

# Desenvolvimento e Análise de Descritores Geométricos para Reconhecimento de Padrões.

Paulo N. S. do Carmo<sup>1</sup>, Wener B. de Sampaio<sup>1</sup>

<sup>1</sup>Engenharia da Computação  
Universidade Federal do Maranhão (UFMA) – São Luís, MA – Brazil

paulonayron.s@hotmail.com, wener.sampaio@ufma.br

**Abstract.** *In Digital Image Processing and Computational Vision, descriptors are often used to extract features from images. This work uses digital image processing techniques to propose new geometric descriptors invariant to scale, rotation and revolution. They are used for the training of machine learning algorithms, which tests presented promising results that reached the correct classification of 99.58% of the study cases.*

**Resumo.** *Em Processamento Digital de Imagens e Visão Computacional os descritores são frequentemente usados para extrair características de imagens. Este trabalho faz uso de técnicas de processamento digital de imagens para propor novos descritores geométricos invariantes à escala, rotação e translação. Os mesmos são utilizados para o treinamento de algoritmos de aprendizado de máquinas, onde os testes apresentaram resultados promissores que alcançaram a classificação correta de 99,58% dos casos de estudo.*

## 1. Introdução

Para o sistema visual humano o reconhecimento de objetos é normalmente uma tarefa fácil, porém, está é uma tarefa não trivial para sistemas computacionais. Estes sistemas geralmente utilizam descritores de textura, geometria e cor. Bons descritores conseguem extrair características das imagens de modo a encontrar a assinatura de uma determinada classe de objetos, permitindo que, algoritmos de aprendizado de máquinas consigam distinguir objetos de diferentes classes. Este trabalho descreve as principais etapas do desenvolvimento e análise de novos descritores baseados em geometria.

De forma geral, um sistema de Visão Computacional pode ser dividido nas etapas de: aquisição, pré-processamento, segmentação, descrição e classificação. A primeira etapa é a aquisição de imagem, que consiste na captura da imagem, podendo ser por meio de câmeras, scanners, tomografias, radiografias etc. O pré-processamento tem como objetivo corrigir imperfeições, aumentando a qualidade da imagem, eliminando ruídos, realçando estruturas, tornando mais fácil as etapas seguintes. Com a segmentação é possível de subdividir a imagem em regiões menores e que possuem atributos similares. Além disso é possível eliminar estruturas indesejadas. Isso permite concentrar processamento em regiões de interesse (do inglês *Region of Interest* - ROI).

A descrição extrai da imagem informações quantitativas que podem melhor ser compreendidas pelos algoritmos de aprendizado de máquinas. Neste caso, uma imagem será representada por um vetor de características, geralmente numéricas.

A classificação utiliza os vetores de características gerados na etapa anterior e com eles, realiza um treinamento, cujo objetivo é conseguir aprender os padrões das classes de dados informados. Após o treinamento, o classificador associa essas classes a um novo vetor ainda não treinado, chamamos essa associação de classificação.

## 2. Trabalhos Relacionados

Pesquisas similares na área vêm sendo desenvolvidas com a finalidade de criar descritores capazes de distinguir formas com precisão. Os trabalhos apresentados a seguir, serviram de base para esse trabalho.

O trabalho apresentado em [Sampaio 2015] propôs uma nova metodologia computacional para detecção de massas mamárias utilizando descritores geométricos e de textura, em conjunto com o classificador Máquina de Suporte Vetorial. O melhor resultado produziu uma sensibilidade de 92,99%, uma taxa de 0,15 falsos positivos por imagem e uma área sob a curva ROC de 0,96 na análise dos seios não densos; e uma sensibilidade de 83,70%, uma taxa de 0,19 falsos positivos por imagem e uma área sob a curva ROC de 0,85. Este trabalho serviu como base para o estudo aqui proposto.

A literatura atual apresenta muitos trabalhos envolvendo descritores geométricos e análise de formas. Dentre eles podemos citar [Majidpour et al. 2015], [Deldjoo et al. 2018], [Carlier et al. 2016], [Deldjoo et al. 2017], [Khalid et al. 2017], [Akimaliev and Demirci 2015].

## 3. Descritores Geométricos

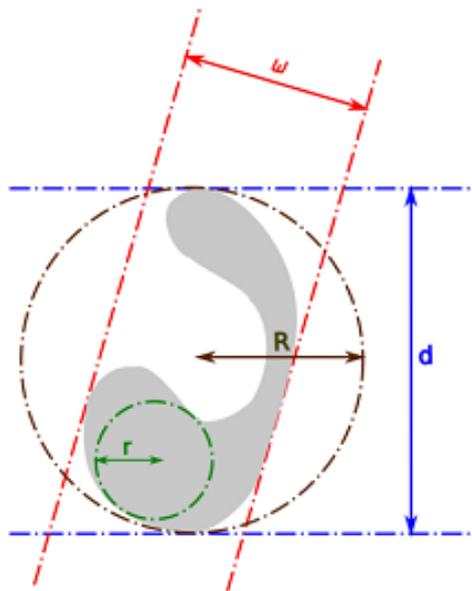
Como dito anteriormente, a etapa de descrição tem por objetivo extrair informações das imagens. Essas informações geralmente estão relacionadas a cor, textura e geometria. De forma geral, bons descritores geométricos devem ser invariantes às transformações de escala (variação de tamanho), translação (variação de posição) e rotação (variação de orientação). Os descritores geométricos extraem informações de imagens considerando apenas a forma, não se importando com a textura ou cor. Alguns exemplos de descritores de formas são: compacidade, excentricidade, circularidade, descritores baseados em momentos e curvatura [Gonzalez and Woods 2007].

Tomando uma forma genérica em 2D, apresentada na Figura 1, é possível extrair as informações como área ( $A$ ), perímetro ( $P$ ),  $r$  (raio da maior circunferência interna da forma),  $R$  (raio da menor circunferência externa),  $w$  (diâmetro mínimo de Feret). Com essas informações é possível criar um conjunto de descritores de geometria, conforme o trabalho apresentado em [Sampaio 2015].

Outra técnica utilizada para o desenvolvimento de descritores é a esqueletização, que trata-se de um algoritmo de afinamento da representação estrutural de uma região [Gonzalez and Woods 2007]. A visualização do esqueleto de uma forma pode facilitar o reconhecimento das assinaturas das geometrias estudadas. A Figura 2 apresenta o resultado da processo de esqueletização.

## 4. Metodologia

A Figura 3 apresenta o fluxo das etapas utilizadas na metodologia proposta. Na aquisição, utilizou-se um banco de imagens previamente rotuladas, onde são aplicados os descri-



**Figura 1. Informações de geometria:  $r$  (raio da maior circunferência interna da forma),  $R$  (raio da menor circunferência externa),  $w$  (diâmetro mínimo de Feret),  $d$  (diâmetro). Fonte: [Sampaio 2015]**

tores geométricos para a extração dos vetores de características geométricas. Em seguida, é realizado o treinamento dos algoritmos de aprendizado de máquina e finalmente a avaliação/validação dos modelos gerados.

#### 4.1. Banco de Imagens

O banco de imagens utilizado neste trabalho é o *ETU10 Silhouette Dataset* [Akimaliev and Demirci 2015], constituído de um conjunto de dados que consiste em 720 silhuetas de 10 objetos, com 72 visualizações por classe. A Figura 4 contém uma amostra de cada classe.

#### 4.2. Descritores Geométricos

Os vinte primeiros descritores geométricos apresentados na Tabela 1 foram utilizados no trabalho de [Sampaio 2015]. Os descritores 21 a 28 são novos descritores geométricos aqui propostos.

O descritor de número 21 relaciona a área da forma e os diâmetros mínimo e máximo de Feret.

O descritor de número 22 relaciona a distância média entre os pontos do esqueleto ao centro da menor circunferência interna com a quantidade de pontos do esqueleto.

O descritor de número 23 relaciona a distância média entre os pontos do esqueleto ao centro de massa da forma com a quantidade de pontos do esqueleto.

O descritor de número 24 relaciona a distância mínima entre os pontos do esqueleto aos pontos da borda da forma com a quantidade de pontos da borda.

No descritor de número 25 é tomada a relação entre a área do círculo externo multiplicado pelo seu raio e a área da forma multiplicada pelo diâmetro mínimo de Feret.

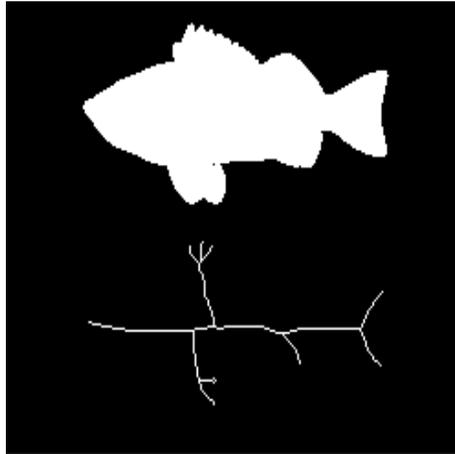


Figura 2. Esqueleto de uma forma. Fonte: O próprio autor.



Figura 3. Etapas aplicadas na metodologia.

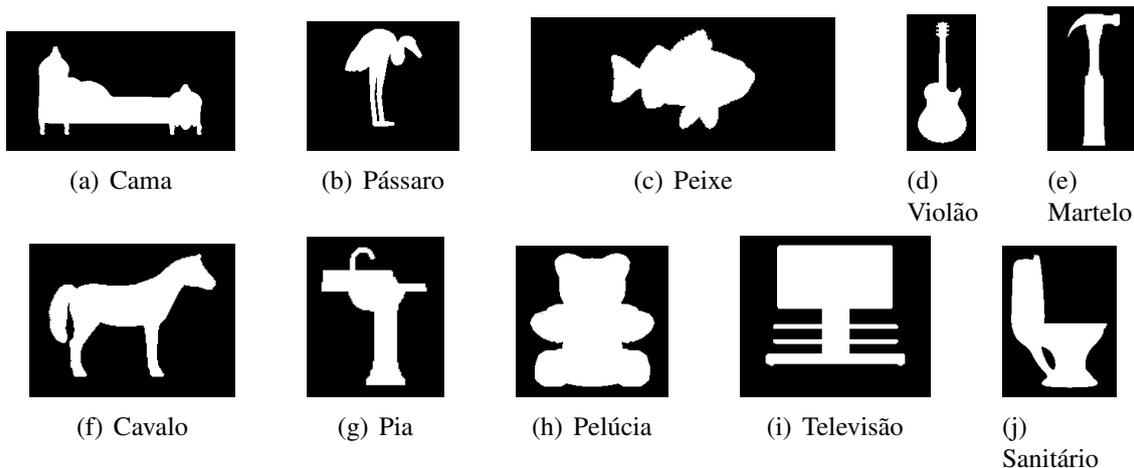
No descritor de número 26 é tomada a relação entre a área do círculo externo e a área do círculo cujo o diâmetro seja o diâmetro mínimo de Feret.

O descritor de número 27 é caracterizado pela diferença entre a área da forma estudada e a área do círculo cujo o raio é a distância entre os centros da maior circunferência interna e a menor circunferência externa.

Para o descritor 28 é proposta uma relação entre a área do retângulo, cuja a base é o  $w$  e altura igual ao  $d$ , e a área resultante da diferença entre o círculo formado pelo raio  $R$  e o círculo cujo o raio é o  $r$  da forma.

Tabela 1. Descritores geométricos

1.	$\frac{r}{R}$	8.	$\frac{\sqrt{3} \cdot R}{d}$	15.	$\frac{2.0 \cdot r}{w}$	22.	$\frac{distMed(esqueleto,centro)}{tamanho(esqueleto)}$
2.	$\frac{w}{2.0 \cdot R}$	9.	$\frac{2.0 \cdot \pi \cdot r}{P}$	16.	$\frac{d}{tamanho(esqueleto)}$	23.	$\frac{distMed(esqueleto,centMassa)}{tamanho(esqueleto)}$
3.	$\frac{A}{2.0 \cdot R^2}$	10.	$\frac{\pi \cdot w}{P}$	17.	$\frac{w}{tamanho(esqueleto)}$	24.	$\frac{distMed(esqueleto,borda)}{perimetro}$
4.	$\frac{d}{2.0 \cdot R}$	11.	$\frac{4.0 \cdot \pi \cdot A}{P^2}$	18.	$\frac{distMed(borda,esqueleto)}{tamanho(esqueleto) \cdot R}$	25.	$\frac{\pi \cdot R^3}{A \cdot w}$
5.	$\frac{2.0 \cdot r}{d}$	12.	$\frac{2.0 \cdot d}{P}$	19.	$\frac{distMedPD(Borda,Esqueleto)}{R}$	26.	$\frac{\pi \cdot R^2}{\pi \cdot \left(\frac{w}{2}\right)^2}$
6.	$\frac{w}{d}$	13.	$\frac{4.0 \cdot R}{P}$	20.	$\frac{dist(centMassa,r)}{dist(centMassa,R)+1}$	27.	$\frac{Área da Forma}{\pi \cdot dist(centror,centroR)^2}$
7.	$\frac{4.0 \cdot A}{\pi \cdot d^2}$	14.	$\frac{\pi \cdot r^2}{A}$	21.	$\frac{A}{w \cdot d}$	28.	$\frac{w \cdot d}{\pi \cdot (R^2 - r^2)}$



**Figura 4. Amostra das classes utilizadas.**

### 4.3. KNN

O algoritmo *K-Nearest Neighbours* (K-vizinhos mais próximos) é um método baseado em distâncias. Como seu próprio nome sugere, a classificação de um novo objeto é feita com base nos exemplos de treinamento que são mais próximos a ele. Essa é uma técnica *lazy* (preguiçosa), onde não há o aprendizado de um modelo compacto para os dados, apenas memorização [Faceli et al. 2011].

A resposta do *KNN*, depende de qual o seu objetivo, podendo ser usado para classificação ou regressão. Este método é escolhido para este trabalho devido a sua simplicidade de implementação, pelo seu alto poder de classificação e por trabalhar bem com valor de ponto flutuante. Neste trabalho utilizou-se o valor de  $k$  igual à 1, que também é conhecido como algoritmo do vizinho mais próximo.

### 4.4. Random Forest

O algoritmo da Floresta Randômica [Breiman 2001] toma aleatoriamente um conjunto de treinamento  $N$ , de modo que, se houver variáveis ou características de entrada  $M$ , um número  $m < M$  é especificado, então, em cada nó,  $m$  variáveis são selecionadas aleatoriamente fora de  $M$ . A melhor divisão nestas  $m$  é usada para dividir o nó. O valor de  $m$  é mantido constante enquanto é cultivada a floresta. Cada árvore cresce na maior extensão possível e sem poda. Depois agregam-se as previsões das árvores e de acordo com a maioria de votos, faz-se a classificação.

A escolha desse algoritmo deve-se a sua capacidade de classificação, e por trabalhar bem com um grande volume de dados e de grandes dimensionalidades.

### 4.5. Perceptron Multicamadas

Uma Rede Neural Artificial (RNA) é um conjunto de técnicas de Aprendizado de Máquinas baseadas em otimização, que tem inspiração nas redes neurais biológicas. O algoritmo de treinamento mais comum das RNAs envolve uma regra de correção de erros. Uma RNA com uma ou mais camadas intermedirárias é chamada de *Perceptron Multicamadas* (MLP, no inglês *Multilayer Perceptron*) [Kovács 2002] [Haykin 2007].

A escolha do treinamento de uma MLP deve-se ao fato do algoritmo lidar muito bem com atributos de entrada do tipo ponto flutuante. A configuração da MLP utilizada foi uma camada oculta com 15 neurônios e uma camada de saída com 10 neurônios, todos ativados pela função sigmoideal.

#### 4.6. MVS

A Máquina de Vetores de Suporte (MVS) [Cristianini and Shawe-Taylor 2000] é um algoritmo de aprendizado supervisionado baseado em otimização e embasada na teoria de aprendizado estatístico que estabelece uma série de princípios que devem ser seguidos na obtenção de classificadores com boa capacidade de generalização. O objetivo de um treinamento de uma MVS é a obtenção de hiperplanos que dividam as amostras de tal maneira que sejam otimizados os limites de separação entre as classes [Faceli et al. 2011].

A arquitetura da MVS treinada utilizou um *kernel* polinomial, o parâmetro de complexidade do modelo e tolerância é igual a 1 e 0,001, respectivamente. O calibrador escolhido é o modelo de regressão logística multinomial com um estimador de cume [Le Cessie and Van Houwelingen 1992] e a variável para suavização é igual a  $1,0E - 12$ .

#### 4.7. Avaliação dos Modelos

A utilização de métodos de amostragem produz estimativas de desempenho preditivo mais confiáveis, definindo subconjuntos de treinamento e de teste [Toussaint 1974]. Essas subdivisões devem ser disjuntas para assegurar que as medidas de desempenho sejam obtidas a partir de um conjunto de exemplos diferentes usados no aprendizado, exemplos de métodos de amostragem são *holdout*, amostragem aleatória, *bootstrap* e a validação cruzada [Faceli et al. 2011]. Este trabalho utiliza o método de validação cruzada.

A validação cruzada *k-folds* (*cross-validation*) é um método baseado em amostragens, onde o conjunto de exemplos é dividido em *k* partições de tamanhos aproximadamente iguais [Kohavi et al. 1995]. As *k-1* subdivisões são utilizadas para treinamento do preditor, e as restantes são usadas para teste. O processo descrito é repetido *k* vezes, onde em cada fase são utilizadas *folds* diferentes, onde a estimativa de quão preciso é o preditor é dada pela média de todas as combinações de teste [Faceli et al. 2011].

Para os testes de desempenho, a quantidade de partições é de *k* igual a 10 para todos os algoritmos de classificação. As medidas de desempenho utilizadas neste trabalho foram o índice Kappa, taxa de Falso Positivos, Sensibilidade, Medida-F e a área sob a curva ROC.

### 5. Resultados

Os resultados obtidos com os algoritmos de aprendizado, foram: O treinamento do KNN classificou corretamente 99,58% das incurssões, o erro médio absoluto chegou a 0,004 e o erro quadrado médio 0,029. O treinamento da RF resultou na classificação correta em 99,44%, com o erro médio absoluto de 0,003 e o erro quadrado médio 0,030. O treinamento da MLP resultou em 99,17% classificadas corretamente, com o erro médio absoluto de 0,005 e erro quadrado médio 0,040. A MVS resultou na classificação de 97,50% corretamente, com o erro médio absoluto de 0,160 e erro quadrado médio de 0,272. Mais índices na Tabela 2.

**Tabela 2. Precisão detalhada**

	Taxa FP	Sensibilidade	Medida-F	Kappa	Área ROC
KNN	0,000	0,996	0,996	0,9954	0,992
RF	0,001	0,994	0,994	0,9938	1,000
MLP	0,001	0,992	0,992	0,9907	1,000
MVS	0,003	0,975	0,975	0,9722	0,993

## 6. Conclusões e Trabalhos Futuros

Os resultados apresentados na Seção 5 mostram que os descritores propostos possuem relevância na classificação das formas. Para trabalhos futuros é proposto fazer a seleção e redução dos mesmos, além de desenvolver descritores baseados em textura, cor etc. Propostas de aplicações: identificação e classificação de doença pulmonar obstrutiva crônica, identificação e classificação de patologias em concreto armado.

## Referências

- Akimaliev, M. and Demirci, M. F. (2015). Improving skeletal shape abstraction using multiple optimal solutions. *Pattern Recognition*, 48(11):3504–3515.
- Breiman, L. (2001). Random forests. In Schapire, R. E., editor, *Machine learning*, volume 45, pages 5–32. Springer.
- Carlier, A., Leonard, K., Hahmann, S., Morin, G., and Collins, M. (2016). The 2d shape structure dataset: A user annotated open access database. *Computers & Graphics*, 58:23–30.
- Cristianini, N. and Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.
- Deldjoo, Y., Elahi, M., Quadrana, M., and Cremonesi, P. (2018). Using visual features based on mpeg-7 and deep learning for movie recommendation. *International journal of multimedia information retrieval*, 7(4):207–219.
- Deldjoo, Y., Quadrana, M., Elahi, M., and Cremonesi, P. (2017). Using mise-en-scène visual features based on mpeg-7 and deep learning for movie recommendation. *arXiv preprint arXiv:1704.06109*.
- Faceli, K., Lorena, A. C., Gama, J., and Carvalho, A. (2011). Inteligência artificial: Uma abordagem de aprendizado de máquina. *Rio de Janeiro: LTC*, 2:192.
- Gonzalez, R. C. and Woods, R. E. (2007). *Digital Image Processing. Ed III*. Pearson Education Asia, New Delhi.
- Haykin, S. (2007). *Redes neurais: princípios e prática. Ed. II*. Bookman Editora, São Paulo.
- Khalid, S., Sabir, B., Jabbar, S., and Chilamkurti, N. (2017). Precise shape matching of large shape datasets using hybrid approach. *Journal of Parallel and Distributed Computing*, 110:16–30.
- Kohavi, R. et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Stanford, CA.

- Kovács, Z. L. (2002). *Redes neurais artificiais: Fundamentos e aplicação. Ed. IV*. Editora Livraria da Física, São Paulo.
- Le Cessie, S. and Van Houwelingen, J. C. (1992). Ridge estimators in logistic regression. *Applied statistics*, pages 191–201.
- Majidpour, J., Khezri, E., Hassanzade, H., and Mohammed, K. S. (2015). Interactive tool to improve the automatic image annotation using mpeg-7 and multi-class svm. In *2015 7th Conference on Information and Knowledge Technology (IKT)*, pages 1–7.
- Sampaio, W. B. (2015). Detecção de massas em imagens mamográficas usando uma metodologia adaptada à densidade da mama. Tese de doutorado - programa de pós-graduação em engenharia de eletricidade, PPGEE, Universidade Federal do Maranhão.
- Toussaint, G. (1974). Bibliography on estimation of misclassification. *IEEE Transactions on information Theory*, 20(4):472–479.