas como limpeza de dados, agregação, filtragem e análise estatística, por exemplo. Sua compatibilidade com formatos de dados como CSV, Excel e SQL a torna uma ferramenta essencial para ciência de dados e análise exploratória (MCKINNEY et al., 2010). Neste trabalho, esta biblioteca foi utilizada na análise de parâmetros.
- *Matplotlib*: O Matplotlib é uma biblioteca utilizada na criação de visualizações de dados 2D, auxiliando na geração de gráficos estáticos, animados e interativos. É uma ferramenta relevante para exploração e análise de dados (HUNTER, 2007). Assim como a biblioteca Pandas, a Matplotlib foi utilizada na análise de parâmetros, auxiliando na visualização dos dados obtidos.

### 3.5 Métodos de classificação de contornos comparados

É de fundamental importância que novos métodos tenham seus resultados comparados com métodos já conhecidos e estabelecidos na literatura. Portanto, na Seção 5.2, os métodos descritos nesta seção serão comparados com a abordagem proposta.

- **Representações Complexas para Análise de Contorno (CRSA)**: Assim como neste trabalho, este método modela o contorno como uma rede complexa. São incorporadas características topológicas da rede por meio de Mapas de Características Topológicas (TFMs), que são gerados a partir da evolução dinâmica da RC. No processamento desses mapas, são utilizadas Redes Neurais Randomizadas (RNNs), adotando uma abordagem de janela deslizante. Os pesos treinados da camada de saída das RNNs são utilizados como vetor de características (RIBAS; BRUNO, 2024).
- **Rede de transformada da distância (DTN)**: Utiliza-se o mapa de distância da imagem modelada como RC, obtido ao aplicar a transformada da distância Euclidiana e limita-se os pixels a um raio de dilatação  $r$ . Descritores de grau para diferentes valores

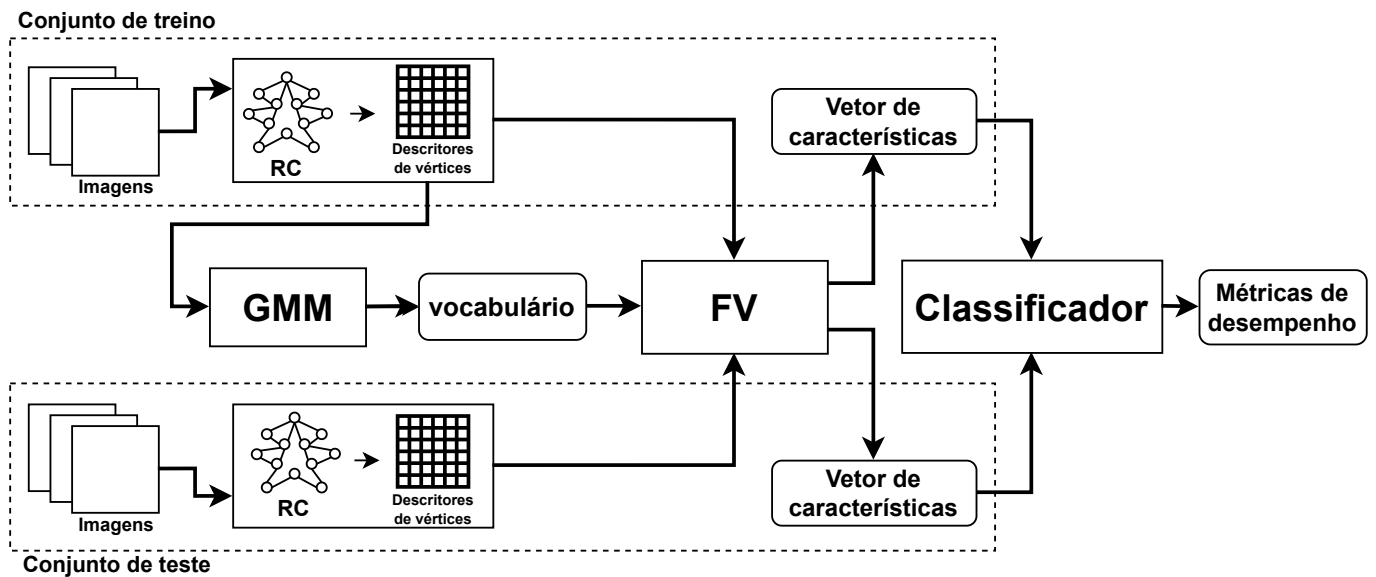
de limiares da evolução dinâmica e raios de dilatação são usados para obter o vetor de características (RIBAS; NEIVA; BRUNO, 2019).

- **C.N. degree:** Neste método, a imagem, também modelada como uma rede complexa, passa pela evolução dinâmica por meio de limiares, tendo seus descritores de grau e grau conjunto (*joint degree*) extraídos e concatenados para formar o vetor de características final (BACKES; CASANOVA; BRUNO, 2009).
- **Descritor de Fourier:** Aplica-se a transformada espectral ao contorno da imagem e utiliza-se os coeficientes resultantes para a extração de características. Nos experimentos realizados, foram selecionados os 20 coeficientes mais relevantes do espectro como representações das características da imagem (OSOWSKI; NGHIA, 2002).
- **Descritores de curvatura:** Este método consiste em analisar o contorno como uma curva, com o objetivo de extrair descritores, como pontos de máximo e mínimo, que capturam características relevantes, como alterações na direção do contorno (WU; WANG, 1993).
- **Momentos de Zernike:** Esta técnica calcula um conjunto de momentos de Zernike da imagem, de ordem 0 a 7, a fim de usá-los como vetor de características. O método se destaca por ser invariante a rotação (ZHENJIANG, 2000).
- **Dimensão fractal multi-escala:** A dimensão fractal, neste método, permite capturar a complexidade do contorno, que é modelado como uma curva. O vetor de características resultante é composto pelos 50 pontos mais significativos do contorno, resultando em uma representação unidimensional do objeto (TORRES; FALCAO; COSTA, 2004).
- **Análise de segmento:** Baseia-se na descrição do contorno por meio de estatísticas de segmentos de linha reta. Ele considera porções de pontos contínuos e calcula o comprimento da linha reta entre os pontos extremos dessa porção, calculando também a média e o desvio padrão para cada segmento (JUNIOR; BACKES, 2015).
- **Descritores angulares de RCs (ADCN):** Também fundamentado em RCs, este método analisa o ângulo formado pelos pontos do contorno. Aplica-se uma evolução dinâmica e cria-se um histograma dos ângulos calculados. Os resultados demonstram robustez a rotação, escala, degradação e ruído (SCABINI et al., 2017).

## 4 Abordagem proposta

Neste capítulo, é apresentado o método proposto para a realização dos experimentos e a obtenção dos resultados, com a descrição detalhada de cada etapa. A metodologia desta abordagem é composta por diversas fases, conforme ilustrado no diagrama da Figura 13.

Figura 13 – Ilustração da metodologia adotada na abordagem proposta.



Fonte: Elaborado pelo autor

### 4.1 Visão Geral

Como ilustrado na Figura 13, primeiramente, o conjunto de imagens é dividido em treino e teste, seguindo a validação cruzada 10-fold estratificada, conforme descrito na Seção 3.3. Para ambos os conjuntos, as imagens são modeladas como RCs (Seção 4.2), e os descritores de vértices das redes são extraídos (Seção 4.3). Em seguida, apenas no conjunto de treino, um GMM é treinado a partir dos descritores, gerando o vocabulário visual.

Com o vocabulário definido, os Fisher Vectors dos descritores das RCs são calculados para os conjuntos de treino e teste, utilizando os parâmetros do GMM, ou seja, o vocabulário (Seção 4.4). O Fisher Vector calculado representa o vetor de características final que é empregado em um classificador (Seção 3.2). Por fim, são calculadas as métricas de desempenho da classificação, permitindo a análise do desempenho do modelo ao utilizar o Fisher Vector como vetor de características (Seções 3.3.2 a 3.3.4).

## 4.2 Modelagem do contorno como RC

Para esta etapa, a metodologia utilizada é a descrita por Backes, Casanova e Bruno (2009).

Sendo  $S$  o contorno de uma imagem, onde  $S = [s_1, s_2, \dots, s_n]$ ,  $s_i$  é um vetor de forma  $s_i = [x_i, y_i]$  com componentes sendo valores discretos que representam as coordenadas do ponto  $i$  do contorno. O objetivo é representar o contorno  $S$  como um grafo  $G = (V, E)$ , para isso, cada pixel do contorno é representado como um vértice na rede, isto é,  $S = V$ . O conjunto de arestas que conectam os vértices é calculado usando a distância Euclidiana:

$$d(s_i, s_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \quad (4.1)$$

Assim, a RC é representada por uma matriz  $N \times N$  ponderada  $W$ :

$$w_{ij} = W([w_i, w_j]) = d(s_i, s_j), \quad (4.2)$$

normalizada para o intervalo  $[0,1]$ ,

$$W = \frac{W}{\max_{w_{ij} \in W}} \quad (4.3)$$

Dessa maneira, inicialmente, a rede é um grafo completo, ou seja, todos os vértices são conectados entre si, possuindo o mesmo número de conexões. Entretanto, um grafo completo não se caracteriza como uma RC, já que não há propriedades relevantes para serem analisadas. Portanto, é necessário transformar essa rede em uma RC.

A maneira escolhida de transformar essa rede é pela aplicação de um *threshold*  $T_l$ , a fim de produzir novos conjuntos de arestas à rede. Para cada *threshold*  $T_l$ , o conjunto de arestas será composto por aquelas que possuem peso menor que o valor  $T_l$ .

## 4.3 Evolução dinâmica da RC

Assim como foi descrito por Backes, Casanova e Bruno (2009), a caracterização de uma rede complexa pode ser feita através da caracterização da evolução dinâmica da rede por meio da aplicação de diferentes valores de  $T_l$  de uma transformação  $\delta$ . Essa operação, representada por  $A = \delta_{T_l}(W)$ , é aplicada a cada elemento da matriz ponderada  $W$ , gerando a matriz de adjacência não ponderada  $A$ :

$$A_{T_l} = \delta_{T_l}(W) = \forall w \in W \begin{cases} a_{ij} = 0 & \text{se } w_{ij} \geq T_l, \\ a_{ij} = 1 & \text{se } w_{ij} < T_l. \end{cases} \quad (4.4)$$

Assim, a caracterização é feita utilizando várias transformações  $\delta$ , com o *threshold*  $T_l$  sendo incrementado em um intervalo regular  $T_{inc}$ , onde, em cada incremento, arestas com peso maior que o *threshold*  $T_l$  são cortadas, gerando uma evolução dinâmica dos vértices.

Em um conjunto de *thresholds*  $T$ , um elemento  $T_l \in T$  é definido por:

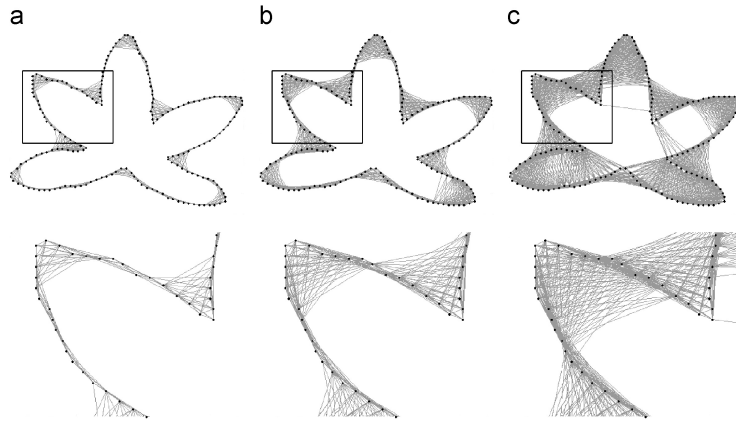
$$\begin{aligned} f : T &\rightarrow T, \text{ onde} \\ T_0 &= T_{ini} \mid 0 < T_{ini} < 1, \\ T_{l+1} &= f(T_l) \quad \text{se } T_l < T_{fim} < 1, \\ f(x) &= x + T_{inc}, \end{aligned} \tag{4.5}$$

onde

$$T_{inc} = \frac{1}{N}, \tag{4.6}$$

sendo  $N - 1$  o número de incrementos que serão realizados no intervalo  $(0, 1)$ , com o valor  $N$  definido pelo usuário. Essa abordagem permite a extração de características da rede em diferentes momentos, como demonstrado na Figura 14.

Figura 14 – Evolução dinâmica da RC por um limiar  $T_l$  e zoom aplicado na área: (a)  $T_l = 0,15$ ; (b)  $T_l = 0,15$ ; (c)  $T_l = 0,2$



Fonte: Backes, Casanova e Bruno (2009)

Portanto, ao final, obtém-se  $n$  vetores de descritores de cada vértice da rede complexa que modela a imagem do contorno. Para cada vértice da evolução dinâmica, são obtidos vetores das medidas de grau, força e coeficiente de clustering dos vértices, todas normalizadas para o intervalo  $(0, 1)$ .

#### 4.3.1 Grau

Para cada vértice  $V_i$  da matriz de adjacência  $A$  descrita em Eq. (4.4), considerando o atual valor de *threshold*  $T_l$ , se calcula o grau  $g$  do vértice, como descrito na Seção 2.1.2.

$$V_g^i = [g_i^{t_1}, g_i^{t_2}, \dots, g_i^{t_n}] \quad (4.7)$$

### 4.3.2 Força

Para cada vértice  $V_i$  da matriz de adjacência  $A$  descrita em Eq. (4.4), considerando o atual valor de *threshold*  $T_l$ , se considera a força  $s$  do vértice, como descrita na Seção 2.1.2, que já estará presente em  $W$ , Eq. (4.2).

$$V_s^i = [s_i^{t_1}, s_i^{t_2}, \dots, s_i^{t_n}] \quad (4.8)$$

### 4.3.3 Coeficiente de *clustering*

Para cada vértice  $V_i$  da matriz de adjacência  $A$  descrita em Eq. (4.4), considerando o atual valor de *threshold*  $T_l$ , se calcula o coeficiente de clustering local  $c$ , como descrito na Seção 2.1.2.

$$V_c^i = [c_i^{t_1}, c_i^{t_2}, \dots, c_i^{t_n}] \quad (4.9)$$

## 4.4 Obtenção do vetor de características

É importante que o processo de obtenção do Fisher Vector seja bem definido, a fim de que o método seja agora compreendido e futuramente reproduzido em mais estudos e aplicações. A Figura 15 demonstra esse processo:

- a) **Contorno:** Inicia-se com a imagem de um contorno, binária, onde os pixels brancos representam o contorno e os pretos representam o fundo. Esta será a entrada ao método.
- b) **Modelagem como RC:** A imagem de entrada será modelada como uma RC e submetida a uma evolução dinâmica de  $N - 1$  limiares, estruturando o comportamento do contorno ao decorrer de cada transformação, assim como descrito na Seção 4.3.
- c) **Extração dos descritores:** Com os contornos da evolução dinâmica estabelecidos, extrai-se seus descritores que, neste trabalho, poderão ser uma combinação  $\phi$  entre grau, força e clustering. E, a fim de que cada linha da matriz reflita a evolução dinâmica de um vértice da RC, com dimensão constante  $D$ , é necessário transpô-la. A matriz transposta de descritores terá uma dimensão  $m \times D$ , onde  $m$  é a quantidade de vértices da rede complexa e  $D$  é a quantidade  $n = N - 1$  de limiares multiplicada pela quantidade  $|\phi|$  de descritores combinados.



$$D = n * |\phi| \quad (4.10)$$

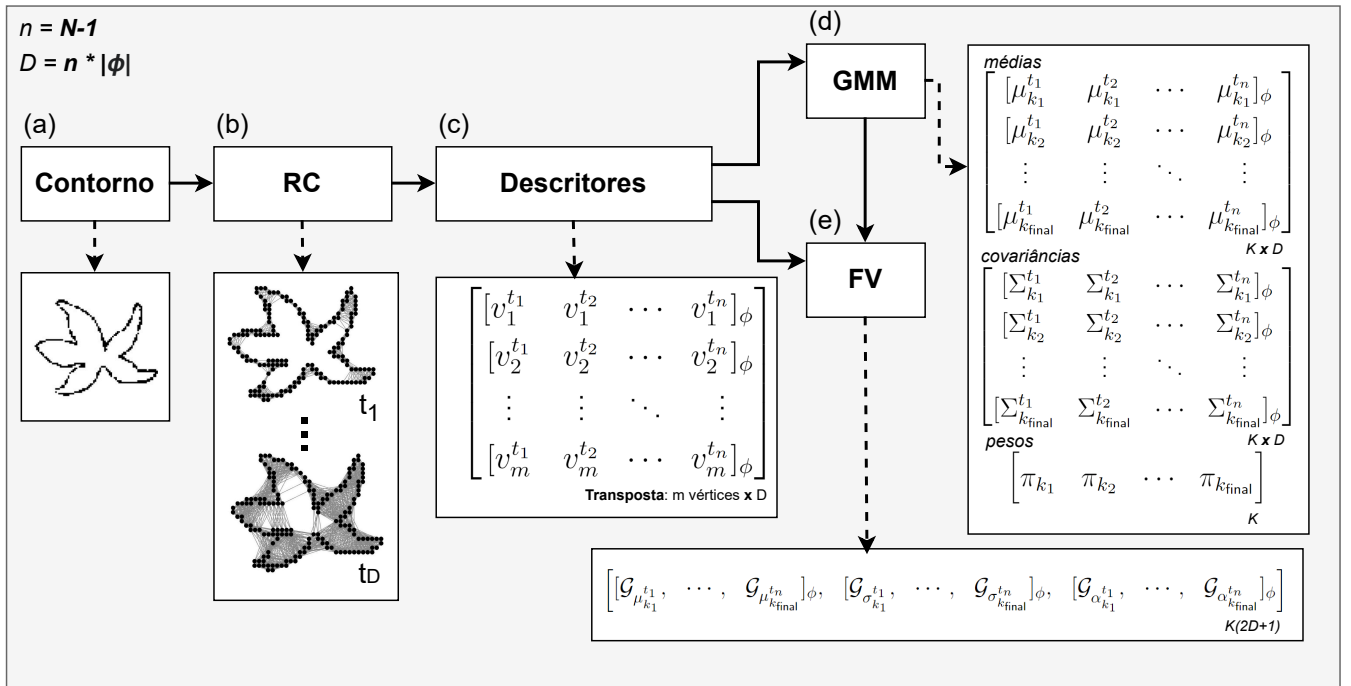
A concatenação das linhas ocorre da seguinte forma, considerando que  $\phi = [\text{grau}, \text{força}]$  neste exemplo:

$$\begin{bmatrix} g_1^{t_1} & g_1^{t_2} & \cdots & g_1^{t_n} & s_1^{t_1} & s_1^{t_2} & \cdots & s_1^{t_n} \\ g_2^{t_1} & g_2^{t_2} & \cdots & g_2^{t_n} & s_2^{t_1} & s_2^{t_2} & \cdots & s_2^{t_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ g_m^{t_1} & g_m^{t_2} & \cdots & g_m^{t_n} & s_m^{t_1} & s_m^{t_2} & \cdots & s_m^{t_n} \end{bmatrix} \quad (4.11)$$

- d) **GMM**: A matriz transposta  $m \times D$  será a entrada para o GMM, que construirá o vocabulário. Nessa etapa, é informada a quantidade  $K$  de gaussianas. Então, assim como descrito em Seção 2.2.1, serão calculadas as médias  $\mu$ , covariâncias  $\Sigma$  e pesos  $\pi$ , nas etapas de *E-step* e *M-step*. As médias e covariâncias são calculadas a partir das linhas da matriz de descritores, que representam as características de um vértice no decorrer dos limiares. Nas matrizes geradas, cada linha representa a distribuição dessas médias e covariâncias em cada componente  $K$ , resultando em matrizes  $K \times D$ . O vetor de pesos, por outro lado tem dimensão  $K$ , já que se trata de uma atribuição de peso a cada gaussiana.
- e) **Fisher Vector**: As matrizes de dimensão  $K \times D$  de médias  $\mu$  e covariâncias  $\Sigma$ , o vetor com dimensão  $K$  de pesos  $\pi$  e a matriz transposta  $m \times D$  de descritores serão a entrada para a codificação do Fisher Vector. Nesta etapa, assim como descrito na Seção 2.2.3, a matriz de covariância  $\Sigma$  é re-parametrizada em um vetor de desvios-padrões  $\sigma$  (SÁNCHEZ et al., 2013) e o vetor de pesos  $\pi$  re-parametrizado em um vetor  $\alpha$  pela função soft-max (KRAPAC; VERBEEK; JURIE, 2011). Após isso, são calculados e concatenados os gradientes de verossimilhança  $\mathcal{G}$  dos descritores, onde cada linha é considerada um descritor, em relação a todos os parâmetros  $\mu_k, \sigma_k, \alpha_k$  em um único vetor de dimensão  $K(2D + 1)$ , que será o vetor de características final.

Ressalta-se que o processo de ajuste do GMM será feito somente com as imagens do conjunto de treino. Assim, o vocabulário gerado será utilizado tanto na codificação dos FVs das imagens do conjunto de treino, quanto das imagens do conjunto de teste.

Figura 15 – Processo de obtenção do Fisher Vector que será o vetor de característica final para descrição da evolução da rede complexa modelando um contorno.



Fonte: Elaborado pelo autor com figuras adaptadas de Ribas e Bruno (2024)

## 5 Resultados

Este capítulo tem como objetivo apresentar a análise dos parâmetros realizada, seguida da comparação dos resultados obtidos com os métodos descritos na Seção 3.5. Ao final, são discutidos os resultados para todas as bases de dados, incluindo a análise de invariância e robustez do método.

### 5.1 Análise de parâmetros

Com o objetivo de fornecer uma avaliação abrangente do desempenho do método, foram realizados experimentos em todas as bases de dados, variando o número de componentes  $K$  no intervalo de 4 a 24, e variando o valor  $N$  de 10 a 70, para definir o incremento dos limiares. Além disso, foram testadas diferentes combinações de descritores para caracterizar as redes complexas e aplicar a GMM, incluindo as métricas de grau, força e coeficiente de clustering.

A Tabela 1 apresenta alguns resultados obtidos com combinações diferentes de parâmetros. Nela, percebe-se que determinar o melhor parâmetro é uma tarefa desafiadora, portanto, faz-se necessária uma análise de parâmetros, a fim de que o comportamento do método seja estudado.

Nesta seção, são investigados os efeitos dos seguintes parâmetros na abordagem proposta: (i) quantidade  $K$  de gaussianas; (ii) valor  $N$  para definição de limiares  $N - 1$  limiares no intervalo (0,1); e (iii) combinação  $\phi$  de descritores de grau ( $g$ ), força ( $s$ ) e coeficiente de clustering ( $c$ ). Para realizar essa análise, as bases foram separadas em três categorias: genéricas, otólitos e folhas, a fim de compreender o comportamento do método em bases com contornos similares. Na análise, foram considerados os resultados obtidos usando a validação cruzada 10-fold estratificada e o classificador LDA.

A análise inicial teve como foco os parâmetros  $K$  e  $N$ , avaliando um conjunto  $K = \{4, 8, 10, 14, 18, 20, 24\}$  e  $N = \{10, 20, 35, 45, 60, 70\}$ , enquanto a concatenação dos descritores grau, força e clustering foi mantida, resultando em um conjunto  $\phi = [g, s, c]$ . Já em uma segunda análise, com três parâmetros  $K$  e  $N$  escolhidos, passa-se a focar no impacto do conjunto de descritores  $\phi$  na performance do método. Sendo assim, serão consideradas todas as combinações entre grau, força e coeficiente de clustering. A análise dos parâmetros  $K$ ,  $N$  e  $\phi$  é de extrema importância pois eles determinam tanto a entrada para o cálculo do Fisher Vector quanto a dimensão de sua saída.

Tabela 1 – Recorte dos experimentos iniciais com diferentes combinações de parâmetros em todas as bases de dados, demonstrando a porcentagem de acurácia para os classificadores LDA e SVM.

K	N	Descritores $\phi$	ETH-80		Kimia		Aforo Otolith	
			LDA	SVM	LDA	SVM	LDA	SVM
4	10	[grau]	80.76	86.68	96.89	95.89	61.67	66.67
4	70	[grau, força, clustering]	88.89	92.26	97.89	96.89	76.11	68.33
8	20	[força]	80.15	86.19	99.00	96.89	64.44	62.22
10	35	[força, clustering]	90.18	92.74	99.00	96.89	71.11	67.22
10	60	[clustering]	90.58	92.29	98.00	98.00	71.11	71.11
14	35	[grau, força]	90.00	90.67	98.89	97.89	70.56	67.22
14	45	[grau, força, clustering]	92.71	93.54	95.89	96.89	73.89	69.44
18	20	[grau]	87.23	90.58	96.89	97.89	75.00	66.11
18	45	[grau, força, clustering]	92.93	93.45	97.89	96.89	71.11	66.11
20	10	[clustering]	86.62	91.13	96.89	98.00	70.00	63.33
20	60	[grau, força]	90.67	91.40	98.89	97.89	70.00	71.67
24	70	[força]	85.34	86.80	97.89	97.89	65.56	62.22
K	N	Descritores $\phi$	Otolith		Portuguese Leaves		USP Leaves	
			LDA	SVM	LDA	SVM	LDA	SVM
4	10	[grau]	90.81	92.55	69.70	76.20	80.50	84.67
4	70	[grau, força, clustering]	94.04	92.55	74.97	76.74	85.83	88.00
8	20	[força]	92.06	92.29	71.17	71.50	80.17	84.17
10	35	[força, clustering]	96.02	95.03	75.59	74.71	8.65	87.67
10	60	[clustering]	96.53	96.02	76.16	76.41	85.83	87.50
14	35	[grau, força]	93.79	90.55	79.67	76.45	86.17	88.83
14	45	[grau, força, clustering]	95.04	94.54	75.30	79.69	85.00	88.50
18	20	[grau]	93.30	93.54	76.79	79.15	84.50	87.33
18	45	[grau, força, clustering]	94.79	94.04	76.45	79.40	87.17	88.17
20	10	[clustering]	94.53	95.28	74.44	72.97	85.33	89.00
20	60	[grau, força]	95.29	92.79	79.66	77.35	87.17	88.17
24	70	[força]	94.79	94.29	74.06	74.70	82.83	85.00

### 5.1.1 Genéricas

Nesta primeira categoria de contornos, estão a base de dados ETH-80 e Kimia. A Figura 16 contém mapas de calor com as acurácias das bases desta categoria, variando de acordo com os diferentes valores de  $K$  e  $N$ . Pode-se observar que, embora a base Kimia possua bons resultados com valores menores de  $K$  e  $N$ , ambas bases possuem boa performance com valores maiores de  $K$  e  $N$ , que geram um Fisher Vector com maior dimensão. Portanto, nesta etapa, foram escolhidos os conjuntos  $(N, K) = \{(60, 24), (70, 24), (70, 20)\}$ .

Agora, investiga-se os efeitos das diferentes combinações  $\phi$  de descritores (grau, força e coeficiente de clustering) na abordagem proposta, considerando os conjuntos de parâmetros  $K$  e  $N$  escolhidos. A Tabela 2 demonstra o comportamento do método nas diferentes combinações de parâmetros, também está presente o número de características obtidas na saída do Fisher Vector para cada imagem.

Para que o desempenho fosse o melhor possível tanto para a base Kimia quanto para a base ETH-80, foi escolhida a combinação de parâmetros  $(N = 70, K = 24, \phi = [s, c])$ . Com

Figura 16 – Mapas de calor das acurácias (%) nas bases de dados (a) Kimia e (b) ETH-80 para diferentes valores de  $K$  e  $N$ , com descritores  $\phi = [g, s, c]$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada.

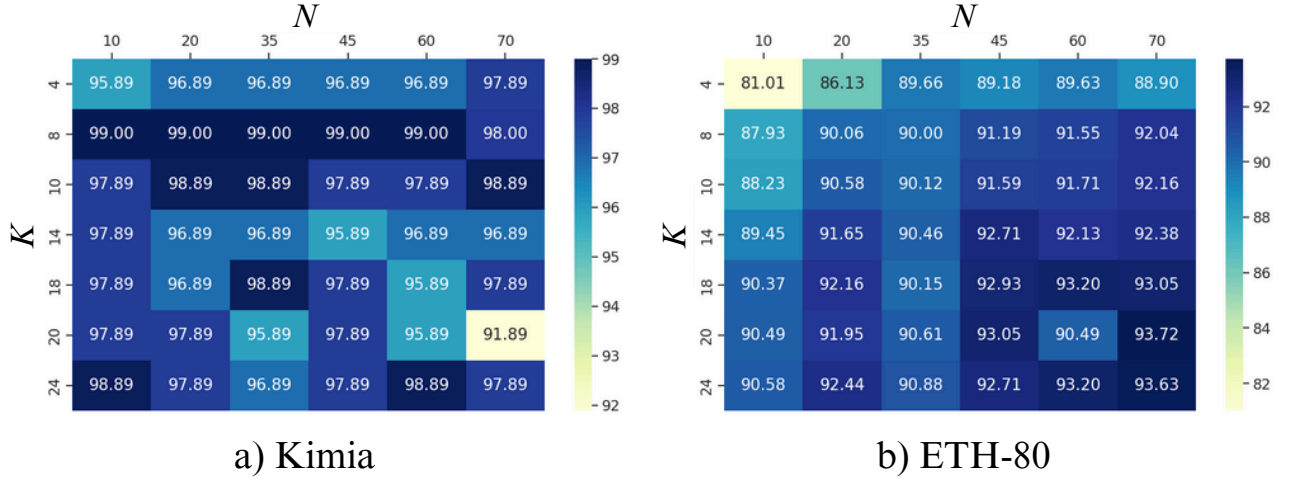


Tabela 2 – Número de features e acurácias (%) nas bases de dados Kimia e ETH-80 para diferentes combinações  $\phi$  de descritores grau, força e clustering, nos conjuntos os conjuntos  $(N, K) = \{(60, 24), (70, 24), (70, 20)\}$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada (Destaque na combinação escolhida).

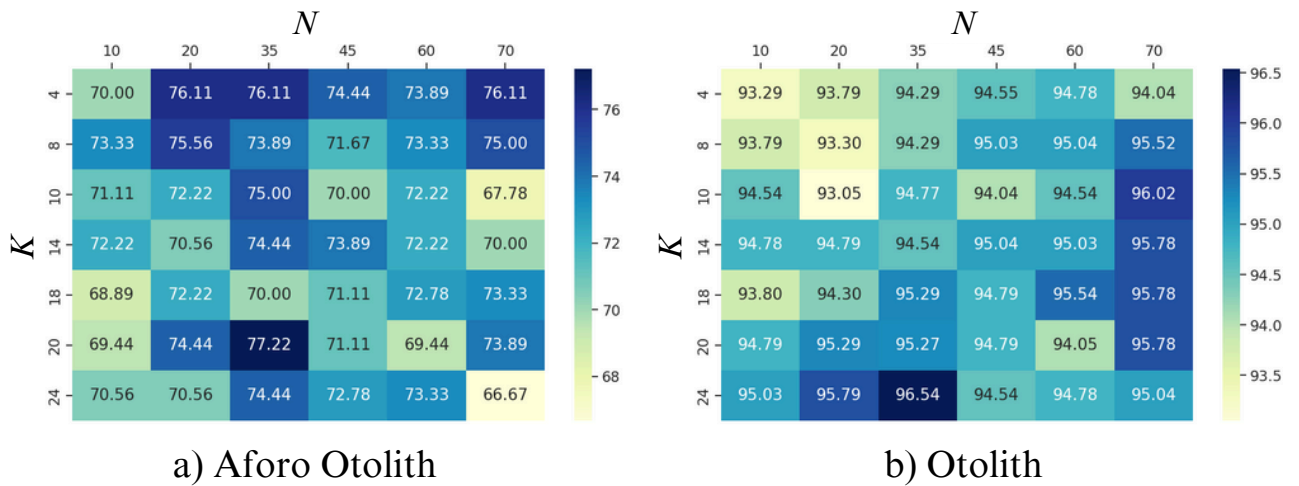
Descritores $\phi$	N	K	Nº de features	ETH-80	Kimia
[grau]	60	24	2856	89.60 ( $\pm 1.80$ )	98.89 ( $\pm 3.33$ )
[força]	60	24	2856	85.40 ( $\pm 2.32$ )	97.89 ( $\pm 4.23$ )
[clustering]	60	24	2856	91.86 ( $\pm 1.19$ )	97.00 ( $\pm 6.40$ )
[grau, força]	60	24	5688	90.40 ( $\pm 1.76$ )	98.89 ( $\pm 3.33$ )
[grau, clustering]	60	24	5688	92.90 ( $\pm 0.97$ )	97.89 ( $\pm 4.23$ )
[força, clustering]	60	24	5688	92.74 ( $\pm 0.95$ )	97.89 ( $\pm 4.23$ )
[grau, força, clustering]	60	24	8520	93.20 ( $\pm 0.93$ )	98.89 ( $\pm 3.33$ )
[grau]	70	24	3336	90.67 ( $\pm 1.19$ )	98.89 ( $\pm 3.33$ )
[força]	70	24	3336	85.34 ( $\pm 2.48$ )	97.89 ( $\pm 4.23$ )
[clustering]	70	24	3336	91.62 ( $\pm 0.95$ )	95.89 ( $\pm 6.74$ )
[grau, força]	70	24	6648	91.25 ( $\pm 1.22$ )	98.89 ( $\pm 3.33$ )
[grau, clustering]	70	24	6648	93.32 ( $\pm 1.05$ )	97.89 ( $\pm 4.23$ )
[força, clustering]	70	24	6648	92.74 ( $\pm 0.87$ )	97.00 ( $\pm 4.58$ )
[grau, força, clustering]	70	24	9960	93.63 ( $\pm 0.97$ )	97.89 ( $\pm 4.23$ )
[grau]	70	20	2780	89.88 ( $\pm 1.58$ )	98.89 ( $\pm 3.33$ )
[força]	70	20	2780	85.43 ( $\pm 2.08$ )	96.89 ( $\pm 4.76$ )
[clustering]	70	20	2780	91.86 ( $\pm 1.51$ )	97.00 ( $\pm 6.40$ )
[grau, força]	70	20	5540	91.49 ( $\pm 0.94$ )	97.89 ( $\pm 4.23$ )
[grau, clustering]	70	20	5540	92.96 ( $\pm 0.97$ )	97.89 ( $\pm 4.23$ )
<b>[força, clustering]</b>	<b>70</b>	<b>20</b>	<b>5540</b>	<b>93.08 (<math>\pm 0.64</math>)</b>	<b>99.00 (<math>\pm 3.00</math>)</b>
[grau, força, clustering]	70	20	8300	93.72 ( $\pm 1.03$ )	91.89 ( $\pm 7.52$ )

esses parâmetros obtém-se uma acurácia de 93.08 ( $\pm 0.64$ ) e 99.00 ( $\pm 3.00$ ) nas respectivas bases, demonstrando a robustez do método na categoria de imagens de contornos genéricas, já que com o mesmo parâmetro é possível obter bons resultados.

### 5.1.2 Otólitos

A segunda categoria de contornos refere-se aos otólitos, que incluem as bases de dados Aforo Otolith e Otolith. Na Figura 17, são apresentados mapas de calor que mostram as acurácias dessas bases, variando conforme os diferentes valores de  $K$  e  $N$ . É possível notar que ambas as bases demonstram um desempenho satisfatório com valores maiores e intermediários de  $N$  e  $K$ , embora a base Aforo Otolith obtenha bons resultados com valores menores de  $K$ . Nesta etapa, portanto, foram selecionados os conjuntos  $(N, K) = \{(70, 8), (35, 24), (35, 14)\}$  para prosseguir na análise.

Figura 17 – Mapas de calor das acurácias (%) nas bases de dados (a) Aforo Otolith e (b) Otolith para diferentes valores de  $K$  e  $N$ , com descritores  $\phi = [g, s, c]$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada.



Com os conjuntos  $(N, K)$  definidos, passa-se a analisar os impactos das diferentes combinações  $\phi$  de descritores (grau, força e coeficiente de clustering) na abordagem proposta. A Tabela 3 ilustra o desempenho do método em relação às diversas combinações de parâmetros, incluindo o número de características geradas na saída do Fisher Vector para cada imagem.

Para otimizar o desempenho nas bases Aforo Otolith e Otolith, foi escolhida a combinação de parâmetros  $(N = 35, K = 14, \phi = [c])$ . Com essa configuração, a acurácia alcançada foi de 75.56 ( $\pm 8.31$ ) e 96.03 ( $\pm 3.37$ ) para as respectivas bases, evidenciando a robustez da abordagem na categoria de otólitos, uma vez que bons resultados foram obtidos com os mesmos parâmetros.

### 5.1.3 Folhas

A terceira categoria de contornos é composta pelas bases de dados de contornos de folhas, sendo elas, Portuguese Leaves e USP Leaves. Na Figura 18, são apresentados mapas de calor que demonstram as acurácias obtidas por essas bases, variando de acordo com diferentes valores de  $K$  e  $N$ . Observa-se que ambas as bases apresentam um desempenho robusto com

Tabela 3 – Número de features e acurácias (%) nas bases de dados Aforo Otolith e Otolith para diferentes combinações  $\phi$  de descritores grau, força e clustering, nos conjuntos os conjuntos  $(N, K) = \{(70, 8), (35, 24), (35, 14)\}$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada (Destaque na combinação escolhida).

Descritores $\phi$	N	K	Nº de features	Aforo otolith	Otolith
[grau]	70	8	1112	75.56 ( $\pm$ 6.85)	94.27 ( $\pm$ 2.99)
[força]	70	8	1112	63.33 ( $\pm$ 4.71)	93.05 ( $\pm$ 2.91)
[clustering]	70	8	1112	72.78 ( $\pm$ 6.71)	95.77 ( $\pm$ 3.35)
[grau, força]	70	8	2216	72.22 ( $\pm$ 7.11)	94.30 ( $\pm$ 3.16)
[grau, clustering]	70	8	2216	75.56 ( $\pm$ 6.85)	96.27 ( $\pm$ 3.38)
[força, clustering]	70	8	2216	73.33 ( $\pm$ 5.27)	96.52 ( $\pm$ 2.78)
[grau, força, clustering]	70	8	3320	75.00 ( $\pm$ 7.82)	95.52 ( $\pm$ 2.45)
[grau]	35	24	1656	73.33 ( $\pm$ 6.67)	95.54 ( $\pm$ 2.43)
[força]	35	24	1656	63.33 ( $\pm$ 6.67)	94.54 ( $\pm$ 2.66)
[clustering]	35	24	1656	71.11 ( $\pm$ 6.57)	96.76 ( $\pm$ 2.50)
[grau, força]	35	24	3288	72.22 ( $\pm$ 8.53)	94.05 ( $\pm$ 2.52)
[grau, clustering]	35	24	3288	73.89 ( $\pm$ 8.75)	96.04 ( $\pm$ 3.69)
[força, clustering]	35	24	3288	69.44 ( $\pm$ 7.24)	95.77 ( $\pm$ 3.72)
[grau, força, clustering]	35	24	4920	74.44 ( $\pm$ 7.24)	96.54 ( $\pm$ 2.53)
[grau]	35	14	966	72.22 ( $\pm$ 7.49)	92.80 ( $\pm$ 1.73)
[força]	35	14	966	63.33 ( $\pm$ 6.24)	94.05 ( $\pm$ 3.36)
<b>[clustering]</b>	<b>35</b>	<b>14</b>	<b>966</b>	<b>75.56</b> ( $\pm$ 8.31)	<b>96.03</b> ( $\pm$ 3.37)
[grau, força]	35	14	1918	70.56 ( $\pm$ 11.17)	93.79 ( $\pm$ 2.02)
[grau, clustering]	35	14	1918	71.67 ( $\pm$ 6.24)	94.54 ( $\pm$ 2.42)
[força, clustering]	35	14	1918	73.89 ( $\pm$ 9.36)	96.51 ( $\pm$ 3.57)
[grau, força, clustering]	35	14	2870	74.44 ( $\pm$ 6.85)	94.54 ( $\pm$ 3.49)

valores elevados de  $N$  e  $K$ , e a base Portuguese Leaves, em particular, também demonstra resultados satisfatórios com valores reduzidos de  $K$ . Dessa forma, para avançar na análise, foram selecionados os conjuntos  $(N, K) = \{(70, 24), (70, 20), (60, 20)\}$ .

Com os conjuntos  $(N, K)$  estabelecidos, inicia-se a investigação dos efeitos das diferentes combinações  $\phi$  de descritores (grau, força e coeficiente de clustering) na metodologia proposta. A Tabela 4 apresenta o desempenho do método em relação às diversas combinações de parâmetros, incluindo o número de características geradas na saída do Fisher Vector para cada imagem.

Visando o melhor desempenho em ambas as bases, foi adotada a combinação de parâmetros  $(N = 60, K = 20, \phi = [g, s, c])$ . Com essa configuração, as acurácias obtidas foram de 81.16 ( $\pm$  2.48) e 88.33 ( $\pm$  4.89) para as respectivas bases, demonstrando a eficácia da abordagem na categoria de folhas, já que bons resultados foram alcançados utilizando os mesmos parâmetros.

Figura 18 – Mapas de calor das acurácias (%) nas bases de dados (a) Portuguese Leaves e (b) USP Leaves para diferentes valores de  $K$  e  $N$ , com descritores  $\phi = [g, s, c]$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada.

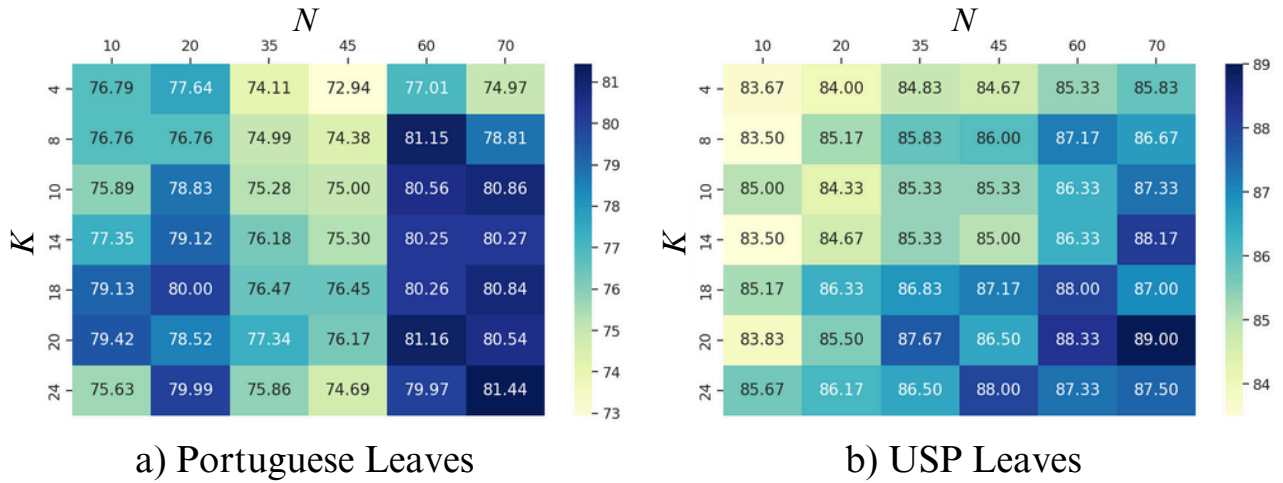


Tabela 4 – Número de features e acurácias (%) nas bases de dados Portuguese Leaves e USP Leaves para diferentes combinações  $\phi$  de descritores grau, força e clustering, nos conjuntos os conjuntos  $(N, K) = \{(70, 24), (70, 20), (60, 20)\}$ , obtidas com o classificador LDA e validação cruzada 10-fold estratificada (Destaque na combinação escolhida).

Descritores $\phi$	N	K	Nº de features	Portuguese Leaves	USP Leaves
[grau]	70	24	3336	79.67 ( $\pm 4.69$ )	86.83 ( $\pm 5.08$ )
[força]	70	24	3336	74.06 ( $\pm 6.00$ )	82.83 ( $\pm 5.27$ )
[clustering]	70	24	3336	75.85 ( $\pm 4.54$ )	87.83 ( $\pm 3.42$ )
[grau, força]	70	24	6648	80.84 ( $\pm 4.63$ )	87.67 ( $\pm 5.73$ )
[grau, clustering]	70	24	6648	79.96 ( $\pm 4.55$ )	89.00 ( $\pm 4.16$ )
[força, clustering]	70	24	6648	80.56 ( $\pm 5.04$ )	88.17 ( $\pm 4.08$ )
[grau, força, clustering]	70	24	9960	81.44 ( $\pm 3.77$ )	87.50 ( $\pm 3.44$ )
[grau]	70	20	2780	79.08 ( $\pm 4.66$ )	86.33 ( $\pm 5.42$ )
[força]	70	20	2780	75.83 ( $\pm 6.15$ )	82.50 ( $\pm 5.88$ )
[clustering]	70	20	2780	77.90 ( $\pm 5.91$ )	86.33 ( $\pm 3.32$ )
[grau, força]	70	20	5540	79.65 ( $\pm 5.84$ )	87.50 ( $\pm 5.34$ )
[grau, clustering]	70	20	5540	82.34 ( $\pm 3.80$ )	87.67 ( $\pm 3.89$ )
[força, clustering]	70	20	5540	79.40 ( $\pm 4.09$ )	87.33 ( $\pm 3.96$ )
[grau, força, clustering]	70	20	8300	80.54 ( $\pm 4.68$ )	89.00 ( $\pm 4.73$ )
[grau]	60	20	2380	79.08 ( $\pm 4.62$ )	85.17 ( $\pm 6.47$ )
[força]	60	20	2380	75.54 ( $\pm 5.83$ )	83.00 ( $\pm 4.93$ )
[clustering]	60	20	2380	77.64 ( $\pm 3.54$ )	86.50 ( $\pm 3.83$ )
[grau, força]	60	20	4740	79.66 ( $\pm 4.29$ )	87.17 ( $\pm 5.53$ )
[grau, clustering]	60	20	4740	82.62 ( $\pm 4.20$ )	87.17 ( $\pm 4.48$ )
[força, clustering]	60	20	4740	79.37 ( $\pm 4.25$ )	88.50 ( $\pm 4.89$ )
<b>[grau, força, clustering]</b>	<b>60</b>	<b>20</b>	<b>7100</b>	<b>81.16 (<math>\pm 2.48</math>)</b>	<b>88.33 (<math>\pm 4.89</math>)</b>



## 5.2 Discussão dos resultados

Esta seção tem o propósito de analisar o desempenho do método proposto em comparação com as abordagens descritas na Seção 3.5. A seguir, são apresentados e discutidos os resultados obtidos para as seis bases de dados de contornos apresentadas na Seção 3.1, bem como a análise de invariância a rotação e escala e robustez a ruído.

Tabela 5 – Comparação de resultados para todas as bases de dados e métodos da literatura utilizando os classificadores LDA e SVM.

Métodos	LDA	SVM	LDA	SVM	LDA	SVM
	Aforo Otolith		Otolith		Portuguese Leaves	
<b>Método proposto</b>	<b>75.56</b> ( $\pm 8.31$ )	<b>72.22</b> ( $\pm 5.33$ )	<b>96.03</b> ( $\pm 3.37$ )	<b>96.77</b> ( $\pm 2.74$ )	81.16 ( $\pm 2.48$ )	<b>79.40</b> ( $\pm 5.92$ )
DTN	75.00 ( $\pm 12.0$ )	66.23 ( $\pm 1.22$ )	94.90 ( $\pm 4.25$ )	93.55 ( $\pm 2.87$ )	85.00 ( $\pm 4.26$ )	77.79 ( $\pm 0.788$ )
C.N. degree	63.89 ( $\pm 7.52$ )	48.08 ( $\pm 1.07$ )	91.08 ( $\pm 4.80$ )	86.33 ( $\pm 5.76$ )	72.35 ( $\pm 7.99$ )	66.08 ( $\pm 0.85$ )
Fourier	52.77 ( $\pm 10.5$ )	50.35 ( $\pm 0.72$ )	82.63 ( $\pm 6.72$ )	85.43 ( $\pm 7.61$ )	53.53 ( $\pm 9.88$ )	55.33 ( $\pm 0.90$ )
Curvature	55.56 ( $\pm 15.0$ )	49.93 ( $\pm 1.31$ )	90.06 ( $\pm 2.40$ )	89.86 ( $\pm 5.51$ )	64.41 ( $\pm 7.52$ )	65.75 ( $\pm 1.07$ )
Zernike	21.67 ( $\pm 4.10$ )	20.04 ( $\pm 1.44$ )	56.09 ( $\pm 8.26$ )	58.13 ( $\pm 8.37$ )	40.00 ( $\pm 10.4$ )	34.85 ( $\pm 0.90$ )
M. S. fractal	53.33 ( $\pm 11.1$ )	46.89 ( $\pm 0.91$ )	53.33 ( $\pm 11.1$ )	46.89 ( $\pm 0.91$ )	63.82 ( $\pm 9.09$ )	53.52 ( $\pm 0.51$ )
Segment analysis	57.78 ( $\pm 8.76$ )	37.78 ( $\pm 0.94$ )	57.78 ( $\pm 8.76$ )	37.78 ( $\pm 0.94$ )	72.35 ( $\pm 5.58$ )	63.47 ( $\pm 0.77$ )
ADCN	65.55 ( $\pm 9.37$ )	36.32 ( $\pm 0.71$ )	65.55 ( $\pm 9.37$ )	36.32 ( $\pm 0.71$ )	11.52 ( $\pm 0.62$ )	17.64 ( $\pm 7.33$ )
	ETH-80		Kimia		USP Leaves	
<b>Método proposto</b>	93.08 ( $\pm 0.64$ )	<b>93.63</b> ( $\pm 1.04$ )	99.00 ( $\pm 3.00$ )	98.00 ( $\pm 4.00$ )	88.33 ( $\pm 4.89$ )	90.33 ( $\pm 3.14$ )
CRSA	93.69	92.32	100.00	97.98	94.67	92.00
DTN	92.02 ( $\pm 1.68$ )	93.47 ( $\pm 1.10$ )	99.00 ( $\pm 3.16$ )	98.00 ( $\pm 4.03$ )	94.16 ( $\pm 4.24$ )	93.16 ( $\pm 2.41$ )
C.N. degree	72.34 ( $\pm 2.34$ )	82.92 ( $\pm 2.58$ )	96.00 ( $\pm 6.99$ )	95.95 ( $\pm 5.16$ )	84.00 ( $\pm 5.67$ )	85.16 ( $\pm 3.63$ )
Fourier	79.26 ( $\pm 1.56$ )	86.12 ( $\pm 1.89$ )	93.88 ( $\pm 7.07$ )	97.97 ( $\pm 4.21$ )	74.66 ( $\pm 6.97$ )	83.16 ( $\pm 4.61$ )
Curvature	62.43 ( $\pm 2.03$ )	72.01 ( $\pm 2.61$ )	78.77 ( $\pm 15.2$ )	84.84 ( $\pm 12.5$ )	77.00 ( $\pm 6.42$ )	81.66 ( $\pm 3.33$ )
Zernike	82.86 ( $\pm 2.01$ )	87.98 ( $\pm 1.34$ )	95.00 ( $\pm 5.27$ )	87.98 ( $\pm 1.34$ )	69.66 ( $\pm 5.76$ )	76.33 ( $\pm 5.07$ )
M. S. fractal	73.96 ( $\pm 1.65$ )	76.73 ( $\pm 1.71$ )	95.00 ( $\pm 9.71$ )	76.73 ( $\pm 1.71$ )	71.16 ( $\pm 4.51$ )	73.66 ( $\pm 5.81$ )
Segment analysis	78.71 ( $\pm 2.53$ )	80.00 ( $\pm 2.46$ )	98.00 ( $\pm 4.21$ )	98.80 ( $\pm 3.16$ )	83.33 ( $\pm 5.44$ )	83.50 ( $\pm 4.26$ )
ADCN	85.39 ( $\pm 2.64$ )	74.35 ( $\pm 1.72$ )	99.00 ( $\pm 3.16$ )	85.78 ( $\pm 9.36$ )	88.50 ( $\pm 6.63$ )	78.83 ( $\pm 3.68$ )
	USP Leaves Rotacionada		USP Leaves Dimensionada			
<b>Método proposto</b>	87.78	91.83	88.08	91.54		
CRSA	93.78	93.39	93.71	93.25		
DTN	89.06	88.67	92.38	93.08		
C.N. degree	80.61	84.28	81.46	85.04		
Fourier	67.53	77.17	63.25	85.29		
Curvature	76.31	83.33	76.92	84.46		
Zernike	65.53	76.64	53.38	71.33		
M. S. fractal	58.17	68.58	65.25	68.50		
Segment analysis	81.47	76.22	82.04	79.75		
ADCN	86.36	77.11	86.13	80.75		

### 5.2.1 Genéricas

Devido a sua simplicidade e tamanho reduzido em relação às outras bases, a base de dados Kimia apresenta um desempenho elevado em todos os métodos comparados, assim como a Tabela 5 demonstra. Entre os métodos comparados, o *CRSA* atingiu a maior acurácia com o classificador LDA, alcançando os 100%. Já no classificador SVM, o algoritmo de *Segment Analysis* alcançou a maior acurácia, com 98,80%. O método proposto obteve a segunda melhor acurácia, tanto no LDA, com 99%, quanto no SVM, com 98%.

A Tabela 5 também mostra a acurácia para o reconhecimento na base ETH-80, considerando todos os métodos testados, nela pode-se notar que o método proposto obteve

uma das maiores acurácias, sendo a segunda maior no classificador LDA, com 93,08%, e a maior no classificador SVM, com 93,63%.

Dessa forma, observa-se que o método proposto alcançou resultados robustos, consistentes e satisfatórios para a categoria de imagens genéricas, se mostrando competitivo entre os métodos presentes na literatura.

### 5.2.2 Otólitos

A base Aforo Otolith não alcança altas acurácias nos métodos comparados, o que demonstra que seu reconhecimento é uma tarefa desafiadora. Porém, ainda é possível melhorar o desempenho, pois, assim como a Tabela 5 revela, o método proposto alcança a maior acurácia em ambos classificadores, com 75,56% no LDA e 72,22% no SVM, superando o método de inspiração, *C.N. degree*, e seu derivado, *DTN*, que alcançou 75% no LDA e 66,23% no SVM.

O método proposto alcançou as maiores acurácias, em ambos classificadores, na base de dados Otolith, com 96,03% no LDA e 96,77% no SVM, como é possível observar na Tabela 5. O método inspirado, *C.N. degree*, obteve a terceira melhor taxa de classificação, com 91,08% ao utilizar o classificador LDA. Já o método *DTN* obteve a segunda melhor acurácia em ambos classificadores, com 94,90% no LDA e 93,55% no SVM.

Portanto, o método proposto alcança resultados extremamente relevantes para a categoria de imagens de otólitos, já que se mostrou capaz de obter as melhores acurácias entre os métodos comparados.

### 5.2.3 Folhas

Na base Portuguese Leaves, o método proposto alcançou 79,40% de acurácia ao utilizar o classificador SVM, mas obteve maior desvio-padrão do que a maioria dos demais métodos, o que o colocaria como o segundo melhor método, com desempenho similar ao método *DTN*. O *DTN* permaneceu com a melhor acurácia, de 85% com o classificador SVM, enquanto a abordagem proposta alcançou a segunda melhor, com 81,16%.

Já a base USP Leaves apresenta características relevantes para a avaliação de métodos de análise de contornos, sendo composta por folhas reais da flora brasileira, abrangendo 30 espécies de plantas (BACKES; CASANOVA; BRUNO, 2009). É uma base de dados bastante desafiadora devido às semelhanças entre folhas de diferentes espécies e à grande variabilidade entre as folhas de uma mesma espécie. O método proposto atinge a quarta melhor acurácia usando o classificador LDA, com 88,33% e a terceira melhor com o classificador SVM, alcançando 90,33%.

O desempenho do método proposto na categoria de imagens de folhas evidencia a complexidade do reconhecimento nessas bases. No entanto, a abordagem mostrou-se compe-

titiva, com resultados satisfatórios, sendo capaz de realizar o reconhecimento de padrões de forma consistente.

#### 5.2.4 Invariância a escala e rotação

A fim de validar as propriedades de invariância à rotação e escala, realizamos experimentos em bases de dados modificadas a partir da USP Leaves, assim como foi apresentado na Seção 3.1. Os resultados obtidos com imagens rotacionadas e dimensionadas demonstram a forte invariância do método proposto a essas transformações. Para validação, foram utilizadas amostras modificadas (rotacionadas ou dimensionadas) do mesmo contorno original para testes, enquanto os demais contornos originais foram utilizados no treinamento, evitando assim que informações sobre as amostras de teste fossem incorporadas na etapa de treinamento.

O método proposto obteve resultados satisfatórios em ambas bases, como demonstrado na Tabela 5, com acurácias similares às alcançadas na USP Leaves original. Além disso, foram alcançadas acurácias semelhantes para ambos os classificadores, o que reflete a insensibilidade aos algoritmos de aprendizado de máquina, propriedade esta que outros métodos não apresentaram. Os resultados desses experimentos auxiliam na comprovação de que o vetor de características obtido com a abordagem proposta é consistente, possuindo invariância a escala e rotação.

#### 5.2.5 Robustez a ruído

Assim como descrito na Seção 3.1, o método proposto foi aplicado em uma versão modificada da base USP Leaves, com adição de ruído. Os resultados do experimento, utilizando o classificador LDA, são apresentados na Tabela 6. O método proposto demonstrou grande capacidade de classificação mesmo em contornos com alta quantidade de ruído, já que a acurácia obtida no decorrer dos níveis sofre pouca alteração, revelando a consistência e robustez do vetor de características proveniente da abordagem proposta.

Tabela 6 – Comparação da acurácia da abordagem proposta com outros métodos na base de dados USP Leaves submetida a diferentes níveis de ruído. Acurácias alcançadas usando o classificador LDA.

Métodos	Nível 1	Nível 2	Nível 3	Nível 4	Todos níveis
CN Degree	79.67	77.50	78.17	75.33	77.58
Fourier	65.50	62.50	59.33	55.67	58.62
Curvature	72.00	69.33	64.33	61.67	68.71
Zernike	67.33	65.17	66.33	62.00	61.17
M.S Fractal	63.33	63.83	62.17	59.17	63.50
ADCN	83.33	81.50	80.17	79.67	81.12
Segment Analysis	75.17	71.83	66.17	69.67	69.58
DTN	87.33	79.50	75.17	72.50	86.54
CRSA	89.66	84.33	83.67	81.50	91.37
<b>Método proposto</b>	85.00	83.50	83.33	83.00	84.46

### 5.2.6 Explicabilidade do vetor e custo computacional

Como apresentado na Seção 4.4, o tamanho do vetor de características sempre terá a dimensão  $K(2D + 1)$ , representando os gradientes dos descritores dos vértices em relação a todos os parâmetros do GMM, que constitui o vocabulário visual. Essa propriedade é interessante, pois permite compreender a composição do vetor de características empregado.

Observou-se que, em relação ao tamanho do vetor, certos contornos exigiram vetores maiores, enquanto outros demandaram vetores menores. Assim, o tamanho do vetor, influenciado pelos parâmetros  $N$ ,  $K$  e  $\phi$ , demonstrou-se relacionado à aplicação; ou seja, diferentes tipos de contornos comportam-se distintamente ao serem submetidos à abordagem proposta.

Em relação ao custo computacional, verificou-se que, quanto maiores os valores dos parâmetros, mais custoso se torna o processo, especialmente no estágio de obtenção do vocabulário. No entanto, é importante destacar que, uma vez obtido o vocabulário, o tempo para a predição de uma única amostra é competitivo.

### 5.2.7 Desempenho geral

A análise dos resultados, conforme apresentado na Tabela 5, revela que o método proposto se destaca em relação aos demais abordagens, especialmente nas bases Aforo Otolith, Otolith, Portuguese Leaves e ETH-80, onde obtém desempenhos notáveis. Com isso, o método se mostra relevante para o reconhecimento de padrões nas três categorias criadas para análise: genéricas, otólitos e folhas. Além disso, a utilização de redes complexas e Fisher Vector se mostra eficaz para lidar com a complexidade desses problemas, permitindo que o método proposto supere diversos desafios como a rotação, escala e ruído (Tabela 6).

Ademais, é importante ressaltar que o método se mostrou consistente em todas as bases de dados, o que demonstra sua robustez e versatilidade, sendo capaz de se adaptar a diferentes características dos dados. Em outra perspectiva, essa consistência também se mantém nos dois tipos de classificadores utilizados, LDA e SVM, o que não é algo comum para a maioria dos métodos, que apresentam alta sensibilidade aos classificadores. Esses resultados confirmam que o vetor de características gerado pelo método é eficaz.

## 6 Conclusão

Este trabalho propôs uma nova abordagem para a caracterização de redes complexas por meio da codificação Fisher Vector, aplicada à análise de contornos. A integração desses dois métodos mostrou-se altamente eficaz para superar a maioria dos desafios clássicos de reconhecimento de padrões. A aplicação dessa abordagem em diferentes bases de dados, como Aforo Otolith, Otolith, Portuguese Leaves e ETH-80, revelou um desempenho notável em comparação com métodos tradicionais, comprovando a robustez e versatilidade da proposta.

### 6.1 Contribuição científica

A análise dos resultados indica que o método se destacou pela alta acurácia em diferentes contextos, mostrando-se adaptável às características específicas de cada base de dados. Em particular, o método proposto alcançou resultados consistentes em situações desafiadoras, como na presença de ruído e variações nas imagens, o que evidencia sua aplicabilidade em cenários do mundo real. A invariância do método proposto a transformações como rotação e escala, bem como sua consistência em relação aos classificadores utilizados (LDA e SVM), reforça a eficácia do vetor de características gerado. Essa robustez é um diferencial importante, já que muitos métodos tradicionais apresentam alta sensibilidade a variações nos dados e nos algoritmos de aprendizado de máquina.

Ressalta-se que um artigo descrevendo a metodologia e resultados está em fase de redação e será submetido para um journal internacional, visto que a combinação de redes complexas com Fisher Vector apresenta uma contribuição importante para a área de análise de dados e reconhecimento de padrões, especialmente em visão computacional.

### 6.2 Trabalhos futuros

Os resultados obtidos neste trabalho abrem novas possibilidades para pesquisas futuras, focadas no aprimoramento das técnicas de classificação e análise de dados, com o objetivo de aumentar a precisão e eficiência em tarefas de reconhecimento de padrões mais desafiadoras, aumentando também a complexidade dos parâmetros empregados nessa abordagem.

É importante ressaltar que o método proposto pode ser aplicado em diversas áreas onde o objeto de estudo possa ser representado como uma RC. Exemplos incluem a análise de redes sociais, redes biológicas, redes de transporte, redes de comunicação e sistemas dinâmicos complexos. A flexibilidade da modelagem por RC permite adaptar o método a diferentes domínios, explorando interações complexas e padrões em dados de natureza diversa.

Assim, conclui-se que a proposta de utilizar a codificação Fisher Vector para a caracterização de redes complexas é não apenas viável, mas também promissora, apresentando resultados que reforçam a relevância desse método em um cenário contemporâneo de crescente complexidade nas análises de dados. Futuros trabalhos podem expandir essa linha de pesquisa, explorando novas aplicações e refinando as técnicas para maximizar ainda mais o desempenho em tarefas de classificação e reconhecimento de padrões.

# Referências

- AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, v. 19, p. 716–723, 1974.
- ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, American Physical Society, v. 74, p. 47–97, Jan 2002.
- AMARI, S.-i.; NAGAOKA, H. *Methods of Information Geometry*. Providence, RI: American Mathematical Society, 2000. v. 191. (Translations of Mathematical Monographs, v. 191). ISBN 9780821843024.
- ASLAN, M. et al. Shape feature encoding via fisher vector for efficient fall detection in depth-videos. *Applied Soft Computing*, v. 37, p. 1023–1028, 2015. ISSN 1568-4946.
- BACH, S. et al. Analyzing classifiers: Fisher vectors and deep neural networks. *arXiv preprint arXiv:1512.00172*, 2015.
- BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. A complex network-based approach for boundary shape analysis. *Pattern Recognition*, v. 42, n. 1, p. 54–67, 2009. ISSN 0031-3203.
- BALAKRISHNAMA, S.; GANAPATHIRAJU, A. Linear discriminant analysis-a brief tutorial. *Institute for Signal and information Processing*, Mississippi, v. 18, n. 1998, p. 1–8, 1998.
- BARABASI, A.-L.; OLTVAI, Z. Network biology: Understanding the cell's functional organization. *Nature reviews. Genetics*, v. 5, p. 101–13, 03 2004.
- BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *Science*, v. 286, n. 5439, p. 509–512, 1999.
- BARABÁSI, A.-L.; PÓSFAL, M. *Network science*. Cambridge: Cambridge University Press, 2016. ISBN 9781107076266 1107076269.
- BAYES, T. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions*, v. 53, p. 370–418, 1763.
- BROPHY, J. K. et al. Classification of bovidae fossils from gladysvale, south africa using elastic shape analysis. *Journal of Archaeological Science*, v. 166, p. 105959, 2024. ISSN 0305-4403.
- BURGES, C. J. C. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, v. 2, p. 121–167, 1998.
- CHANDRASEKHAR, V. et al. A practical guide to cnns and fisher vectors for image instance retrieval. *Signal Processing*, v. 128, p. 426–439, 2016. ISSN 0165-1684.
- CHEN, L. et al. Identifying influential nodes in complex networks via transformer. *Information Processing Management*, v. 61, n. 5, p. 103775, 2024. ISSN 0306-4573.
- CORTES, C.; VAPNIK, V. N. Support-vector networks. *Machine Learning*, v. 20, p. 273–297, 1995.

- COSTA, L. d. F. et al. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. *Advances in Physics*, Informa UK Limited, v. 60, n. 3, p. 329–412, jun. 2011. ISSN 1460-6976.
- COSTA, L. d. F. et al. Characterization of complex networks: A survey of measurements. *Advances in Physics*, Informa UK Limited, v. 56, n. 1, p. 167–242, jan. 2007. ISSN 1460-6976.
- DEISENROTH, M.; FAISAL, A.; ONG, C. *Mathematics for Machine Learning*. [S.l.]: Cambridge University Press, 2020. ISBN 9781108470049.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B*, v. 39, p. 1–38, 1977.
- ERDOS, P.; RENYI, A. On the evolution of random graphs. *Publ. Math. Inst. Hungary. Acad. Sci.*, v. 5, p. 17–61, 1960.
- FISHER, R. A. The use of multiple measurements in taxonomic problems. *Annals of Human Genetics*, v. 7, p. 179–188, 1936.
- GARRIDO, A. A survey on complex networks. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, v. 2, n. 1, p. pp. 63–70, Jan. 2011.
- GERIG, G. et al. Shape versus size: Improved understanding of the morphology of brain structures. In: . [S.l.: s.n.], 2001. v. 2208, p. 24–32. ISBN 978-3-540-42697-4.
- GÓMEZ-SIRVENT, J. L. et al. Defect classification on semiconductor wafers using fisher vector and visual vocabularies coding. *Measurement*, v. 202, p. 111872, 2022. ISSN 0263-2241.
- HARRIS, C. R. et al. Array programming with NumPy. *Nature*, Springer Science and Business Media LLC, v. 585, n. 7825, p. 357–362, set. 2020.
- HUANG, C.-Y.; SUN, C.-T.; LIN, H.-C. Influence of local information on social simulations in small-world network models. *Journal of Artificial Societies and Social Simulation*, v. 8, 10 2005.
- HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing in science & engineering*, IEEE, v. 9, n. 3, p. 90–95, 2007.
- HM, C.; KARJIGI, V.; SREEDEVI, N. Speech intelligibility assessment of dysarthria using fisher vector encoding. *Computer Speech Language*, v. 77, p. 101411, 2023. ISSN 0885-2308.
- JAAKKOLA, T.; HAUSSLER, D. Exploiting generative models in discriminative classifiers. In: KEARNS, M.; SOLLA, S.; COHN, D. (Ed.). *Advances in Neural Information Processing Systems*. [S.l.]: MIT Press, 1998. v. 11.
- Joblib Development Team. *Joblib: running Python functions as pipeline jobs*. 2024. Disponível em: <<https://joblib.readthedocs.io/>>.
- JUNIOR, J. J. d. M. S.; BACKES, A. R. Shape classification using line segment statistics. *Information Sciences*, Elsevier, v. 305, p. 349–356, 2015.
- KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *International Joint Conference on Artificial Intelligence*. [S.l.: s.n.], 1995.



- KRAPAC, J.; VERBEEK, J. J.; JURIE, F. Modeling spatial layout with fisher vectors for image categorization. *2011 International Conference on Computer Vision*, p. 1487–1494, 2011.
- LEIBE, B.; SCHIELE, B. Analyzing appearance and contour based methods for object categorization. In: *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings*. [S.l.: s.n.], 2003. v. 2, p. II–409.
- LOMBARTE, A. et al. A web-based environment for shape analysis of fish otoliths. the aforo database. CSIC-Instituto de Ciencias del Mar (ICM), 2006.
- LONCARIC, S. A survey of shape analysis techniques. *Pattern Recognition*, v. 31, n. 8, p. 983–1001, 1998. ISSN 0031-3203.
- LÉVY Émile B. Sur la résolution numérique des systèmes d'équations linéaires. *Bulletin géodésique*, v. 2, n. 4, p. 67–77, 1924.
- MCKINNEY, W. et al. Data structures for statistical computing in python. In: AUSTIN, TX. *Proceedings of the 9th Python in Science Conference*. [S.l.], 2010. v. 445, p. 51–56.
- MIRANDA, G. H.; MACHICAO, J.; BRUNO, O. M. An optimized shape descriptor based on structural properties of networks. *Digital Signal Processing*, Elsevier, v. 82, p. 216–229, 2018.
- NEWMAN, M.; BARABÁSI, A.-L.; WATTS, D. J. *The Structure and Dynamics of Networks*. Princeton: Princeton University Press, 2006. ISBN 9781400841356.
- NEWMAN, M. E. J.; PARK, J. Why social networks are different from other types of networks. *Phys. Rev. E*, American Physical Society, v. 68, p. 036122, Sep 2003.
- OSOWSKI, S.; NGHIA, D. D. Fourier and wavelet descriptors for shape recognition using neural networks—a comparative study. *Pattern Recognition*, v. 35, n. 9, p. 1949–1957, 2002. ISSN 0031-3203.
- PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, v. 12, n. Oct, p. 2825–2830, 2011.
- PERRONNIN, F.; DANCE, C. Fisher kernels on visual vocabularies for image categorization. In: *2007 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2007. p. 1–8.
- PERRONNIN, F. et al. Large-scale image retrieval with compressed fisher vectors. In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2010. p. 3384–3391.
- PERRONNIN, F.; SÁNCHEZ, J.; MENSINK, T. Improving the fisher kernel for large-scale image classification. In: DANIILIDIS, K.; MARAGOS, P.; PARAGIOS, N. (Ed.). *Computer Vision – ECCV 2010*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. p. 143–156. ISBN 978-3-642-15561-1.
- PLOTZE, R. et al. Leaf shape analysis using the multiscale minkowski fractal dimension, a new morphometric method: A study with pass flora (passifloraceae). *Canadian Journal of Botany*, v. 83, p. 287–301, 03 2005.

- PURNOMO, V. et al. New approach for particle size and shape analysis of iron-based oxygen carriers at multiple oxidation states. *Particuology*, v. 90, p. 493–503, 2024. ISSN 1674-2001.
- RIBAS, L. C. *Aprendizado de representações e caracterização de redes complexas com aplicações em visão computacional*. Tese (Doutorado) — Universidade de São Paulo, 2021.
- RIBAS, L. C.; BRUNO, O. M. Learning a complex network representation for shape classification. *Pattern Recognition*, v. 154, p. 110566, 2024. ISSN 0031-3203.
- RIBAS, L. C.; NEIVA, M. B.; BRUNO, O. M. Distance transform network for shape analysis. *Information Sciences*, v. 470, p. 28–42, 2019. ISSN 0020-0255.
- RIBAS, L. C. et al. A complex network based approach for knee osteoarthritis detection: Data from the osteoarthritis initiative. *Biomedical Signal Processing and Control*, v. 71, p. 103133, 2022. ISSN 1746-8094.
- RIBAS, L. C. et al. Fusion of complex networks and randomized neural networks for texture analysis. *Pattern Recognition*, v. 103, p. 107189, 2020. ISSN 0031-3203.
- ROSSUM, G. V.; DRAKE, F. L. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009. ISBN 1441412697.
- SALIMI, N. et al. Fully-automated identification of fish species based on otolith contour: using short-time fourier transform and discriminant analysis (stft-da). 11 2015.
- SÁNCHEZ, J. et al. Image classification with the fisher vector: Theory and practice. *International journal of computer vision*, Springer, v. 105, p. 222–245, 2013.
- SCABINI, L. F. et al. Angular descriptors of complex networks: A novel approach for boundary shape analysis. *Expert Systems with Applications*, v. 89, p. 362–373, 2017. ISSN 0957-4174.
- SCABINI, L. F.; RIBAS, L. C.; BRUNO, O. M. Spatio-spectral networks for color-texture analysis. *Information Sciences*, v. 515, p. 64–79, 2020. ISSN 0020-0255.
- SCABINI, L. F. S.; GONÇALVES, W. N.; CASTRO, A. A. Texture analysis by bag-of-visual-words of complex networks. In: PARDO, A.; KITTLER, J. (Ed.). *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Cham: Springer International Publishing, 2015. p. 485–492. ISBN 978-3-319-25751-8.
- SCHWARZ, G. Estimating the Dimension of a Model. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 6, n. 2, p. 461 – 464, 1978.
- SEBASTIAN, T. B.; KLEIN, P. N.; KIMIA, B. B. Recognition of shapes by editing their shock graphs. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 26, n. 5, p. 550–571, 2004.
- SILVA, P. F.; MARCAL, A. R.; SILVA, R. M. A. da. Evaluation of features for leaf discrimination. In: SPRINGER. *International Conference Image Analysis and Recognition*. [S.l.], 2013. p. 197–204.
- TORRES, R. d. S.; FALCAO, A. X.; COSTA, L. d. F. A graph-based approach for multiscale shape analysis. *Pattern Recognition*, Elsevier, v. 37, n. 6, p. 1163–1174, 2004.

- TSAI, A. et al. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE transactions on medical imaging*, v. 22, p. 137–54, 03 2003.
- WALT, S. Van der et al. scikit-image: image processing in python. *PeerJ*, PeerJ Inc., v. 2, p. e453, 2014.
- WANG, X.; CHEN, G. Complex networks: Small-world, scale-free and beyond. *Circuits and Systems Magazine, IEEE*, v. 3, p. 6 – 20, 02 2003.
- WATTS, D. J.; STROGATZ, S. H. Collective dynamics of 'small-world' networks. *Nature*, v. 393, n. 6684, p. 440–442, 1998.
- WU, W.; WANG, M. Detecting the dominant points by the curvature-based polygonal approximation. *CVGIP: Graphical Models and Image Processing*, v. 55, n. 2, p. 79–88, 1993. ISSN 1049-9652.
- WU, X.; JING, Z.; WANG, X. The security of iot from the perspective of the observability of complex networks. *Heliyon*, v. 10, n. 5, p. e27104, 2024. ISSN 2405-8440.
- YAGNAVAJJULA, M. K. et al. Detection of neurogenic voice disorders using the fisher vector representation of cepstral features. *Journal of Voice*, 2022. ISSN 0892-1997.
- YU, M. su et al. Identification of konosirus punctatus (temminic and schlegel, 1846) stocks in korean waters via otolith shape analysis and comparison of growth parameters. *Regional Studies in Marine Science*, v. 76, p. 103581, 2024. ISSN 2352-4855.
- ZHANG, Q. et al. Structural analysis and the sum of nodes' betweenness centrality in complex networks. *Chaos, Solitons Fractals*, v. 185, p. 115158, 2024. ISSN 0960-0779.
- ZHENJIANG, M. Zernike moment-based image shape analysis and its application. *Pattern Recognition Letters*, Elsevier, v. 21, n. 2, p. 169–177, 2000.