



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE TECNOLOGIA**  
**DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA**  
**MESTRADO EM ENGENHARIA DE TELEINFORMÁTICA**

**ALLAN CORDEIRO CARNEIRO**

**O PAPEL DA FUNÇÃO CUSTO NA OTIMIZAÇÃO DE UM DESCRITOR DE  
FORMAS MULTIESCALA**

**FORTALEZA**

**2018**

ALLAN CORDEIRO CARNEIRO

O PAPEL DA FUNÇÃO CUSTO NA OTIMIZAÇÃO DE UM DESCRITOR DE FORMAS  
MULTIESCALA

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Teleinformática da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Engenharia de Teleinformática. Área de concentração: Sinais e Sistemas

Orientadora: Prof. Dra. Fátima Nelsi-zeuma Sombra de Medeiros

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- C287p Carneiro, Allan Cordeiro.  
O papel da função custo na otimização de um descritor de formas multiescala / Allan Cordeiro Carneiro. – 2018.  
58 f. : il. color.
- Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2018.  
Orientação: Profa. Dra. Fátima Nelsizeuma Sombra de Medeiros.
1. Descrição de formas. 2. Otimização meta-heurística. 3. Índices de validação de agrupamentos. 4. Recuperação de formas. 5. Classificação de formas. I. Título.
- CDD 621.38
-

ALLAN CORDEIRO CARNEIRO

O PAPEL DA FUNÇÃO CUSTO NA OTIMIZAÇÃO DE UM DESCRITOR DE FORMAS  
MULTIESCALA

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Teleinformática da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Engenharia de Teleinformática. Área de concentração: Sinais e Sistemas

Aprovada em: 19 de Fevereiro de 2018.

BANCA EXAMINADORA

---

Prof. Dra. Fátima Nelsizeuma Sombra de Medeiros (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Marcelo Marques Simões de Souza  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Francisco Nivando Bezerra  
Instituto Federal de Educação, Ciência e Tecnologia do Ceará  
(IFCE)

---

Prof. Dr. Régis Cristiano Pinheiro Marques  
Instituto Federal de Educação, Ciência e Tecnologia do Ceará  
(IFCE)

---

Prof. Dr. Pedro Pedrosa Rebouças Filho  
Instituto Federal de Educação, Ciência e Tecnologia do Ceará  
(IFCE)

Aos meus avós Raimundo Pinto Carneiro (*in memoriam*) e Rita Pinto Carneiro, e ao meu padrinho José Mendes de Lima (*in memoriam*), que tanto contribuíram para meus estudos.

## **AGRADECIMENTOS**

Primeiramente, dedico este espaço para agradecer a Deus por estar sempre me conduzindo.

Agradeço a minha família pelo apoio e dedicação, em especial minha mãe que tanto se esforça pelo bem-estar de nossa família. Sou grato aos meus amigos por ajudarem na minha caminhada.

Dedico um agradecimento notório a toda equipe do LABVIS, com os quais tanto venho aprendendo.

Também expresso minha gratidão a todo o corpo docente do DETI, em especial à Prof<sup>a</sup>. Dr<sup>a</sup>. Fátima Nelsizeuma Sombra de Medeiros pela oportunidade e confiança.

Por fim, meu agradecimento a todos cidadãos brasileiros, que por meio do CNPq, apoiaram financeiramente este trabalho, embora de forma coercitiva.

“Ninguém gasta o dinheiro dos outros com tanto cuidado como gasta o seu próprio. Se quisermos eficiência e eficácia, se quisermos que o conhecimento seja bem usado, isso precisa ser feito por meio da iniciativa privada.”

(Milton Friedman)

## RESUMO

O ajuste de parâmetros de descritores de formas é uma tarefa desafiadora em visão computacional e processamento de imagens. Ao contrário de esquemas manuais, que são custosos e tediosos, há uma tendência de aplicação de otimização meta-heurística para a configuração dos parâmetros desses descritores. No entanto, a escolha da função custo, a qual desempenha um importante papel no processo de otimização, ainda é um problema em aberto. Este trabalho investiga a influência da função custo no desempenho de um descritor multiescala de formas, o qual foi otimizado usando três índices de validação de agrupamentos: a *Silhouette*, o índice de Davies-Bouldin e o índice de Calinski-Harabasz. Para a otimização dos parâmetros de escala do descritor denominado energia de dobramento multiescala normalizada, utilizou-se a meta-heurística de recozimento simulado, realizando experimentos de classificação e recuperação com uma base de formas sintéticas (Kimia 99), duas bases de folhas de plantas (ShapeCN e Swedish) e uma base de pílulas da *United States National Library of Medicine* (base NLM Pills). Na avaliação de desempenho dos experimentos de recuperação e classificação, utilizou-se a medida *Bulls-eye* e a Acurácia da classificação, respectivamente. Os resultados mostraram que o descritor otimizado com o índice de Calinski-Harabasz no papel de função custo apresentou desempenho inferior aos resultados obtidos com as outras funções nas bases com maior dissimilaridade intraclasse. Particularmente para base NLM Pills, onde cada classe tem um padrão bem característico e portanto não há muitas diferenças entre os elementos de uma mesma classe, o descritor de energia de dobramento multiescala normalizada não se beneficiou da metodologia de otimização.

**Palavras-chave:** Descrição de formas, Otimização meta-heurística, Índices de validação de agrupamentos, Recuperação de formas, Classificação de formas.

## ABSTRACT

Parameter adjustment of shape descriptors is a challenging task in computer vision and image processing. Rather than manual schemes, which are time consuming and tedious, there is a trend towards applying metaheuristic optimization to fully support parameter settings of shape descriptors. Nevertheless, the problem of cost function selection, which plays an important role in the optimization process, has not yet been fully addressed. This work investigates the influence of the cost function on the performance of an optimized multiscale shape descriptor using three distinct clustering validation indices: the Silhouette, Davies-Bouldin and Calinski-Harabasz indices. Here, we optimize the scale parameters of the normalized multiscale bending energy descriptor using the simulated annealing metaheuristic; both classification and retrieval experiments are conducted using a synthetic shape dataset (Kimia 99), two plant leaf datasets (ShapeCN and Swedish) and the National Library of Medicine (NLM) pill image dataset (NLM Pills). The performance evaluation, in terms of the Bulls-eye and Accuracy measures, showed that optimized descriptor with the Calinski-Harabasz cost function underperformed the other functions in datasets where there is high level of dissimilarity between classes. Particularly for the NLM Pills, where each class has a well-defined pattern and furthermore the differences within pill classes are quite small. Thus, the Normalized Multiscale Bending Energy descriptor did not benefit from the optimization methodology.

**Keywords:** Shape description, Metaheuristic optimization, Clustering validation indices, Shape retrieval, Shape classification.

## LISTA DE FIGURAS

Figura 1 – Arquitetura de um sistema CBIR . . . . .	20
Figura 2 – Contorno de uma mão. . . . .	22
Figura 3 – Contorno parametrizado. . . . .	22
Figura 4 – Distância ao centróide. . . . .	23
Figura 5 – Curvatura para diferentes escalas de suavização do contorno. . . . .	27
Figura 6 – Metodologia de otimização para um descritor de formas multiescala. . . . .	29
Figura 7 – Exemplos de formas da base ShapeCN. . . . .	30
Figura 8 – Exemplos de formas da base Swedish. . . . .	30
Figura 9 – Exemplos de formas da base Kimia 99. . . . .	31
Figura 10 – Exemplos de formas da base NLM Pills. . . . .	31
Figura 11 – Convergência do SA para a função de Rastringi . . . . .	36
Figura 12 – Rede SOM. . . . .	36
Figura 13 – Matriz U. . . . .	38
Figura 14 – Exemplos das bases sintéticas A e B . . . . .	42
Figura 15 – Resultados dos experimentos de recuperação para um exemplar da base Swedish usando a otimização com a <i>Silhouette</i> (SI), o Índice de Davies-Bouldin (DB) e o Índice de Calinski-Harabasz (CH), além do método não otimizado proposto por Cesar Jr e Costa (1997). . . . .	47
Figura 16 – Matrizes U para a base ShapeCN usando a otimização com três funções custo diferentes e o método não otimizado proposto por Cesar Jr e Costa (1997). . . . .	49
Figura 17 – Matrizes U para a base Swedish usando a otimização com três funções custo diferentes e o método não otimizado proposto por Cesar Jr e Costa (1997). . . . .	50
Figura 18 – Matrizes U para a base Kimia 99 usando a otimização com três funções custo diferentes e o método não otimizado proposto por Cesar Jr e Costa (1997). . . . .	51
Figura 19 – Matrizes U para a base NLM Pills usando a otimização com três funções custo diferentes e o método não otimizado proposto por Cesar Jr e Costa (1997). . . . .	52

## LISTA DE TABELAS

Tabela 1 – Índices de validação de agrupamento . . . . .	40
Tabela 2 – Vetores de média das distribuições normais que geram as bases A e B. . . . .	41
Tabela 3 – Índices de agrupamento para as bases A e B . . . . .	41
Tabela 4 – Bulls-eye(%) . . . . .	45
Tabela 5 – Acurácia (%) . . . . .	45

## LISTA DE ALGORITMOS

Algoritmo 1 – Algoritmo de recozimento simulado (SA) . . . . .	33
Algoritmo 2 – Bulls-eye. . . . .	42

## LISTA DE ABREVIATURAS E SIGLAS

CBIR	Recuperação de imagens baseada em conteúdo
CH	Índice de Calinski-Harabasz
DB	Índice de Davies-Bouldin
DE	Evolução diferencial
GUI	Interface gráfica do usuário
kNN	$k$ vizinhos mais próximos
NMBE	Energia de dobramento multiescala normalizada
PSO	Otimização por enxame de partículas
SA	Recozimento simulado
SI	<i>Silhouette</i>
SOM	Mapa auto-organizável

## LISTA DE SÍMBOLOS

$N$	Número de pontos do contorno
$u$	Contorno parametrizado
$\kappa$	Curvatura
$\text{Im}\{\cdot\}$	Parte imaginária de um número complexo
$\mathcal{F}\{\cdot\}$	Transformada de Fourier
$\mathcal{F}^{-1}\{\cdot\}$	Transformada de Fourier inversa
$ \cdot $	Módulo de um número complexo
$\ \cdot\ $	Norma euclidiana
$E_n$	Energia de dobramento normalizada
$g_\sigma$	Filtro passa-baixas gaussiano
$\sigma$	Desvio padrão de um filtro passa-baixas gaussiano
$L$	Perímetro do contorno original
$L_\sigma$	Perímetro do contorno suavizado
$E_n$	Energia de dobramento normalizada
$N_S$	Número de escalas do descritor energia de dobramento normalizada
$t$	Índice que indica a $t$ -ésima época do algoritmo recozimento simulado
$s_a(t)$	Solução atual da $t$ -ésima época do recozimento simulado
$s_c(t)$	Solução candidata da $t$ -ésima época do recozimento simulado
$T(t)$	Temperatura do algoritmo recozimento simulado para a $t$ -ésima época
$N_e$	Número de épocas do algoritmo recozimento simulado
$\arg \min\{\cdot\}$	Argumento do mínimo
$w_{ij}$	Vetor de pesos do neurônio com coordenadas $(i, j)$
$r$	Raio de vizinhança de um neurônio do mapa auto-organizável (do inglês <i>self organized map</i> )
$V_r(i, j)$	Vizinhança do neurônio com coordenadas $(i, j)$ para o raio de vizinhança $r$
$\beta$	Ganho de adaptação do vetor de peso de um dado neurônio do mapa auto-organizável

$\beta_0$	Ganho de adaptação inicial do vetor de peso de um dado neurônio do mapa auto-organizável
$q_n$	Tamanho do mapa auto-organizável (quantidade de neurônios)
$T_{tr}$	Duração do treinamento do mapa auto-organizável
$\max\{\cdot\}$	Resulta no valor máximo dos operandos
$S(i)$	Silhouette do elemento $i$
$a(i)$	Dissimilaridade média entre o elemento $i$ e aqueles que pertencem à sua classe
$b(i)$	Dissimilaridade média entre o elemento $i$ e a classe vizinha mais próxima, excluindo a classe que $i$ pertence
$DB$	Índice de Davies-Bouldin
$d(r, s)$	Distância entre os vetores $r$ e $s$
$A_i$	$i$ -ésimo agrupamento de uma da base de dados
$c$	Centróide da base de dados
$c_i$	Centróide do $i$ -ésimo agrupamento de uma da base de dados
$\eta$	Número de objetos da base de dados
$N_A$	Número de agrupamentos de uma da base de dados
$\eta_i$	Número de objetos do $i$ -ésimo agrupamento de uma da base de dados
$CH$	Índice de Calinski-Harabasz
$BE$	<i>Bulls-eye</i>
$I$	Matriz identidade
$\text{diag}\{\cdot\}$	Resulta em uma matriz diagonal cujos elementos são os componentes do vetor utilizado como operando
$Q$	Número de amostras do conjunto de treinamento do algoritmo $k$ vizinhos mais próximos
$k$	Número de vizinhos mais próximos do algoritmo $k$ vizinhos mais próximos
$C_T$	Conjunto de treinamento do algoritmo $k$ vizinhos mais próximos

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>16</b>
<b>1.1</b>	<b>Motivação</b>	<b>17</b>
<b>1.2</b>	<b>Objetivos</b>	<b>17</b>
<b>1.3</b>	<b>Produção científica</b>	<b>17</b>
<b>1.4</b>	<b>Organização da dissertação</b>	<b>18</b>
<b>2</b>	<b>FUNDAMENTOS TEÓRICOS</b>	<b>19</b>
<b>2.1</b>	<b>Recuperação de imagens baseada em conteúdo (CBIR)</b>	<b>19</b>
<b>2.2</b>	<b>Extração de características</b>	<b>20</b>
<b>2.2.1</b>	<i>Processamento do Contorno</i>	<b>20</b>
<b>2.2.2</b>	<i>Curvatura</i>	<b>23</b>
<b>2.2.3</b>	<i>Energia de dobramento multiescala normalizada (NMBE)</i>	<b>26</b>
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>29</b>
<b>3.1</b>	<b>Bases de imagens</b>	<b>29</b>
<b>3.2</b>	<b>Algoritmo de otimização</b>	<b>32</b>
<b>3.3</b>	<b>Mapa auto-organizável (SOM)</b>	<b>35</b>
<b>3.4</b>	<b>Índices de validação de agrupamento</b>	<b>38</b>
<b>3.5</b>	<b>Recuperação</b>	<b>41</b>
<b>3.6</b>	<b>Classificação</b>	<b>43</b>
<b>3.7</b>	<b>Experimentos</b>	<b>44</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÕES</b>	<b>45</b>
<b>5</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS</b>	<b>53</b>
	<b>REFERÊNCIAS</b>	<b>54</b>

## 1 INTRODUÇÃO

Vários descritores de forma estão disponíveis na literatura, sendo que parte desses são paramétricos (LING; JACOBS, 2007; WANG *et al.*, 2012; WANG *et al.*, 2015). Em geral, essas abordagens ajustam os parâmetros do descritor manualmente através de uma busca exaustiva, conduzindo particularmente a resultados sub-ótimos. Um exemplo disso é o descritor *inner-distance shape context* (LING; JACOBS, 2007), quando Wang *et al.* (2015) utilizaram esse descritor na identificação de espécies de plantas, esses autores adotaram valores para os parâmetros previamente ajustados para a base de imagens MPEG-7 (LATECKI *et al.*, 2000). De fato, o ajuste de parâmetros depende da aplicação, o que implica em uma busca do melhor conjunto de parâmetros para cada aplicação específica.

Souza *et al.* (2016) introduziram uma metodologia de otimização que permite a escolha da função custo de acordo com o objetivo da aplicação. Essa metodologia aplica técnicas de otimização meta-heurística utilizando o índice de validação de agrupamento *Silhouette* (ROUSSEEUW, 1987) como uma função custo para a otimização dos parâmetros de um descritor multiescala. Inspirado por Souza *et al.* (2016), este trabalho investiga a influência da função custo na otimização dos parâmetros das escalas do descritor de formas energia de dobramento multiescala normalizada (Energia de dobramento multiescala normalizada (NMBE)) (CESAR JR; COSTA, 1997). Realizou-se experimentos de otimização com a meta-heurística reconhecimento simulado (Reconhecimento simulado (SA)) (KIRKPATRICK *et al.*, 1982; CERNY, 1985) e três diferentes funções custo: a *Silhouette* (SI), o índice de Davies-Bouldin (DB) (DAVIES; BOULDIN, 1979) e o índice de Calinski-Harabasz (CH) (CALINSKI; HARABASZ, 1974). As bases de imagens utilizadas nos testes foram duas bases de folhas de plantas - ShapeCN (BACKES *et al.*, 2009) e Swedish (SÖDERKVIST, 2001), a base sintética Kimia 99 (SEBASTIAN *et al.*, 2004) e a base de pílulas da *National Library of Medicine* - NLM Pills<sup>1</sup> (USHIZIMA *et al.*, 2015)

Avaliou-se o desempenho dos descritores otimizados na recuperação de imagens baseada em conteúdo (Recuperação de imagens baseada em conteúdo (CBIR)) e em experimentos de classificação supervisionada, sendo utilizado o algoritmo de classificação  $k$  vizinhos mais próximos ( $k$  vizinhos mais próximos (kNN)) (FIX; HODGES JR, 1951). Nos experimentos de recuperação de formas, considerou-se o *Bulls-eye* (LATECKI *et al.*, 2000) como uma medida do desempenho do descritor e em classificação, a Acurácia de classificação. A Matriz U (ULTSCH; SIEMON, 1990) foi utilizada como ferramenta de visualização qualitativa dos dados.

<sup>1</sup> Disponível em <<https://www.nlm.nih.gov/news/nlm-pill-image-challenge-2016.html>>

## 1.1 Motivação

Os índices de agrupamentos baseiam seus cálculos em conceitos distintos, chegando mesmo a ter resultados divergentes, o que fica evidente quando seus valores são comparados para diferentes bases de dados (CHOU *et al.*, 2004; LIU *et al.*, 2010; BAARSCH; CELEBI, 2012). Assim, quando se utiliza esses índices como funções custo para a otimização de descritores de formas, espera-se que alguns desses sejam mais adequados para uma determinada base de imagens.

## 1.2 Objetivos

Esta dissertação possui três objetivos principais:

- Modificar a metodologia de otimização proposta por Souza *et al.* (2016), utilizando diferentes funções custo para guiar a meta-heurística recozimento simulado (SA) na otimização dos parâmetros das escalas do descritor de formas energia de dobramento multiescala normalizada (NMBE);
- Analisar a influência que a escolha da função custo exerce no ajuste por otimização dos parâmetros do NMBE;
- Comparar os resultados obtidos na otimização com os que foram alcançados utilizando os parâmetros sugeridos pelos autores do NMBE.

## 1.3 Produção científica

- ARAÚJO, F. H. D.; CARNEIRO, A. C.; SILVA, R. R. V.; MEDEIROS, F. N. S.; USHIZIMA, D. M. Redes neurais convolucionais com tensorflow: Teoria e prática. In: SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. **III Escola Regional de Informática do Piauí. Livro Anais - Artigos e Minicursos**. [S.l.]: Sociedade Brasileira de Computação, 2017. v. 1, p. 382–406. ISBN 978-85-7669-395-6.
- CARNEIRO, A. C.; LOPES, J. G. F.; ARAÚJO, F. H. D.; SILVA, R. R. V.; PASSARINHO, C. J. P.; ROCHA NETO, J. F. S.; MEDEIROS, F. N. S. Análise de fotografias de pílulas por redes neurais convolucionais. In: UNIVERSIDADE FEDERAL DO ABC. **8º Simpósio de Instrumentação e Imagens Médicas (SIIM) e o 7º Simpósio de Processamento de Sinais (SPS)**. [S.l.], 2017.
- CARNEIRO, A. C.; LOPES, J. G. F.; ROCHA NETO, J. F. S.; SOUZA, M. M. S.;

MEDEIROS, F. N. S.; BEZERRA, F. N. On the evaluation of cost functions for parameter optimization of a multiscale shape descriptor. In: IEEE. **International Symposium on Signal Processing and Information Technology (ISSPIT)**. [S.l.], 2017.

#### 1.4 Organização da dissertação

Este trabalho foi elaborado conforme a seguinte estrutura:

- **Fundamentos teóricos** – são discutidos os conceitos fundamentais do trabalho, sendo esses: a recuperação de imagens baseada em conteúdo e o processo de extração de características das imagens.
- **Materiais e Métodos** – nessa parte do trabalho, são apresentadas as quatro bases públicas que foram utilizadas com o descritor para a obtenção dos resultados, a metodologia de otimização, as funções custo empregadas, a técnica de visualização Matriz U e os métodos de avaliação dos resultados.
- **Resultados e Discussão** – os resultados obtidos são exibidos e comentados.
- **Conclusões e Trabalhos Futuros** – são apresentadas as conclusões finais deste trabalho e sugestões de trabalhos futuros.

## 2 FUNDAMENTOS TEÓRICOS

Neste capítulo, são apresentados os conceitos necessários para o entendimento deste trabalho, destacando-se a recuperação de imagens baseada em conteúdo e o processo de extração de características das imagens.

### 2.1 Recuperação de imagens baseada em conteúdo (CBIR)

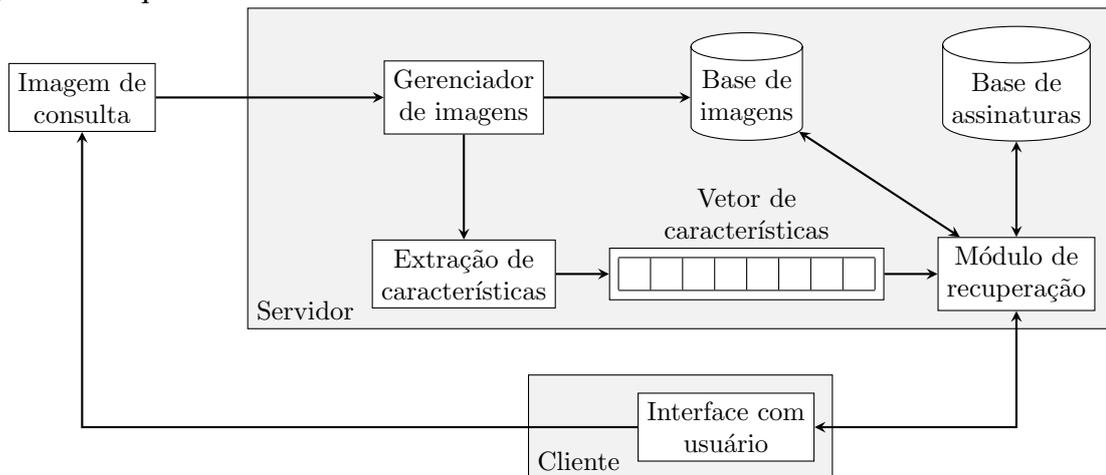
A recuperação de imagens baseada em conteúdo, do inglês *content-based image retrieval* (CBIR), consiste em um conjunto de técnicas para a recuperação de determinadas imagens de uma base utilizando as características das imagens que foram extraídas de forma automática (WANG, 2001).

Um típico sistema CBIR é exibido na Figura 1. Os principais módulos que o compõem são:

- Cliente – é composto por apenas um módulo de interface gráfica do usuário, do inglês *graphical user interface* (Interface gráfica do usuário (GUI)). Esse componente do sistema CBIR consiste em um programa que provê a interatividade entre o usuário e os módulos do servidor do sistema CBIR, possibilitando ao usuário realizar consultas e visualizar resultados (MARQUES; FURHT, 2002).
- Servidor – é formado por módulos que gerenciam e processam as consultas e as bases de dados do sistema CBIR. O servidor aceita consultas e fornece seus resultados ao cliente, sendo formado pelos seguintes módulos:
  - Gerenciador de imagens – gerencia o acesso dos arquivos de imagens, faz conversões de formatos e retorna informações sobre os arquivos (WANG, 2001).
  - Extração de características – módulo que utiliza um conjunto de descritores para obter as assinaturas das imagens consultadas, as quais são representadas por vetores de características.
  - Módulo de recuperação – conjunto de algoritmos que buscam e recuperam imagens, cujas características associadas, que estão na base de assinaturas, são mais próximas das observadas no vetor de características da imagem de consulta (MARQUES; FURHT, 2002).

Como um sistema CBIR necessita das características das imagens para comparar quais são semelhantes e assim efetuar a recuperação, a etapa de extração de características é

Figura 1 – Arquitetura de um sistema CBIR



Fonte: adaptado de (WANG, 2001).

fundamental, sendo utilizados atributos das imagens, como cor, textura, forma, dentre outros. Nesse contexto, este trabalho enfatiza a extração de características das formas de objetos presentes em uma imagem, particularmente por meio do contorno. Isso porque a forma é um atributo suficiente no reconhecimento de objetos em uma ampla gama de problemas de visão computacional (COSTA; CESAR JR, 2009). Ademais, grande parte das informações das formas dos objetos encontram-se representadas no contorno (GONZALEZ; WOODS, 2007).

## 2.2 Extração de características

Nesta seção, serão abordadas técnicas de extração de características baseadas no contorno de formas, dando destaque ao descritor de formas energia de dobramento multiescala normalizada (NMBE).

### 2.2.1 Processamento do Contorno

Um desafio para a descrição de formas é a escolha de características que as especifiquem apropriadamente, o que é necessário para que um sistema computacional possa distingui-las, agrupá-las ou classificá-las.

As abordagens de descrição de formas se dividem em métodos baseados em contorno e baseados em região (ZHANG; LU, 2004). O descritor de formas abordado neste trabalho (NMBE) utiliza em seus cálculos relações entre os elementos do contorno, o qual pode ser definido como o conjunto de pontos que delimitam uma forma. Esse trabalho considera apenas o contorno externo, ou seja, o contorno que separa um objeto bidimensional de seu exterior.

### ***Contorno paramétrico***

Para extrair os pontos do contorno de um objeto, define-se um ponto inicial em sua borda e a partir desse, percorre-se todo o contorno em sentido horário (ou anti-horário). Desse processo, resultam pontos ordenados que podem ser representados como dois vetores (COSTA; CESAR JR, 2009):

$$x = [x(1), x(2), \dots, x(N)], \quad y = [y(1), y(2), \dots, y(N)], \quad (2.1)$$

para  $n \in [1, 2, \dots, N]$ , sendo  $N$  o número de pontos do contorno.

Cada par  $(x(n), y(n))$  corresponde a um ponto do contorno e  $n$ , a posição de dado elemento. Assim, obtém-se o contorno parametrizado pela posição em que seus elementos foram extraídos. Outra forma de representação do par  $(x(n), y(n))$  é através da parametrização complexa (COSTA; CESAR JR, 2009):

$$u(n) = x(n) + jy(n), \quad (2.2)$$

em que  $u$  é o contorno parametrizado.

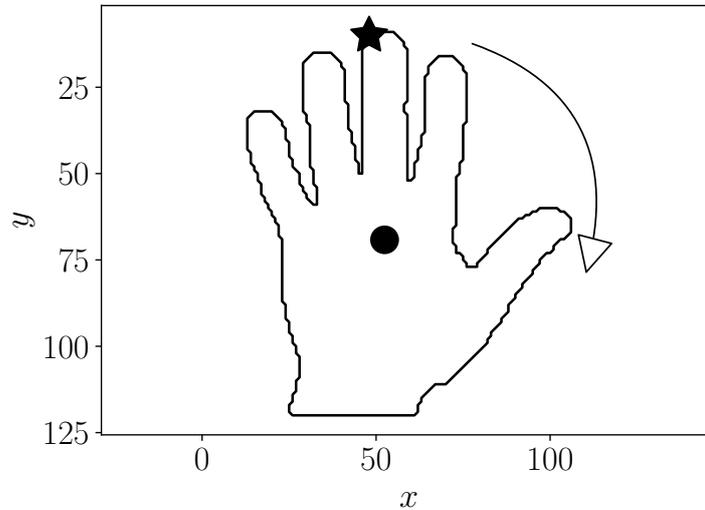
A Figura 2 exemplifica o processo de extração do contorno. Partindo do ponto inicial, marcado como estrela, percorre-se os outros pontos do contorno no sentido horário. Os elementos dos vetores  $x$  e  $y$ , resultantes da Figura 2, são plotados versus a posição que foram extraídos na Figura 3, aonde a linha tracejada corresponde aos elementos de  $x$  e a contínua aos de  $y$ .

### ***Assinatura***

As assinaturas são representações unidimensionais do contorno das formas dos objetos que permitem a aplicação de técnicas de processamento de sinais unidimensionais na análise de formas (COSTA; CESAR JR, 2009).

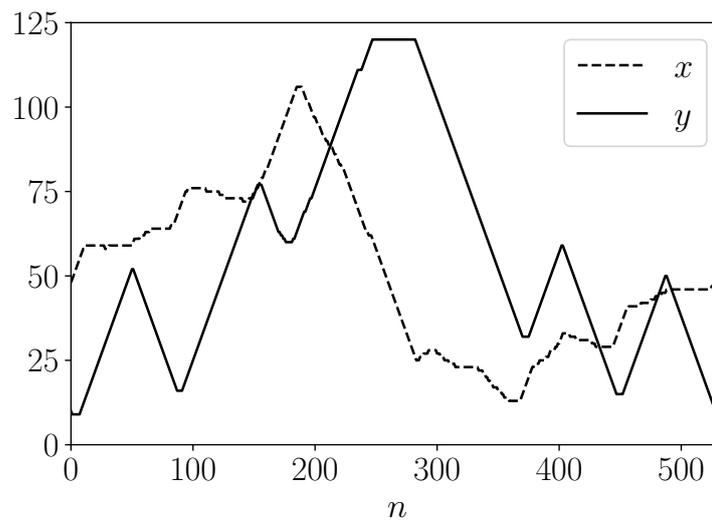
Um exemplo de assinatura é a distância ao centróide. Essa é obtida percorrendo os pontos do contorno e calculando a distância entre eles e o centróide da forma. De posse do contorno extraído da forma da Figura 2, obtém-se o centróide (marcado como um círculo preto na Figura 2) e calcula-se a distância ao centróide exibida na Figura 4 (PEDRINI; SCHWARTZ, 2007). A distância considerada no cálculo da assinatura foi a euclidiana. Essa última é definida

Figura 2 – Contorno de uma mão.



Fonte: o autor.

Figura 3 – Contorno parametrizado.



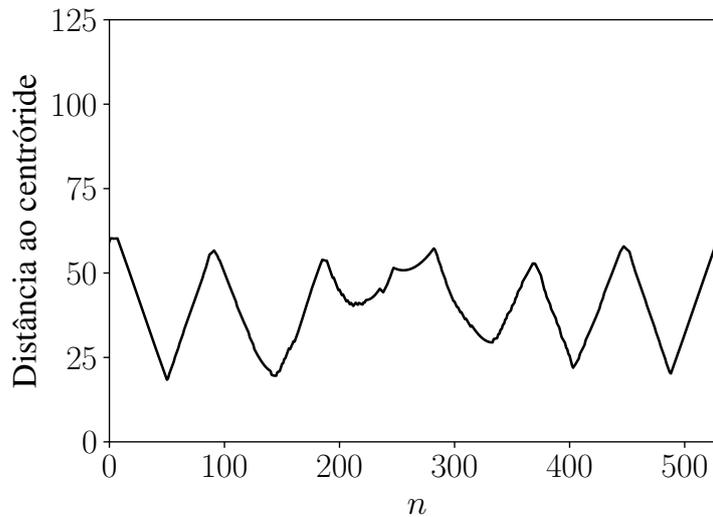
Fonte: o autor.

como:

$$d_{euclidiana} = \sqrt{\sum_{i=1}^{N_v} (a_i - b_i)^2}, \quad (2.3)$$

onde  $a$  e  $b$  são vetores de dimensão  $N_v$ , sendo  $a_i$  e  $b_i$  suas  $i$ -ésimas componentes, respectivamente. No caso da distância ao centróide, os vetores  $a$  e  $b$  representam as coordenadas de um determinado ponto do contorno e do centróide, respectivamente.

Figura 4 – Distância ao centróide.



Fonte: o autor.

### 2.2.2 Curvatura

A curvatura é uma assinatura que indica a variação local do contorno, pode ser definida através das coordenadas paramétricas do contorno como (CESAR JR; COSTA, 1996):

$$\kappa(n) = \frac{\dot{x}(n)\ddot{y}(n) - \ddot{x}(n)\dot{y}(n)}{(\dot{x}(n)^2 + \dot{y}(n)^2)^{\frac{3}{2}}}, \quad (2.4)$$

em que (  $\dot{\quad}$  ) representa a primeira derivada e (  $\ddot{\quad}$  ) representa a segunda derivada.

A equação 2.4 resulta em valores positivos ou negativos dependendo da concavidade local. Além disso, a curvatura é muito influenciada pela geometria do contorno. Por exemplo, a curvatura de uma reta é zero, de uma circunferência é constante e a de um objeto irregular repleto de cantos e junções apresenta grande variação.

Considerando a parametrização da equação 2.2, pode-se definir a primeira e a segunda derivadas de  $u(n)$  em relação ao parâmetro  $n$ , respectivamente, como:

$$\dot{u}(n) = \dot{x}(n) + j\dot{y}(n) \text{ e} \quad (2.5)$$

$$\ddot{u}(n) = \ddot{x}(n) + j\ddot{y}(n). \quad (2.6)$$

Denotando  $\ddot{u}^*(n)$  como o complexo conjugado de  $\ddot{u}(n)$  e calculando  $\dot{u}(n)\ddot{u}^*(n)$ ,

tem-se:

$$\begin{aligned}
 \dot{u}(n)\ddot{u}^*(n) &= [\dot{x}(n) + \dot{y}(n)][\ddot{x}(n) - \ddot{y}(n)] \\
 \dot{u}(n)\ddot{u}^*(n) &= \dot{x}(n)\ddot{x}(n) - j\dot{x}(n)\ddot{y}(n) + j\dot{x}(n)\ddot{y}(n) + \dot{y}(n)\ddot{y}(n) \\
 \dot{u}(n)\ddot{u}^*(n) &= \dot{x}(n)\ddot{x}(n) + \dot{y}(n)\ddot{y}(n) - j[\dot{x}(n)\ddot{y}(n) - \dot{y}(n)\ddot{x}(n)].
 \end{aligned} \tag{2.7}$$

Denotando  $|\cdot|$  como o módulo de um número complexo e calculando  $|\dot{u}(n)|^3$ ,

tem-se:

$$\begin{aligned}
 |\dot{u}(n)|^3 &= \left[ \sqrt{\dot{x}^2(n) + \dot{y}^2(n)} \right]^3 \\
 |\dot{u}(n)|^3 &= [\dot{x}^2(n) + \dot{y}^2(n)]^{\frac{3}{2}}.
 \end{aligned} \tag{2.8}$$

Nota-se que a parte imaginária da equação 2.7 é o oposto do numerador da equação 2.4 e substituindo a equação 2.8 na equação 2.4, tem-se:

$$\kappa(n) = \frac{-\text{Im}\{\dot{u}(n)\ddot{u}^*(n)\}}{|\dot{u}(n)|^3}, \tag{2.9}$$

em que  $\text{Im}\{\cdot\}$  denota a parte imaginária de um número complexo.

### ***Filtragem gaussiana e abordagem multiescala***

Observa-se na equação 2.4 que a curvatura depende diretamente da primeira e segunda derivadas das coordenadas paramétricas do contorno. Conseqüentemente, a curvatura é contaminada, de forma amplificada, pelos ruídos inerentes aos processos de aquisição, amostragem e quantização dos contornos dos objetos das imagens, o que restringe o seu uso direto como descritor. Uma solução possível é a suavização do contorno  $u(n)$ , por exemplo, por um filtro passa-baixas gaussiano (COSTA; CESAR JR, 2009):

$$g_\sigma(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{n^2}{2\sigma^2}\right). \tag{2.10}$$

Portanto, pode-se definir a versão filtrada de  $u(n)$  por  $g_\sigma(n)$  como:

$$\hat{u}(n, \sigma) = u(n) * g_\sigma(n), \tag{2.11}$$

onde  $*$  denota a operação de convolução.

O parâmetro  $\sigma$  controla o nível de detalhes do contorno em sua descrição. Assim, quanto maior o valor de  $\sigma$ , maior será o grau de suavização do contorno, e portanto mais grosseira

será a representação do contorno suavizado, ou seja, menos detalhes serão representados pelo descritor.

A primeira e a segunda derivadas de  $\hat{u}(n, \sigma)$  em relação ao parâmetro  $n$  são definidas, respectivamente, como:

$$\hat{u}'(n, \sigma) = \dot{u}(n) * g_\sigma(n) \text{ e} \quad (2.12)$$

$$\hat{u}''(n, \sigma) = \ddot{u}(n) * g_\sigma(n). \quad (2.13)$$

O cálculo de  $\hat{u}(n, \sigma)$  e de suas primeira e segunda derivadas pode ser feito mais eficientemente no domínio da frequência (CESAR JR; COSTA, 1997), conforme adotado nesta dissertação. Seguindo essa abordagem e aplicando a propriedade da derivada da transformada de Fourier tem-se:

$$\hat{u}(n, \sigma) = \mathcal{F}^{-1} \left\{ U(f) G_{\frac{1}{\sigma}}(f) \right\}, \quad (2.14)$$

$$\hat{u}'(n, \sigma) = \mathcal{F}^{-1} \left\{ j2\pi f U(f) G_{\frac{1}{\sigma}}(f) \right\} \text{ e} \quad (2.15)$$

$$\hat{u}''(n, \sigma) = \mathcal{F}^{-1} \left\{ -(2\pi f)^2 U(f) G_{\frac{1}{\sigma}}(f) \right\}, \quad (2.16)$$

em que

$$U(f) = \mathcal{F} \{ u(n) \}, \quad (2.17)$$

$$G_{\frac{1}{\sigma}}(f) = \mathcal{F} \{ g_\sigma(n) \} = \exp \left( -\frac{(2\pi f)^2 \sigma^2}{2} \right), \quad (2.18)$$

$\mathcal{F} \{ \cdot \}$  e  $\mathcal{F}^{-1} \{ \cdot \}$  denotam a transformada de Fourier e sua inversa, respectivamente.

### ***Efeito do encolhimento***

O processo de filtragem gaussiana modifica a amplitude da representação do sinal dos contornos incorrendo no efeito do encolhimento dos contornos filtrados. Dessa forma, é necessário normalizar os contornos suavizados para que a curvatura não seja afetada (COSTA; CESAR JR, 2009). Uma solução é definir como constante de normalização ( $P_\sigma$ ) a razão entre o perímetro do contorno original ( $L$ ) e o perímetro do contorno suavizado ( $L_\sigma$ ), como mostrado na seguinte equação:

$$P_\sigma = \frac{L}{L_\sigma}, \quad (2.19)$$

sendo que  $L$  e  $L_\sigma$  são estimados por:

$$L = \frac{2\pi}{N} \sum_{n=0}^{N-1} |\dot{u}(n)| \quad \text{e} \quad (2.20)$$

$$L_\sigma = \frac{2\pi}{N} \sum_{n=0}^{N-1} |\hat{u}(n, \sigma)|. \quad (2.21)$$

O contorno e suas primeira e segunda derivadas normalizados são dados, respectivamente, por:

$$u(n, \sigma) = \hat{u}(n, \sigma)P_\sigma, \quad (2.22)$$

$$\dot{u}(n, \sigma) = \hat{u}(n, \sigma)P_\sigma \quad \text{e} \quad (2.23)$$

$$\ddot{u}(n, \sigma) = \hat{u}(n, \sigma)P_\sigma. \quad (2.24)$$

### ***Curvatura multiescala***

Com as equações 2.22, 2.23 e 2.24, pode-se definir a curvatura multiescala como:

$$\kappa(n, \sigma) = \frac{-\text{Im}\{\dot{u}(n, \sigma)\dot{u}^*(n, \sigma)\}}{|\dot{u}(n, \sigma)|^3}, \quad (2.25)$$

onde cada valor  $\sigma$  corresponde a uma escala diferente de detalhes.

A Figura 5 ilustra o efeito do processo de suavização sobre o contorno da forma e a sua curvatura. Nota-se que à medida que o contorno perde detalhes ao ser suavizado, a curvatura também fica mais suave, indicando a diminuição do ruído, mas também a perda de informações em relação à forma inicial.

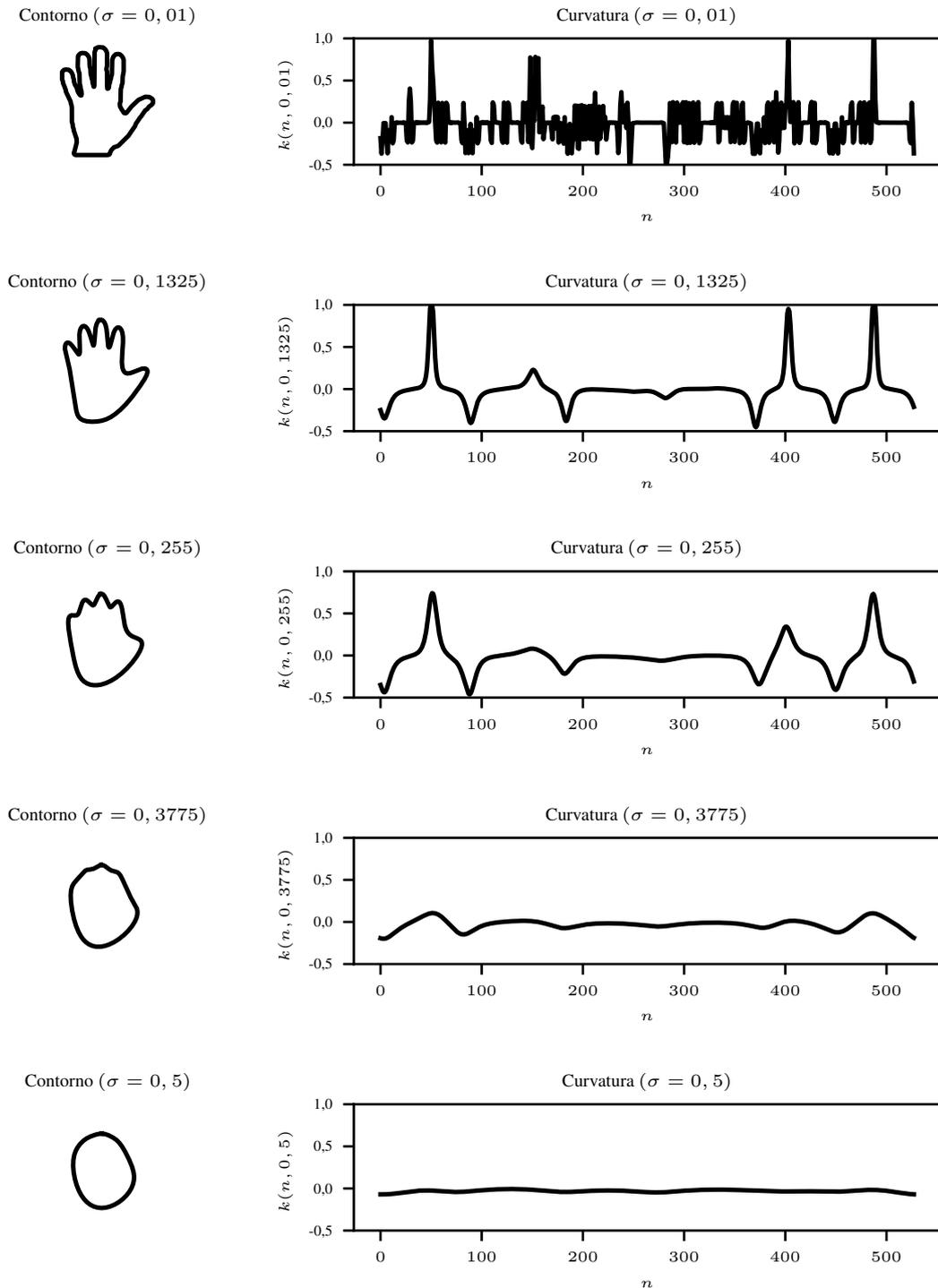
### **2.2.3 Energia de dobramento multiescala normalizada (NMBE)**

A energia de dobramento, do inglês *bending energy*, mede a quantidade de energia necessária para deformar um determinado contorno fechado para torná-lo um contorno de menor energia equivalente, ou seja, um círculo de perímetro igual ao do contorno original (YOUNG *et al.*, 1974). Para o caso discreto, a energia de dobramento é definida por (CESAR JR; COSTA, 1997):

$$E = \frac{1}{N} \sum_{n=1}^N \kappa(n)^2, \quad (2.26)$$

sendo  $N$  o número de pontos do contorno.

Figura 5 – Curvatura para diferentes escalas de suavização do contorno.



Fonte: o autor.

Para que a energia de dobramento seja invariante à escala, é necessária a sua normalização, incluindo o termo  $L^2$  na equação 2.26 em que  $L$  é o perímetro da forma. A energia de

dobramento normalizada é então definida por:

$$E_n = \frac{L^2}{N} \sum_{n=1}^N \kappa(n)^2. \quad (2.27)$$

Utilizando a definição multiescala da curvatura, pode-se definir a energia de dobramento multiescala normalizada, do inglês *normalized multiscale bending energy* (NMBE) (CESAR JR; COSTA, 1997), como:

$$E_n(\sigma) = \frac{L^2}{N} \sum_{n=1}^N \kappa(n, \sigma)^2, \quad (2.28)$$

onde  $E_n(\sigma)$  é a energia de dobramento normalizada na escala  $\sigma$ .

Uma vez que  $E_n(\sigma)$  se trata de uma energia, Souza *et al.* (2016) aplicaram a operação  $\log(\cdot)$  para cada escala da equação 2.29, definindo o descritor NMBE como:

$$\text{NMBE} = [\log E_n(\sigma_1), \log E_n(\sigma_2), \dots, \log E_n(\sigma_M)], \quad (2.29)$$

em que  $M$  indica o número de escalas e  $\sigma_l$  é o  $l$ -ésimo parâmetro de escala.

### ***Ajuste de escalas***

Cesar Jr e Costa (1997) propuseram um método de ajuste das escalas ( $\sigma_l$ ) do descritor NMBE pelas seguintes equações:

$$\sigma_l = \left[ \left( \frac{\sigma_{min}^{-1} - \sigma_{max}^{-1}}{oct_{max} - \sqrt{2}} \right) (oct_l - \sqrt{2}) + \sigma_{max}^{-1} \right]^{-1}, \quad (2.30)$$

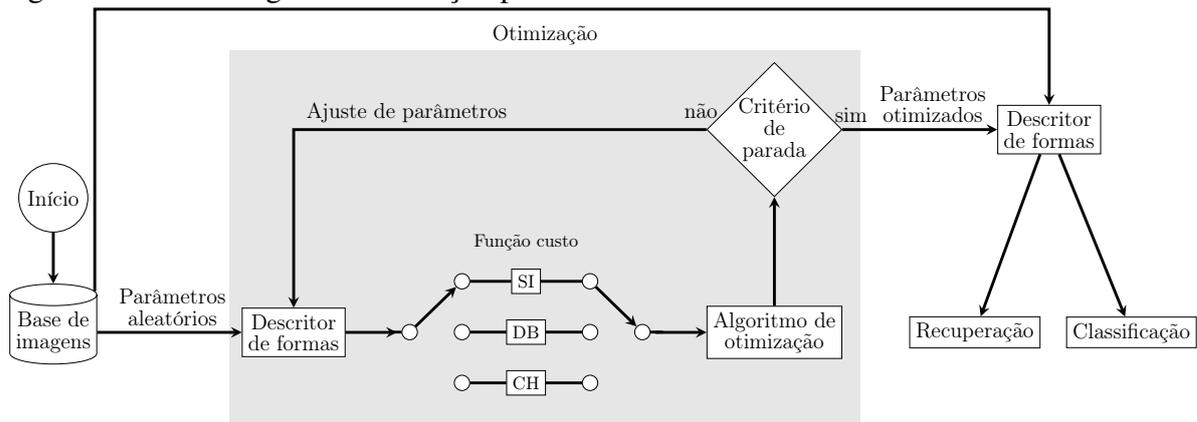
$$oct_l = (\sqrt{2})^l; l = [1, 2, \dots, N_S]. \quad (2.31)$$

Nas equações 2.31 e 2.30, o índice  $l$  indica a escala, sendo o número de escalas definido como  $N_S$ . O índice auxiliar  $oct$  é incrementado em oitavas, e os termos  $\sigma_{min}$  e  $\sigma_{max}$  correspondem aos valores mínimos e máximos que o parâmetro de escala  $\sigma$  pode assumir. O termo  $oct_{max}$  é o valor máximo que  $oct$  pode assumir. Cesar Jr e Costa (1997) obtiveram as escalas adotando os seguintes valores para os parâmetros:  $N_S = 32$ ,  $\sigma_{min} = \frac{1}{64}$  e  $\sigma_{max} = 2$ .

### 3 MATERIAIS E MÉTODOS

A Figura 6 ilustra a metodologia de otimização seguida neste trabalho. Modificou-se a metodologia introduzida por Souza *et al.* (2016) empregando três diferentes funções custo: Silhouette (SI), índice de Davies-Bouldin (DB) e índice de Calinski-Harabasz (CH). Primeiramente, o algoritmo de otimização seleciona aleatoriamente os parâmetros do descritor de formas. Em seguida, extrai-se as características das formas da base de imagens e calcula-se a função custo. Esse processo é repetido até se atingir um critério de parada, neste caso, o número de rodadas. A função custo guia o algoritmo de otimização na busca de um conjunto de parâmetros para o descritor de formas que maximize a separação interclasse e minimize a separação intraclasse. Quando o critério de parada é atingido, os parâmetros (escalas do descritor) ajustados são utilizados para uma nova etapa de descrição da base de imagens, gerando uma base de assinaturas, a qual é submetida à experimentos de recuperação e classificação.

Figura 6 – Metodologia de otimização para um descritor de formas multiescala.



Fonte: o autor.

Nas próximas seções, as etapas mostradas na Figura 6 serão comentadas, com exceção do descritor de formas, o qual foi detalhado no Capítulo 2. Além disso, a condução dos experimentos será discutida na última seção deste capítulo.

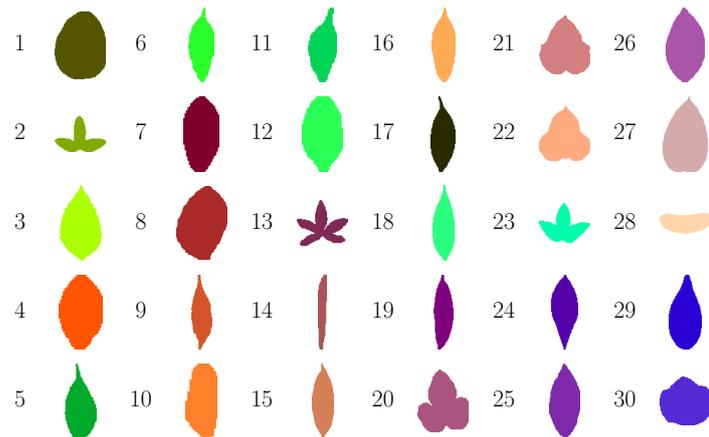
#### 3.1 Bases de imagens

Esta subseção apresenta as bases de imagens mostrando um exemplar de cada classe com a cor que será empregada na apresentação dos resultados, já que as próprias formas indicarão sua posição correspondente na estrutura da Matriz U (ULTSCH; SIEMON, 1990).

### *ShapeCN*

A ShapeCN (BACKES *et al.*, 2009) é uma base com 30 classes de folhas, sendo 20 exemplares por classe. Um exemplo de forma para cada classe é exibido na Figura 7.

Figura 7 – Exemplares de formas da base ShapeCN.

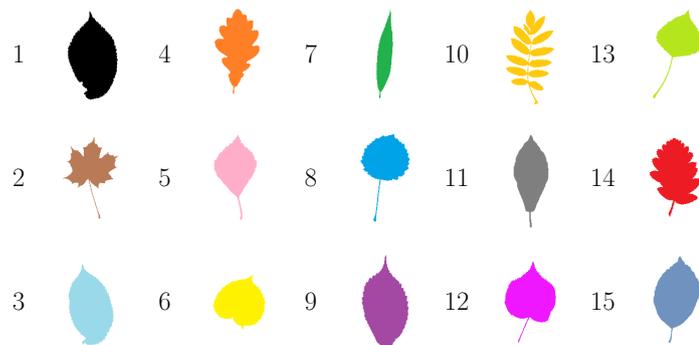


Fonte: o autor.

### *Swedish*

A Swedish (SÖDERKVIST, 2001) é uma base de folhas de espécies vegetais suecas composta por 15 classes com 75 elementos cada. Um elemento de cada classe é mostrado na Figura 8. As imagens da base foram reduzidas a 20% do tamanho original, a fim de reduzir o tempo de computação dos algoritmos.

Figura 8 – Exemplares de formas da base Swedish.

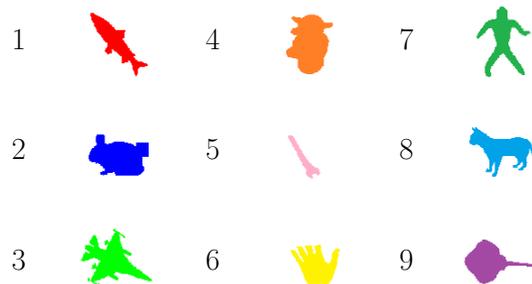


Fonte: o autor.

### ***Kimia 99***

A Kimia 99 (SEBASTIAN *et al.*, 2004) é uma base sintética de 99 formas, que possui 9 classes, cada qual com 11 formas. A Figura 9 exibe exemplares de cada classe da base Kimia 99.

Figura 9 – Exemplares de formas da base Kimia 99.

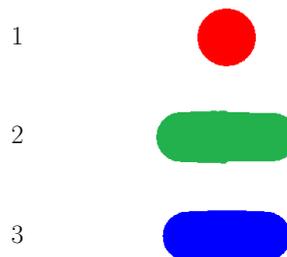


Fonte: o autor.

### ***NLM Pills***

Utilizou-se um subconjunto da base NLM Pills (USHIZIMA *et al.*, 2015), excluindo a classe "Others", uma vez que esta classe contém apenas formas com diferentes padrões que não coincidem com os das outras classes. Além disso, selecionou-se aleatoriamente 166 formas de cada uma das três classes restantes da base NLM Pills para formar uma base balanceada. Esse subconjunto consiste de 3 classes com 166 elementos cada, sendo um exemplar de cada classe exibido na Figura 10.

Figura 10 – Exemplares de formas da base NLM Pills.



Fonte: o autor.

### 3.2 Algoritmo de otimização

Souza *et al.* (2016) testaram três algoritmos para otimização do descritor NMBE: o recozimento simulado (SA), a evolução diferencial (Evolução diferencial (DE)) (STORN; PRICE, 1997) e a otimização por enxame de partículas (Otimização por enxame de partículas (PSO)) (KENNEDY; EBERHART, 1995), sendo que o SA e a DE apresentaram melhores resultados nessa tarefa. Além disso, os referidos autores argumentaram que o SA é menos dispendioso que o DE, uma vez que o primeiro algoritmo tem um custo computacional significativamente menor. Seguindo essa abordagem, a otimização do descritor NMBE será realizada com o algoritmo SA.

#### *Recozimento simulado (SA)*

O algoritmo de recozimento simulado (KIRKPATRICK *et al.*, 1982; CERNY, 1985), do inglês *simulated annealing* (SA), é uma analogia entre a simulação do recozimento de sólidos e o problema de otimização combinatória (LAARHOVEN; AARTS, 1987).

Em física da matéria condensada, recozimento é um processo no qual um sólido é aquecido, sendo a temperatura aumentada até um valor máximo no qual as partículas do sólido se encontram em elevado grau de agitação térmica. Em seguida, o sólido passa por um lento processo de resfriamento, diminuindo assim a agitação térmica e, conseqüentemente, o grau de mobilidade das partículas do material. Se a temperatura máxima for suficientemente alta e o resfriamento lento o bastante, as partículas do material tenderão a formar uma estrutura cristalina, que é um estado de mínima energia para o sistema (LAARHOVEN; AARTS, 1987; ENGELBRECHT, 2007).

Considerando um problema de minimização, o SA emprega uma estratégia de busca aleatória que não aceita somente novas soluções que diminuem a função custo, mas também admite aleatoriamente soluções com maior valor de função custo seguindo a distribuição de probabilidades de Boltzmann–Gibbs (ENGELBRECHT, 2007). Definindo a solução atual da  $t$ -ésima época do SA como  $s_a(t)$  e considerando  $s_c(t)$  a solução candidata obtida com a perturbação de  $s_a(t)$ , a probabilidade de aceitar  $s_c(t)$  no lugar de  $s_a(t)$  é dada por:

$$P(t) = \begin{cases} 1 & \text{se } f(s_c(t)) < f(s_a(t)) \\ e^{-\frac{f(s_c(t)) - f(s_a(t))}{\theta T(t)}} & \text{caso contrário} \end{cases} \quad (3.1)$$

em que  $f(\cdot)$  denota a função custo,  $\theta$  é uma constante e  $T(t)$  é a temperatura do sistema na

$t$ -ésima época. O Algoritmo 1 descreve o método de otimização SA.

---

**Algoritmo 1:** Algoritmo de recozimento simulado (SA)

---

```

Cria solução inicial,  $s_a(0)$ ;
Calcula a função custo da solução inicial,  $f(s_a(0))$ ;
 $t = 1$ ;
Define o valor inicial da temperatura,  $T(t)$ ;
repita
|    $contador = 1$ ;
|   enquanto  $contador \leq L_e$  faça
|   |   Gera solução candidata,  $s_c(t)$ ;
|   |   Calcula a função custo da solução candidata,  $f(s_c(t))$ ;
|   |   Calcula a probabilidade de aceitação usando a equação 3.1;
|   |   se  $\text{rand}\{[0, 1]\} \leq \textit{probabilidade de aceitação}$  então
|   |   |   Solução candidata se torna a atual,  $s_a(t) = s_c(t)$ ;
|   |   senão
|   |   |    $s_a(t) = s_a(t - 1)$ ;
|   |   fim
|   |    $contador = contador + 1$ ;
|   fim
|    $t = t + 1$ ;
|   Atualiza a temperatura  $T(t)$ ;
até critério de parada é verdadeiro;
retorna  $s_a(t)$  como solução;

```

---

Fonte: adaptado de (ENGELBRECHT, 2007).

A função  $\text{rand}\{[0, 1]\}$  retorna uma amostra de uma distribuição uniforme no intervalo  $[0, 1)$ . O parâmetro *epoch length* (PARK; KIM, 1998) ( $L_e$ ) é o número de vezes que a solução atual é testada por diferentes soluções candidatas em uma determinada época. Pode-se definir época como cada rodada do algoritmo SA na qual a temperatura é atualizada. Como critério de parada, pode-se estabelecer o número de épocas ( $N_e$ ), por exemplo. O índice  $t$  indica a  $t$ -ésima época na qual a temperatura é atualizada.

Segundo Engelbrecht (2007), o algoritmo SA requer a definição dos seguintes componentes:

- Um modo de representar as possíveis soluções, que usualmente é feito com um vetor de valores de ponto flutuante com cada componente correspondendo a um parâmetro para se otimizar;
- Um critério para perturbar a solução atual gerando soluções candidatas. Se for permitido perturbações suficientemente grandes em relação ao espaço de busca durante todo o processo de otimização, a convergência do algoritmo será prejudicada, já que será difícil concentrar a busca em uma determinada região com o vetor da solução se movendo para lugares distantes. Há diversas formas de perturbar a solução atual, como por exemplo, adicionar um vetor aleatório, como mostrado a seguir:

$$s_c(t) = s_a(t) + r(t)D(t), \quad (3.2)$$

em que  $s_c$  é o vetor linha que indica a solução candidata,  $s_a$  é o vetor linha que corresponde a solução atual. O índice  $t$  denota a  $t$ -ésima época,  $D(t)$  é uma matriz diagonal que define o máximo de mudança permitida para cada parâmetro que se deseja otimizar e  $r(t)$  é um vetor aleatório de mesma dimensão da solução com elementos uniformemente distribuídos no intervalo de  $[-1, 1]$ ;

- Um método para avaliar as soluções, o que é feito em geral com uma função custo;
- Um critério de resfriamento que defina a temperatura inicial e uma regra para sua diminuição ao longo das épocas. Esse procedimento tem como efeito a diminuição da probabilidade do algoritmo aceitar soluções com maior valor de função custo que a solução atual, o que determina o grau de mobilidade durante a busca. Uma exemplo desse critério seria o resfriamento exponencial dado por:

$$T(t+1) = \alpha T(t), \quad (3.3)$$

sendo  $\alpha$  a taxa de resfriamento ( $\alpha \in [0, 1]$ ) e  $t$  indica a época atual.

Adicionalmente, é necessário indicar os limites do espaço de busca, bem como um método para evitar que a solução os ultrapasse. Por exemplo, caso a solução saia do espaço de busca, pode-se atribuir o valor do limite violado à solução.

A implementação do SA seguiu o Algoritmo 1, adotando o critério de resfriamento exponencial, que atualiza a temperatura conforme a equação 3.3, e o número de épocas ( $N_e$ ) como critério de parada. Foram ajustados empiricamente os seguintes parâmetros do SA: a constante da equação 3.1 ( $\theta = 1$ ), a temperatura inicial ( $T(1) = 100$ ), taxa de resfriamento ( $\alpha = 0,9$ ) e o número de épocas ( $N_e = 100$ ). Se a solução candidata ultrapassar os limites do

espaço de busca, essa receberá o valor correspondente ao limite que ultrapassou. Como critério de perturbação da solução atual para geração de soluções candidatas, utilizou-se uma versão modificada da equação 3.2, a qual mostrou uma melhor convergência:

$$s_c(t) = s_a(t) [I + \text{diag} \{r(t)D(t)\}], \quad (3.4)$$

em que  $s_c$  indica a solução candidata (representada por um vetor linha com  $N_S$  elementos, sendo  $N_S$  o número de escalas do descritor NMBE),  $s_a$  é a solução atual (vetor linha com a mesma dimensão de  $s_c$ ), o índice  $t$  denota a  $t$ -ésima época,  $D(t)$  é uma matriz diagonal de dimensão  $N_S \times N_S$  que define o máximo de mudança permitida para cada parâmetro que se deseja otimizar,  $r(t)$  é um vetor aleatório de mesma dimensão da solução com os elementos uniformemente distribuídos no intervalo de  $(-1, 1)$ , o operador  $\text{diag}\{\cdot\}$  resulta em uma matriz diagonal cujos elementos são os componentes do vetor utilizado como operando, e  $I$  é a matriz identidade de dimensão  $N_S \times N_S$ .

Ao contrário da equação 3.2, que adiciona uma perturbação aleatória diretamente à solução atual, a equação 3.4 acrescenta à solução atual uma versão modificada dela própria. Esse procedimento garante uma melhor convergência do algoritmo, já que a influência da componente aleatória é mais limitada. Considerando os mesmos parâmetros que foram utilizados para a otimização do descritor NMBE, o desempenho do SA utilizando as equações 3.2 e 3.4 foi testado no problema de minimização da função de Rastringi para duas dimensões, a qual é dada por:

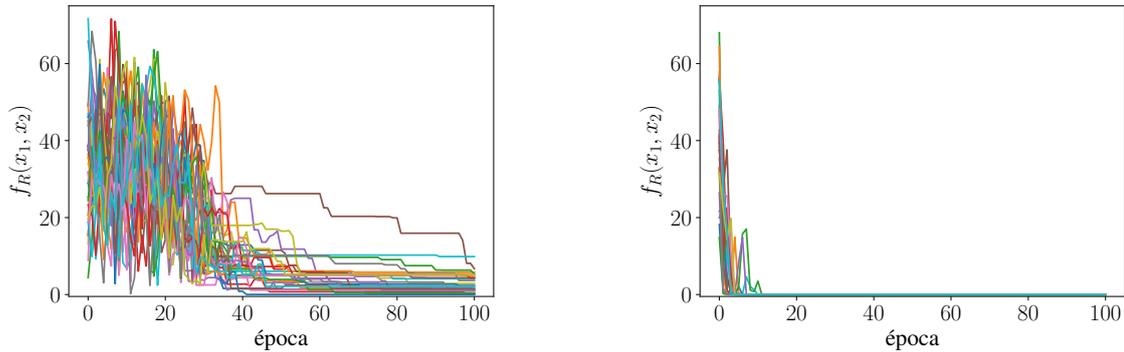
$$f_R(x_1, x_2) = 20 + x_1^2 + x_2^2 - 10 [\cos(2\pi x_1) + \cos(2\pi x_2)], \quad (3.5)$$

onde  $x_i \in [-5, 12, 5, 12]$  para  $i = 1, 2$ . Essa função possui um mínimo global em  $(0, 0)$ , sendo  $f_R(0, 0) = 0$ . A função de Rastringi é plotada versus as épocas de 30 realizações do SA (cada uma indicada por uma linha de uma cor distinta) nas Figuras 11a e 11b com o critério de perturbação dado pelas equações 3.2 e 3.4, respectivamente. Nota-se que todas as realizações do SA da Figura 11b atingem valores próximos ao mínimo da função em bem menos épocas que as da Figura 11a, o que indica uma melhor convergência do SA com a utilização da equação 3.4.

### 3.3 Mapa auto-organizável (SOM)

Para avaliar qualitativamente os resultados do descritor abordado neste trabalho, utilizou-se o mapa auto-organizável (KOHONEN, 1982), do inglês *self organized map* (Mapa auto-organizável (SOM)).

Figura 11 – Convergência do SA para a função de Rastringi



(a) Critério de perturbação do SA dado pela equação 3.2.

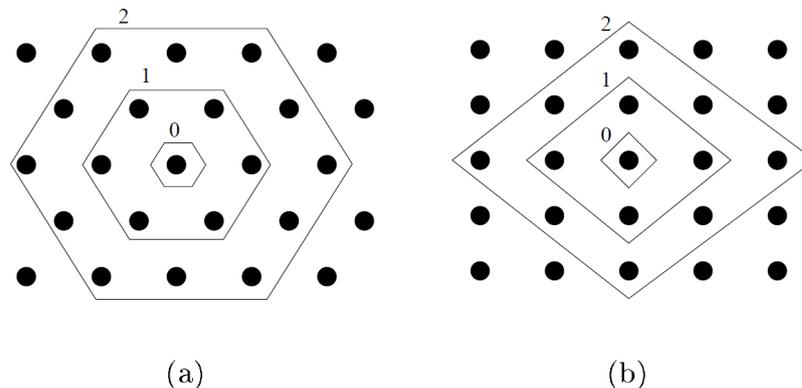
(b) Critério de perturbação do SA dado pela equação 3.4.

Fonte: o autor.

O mapa auto-organizável tem como objetivo representar elementos multidimensionais em um mapa de dimensão reduzida, preservando os relacionamentos topológicos desses elementos. Dada esta característica, o SOM pode ser usado para visualização de dados de alta dimensionalidade.

O mapa SOM é uma rede neural não supervisionada, na qual para o caso bidimensional, os neurônios são agrupados em uma grade e conectados seguindo certo critério de vizinhança. A vizinhança é definida como o conjunto dos neurônios cujos índices não variam mais que  $r$  unidades (MARQUES, 2005). Na Figura 12, são exibidas duas grades bidimensionais da rede SOM, nas quais os neurônios são representados por círculos pretos. Na Figura 12a e na Figura 12b, foram adotadas uma grade hexagonal e retangular, respectivamente. Note que os números indicam o valor de  $r$  e as linhas delimitam a vizinhança a um raio  $r$  do neurônio central (VESANTO *et al.*, 1999).

Figura 12 – Rede SOM.



Fonte: adaptado de (VESANTO *et al.*, 1999).

### ***Treinamento da rede SOM***

Cada neurônio da rede está associado a um vetor de pesos. As entradas são sorteadas aleatoriamente e apresentadas à rede. Os vetores de pesos são atualizados de acordo com as equações a seguir:

$$(i, j) = \arg \min_{p,q} \{ \| x - w_{pq} \| \}, \quad (3.6)$$

$$w_{pq} = w_{pq} + \Delta w_{pq} \quad \forall (p, q) \in V_r(i, j), \quad (3.7)$$

$$\Delta w_{pq} = \beta (x - w_{pq}), \quad (3.8)$$

sendo que  $\| \cdot \|$  representa a norma euclidiana, o par  $(p, q)$  denota as coordenadas do neurônio associadas ao vetor de peso considerado, e  $\arg \min_{p,q} \{ \cdot \}$  denota o argumento do mínimo, que é o conjunto de valores dos argumentos de determinada função para os quais ela atinge seu valor mínimo.

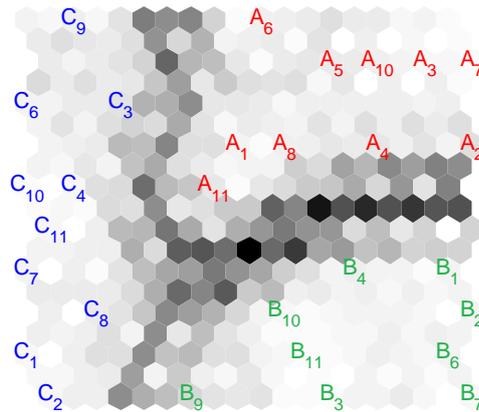
O vetor de pesos que mais se aproxima da entrada apresentada é escolhido conforme a equação 3.6, sendo que a célula correspondente, chamada de neurônio ativo, irá representar a entrada topologicamente. Os vetores de peso do neurônio ativo e de sua vizinhança, denotada por  $V_r(i, j)$ , são atualizados de acordo com as equações 3.7 e 3.8. Os parâmetros  $r$  (raio de vizinhança) e  $\beta$  (ganho de adaptação), normalmente são inicializados com valores elevados que diminuem ao longo do treino. O processo é repetido até que um critério de parada seja observado (MARQUES, 2005).

### ***Matriz U***

A Matriz U (ULTSCH; SIEMON, 1990) é uma representação gráfica da rede SOM, onde as distâncias médias entre vetores de pesos de determinada vizinhança da rede são representadas por tons de cinzas ou escalas de cores (KOHONEN, 2001).

A Figura 13 exibe a Matriz U de uma rede SOM bidimensional considerando uma estrutura local hexagonal. Nesse exemplo, tons mais claros representam distâncias menores e tons mais escuros, distâncias maiores. Os rótulos representam os diferentes elementos das entradas. A letra e a cor indicam a classe, já o índice numérico diferencia os elementos que pertencem à mesma classe. Nota-se a formação de três grupos separados por células escuras, as quais expressam distâncias maiores, enquanto que dentro de determinado grupo, as células possuem tons mais claros, indicando semelhança.

Figura 13 – Matriz U.



Fonte: o autor.

As Matrizes U deste trabalho foram geradas utilizando o SOM Toolbox (VESANTO *et al.*, 1999), sendo adotada uma rede SOM bidimensional com topologia local hexagonal. Foi utilizada a configuração padrão da SOM Toolbox: a inicialização é linear com os valores iniciais dos vetores de pesos dos neurônios iguais aos autovetores da matriz de covariância dos dados com os maiores autovalores correspondentes. Para o treino da rede na configuração padrão, o parâmetro  $r$  (raio de vizinhança) varia linearmente durante o treino, enquanto o parâmetro  $\beta$  (ganho de adaptação) obedece a seguinte equação:

$$\beta = \frac{\beta_0}{\left(1 + \frac{100m}{T_{tr}}\right)}, \quad (3.9)$$

sendo  $\beta_0$  o ganho de adaptação inicial, o qual depende do tamanho da rede ( $q_n$ ) (quantidade de neurônios),  $T_{tr}$  é a duração do treinamento e  $m$  denota uma dada iteração do treinamento (VESANTO *et al.*, 2000). O critério de parada adotado foi a duração do treinamento que é dada por:

$$T_{tr} = \max\left(\frac{20q_n}{\eta}, 1\right), \quad (3.10)$$

onde  $\eta$  é o número de objetos da base de dados, os quais servem de entradas no treinamento da rede SOM, e  $\max\{\cdot\}$  resulta no valor máximo dos operandos.

### 3.4 Índices de validação de agrupamento

Nesta seção, são apresentados índices de validação de agrupamento, os quais serão utilizados como função custo no processo de otimização do descritor de formas.

### ***Silhouette (SI)***

A *Silhouette* (SI) é uma medida da qualidade de agrupamentos definida pela seguinte equação:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}, \quad (3.11)$$

em que  $a(i)$  é a dissimilaridade média entre o elemento  $i$  e aqueles que pertencem à sua classe, e  $b(i)$  é a dissimilaridade média entre o elemento  $i$  e aqueles da classe vizinha mais próxima, que é a classe com menor média das dissimilaridades entre todos seus elementos e o elemento  $i$ , excluindo a classe a que  $i$  pertence (ROUSSEEUW, 1987). A dissimilaridade pode ser calculada através de uma medida de distância, sendo adotada a distância euclidiana neste trabalho.

A SI avalia o grau de pertencimento de certo elemento à sua classe apresentando valores no intervalo  $(-1, 1)$ . Um valor próximo de 1 indica que o elemento está bem classificado e próximo de  $-1$ , que está mal classificado. Um valor de SI próximo de 0 seria um caso intermediário, onde o elemento é praticamente tão semelhante aos elementos da classe ao qual foi atribuído, quanto aos da classe vizinha mais próxima (ROUSSEEUW, 1987).

### ***Índice de Davies-Bouldin (DB)***

O índice de Davies-Bouldin (DB) avalia a compacidade e a separação de agrupamentos. Essa métrica é baseada na razão da soma da dispersão intragrupo e a separação intergrupo, considerando as amostras da base de dados e os centróides formados por elementos da mesma classe (CHOU *et al.*, 2004). O DB é um índice não negativo no qual valores mínimos indicam grupos bem definidos (pequena dispersão intragrupo e grande separação intergrupo). Davies e Bouldin (1979) definem o DB como:

$$DB = \frac{1}{N_A} \sum_{i=1}^{N_A} \max_{j, j \neq i} \left\{ \left[ \frac{1}{\eta_i} \sum_{x \in A_i} d(x, c_i) + \frac{1}{\eta_j} \sum_{x \in A_j} d(x, c_j) \right] / d(c_i, c_j) \right\}, \quad (3.12)$$

em que  $N_A$  é o número de agrupamentos;  $A_i$  indica o  $i$ -ésimo agrupamento;  $\eta_i$  é definido como o número de objetos em  $A_i$ ;  $c_i$  é centróide de  $A_i$ ;  $d(r, s)$  identifica a distância entre os vetores  $r$  e  $s$ .

### ***Índice de Calinski-Harabasz (CH)***

O índice de Calinski-Harabasz (CH) mede a qualidade de grupos tendo por base a média da soma dos quadrados das distâncias intra e intergrupos (CALINSKI; HARABASZ,

1974). Assim como o DB, o CH considera as distâncias entre cada amostra e os centróides formados por elementos de determinada classe. No entanto, o CH também mede a separação das amostras em relação ao centróide da base inteira (BAARSCH; CELEBI, 2012). Define-se o CH como (CALINSKI; HARABASZ, 1974):

$$CH = \frac{\sum_i \eta_i d^2(c_i, c) / (N_A - 1)}{\sum_i \sum_{x \in A_i} d^2(x, c_i) / (\eta - N_A)}, \quad (3.13)$$

onde  $\eta$ : número de objetos da base de dados;  $N_A$ : número de agrupamentos;  $A_i$ :  $i$ -ésimo agrupamento;  $\eta_i$ : número de objetos em  $A_i$ ;  $c_i$ : centróide de  $A_i$ ;  $c$ : centróide da base de dados,  $d(r, s)$ : distância entre os vetores  $r$  e  $s$ .

Na Tabela 1, são mostrados o limite, o valor ótimo e a complexidade computacional de cada índice de validação de agrupamento comentado anteriormente.

Tabela 1 – Índices de validação de agrupamento

Índice	Limite	Valor ótimo	Complexidade computacional
SI	$(-1, 1)$	Máximo	$O(\eta^2)$ (SAITTA <i>et al.</i> , 2008)
DB	$(0, \infty)$	Mínimo	$O(\eta)$ (SAITTA <i>et al.</i> , 2008)
CH	$(0, \infty)$	Máximo	$O(\eta)$ (SAITTA <i>et al.</i> , 2008)

sendo que  $\eta$  denota o número de objetos da base de dados.

Como o CH leva em consideração a distância entre o centróide de cada grupo e o centróide da base inteira, certas características da base de dados podem fazer com que a qualidade dos agrupamentos indicada pelo CH seja bem distinta que a da SI e a do DB. Por exemplo, considerando duas bases de dados A e B cada uma com cinco classes, sendo que em cada classe há 100 elementos. As classes são formadas por elementos, os quais consistem de pontos no espaço bidimensional, cujas coordenadas seguem distribuições normais multivariadas com vetores de médias mostrados na Tabela 2 e matrizes de covariância dadas por  $I_2$ , sendo que  $I_2$  é a matriz identidade de dimensão  $2 \times 2$ . As Figuras 14a e 14b mostram os exemplares das bases sintéticas A e B, respectivamente, onde as diferentes classes são representadas por cores distintas e o centróide de cada base é indicado por um asterisco.

A base A simula uma base cujas classes possuem elementos com características bem distintas, sendo assim não há sobreposição entre seus elementos, já na base B, há dois pares de classes com bastante sobreposição, o que simula classes com características muito semelhantes, e

uma classe (Vermelha) com elementos bem distantes do resto da base. Na Tabela 3, são exibidos os valores da SI, DB e CH para as bases A e B, sendo que os melhores resultados para a SI e o CH são os maiores, enquanto para o DB são os mínimos. A distância significativa entre os elementos da classe Vermelha em relação aos outros afeta bastante a posição do centróide da base e contribui consideravelmente para um maior valor de CH que o da base A, mesmo a base B apresentando sobreposição entre seus elementos, o que não ocorre com a base A. A Silhouette e o DB apresentam resultados inferiores para a base B em comparação à base A, já que esses índices não contabilizam a média da distância entre o centróide de cada grupo e o centróide da base inteira e, assim, não são tão afetados pelo posicionamento da classe Vermelha.

Tabela 2 – Vetores de média das distribuições normais que geram as bases A e B.

Classe	Base A	Base B
Vermelha	[10, 20]	[70, 80]
Azul	[20, 50]	[22, 52]
Verde	[30, 40]	[20, 50]
Laranja	[40, 30]	[32, 42]
Roxa	[50, 20]	[30, 40]

Fonte: o autor.

Tabela 3 – Índices de agrupamento para as bases A e B

Índice	Base A	Base B
SI	0,88	0,51
DB	0,16	0,76
CH	20579,80	33156,48

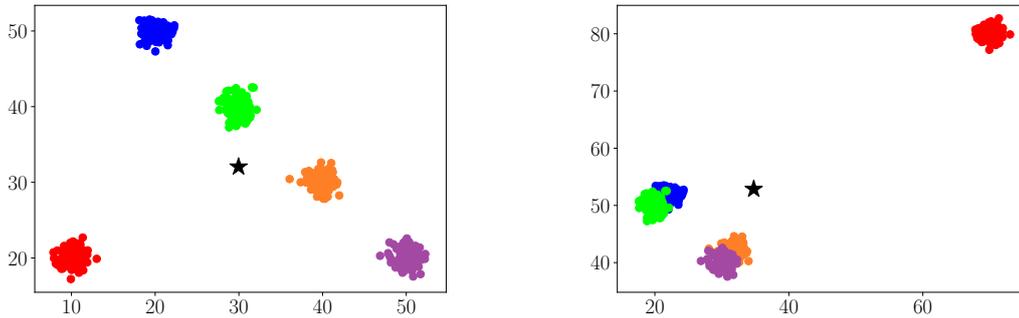
Fonte: o autor.

### 3.5 Recuperação

A medida *Bulls-eye* foi definida na literatura para avaliar descritores de formas em experimentos recuperação de formas baseada em conteúdo (CBIR) com a base de imagens MPEG-7 (LATECKI *et al.*, 2000), no entanto pode ser generalizada para outras bases. Nesse trabalho, o *Bulls-eye* será utilizado para quantificar o desempenho do descritor NMBE nos experimentos recuperação.

Considerando uma base de imagens balanceada, ou seja, com a mesma quantidade de objetos em todas as suas classes, pode-se definir  $N_B$  como o número de imagens da base e  $N_C$

Figura 14 – Exemplos das bases sintéticas A e B



(a) Base A.

(b) Base B.

Fonte: o autor.

como o número de elementos de determinada classe. Realizando um experimento CBIR nessa base, recupera-se as  $2N_C$  imagens mais semelhantes a uma imagem de consulta e contabiliza-se quantas foram recuperadas pertencentes à classe da imagem de consulta utilizada. Esse processo é executado para todas as  $N_B$  imagens da base, tomando-se como imagem de consulta cada uma das imagens da própria base. O número de imagens recuperadas pertencentes à mesma classe das imagens de consulta, em todas as iterações, é denotado por  $Q_A$  (a quantidade de acertos). Assim a medida *Bulls-eye* ( $BE$ ) é definida por:

$$BE = \frac{Q_A}{N_B N_C}. \quad (3.14)$$

O Algoritmo 2 descreve o cálculo do *Bulls-eye*, sendo considerado que  $I(i)$  é a imagem de consulta da iteração  $i$ .

---

**Algoritmo 2:** Bulls-eye.

---

$Q_A = 0$ ;

$i = 1$ ;

**repita**

Recupera-se as  $2N_C$  imagens mais semelhantes a  $I(i)$ ;

$Q_A = Q_A +$  quantidade de imagens recuperadas da mesma classe de  $I(i)$ ;

**até**  $i \leq N_B$ ;

$BE = \frac{Q_A}{N_B N_C}$ ;

---

A medida *Bulls-eye* apresenta valores no intervalo  $(0, 1]$ , valores próximos a 0 indicam que o descritor utilizado no experimento CBIR tem um desempenho ruim na descrição das imagens da base considerada, e valores próximo de 1 indicam um bom desempenho.

No cálculo do *Bulls-eye*, foi adotada como medida de dissimilaridade entre os vetores de características das imagens a distância euclidiana, a qual é dada pela equação 2.3.

### 3.6 Classificação

Para quantificar o desempenho do descritor nos experimentos de classificação foi utilizada a Acurácia da classificação do algoritmo  $k$  vizinhos mais próximos. Ambos são detalhados a seguir.

#### *k vizinhos mais próximos (kNN)*

O algoritmo  $k$  vizinhos mais próximos (FIX; HODGES JR, 1951), do inglês *k-nearest neighbors* (kNN), classifica um padrão de teste  $z$  com base nas classes dos vizinhos mais próximos do conjunto de treinamento  $C_T$ .

Considerando  $C_T$  dado por:

$$C_T = (x_1, y_1), (x_2, y_2), \dots, (x_Q, y_Q), \quad (3.15)$$

sendo  $x_i$  a  $i$ -ésima amostra de treinamento de dimensão  $M$ ,  $y_i$  é a classe de  $x_i$  e  $Q$  é a número de amostras do conjunto de treinamento  $C_T$ . O kNN atribui a  $z$  a classe de maior frequência nas  $k$  amostras pertencentes a  $C_T$  com menor distância a  $z$ . Dessa forma, para a utilização do kNN é necessário amostras rotuladas para formarem o conjunto de treinamento, a escolha de uma determinada distância e a definição do número de vizinhos  $k$  (MURTY; DEVI, 2011). Geralmente, valores ímpares são utilizados para  $k$  a fim de evitar empates (RUSSELL; NORVIG, 2009). Neste trabalho, será utilizado o classificador kNN com  $k = 3$  e a distância euclidiana.

O kNN é um método de aprendizagem baseado em instâncias (ou aprendizagem preguiçosa) porque não constrói um modelo interno geral utilizando o conjunto de treinamento, mas simplesmente compara o padrão de teste com as amostras de treinamento (RAJAGURU; PRABHAKAR, 2017).

#### *Acurácia da classificação*

A Acurácia da classificação é definida pela razão entre a quantidade de elementos de uma determinada base de dados que tiveram a classe corretamente predita e o número total dos elementos da base. Os valores da Acurácia estão no intervalo  $[0, 1]$ .

### 3.7 Experimentos

Realizou-se experimentos de classificação e recuperação de imagens baseada em conteúdo (CBIR) com quatro bases de imagens públicas: ShapeCN, Swedish, Kimia 99 e NLM Pills.

Avaliou-se os experimentos de recuperação com a métrica Bulls-eye. Nos experimentos de classificação, dividiu-se cada base de assinaturas (formada pelas características extraídas de todas as formas da base de imagens pelos descritores) em dois subconjuntos aleatórios para treino (70%) e teste (30%) do classificador kNN ( $k = 3$ ), repetindo-se esse processo 100 vezes para se calcular a Acurácia média de classificação.

Utilizou-se o descritor NMBE com cinco escalas em todos os experimentos com as quatro bases de imagens. Cada escala corresponde a um valor do parâmetro  $\sigma$  (equação 2.25) a ser otimizado pelo SA dentro do seguinte espaço de busca:  $[0, 4, 100]$ . O processo de otimização do descritor NMBE é repetido 30 vezes para cada função custo e cada base de imagens. De posse dos resultados dessas 30 realizações, calcula-se a média e a variância do Bulls-eye e da Acurácia média (obtida pelas 100 rodadas de classificação do kNN com subconjuntos de treino e teste aleatórios para uma única base de assinaturas). É importante notar que, apesar de se trabalhar com a média das Acurácias médias do kNN, essa medida será indicada nos resultados apenas como Acurácia.

Além da avaliação quantitativa, realizada com os experimentos de classificação e recuperação, foi feita a avaliação qualitativa do descritor NMBE otimizado por meio da Matriz U, sendo gerada uma Matriz U relativa a um conjunto de assinaturas de certa base de imagens otimizado por uma determinada função custo.

Todos os resultados de classificação, recuperação e da Matriz U obtidos com a otimização serão confrontados com o método de ajuste de escalas proposto por Cesar Jr e Costa (1997). Os valores mínimo e máximo das escalas desse método ( $\sigma_{min}$  e  $\sigma_{max}$  da equação 2.30, respectivamente) foram definidos iguais aos valores mínimo e máximo do espaço de busca do método otimizado (0, 4 e 100, respectivamente), a fim de que a comparação entre eles seja justa.

## 4 RESULTADOS E DISCUSSÕES

As Tabelas 4 e 5 exibem os resultados dos experimentos de recuperação e classificação conduzidos para bases de assinaturas do descritor NMBE com as escalas ajustadas por otimização e pelo método proposto por Cesar Jr e Costa (1997). Como métricas para avaliação dos resultados dos experimentos, foram utilizadas as medidas Bulls-eye e Acurácia da classificação, respectivamente, na recuperação e classificação de formas. No ajuste do descritor NMBE por otimização, foram testadas três funções custos: SI, DB e CH. Como índices de validação da qualidade de agrupamentos, espera-se que tais funções guiem a otimização para que o ajuste das escalas seja tal que o descritor represente os detalhes das formas de cada base, maximizando a separação entre aquelas pertencentes a classes distintas e minimizando a separação entre formas de uma mesma classe.

Tabela 4 – Bulls-eye(%)

Base de imagens	SI	DB	CH	Método proposto em (CESAR JR; COSTA, 1997)
ShapeCN	70,91 ± 1,81	71,67 ± 1,61	41,98 ± 5,86	63,39 ± 0,00
Swedish	86,53 ± 1,48	76,70 ± 3,38	63,90 ± 3,07	76,08 ± 0,00
Kimia 99	90,20 ± 1,30	90,99 ± 0,92	63,54 ± 0,00	89,26 ± 0,00
NLM Pills	99,95 ± 0,01	99,95 ± 0,09	99,96 ± 0,00	97,57 ± 0,00

Fonte: o autor.

Tabela 5 – Acurácia (%)

Base de imagens	SI	DB	CH	Método proposto em (CESAR JR; COSTA, 1997)
ShapeCN	71,00 ± 2,80	75,16 ± 1,93	39,28 ± 6,27	67,67 ± 0,00
Swedish	86,53 ± 1,97	82,98 ± 1,94	62,23 ± 8,38	85,99 ± 0,00
Kimia 99	84,53 ± 3,38	85,83 ± 2,42	39,24 ± 0,90	76,67 ± 0,00
NLM Pills	99,24 ± 0,13	99,24 ± 0,09	99,34 ± 0,04	99,36 ± 0,00

Fonte: o autor.

Os resultados de recuperação e classificação para as duas bases de folhas e a Kimia 99 demonstram que o descritor NMBE otimizado com a função custo CH gerou características que alcançaram valores de Bulls-eye e Acurácia inferiores àqueles obtidos pela metodologia não otimizada de Cesar Jr e Costa (1997). Para essas mesmas bases, os resultados obtidos com o descritor NMBE otimizado com as funções custo SI e DB são melhores ou similares aos obtidos com o método proposto por Cesar Jr e Costa (1997). Quanto à base NLM Pills, que possui

apenas três classes de formas com padrões bem definidos e com pequenas variações entre as classes, os resultados obtidos nos experimentos foram similares em todas as abordagens para ajuste dos parâmetros do descritor. Portanto, as escalas otimizadas tendo CH como função custo não conseguiram capturar adequadamente os detalhes globais e locais das formas nas bases de folhas e na base Kimia 99. Dessa forma, conclui-se que para se obter resultados satisfatórios com a aplicação da metodologia proposta por Souza *et al.* (2016), deve-se escolher criteriosamente uma função custo que melhor se adeque ao domínio da aplicação.

A desvantagem da abordagem de Cesar Jr e Costa (1997) é não customizar o descritor a uma aplicação específica, já que o método considera apenas o número de escalas e seus valores mínimo e máximo, não levando em conta a base na qual será aplicado. Já a metodologia de otimização proposta por Souza *et al.* (2016) customiza os parâmetros de escala do descritor à aplicação dentro de um espaço de busca e um número pré-estabelecido de escalas, no entanto necessita-se escolher uma função custo adequada.

As diferenças significativas encontradas na maioria dos resultados de recuperação e classificação indicam que escalas distintas do mesmo descritor são necessárias para revelar os detalhes específicos das formas para cada base. Investigações relacionadas a esse trabalho são encontradas em (MILLIGAN; COOPER, 1985) e (SAITTA *et al.*, 2008), onde índices de validação de agrupamento foram testados para determinar o número de agrupamentos de bases de dados. Esses trabalhos concluíram que um índice adequado para uma determinada base pode não ser apropriado para outras.

Como já mostrado na Tabela 1, a complexidade computacional da SI, do DB e do CH são  $O(\eta^2)$ ,  $O(\eta)$  e  $O(\eta)$ , respectivamente. Nota-se que o DB oferece um interessante compromisso entre a complexidade computacional e o desempenho nos experimentos de recuperação e classificação. As Tabelas 4 e 5 mostram que os resultados do descritor otimizado usando o DB e a SI superaram os obtidos com o CH; no entanto, a complexidade computacional do DB é significativamente menor que a da SI. Embora o CH e o DB sejam computacionalmente equivalentes, o CH apresentou resultados inferiores nos experimentos de recuperação e classificação com as bases ShapeCN, Swedish e Kimia 99. O CH, diferente da SI e do DB, considera em seu cálculo a distância entre o centróide de cada grupo e o centróide da base inteira. Dessa forma, determinadas características da base de dados podem implicar que a qualidade dos agrupamentos indicada pelo CH seja bem distinta que a da SI e a do DB, como já comentado na Seção 3.4.

A Figura 15 mostra exemplos de experimentos de recuperação usando as metodo-

logias não otimizada e otimizada para a base Swedish. Cada linha identifica, em sequência, a função custo ou o método não otimizado, a imagem da consulta e as 9 imagens recuperadas cujas características são mais próximas às da imagem de consulta, com exceção dela própria. Para facilitar a análise dos resultados, as folhas de uma mesma classe foram mostradas na mesma cor, conforme o padrão mencionado no Capítulo 3. Esses resultados ilustram que a seleção de uma função custo adequada é essencial para a otimização dos parâmetros de um descritor, já que o descritor NMBE otimizado com a SI e o DB recuperou corretamente a maioria das formas, com apenas uma correspondência incorreta para cada, enquanto o método não otimizado recuperou somente duas formas da mesma classe e o CH, apenas uma.

Figura 15 – Resultados dos experimentos de recuperação para um exemplar da base Swedish usando a otimização com a SI, o DB e o CH, além do método não otimizado proposto por Cesar Jr e Costa (1997).

SI										
DB										
CH										
Método não otimizado										

Fonte: o autor.

As Figuras 16, 17, 18 e 19 exibem as Matrizes U para as bases ShapeCN, Swedish, Kimia 99 e NLM Pills. É importante notar que os resultados otimizados da Figura 15 e das Matrizes U são obtidos com uma única rodada do processo de otimização para cada função custo e para o método proposto por Cesar Jr e Costa (1997), enquanto os resultados mostrados nas Tabelas 4 e 5 são resultados médios de 30 rodadas do SA. Além disso, todas as Matrizes U exibem os valores da medida Bulls-eye correspondente às suas características nas legendas.

As Figuras 16a, 16b, 16c e 16d mostram as Matrizes U com o descritor NMBE otimizado pelo SA com as funções custo SI, DB, CH e o método proposto por Cesar Jr e Costa (1997) para a base ShapeCN, respectivamente. Os arranjos de agrupamentos nas Figuras 16a,

16b e 16d são mais organizados do que na Figura 16c. Note que as formas estão com uma cor particular para cada classe, conforme o padrão mencionado no Capítulo 3.

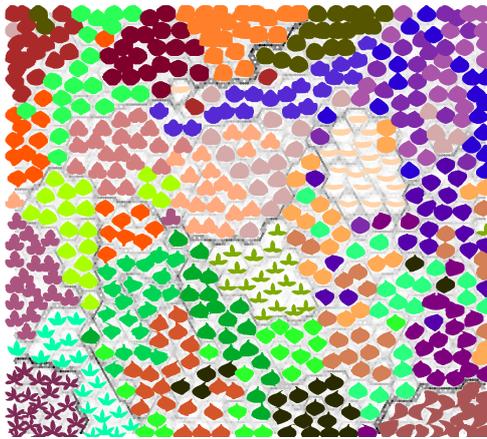
Seguindo o mesmo padrão, as Figuras 17a, 17b, 17c e 17d apresentam as Matrizes U para a base Swedish. Pode-se notar que os resultados provenientes da otimização com o CH, novamente, são inferiores tanto aos que foram otimizados com a SI e o DB como o método não otimizado.

Da mesma forma, a Figura 18 exibe as quatro Matrizes U para a base Kimia 99. Na Figura 18c, os agrupamentos misturam elementos de diversas classes, sendo esses resultados inferiores aos exibidos nas Figuras 18a, 18b e 18d. Nota-se que as Matrizes U usando a SI e o DB alcançaram arranjos semelhantes e a disposição dos agrupamentos do CH foi pior que a do método proposto por Cesar Jr e Costa (1997).

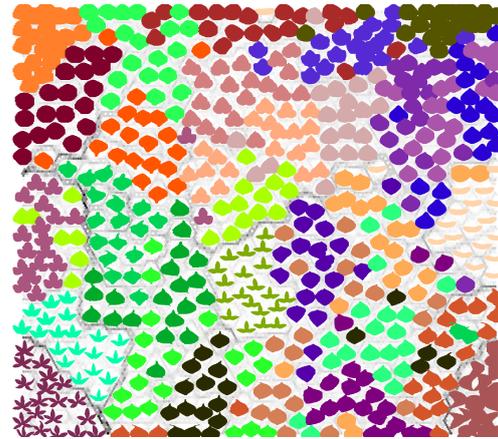
As Figuras 16c, 17c e 18c apresentam regiões sem bordas com um número considerável de elementos de diferentes classes misturados. Isso indica que a função custo CH não é adequada para otimizar os parâmetros de escala do descritor NMBE para fornecer uma representação satisfatória das formas das bases ShapeCN, Swedish e Kimia 99. Assim, mostra-se o SA requer uma função custo adequada para ajustar as escalas do descritor.

As Matrizes U apresentadas na Figura 19 demonstram que as diferentes abordagens alcançaram resultados similares; isso pode ser devido à similaridade intraclasse e à geometria das formas na base NLM Pills. Nesse caso, a metodologia de otimização obteve resultados equivalentes, independente da função custo, e, portanto, essa base não se beneficiou com a otimização. Nota-se que as Matrizes U das Figuras 16, 17, 18 e 19 estão de acordo com os resultados apresentados nas Tabelas 4 e 5.

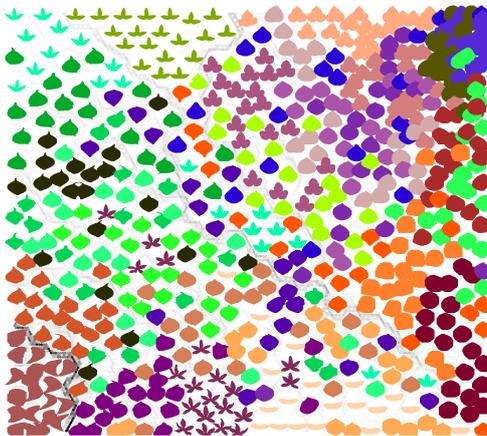
Figura 16 – Matrizes U para a base ShapeCN usando a otimização com três funções custo diferentes e o método não otimizado proposto por Cesar Jr e Costa (1997).



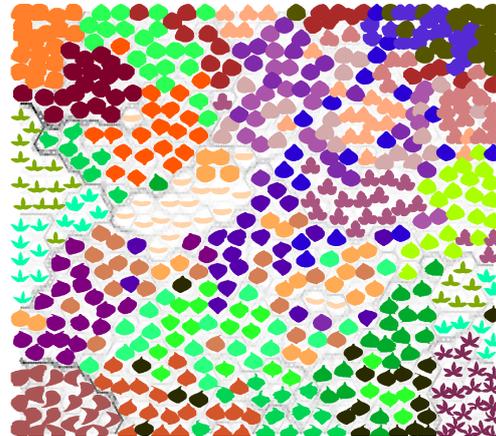
(a) Otimizado com a função custo SI (Bulls-eye = 74,43%).



(b) Otimizado com a função custo DB (Bulls-eye = 72,06%).



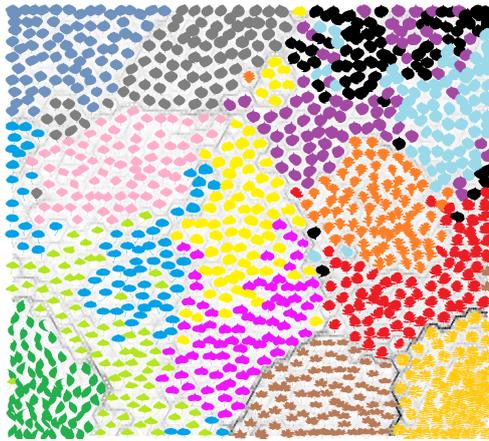
(c) Otimizado com a função custo CH (Bulls-eye = 39,20%).



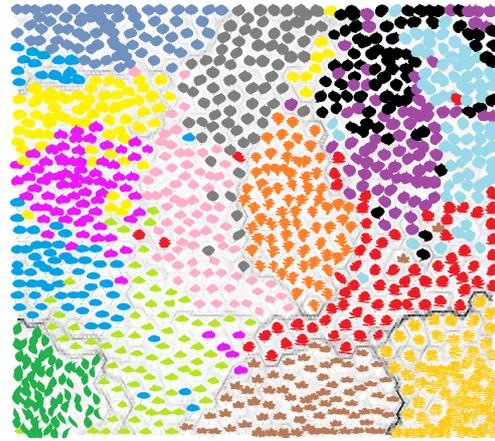
(d) Método não otimizado proposto por Cesar Jr e Costa (1997) (Bulls-eye = 63,39%).

Fonte: o autor.

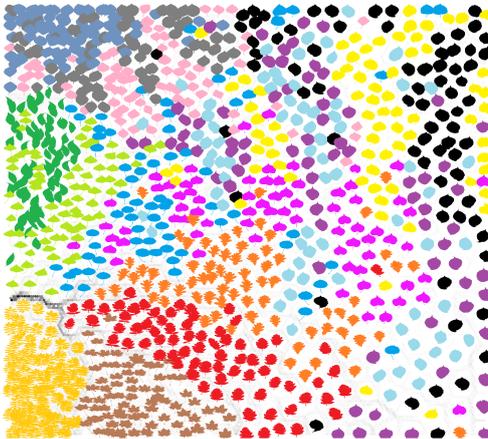
Figura 17 – Matrizes U para a base Swedish usando a otimização com três funções custo diferentes e o método não otimizado proposto por Cesar Jr e Costa (1997).



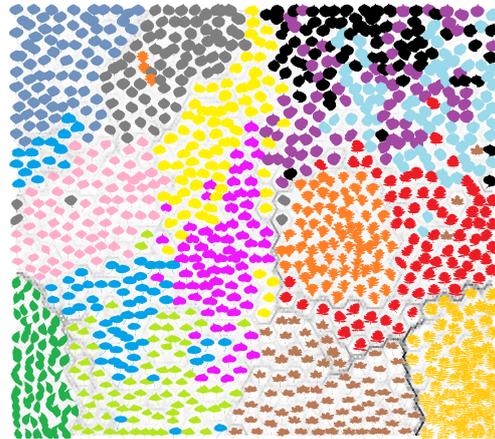
(a) Otimizado com a função custo SI (Bulls-eye = 82,26%).



(b) Otimizado com a função custo DB (Bulls-eye = 78,29%).



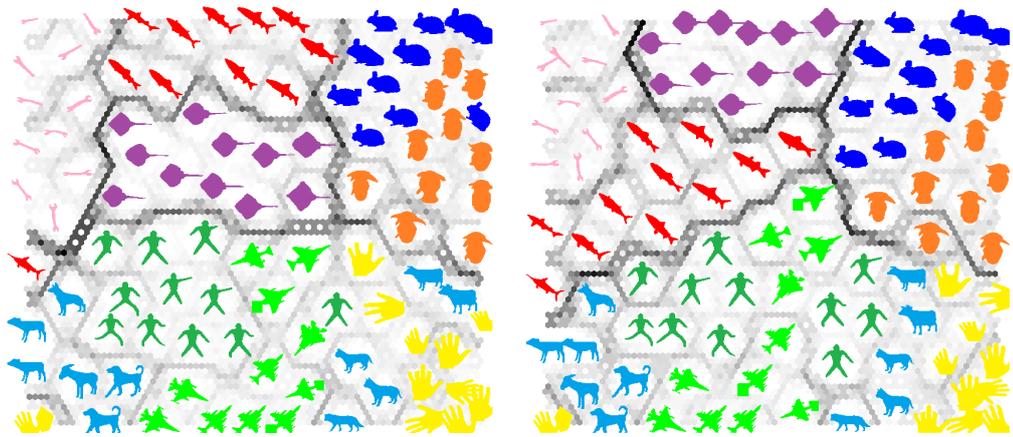
(c) Otimizado com a função custo CH (Bulls-eye = 64,31%).



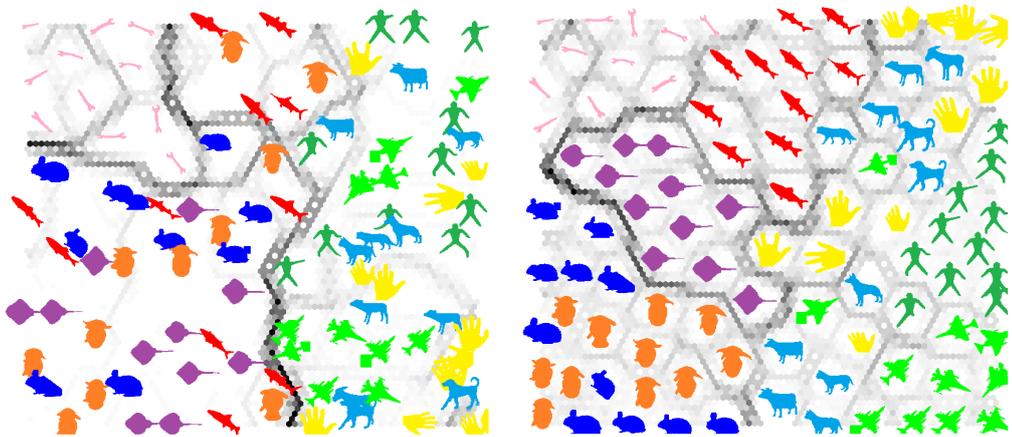
(d) Método não otimizado proposto por Cesar Jr e Costa (1997) (Bulls-eye = 76,08%).

Fonte: o autor.

Figura 18 – Matrizes U para a base Kimia 99 usando a otimização com três funções custo diferentes e o método não otimizado proposto por Cesar Jr e Costa (1997).



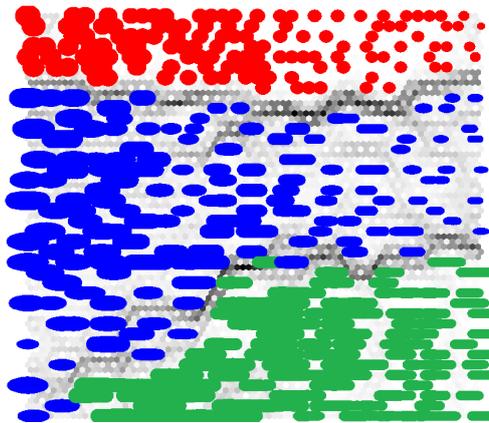
(a) Otimizado com a função custo SI (Bulls-eye = 91,55%). (b) Otimizado com a função custo DB (Bulls-eye = 92,56%).



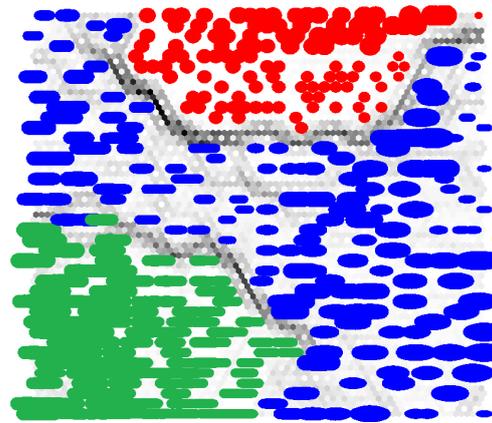
(c) Otimizado com a função custo CH (Bulls-eye = 63,54%). (d) Método não otimizado proposto por Cesar Jr e Costa (1997) (Bulls-eye = 89,26%).

Fonte: o autor.

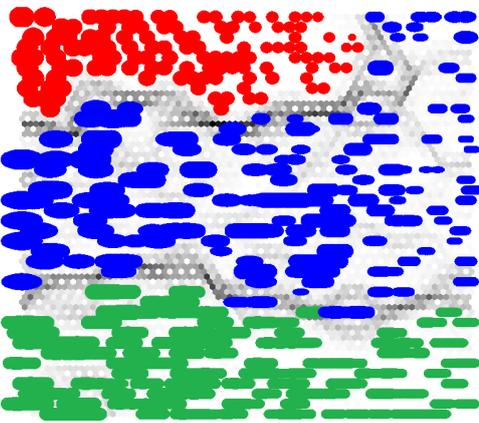
Figura 19 – Matrizes U para a base NLM Pills usando a otimização com três funções custo diferentes e o método não otimizado proposto por Cesar Jr e Costa (1997).



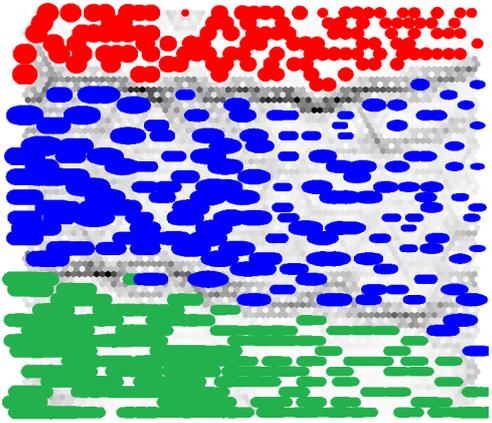
(a) Otimizado com a função custo SI (Bulls-eye = 99,96%).



(b) Otimizado com a função custo DB (Bulls-eye = 99,96%).



(c) Otimizado com a função custo CH (Bulls-eye = 99,96%).



(d) Método não otimizado proposto por Cesar Jr e Costa (1997) (Bulls-eye = 97,57%).

Fonte: o autor.

## 5 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho, investigou-se uma metodologia para otimização de parâmetros de um descritor de formas multiescala, na qual três funções custo foram aplicadas em quatro bases de imagens. O objetivo foi mostrar que uma certa base requer uma função custo apropriada para otimização dos parâmetros do descritor de formas e, portanto, extrair o melhor conjunto de características multiescalas. Foram alcançados resultados semelhantes de recuperação e classificação tanto para o descritor otimizado com as três funções custo consideradas como para o método não otimizado na base NLM Pills, cujas formas de uma mesma classe são notavelmente similares e cada classe tem um padrão bem característico. Isso sugere que, para esses padrões de forma, o ajuste dos parâmetros não é crucial, porque a discriminação de diferentes classes não é uma tarefa complicada para o descritor multiescala. Quanto às bases de folhas e a Kimia 99, observa-se que a SI e o DB apresentaram melhores resultados do que o método não otimizado. No entanto, a função custo SI apresenta uma ordem de complexidade maior do que o DB e o CH. Quando empregou-se como função custo o CH na otimização do descritor para as bases de folhas e a Kimia 99, obteve-se resultados inferiores aos de todas as outras abordagens, indicando que sua utilização não é adequada para esse problema. Nota-se portanto que o papel da função de custo na metodologia de otimização de parâmetros é crucial para alcançar o melhor desempenho de um descritor de formas.

Como perspectivas de trabalhos para a continuidade desse projeto, pode-se citar:

- Investigar a otimização em outros descritores multiescalas;
- Otimizar o descritor com funções custo que não são índices de validação de agrupamentos, por exemplo o próprio Bulls-eye e a Acurácia da classificação.

## REFERÊNCIAS

- ARAÚJO, F. H. D.; CARNEIRO, A. C.; SILVA, R. R. V.; MEDEIROS, F. N. S.; USHIZIMA, D. M. Redes neurais convolucionais com tensorflow: Teoria e prática. In: SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. **III Escola Regional de Informática do Piauí. Livro Anais - Artigos e Minicursos**. [S.l.]: Sociedade Brasileira de Computação, 2017. v. 1, p. 382–406. ISBN 978-85-7669-395-6.
- BAARSCH, J.; CELEBI, M. E. Investigation of internal validity measures for k-means clustering. In: **Proceedings of the International MultiConference of Engineers and Computer Scientists**. [S.l.: s.n.], 2012. v. 1, p. 14–16.
- 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.
- CALINSKI, T.; HARABASZ, J. A dendrite method for cluster analysis. **Communications in Statistics-theory and Methods**, Taylor & Francis, v. 3, n. 1, p. 1–27, 1974.
- CARNEIRO, A. C.; LOPES, J. G. F.; ARAÚJO, F. H. D.; SILVA, R. R. V.; PASSARINHO, C. J. P.; ROCHA NETO, J. F. S.; MEDEIROS, F. N. S. Análise de fotografias de pílulas por redes neurais convolucionais. In: UNIVERSIDADE FEDERAL DO ABC. **8º Simpósio de Instrumentação e Imagens Médicas (SIIM) e o 7º Simpósio de Processamento de Sinais (SPS)**. [S.l.], 2017.
- CARNEIRO, A. C.; LOPES, J. G. F.; ROCHA NETO, J. F. S.; SOUZA, M. M. S.; MEDEIROS, F. N. S.; BEZERRA, F. N. On the evaluation of cost functions for parameter optimization of a multiscale shape descriptor. In: IEEE. **International Symposium on Signal Processing and Information Technology (ISSPIT)**. [S.l.], 2017.
- CERNY, V. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. **Journal of optimization theory and applications**, Springer, v. 45, n. 1, p. 41–51, 1985.
- CESAR JR, R. M.; COSTA, L. F. Towards effective planar shape representation with multiscale digital curvature analysis based on signal processing techniques. **Pattern Recognition**, Elsevier, v. 29, n. 9, p. 1559–1569, 1996.
- CESAR JR, R. M.; COSTA, L. F. Application and assessment of multiscale bending energy for morphometric characterization of neural cells. **Review of Scientific Instruments**, AIP Publishing, v. 68, n. 5, p. 2177–2186, 1997. ISSN 1089-7623.
- CHOU, C. H.; SU, M. C.; LAI, E. A new cluster validity measure and its application to image compression. **Pattern Analysis and Applications**, Springer, v. 7, n. 2, p. 205–220, 2004.
- COSTA, L. F.; CESAR JR, R. M. **Shape Classification and Analysis: Theory and Practice**. 2. ed. Boca Raton, Florida, FL, USA: CRC Press Taylor & Francis Group, 2009. ISBN 978-0-8493-7929-1.
- DAVIES, D. L.; BOULDIN, D. W. A cluster separation measure. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, PAMI-1, n. 2, p. 224–227, 1979.
- ENGELBRECHT, A. P. **Computational Intelligence: An Introduction**. 2. ed. [S.l.]: John Wiley & Sons, 2007. ISBN 978-0-470-03561-0.

FIX, E.; HODGES JR, J. L. **Discriminatory analysis-nonparametric discrimination: consistency properties**. [S.l.], 1951.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 3. ed. Upper Saddle River, NJ 07458: Prentice Hall, 2007. ISBN 013168728X.

KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: IEEE. **Proceedings of the IEEE international conference on neural networks IV**. [S.l.], 1995. p. 1942–1948.

KIRKPATRICK, S.; GELATT, C. D.; VECCHI, M. P. **Optimization by Simulated Annealing**. Yorktown Heights, NY, USA, 1982.

KOHONEN, T. Analysis of a simple self-organizing process. **Biological cybernetics**, Springer, v. 44, n. 2, p. 135–140, 1982.

KOHONEN, T. **Self-organizing Maps**. 3. ed. Berlin, GER: Springer Series in Information Sciences, 2001. v. 30. ISBN 3-540-67921-9.

LAARHOVEN, P. J. M.; AARTS, E. H. L. (Ed.). **Simulated Annealing: Theory and Applications**. Norwell, MA, USA: Springer, 1987. (Mathematics and its applications). ISBN 978-90-481-8438-5.

LATECKI, L. J.; LAKAMPER, R.; ECKHARDT, T. Shape descriptors for non-rigid shapes with a single closed contour. In: **Proceedings of IEEE Conference on Computer Vision and Pattern Recognition**. [S.l.]: IEEE, 2000. v. 1, p. 424–429.

LING, H.; JACOBS, D. W. Shape classification using the inner-distance. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, v. 29, n. 2, p. 286–299, 2007.

LIU, Y.; LI, Z.; XIONG, H.; GAO, X.; WU, J. Understanding of internal clustering validation measures. In: IEEE. **2010 IEEE International Conference on Data Mining**. [S.l.], 2010. p. 911–916.

MARQUES, J. S. **Reconhecimento de padrões: métodos estatísticos e neuronais**. 2. ed. Lisboa, POR: IST (Instituto Superior Técnico) PRESS, 2005. ISBN 9789728469085.

MARQUES, O.; FURHT, B. **Content-based image and video retrieval**. Boca Raton, FL, USA: Springer Science & Business Media, 2002. v. 21. ISBN 978-1-4613-5338-6.

MILLIGAN, G. W.; COOPER, M. C. An examination of procedures for determining the number of clusters in a data set. **Psychometrika**, Springer, v. 50, n. 2, p. 159–179, 1985.

MURTY, M. N.; DEVI, V. S. **Pattern Recognition: An Algorithmic Approach**. 1. ed. [S.l.]: Springer, 2011. (Undergraduate Topics in Computer Science). ISBN 978-0-85729-494-4.

PARK, M.-W.; KIM, Y.-D. A systematic procedure for setting parameters in simulated annealing algorithms. **Computers & Operations Research**, Elsevier, v. 25, n. 3, p. 207–217, 1998.

PEDRINI, H.; SCHWARTZ, W. R. **Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações**. 1. ed. São Paulo, SP, BRA: Editora Thomson Learning Edições Ltda, 2007. ISBN 978-85-221-0595-3.

RAJAGURU, H.; PRABHAKAR, S. K. **KNN Classifier and K-Means Clustering for Robust Classification of Epilepsy from EEG Signals: A Detailed Analysis**. [S.l.]: Anchor Academic Publishing, 2017.

- ROUSSEEUW, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of Computational and Applied Mathematics**, Elsevier, v. 20, p. 53–65, 1987.
- RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 3. ed. Upper Saddle River, NJ, USA: Pearson, 2009. (Prentice Hall Series in Artificial Intelligence). ISBN 978-0-13-604259-4.
- SAITTA, S.; RAPHAEL, B.; SMITH, I. F. A comprehensive validity index for clustering. **Intelligent Data Analysis**, IOS Press, v. 12, n. 6, p. 529–548, 2008.
- SEBASTIAN, T.; KLEIN, P.; KIMIA, B. Recognition of shapes by editing their shock graphs. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 26, n. 5, p. 550–571, 2004.
- SÖDERKVIST, O. **Computer vision classification of leaves from swedish trees**. Dissertação (Mestrado), 2001.
- SOUZA, M. M.; MEDEIROS, F. N.; RAMALHO, G. L.; PAULA JR, I. C.; OLIVEIRA, I. N. Evolutionary optimization of a multiscale descriptor for leaf shape analysis. **Expert Systems with Applications**, Elsevier, v. 63, p. 375–385, 2016.
- STORN, R.; PRICE, K. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. **Journal of global optimization**, Springer, v. 11, n. 4, p. 341–359, 1997.
- ULTSCH, A.; SIEMON, H. P. Kohonen’s self organizing feature maps for exploratory data analysis. In: **Proceedings of International Neural Networks Conference**. Paris: Kluwer Academic Press, 1990. v. 1, p. 305–308.
- USHIZIMA, D.; CARNEIRO, A.; SOUZA, M.; MEDEIROS, F. Investigating pill recognition methods for a new national library of medicine image dataset. In: \_\_\_\_\_. **International Symposium on Visual Computing**. [S.l.]: Springer, 2015. v. 9475, p. 410–419.
- VESANTO, J.; HIMBERG, J.; ALHONIEMI, E.; PARHANKANGAS, J. Self-organizing map in matlab: the som toolbox. In: **Proceedings of the Matlab DSP conference**. [S.l.: s.n.], 1999. v. 99, p. 16–17.
- VESANTO, J.; HIMBERG, J.; ALHONIEMI, E.; PARHANKANGAS, J. **SOM Toolbox for Matlab 5**. [S.l.], 2000.
- WANG, B.; BROWN, D.; GAO, Y.; SALLE, J. L. March: Multiscale-arch-height description for mobile retrieval of leaf images. **Information Sciences**, Elsevier, v. 302, p. 132–148, 2015.
- WANG, J.; BAI, X.; YOU, X.; LIU, W.; LATECKI, L. J. Shape matching and classification using height functions. **Pattern Recognition Letters**, Elsevier, v. 33, n. 2, p. 134–143, 2012.
- WANG, J. Z. **Integrated Region-Based Image Retrieval**. [S.l.]: Springer Science & Business Media, 2001. v. 11. ISBN 978-1-4615-1641-5.
- YOUNG, I. T.; WALKER, J. E.; BOWIE, J. E. An analysis technique for biological shape. i. **Information and control**, Elsevier, v. 25, n. 4, p. 357–370, 1974.

ZHANG, D.; LU, G. Review of shape representation and description techniques. **Pattern recognition**, Elsevier, v. 37, n. 1, p. 1–19, 2004.