

CLASSIFICADOR DE FORMAS UTILIZANDO DESCRITOR DE FOURIER USANDO DISTÂNCIA DO CENTROIDE

Vinicius Santos Andrade¹; Renan Caldeira Menechelli²; Elvio Gilberto da Silva³; Patrick Pedreira Silva⁴

RESUMO

A quantidade de aplicações que fazem uso de imagens aumenta a cada dia. No âmbito social, existem diversas plataformas como o Instagram, que faz uso massivo de processamento e análise de imagens. Na indústria, não é diferente, a aplicação de técnicas para processar e analisar possui vasta aplicabilidade. Neste sentido, a pesquisa visa contribuir avaliando o desempenho de uma aplicação que faz uso de um conjunto de técnicas de processamento de imagens, a fim de possibilitar a aplicação de um descritor (Fourier) e analisar os dados obtidos com o classificador kNN. Com os testes realizados, foi verificado que a sequência de experimentos obteve 60% de acurácia no melhor caso, também foi possível observar queda na precisão versus revocação no decorrer de cada teste, em todos os experimentos feitos neste estudo.

Palavras-chave: Análise de imagens; processamento de imagens; Fourier. KNN.

ABSTRACT

The number of applications that make use of images increases every day. In the social sphere, there are several platforms such as Instagram, which makes massive use of image processing and analysis. In industry, it is no different, the application of techniques to process and analyze has wide applicability. In this sense, a research aims to contribute to evaluate the performance of an application that makes use of a set of image processing techniques, in order to enable the application of a descriptor (Fourier) and analyze the data obtained with the kNN classifier. With the tests carried out, it was found that the sequence of experiments obtained 60% accuracy in the best case, it was also possible to observe a decrease in precision versus recall during each test, in all experiments carried out in this study.

Key Words: Image analysis; image processing; Fourier. kNN.

1 INTRODUÇÃO

Reconhecimento de formas é uma técnica muito utilizada pelas indústrias de diversos segmentos. Com ela é possível identificar diversas falhas que podem acontecer no decorrer de uma linha de produção, seja por embalagens defeituosas ou mesmo por quantidades inadequadas de produtos dentro dessas embalagens.

¹ Doutorando em Ciências da Reabilitação pela Universidade de São Paulo- USP, professor do Centro Universitário Sagrado Coração (UNISAGRADO /Bauru). E-mail: viniciusarmani@gmail.com.

² Mestrado em Engenharia Elétrica pela Universidade de São Paulo-USP, professor e coordenador dos cursos Engenharia de Computação e Jogos Digitais do Centro Universitário Sagrado Coração (UNISAGRADO /Bauru). E-mail: renan.menechelli@unisagrado.edu.br.

³ Doutor Agronomia (Energia na Agricultura) pela Universidade Estadual Paulista UNESP Botucatu, professor do Centro Universitário Sagrado Coração (UNISAGRADO /Bauru). E-mail: egilberto@uol.com.br.

⁴ Doutorando em Ciências da Reabilitação pela Universidade de São Paulo- USP, professor do Centro Universitário Sagrado Coração (UNISAGRADO /Bauru). E-mail: patrickpsilva@gmail.com.

A maioria das assinaturas relacionadas com a forma dos objetos, como coordenadas complexas, requer quantidade considerável de recursos computacionais por conta dos inúmeros cálculos de similaridade, devido à forte normalização da invariância de rotação, sendo necessário o uso de transformações como a de Fourier para normalização dos dados. Em descritores espectrais, como o de Fourier, as formas globais são definidas pelos primeiros termos de baixa frequência, enquanto os termos de frequência mais alta capturam características mais refinadas da forma (GOPAL; PRASAD, 2008; TOM; JIDESH, 2013).

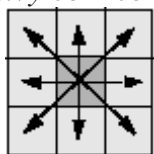
Diante deste contexto, este trabalho tem como objetivo identificar e classificar objetos, utilizando como descritor a *fast Fourier transform* (FFT), mostrando através de métricas de acurácia, precisão e revocação o desempenho do algoritmo. Os dados obtidos através do FFT, foram submetidos ao classificador *k nearest neighbors* (kNN).

2 Processamento De Imagens

Inicialmente é necessário fazer a binarização da imagem. Este problema pode ser facilmente resolvido com a aplicação de um *thresholding*, que irá converter a imagem de níveis de cinza para binária.

Com a binarização feita, é possível iniciar o processo de remoção de ruídos. O estágio inicial foi feito utilizando a técnica *m-connectivity* (GONZALEZ; WOODS, 1992), a fim de conectar os pontos e definir a forma (com padrão de 8 conectividade). A conectividade define em que direção o pixel pode tocar para se conectar a outro pixel. A Figura 1 exemplifica o funcionamento da técnica.

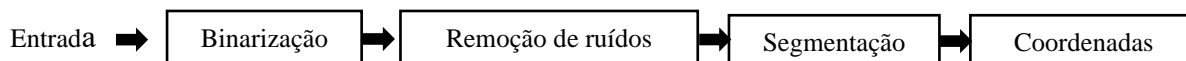
Figura 1 – Funcionamento do *m-connectivity* com conectividade 8



Fonte: Matlab (2020).

Após a remoção dos ruídos, é aplicada uma técnica de rastreamento de contorno (PAVLIDIS, 1982; SUEN; WANG, 1994), para obter coordenadas da borda da forma. O método também auxilia na remoção de ruídos uma vez que, ao efetuar o contorno do objeto, irá utilizar apenas os pixels correlacionados de acordo com as definições do algoritmo proposto por Pavlidis. A Figura 2 representa os processos feitos antes de criar o vetor de características.

Figura 2 – Etapas de criação do vetor de características



Fonte: Elaborada pelo autor (2020).

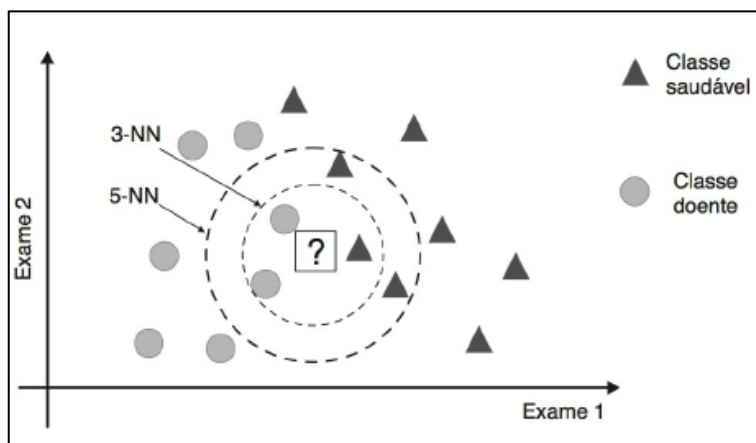
2.1 O modelo kNN

O kNN é um modelo relativamente simples utilizado para classificar dados. O classificador foi proposto por Fukunga e Narendra (1975), e é baseado na classificação de um novo objeto com base nos exemplos rotulados previamente. A função é expressa por:

$$d(X_i, Y_i) = \sqrt[r]{\sum_{i=1}^n |X_i - Y_i|^r} \quad (1)$$

O modelo consegue determinar um rótulo de classificação para uma amostra baseado nas vizinhanças advindas do conjunto de dados. A Figura 3 exemplifica o funcionamento do modelo.

Figura 3 – Impacto do valor de k no algoritmo kNN



Fonte: FACELI *et al.* (2011).

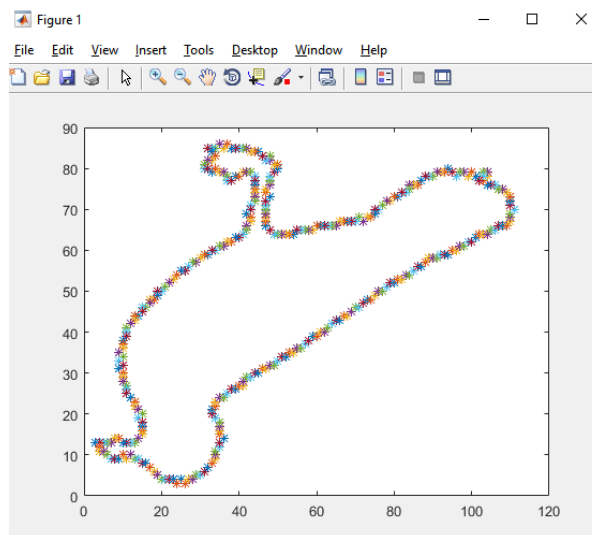
No exemplo da Figura 3, uma nova amostra recebe duas possíveis distâncias. Em uma primeira situação, onde k recebe o valor 3, o que significa que são considerados os 3 objetos mais próximos dele pra efeitos de classificação, assim, o novo objeto seria classificado como pertencendo à classe “doente”, enquanto para $k=5$ (considerando, agora, 5 objetos mais próximos) o novo objeto seria classificado como pertencendo à classe “saudável”.

3 Materiais e Método

A Figura 2, exibida anteriormente, exemplifica todo processo feito até a fase de obtenção das coordenadas da forma.

A Figura 4 mostra o processo de obtenção das coordenadas da forma. O exemplo em questão é uma das 12 representações de uma das formas que faz parte da base de dados, que será comentada adiante. A posição de cada coordenada é expressa pelo caractere asterisco.

Figura 4 – Mapeamento das coordenadas da forma



Fonte: Elaborada pelos autores (2020).

Após essa etapa, é feito o cálculo da distância do centroide para cada ponto. O centroide, conhecido como centro geométrico, tem a sua função $r(t)$ expressa pela distância dos pontos de contorno do centroide (x_c, y_c) da forma (Yan, 2011):

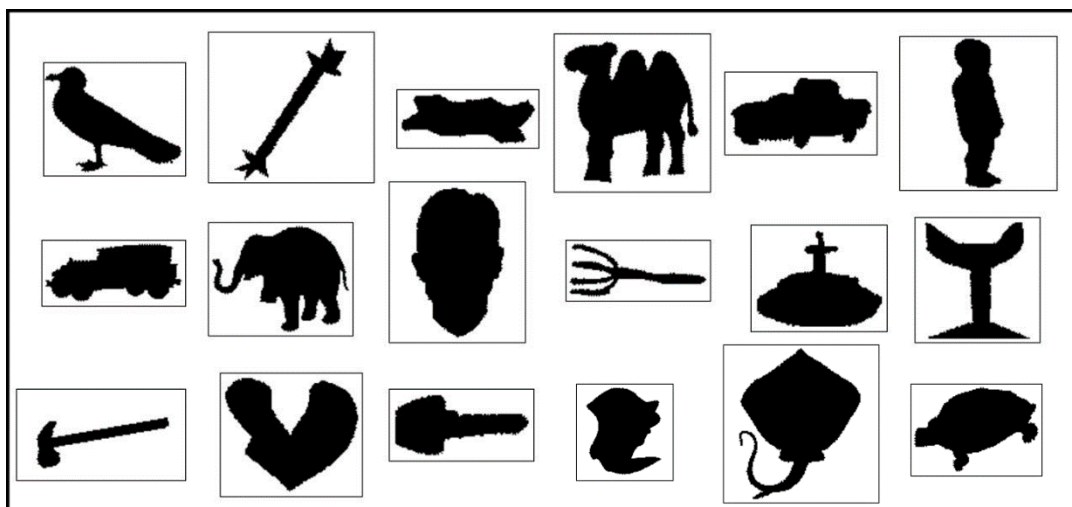
$$r(t) = ([x(t) - x_c]^2 + [y(t) - y_c]^2)^{1/2} \quad (2)$$

Aplicou-se o FFT no vetor obtido com o processo descrito no parágrafo anterior, com o objetivo de obter o vetor de características. O vetor possui os dados obtidos através da aplicação do FFT, sendo o resultado, as distâncias do centroide. Este vetor é submetido à classificação utilizando o algoritmo *k-nearest neighbors* com vizinhança um (padrão). Foram feitos testes considerando outras vizinhanças, porém, os melhores resultados foram obtidos com a vizinhança um.

Para programação foi utilizado o Matlab versão R2015b (MATLAB, 2020).

A base de dados possui 216 formas, sendo 18 formas distintas e cada forma possui 12 representações. A Figura 5 exibe uma representação de cada uma das formas presentes no banco de dados, enquanto as outras 11 representações, não expressas na figura, possuem variação de ângulo.

Figura 5 – Alguns exemplos de formas utilizadas na base de dados



Fonte: Elaborada pelos autores (2020).

O contorno presente em cada imagem na cor preta foi acrescentado pelos autores a fim de melhorar a visualização de cada uma das formas bem como a suas dimensões.

Para o treinamento foram utilizados 50% dos dados (de forma balanceada) obtidos através do descritor. A Figura 6 exibe como foi feita a divisão dos dados.

Figura 6 - Divisão entre dados de treinamento e dados para amostra

Forma 1	Forma 1.1	Forma 2	Forma 2.1
Treinamento	Amostra	Treinamento	Amostra

Fonte: Elaborada pelos autores (2020).

No exemplo apresentado na Figura 6 são especificados dois modelos de forma (Forma 1 e Forma 2) onde cada modelo possui duas representações (Forma 1 e Forma 1.1; Forma 2 e Forma 2.1), em que foram divididos em treinamento e amostra. A Tabela 1 exibe como foi feita a divisão de formas de objetos e suas respectivas quantidades de exemplares para cada teste.

Tabela 1 – Número de formas e representações utilizadas em cada experimento

Experimento	Nº de Formas utilizadas	Nº de representações para cada forma	Nº total de imagens utilizadas
Primeiro	5	4	20
Segundo	10	4	40
Terceiro	18	4	72
Quarto	18	10	180

Fonte: Elaborada pelos autores (2020).

Visto que as formas utilizadas não possuem um tamanho padronizado, foi necessário definir um valor (x) *específico* de elementos a serem utilizados na transformada de Fourier, com o objetivo de gerar a matriz de características de tamanho homogêneo. Para os experimentos atribuiu-se $x=244$ que é o menor número de coordenadas encontrado nas formas utilizadas para os experimentos.

Teoricamente, os elementos com maiores valores possuem características mais importantes em relação à forma, sendo assim, nos casos em que foram necessários reduzir o número de coordenadas para alcançar $x=244$, foi utilizado os maiores valores de cada forma para gerar o vetor de características. Para classificação foi utilizado apenas a parte real dos dados gerados a partir da transformada de Fourier.

4 Resultados

A Tabela 2 expressa os resultados obtidos com os experimentos feitos utilizando FFT, considerando a métrica de acurácia.

Tabela 2 – Resultados obtidos com os experimentos

Experimento	Nº de imagens utilizadas	Acurácia (%)
Primeiro	20	40,00
Segundo	40	60,00
Terceiro	72	52,78
Quarto	180	31,11

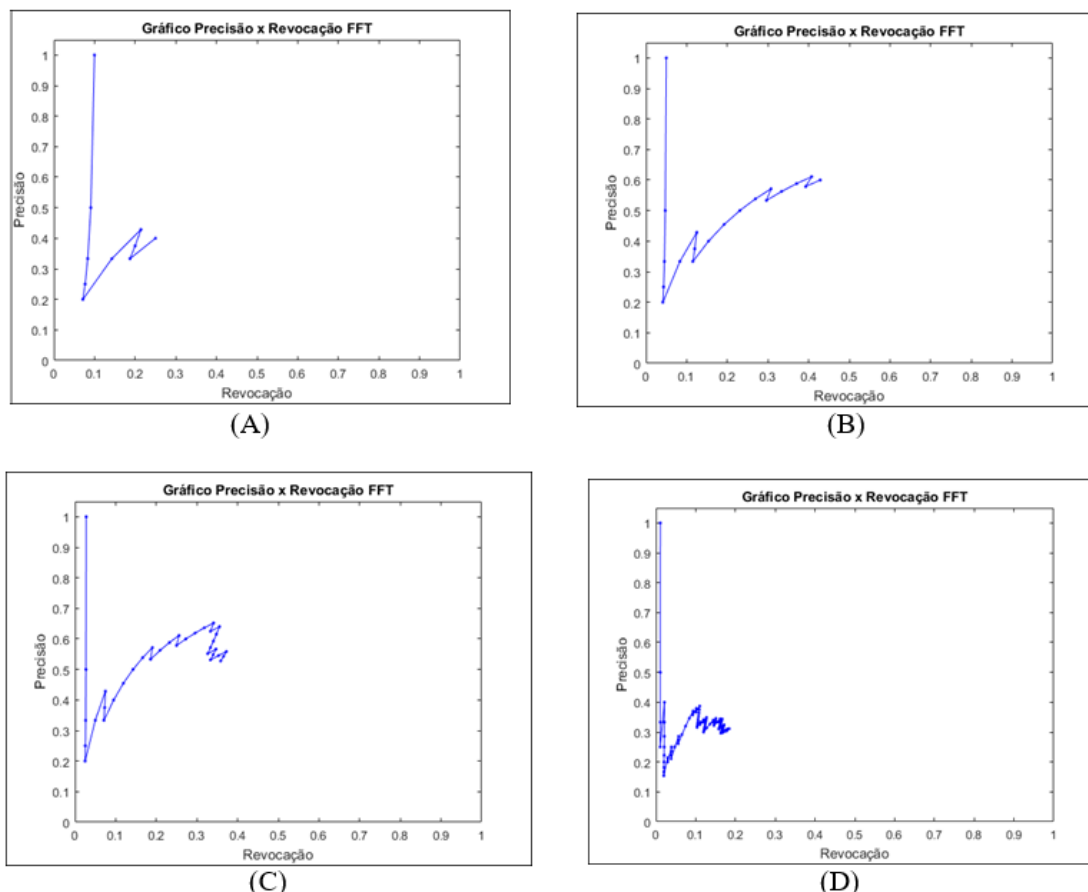
Fonte: Elaborada pelos autores (2020).

Os resultados obtidos são um tanto quanto peculiares, uma vez que, para o primeiro experimento, por possuir menos modelos de formas, esperava-se melhores resultados frente aos demais. No segundo esperava-se redução na acurácia, porém, ocorreu o contrário. No terceiro experimento houve redução de 7,22% na acurácia, mesmo tendo sido o número de formas quase que o dobro do experimento anterior. Da mesma forma, no último experimento, quando comparado ao terceiro, também houve redução (24,67%) à medida que foi acrescida quase que o dobro de imagens. A acurácia média nos experimentos foi de cerca de 46%.

Os experimentos sugerem que o algoritmo não consegue distinguir bem entre grupo de objetos que possuem formas similares. Existem grupos de exemplares semelhantes, como por exemplo, o osso e o martelo. Dependendo da angulação e dimensão, podem possuir características de centroides semelhantes. Outro exemplo de grupos semelhantes é o da “cabeça” e da “criança”, que assim como no exemplo anterior, estes também podem possuir características de centroides semelhantes.

A Figura 7-A representa o gráfico de precisão x revocação para o experimento 1, a Figura 7-B representa o gráfico de precisão x revocação para o experimento 2, e assim sucessivamente. Lembrando que todos os experimentos utilizaram FFT e classificador *k-nearest neighbors* com vizinhança 1 (padrão).

Figura 3 – Impacto do valor de k no algoritmo kNN



Fonte: Elaborada pelos autores (2020).

5 Conclusão e trabalhos futuros

Considerando a base de treinamento com poucas imagens (máximo de 180) o maior valor de acurácia de 60% foi obtido quando foram utilizadas 40 formas. O aumento deste número de formas ou a sua redução provocou quedas no valor da acurácia. Considerando todos os experimentos, o valor médio de acurácia ficou em cerca de 46%. Com base nos dados de acurácia quanto e a curva precisão x revocação, pode-se concluir que a detecção de formas não obteve um bom resultado, mas não é possível afirmar se o real motivo destes resultados é por conta do FFT ou do uso do kNN ou até mesmo a distância do centroide. Como trabalhos futuros sugere-se aplicar outro classificador; utilizar descritor como por exemplo, o de Wavelet ou *Beam Angle Statistics* (BAS); ou empregando outro parâmetro como a função de curvatura, em alternativa ao centroide.

Referências

FACELI, Katti et al. **Inteligência Artificial: Uma abordagem de aprendizado de máquina**. 2011.

GOPAL, T. Venu; PRASAD, V. A novel approach to shape based image retrieval integrating adapted Fourier descriptors and freeman code. **Int. J. Comput. Sci. Netw. Security**, v. 8, n. 6, p. 293-302, 2008.

GONZALEZ, Rafael C.; WOODS, Richard E. Digital image processing. Addison-Wesley, 1992.

Matlab – Documentation. Disponível em: <https://www.mathworks.com/help/>. Acesso em 22 de maio de 2020.

YAN, Li (Ed.). **Intelligent Multimedia Databases and Information Retrieval: Advancing Applications and Technologies: Advancing Applications and Technologies**. IGI Global, 2011.

PAVLIDIS, Theo. An asynchronous thinning algorithm. **Computer Graphics and Image Processing**, v. 20, n. 2, p. 133-157, 1982.

SUEN, Ching Y.; WANG, Patrick Shen-pei. **Thinning methodologies for pattern recognition**. World Scientific, 1994.

TOM, Arun; JIDESH, P. Geometric transform invariant Brain-MR image analysis for tumor detection. In: **2013 International conference on Circuits, Controls and Communications (CCUBE)**. IEEE, 2013. p. 1-6.