

Reconhecimento de Objetos em Imagens

Oeslle A. S. Lucena, Luciana R. Veloso e Waslon T. A. Lopes

Resumo—Uma das áreas mais desafiadoras no processamento de dados em sistemas inteligentes é reconhecer objetos em imagens. Por conseguinte, existe uma necessidade de desenvolver sistemas eficazes com taxas de reconhecimento elevadas, que sejam invariantes à rotação, ruído, luz, entre outros. Tal tarefa pode ser definida como sendo a identificação de um objeto individual como um membro de uma classe particular que contém objetos com características semelhantes. Para implementar esse algoritmo, a discriminação de características é realizada pelos descritores e o reconhecimento pelos classificadores. Nesse contexto, este trabalho envolveu uma implementação de um dos seus possíveis algoritmos, de forma a validar um sistema de identificação de objetos. O algoritmo implementado utilizou o Histograma de Orientações dos Gradientes (*Histogram of Oriented Gradient - HOG*) como descritor e as Máquinas de Vetor Suporte (*Support Vector Machines - SVM*) como o classificador. O sistema foi avaliado utilizando como métrica a taxa de reconhecimento em duas bases de dados distintas: *Caltech-101* e *MSRC v1*.

Palavras-chave: Reconhecimento de Objetos, descritor, classificador, HOG, SVM, taxa de reconhecimento.

I. INTRODUÇÃO

Um dos importantes passos para manipulação e processamento de imagens digitais em sistemas inteligentes em visão computacional é reconhecer objetos [1], [2]. As áreas de biometria, vigilância residencial e comercial, separação de frutas e interação homem-máquina são alguns dos exemplos de atividades que requerem processamento digital de imagens. Assim, há a necessidade de desenvolver sistemas eficazes com altas taxas de reconhecimento, que sejam invariantes a rotação, ruído, luminosidade, entre outros [3], [4].

Este trabalho tem como objetivo a avaliação de uma técnica de reconhecimento de objetos, que se baseia no uso do descritor Histograma de Orientações dos Gradientes (HOG) e das Máquinas de Vetor Suporte (SVM). Para isto, foi realizado uma validação do sistema utilizando as bases de dados *Caltech-101* [5] e *MSRC v1* [6].

O artigo está dividido como se segue: a Seção II encontra-se a fundamentação teórica para o reconhecimento de objetos, como também o descritor e o classificador utilizados. A Seção III detalha os experimentos realizados e os resultados do sistema de reconhecimento de objetos. Por fim, a Seção IV apresenta as conclusões e propostas de trabalhos futuros.

II. RECONHECIMENTO DE OBJETOS

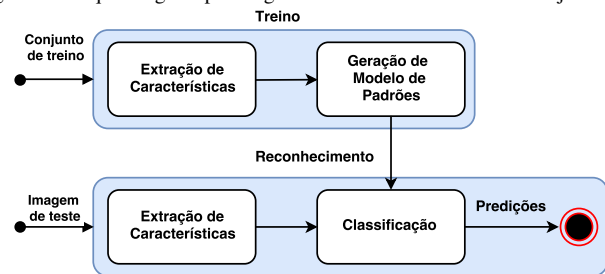
Geralmente, as atividades de reconhecimento de objetos são realizadas em duas etapas: extração de características e

classificação, com duas fases distintas de operação: treinamento e reconhecimento.

A fase de treinamento do algoritmo consiste na busca de padrões para representar os objetos de interesse. Esses padrões são discriminados por um conjunto de características extraídas usando um descritor que pode ser local ou global, ou seja, o descritor atua ao redor de pontos de interesse ou em toda a imagem [1]. Ou seja, dado um conjunto de imagens para treino, o descritor é aplicado, em seguida, as características extraídas são submetidas a um algoritmo de aprendizado de máquina, que por meio de algoritmos de agrupamento usando funções matemáticas, buscam separar as características provenientes dos descritores que diferenciam o objeto de interesse dos demais, gerando um modelo dos padrões [2].

Na fase de reconhecimento, é realizado um casamento dos padrões que correlaciona o objeto de interesse com os de uma classe específica. Para isto, dada uma imagem, os vetores de características (descritores) são extraídos. Na sequência, é realizada a classificação, através de uma medida de similaridade entre os descritores da imagem analisada e de todos os modelos de descritores que descrevem os objetos das classes aprendidas. A Figura 1 mostra o esquema geral para um algoritmo de reconhecimento de objetos.

Figura 1. Esquema geral para algoritmo de reconhecimento de objetos.



A. Histograma de Orientações dos Gradientes

O HOG é um descritor de características introduzido em [7]. Esse descritor utiliza o fato de que a aparência de um objeto local e seu formato podem ser caracterizados pela intensidade de gradientes e pelas direções. Para o cálculo dos atributos HOG, a imagem é dividida em células, sendo elas agrupadas em um conjunto maior denominado bloco. Para cada célula é calculado um histograma de orientações do gradiente que é dividido em 9 partes (*bins*), igualmente espaçados entre 0 e 180 graus.

B. Máquinas de Vetor Suporte

As SVMs são uma técnica de aprendizado de máquina supervisionado. Ou seja, tem-se a figura de um professor externo,

Oeslle A. S. Lucena, Aluno de Mestrado, Universidade Estadual de Campinas (Unicamp), e-mail: oeslle.lucena@gmail.com.

Luciana R. Veloso, D.Sc, Universidade Federal de Campina Grande (UFCG), e-mail: luciana.veloso@dee.ufcg.edu.br.

Waslon T. A. Lopes, D.Sc, Universidade Federal da Paraíba (UFPB), e-mail: waslon@ieee.org.

o qual apresenta o conhecimento do ambiente por conjuntos de exemplos na forma: entrada e saída desejada [8]. Elas são classificadores discriminativos que atuam na separação de padrões por meio de hiperplanos. Em outras palavras, dado um conjunto de amostras rotuladas para treino, o algoritmo busca hiperplanos ótimos para categorizar as amostras de teste [9]. Para isto, o algoritmo SVM baseia-se em encontrar o hiperplano que gera a maior distância entre amostras de classes diferentes para o conjunto de treino. Essa distância é denominada de margem.

III. RESULTADOS

Inicialmente, no treinamento, as imagens são redimensionadas para um tamanho específico e aplica-se o descritor, em seguida, o SVM é utilizado para gerar os modelos de padrões que discriminam cada uma das classes de objetos. Na fase de reconhecimento, para cada imagem no subconjunto de teste, é realizado o redimensionamento e a discriminação de suas características, em seguida, o SVM, com o uso do modelo de padrões aprendido, prediz o rótulo da classe a que o objeto corresponde.

Os experimentos foram realizados com redimensionamentos de imagens de 64×128 pixels (proposto por [7]) e 128×256 pixels (sugestão empírica), como também, blocos de tamanho 2×2 células, com células variando de 8×8 , 16×16 e 32×32 pixels. Por fim, a classificação do SVM é realizada para cada base de dados e suas taxas de reconhecimento são avaliadas. Esse resultado é calculado pela métrica acurácia que representa a taxa de reconhecimento classificador [10].

As bases de dados utilizadas foram *Caltech-101* e *MSRC v1*. Do banco de dados *Caltech-101* foram selecionadas 12 classes de objetos (avião, motocicleta, face, relógio de pulso, leopardo, bonsai, lateral de carro, barco, candelabro, tartaruga, piano e cérebro), contendo cada uma 98 imagens distintas, sendo subdividida nos subconjuntos de treino e teste, em que 80% de suas imagens foram alocadas para o conjunto de treino, e 20% foram alocadas para o conjunto de teste. Para a base de dados *MSRC v1* utilizou-se a mesma metodologia para a criação dos subconjuntos de treino e teste. Esse conjunto possui 240 imagens dividido em 7 classes (árvore, prédio, avião, vaca, face, carro, bicicleta).

Na realização experimentos foi utilizado um computador com processador Intel® i7-3632QM de 2,2 GHz, 8 GB de memória principal e sistema operacional *Windows 10 Home* de 64 bits. Todas as rotinas foram implementadas na Linguagem C++ e com uso da biblioteca OpenCV (Disponível em: <http://opencv.org/>).

Na Tabela I são apresentados os melhores resultados avaliados nos experimentos para as taxas de reconhecimento do classificador quando submetido as duas bases de dados, com variação dos parâmetros de redimensionamento e tamanho de célula. Os resultados da tabela são para células de tamanho 32×32 pixels.

IV. CONCLUSÃO

Nesse trabalho foi implementado um sistema de reconhecimento de objetos em imagens. Para isto, foi empregado um

Tabela I
TAXAS DE RECONHECIMENTO PARA AS BASES DE DADOS *Caltech-101* E *MSRC v1*.

Base de Dados	Redimensionamento (pixels)	Taxa de Reconhecimento (%)
MSRC-V1	64×128	77,08
Caltech-101	128×256	84,21

sistema de identificação usando o descritor HOG como global, um classificador SVM e o sistema foi avaliado em duas bases de dados distintas.

Quando submetido as base de dados da *Caltech-101* e *MSRC v1*, o classificador obteve os valores de taxas de reconhecimento iguais a 84% e 77%, respectivamente, validando o sistema de reconhecimento de objetos empregado. As falhas ocorridas devem-se, em grande parte, ao fato de que havia grande quantidade de informação de fundo nas imagens, possibilitando assim, erros de classificação inter-classes.

Como trabalhos futuros, sugere-se um estudo de outras técnicas de reconhecimento de objetos, bem como, avaliar a possibilidade de combinar técnicas já existentes e o uso de redes neurais artificiais para extração e/ou classificação.

AGRADECIMENTOS

Os autores agradecem ao Laboratório de Percepção Computacional (LPC) e ao Laboratório de Automação, Processamento de Sinais e Imagens (LAPSI), ambos da UFCG, por propiciar a execução desse trabalho.

REFERÊNCIAS

- [1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., 2006.
- [2] S. Addicam, S. Malik, and P. Tian, *Building Intelligent Systems*. Intel Press., 2012.
- [3] K. E. A. van de Sande, J. R. R. Uijlings, T. Gevers, and A. W. M. Smeulders, "Segmentation as selective search for object recognition," in *2011 International Conference on Computer Vision*, Nov 2011, pp. 1879–1886.
- [4] Y. Li, S. Wang, Q. Tian, and X. Ding, "A survey of recent advances in visual feature detection," *Neurocomputing*, vol. 149, Part B, pp. 736 – 751, 2015.
- [5] L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories," in *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04. Conference on*, June 2004, pp. 178–178.
- [6] J. Winn and N. Jojic, "Locus: learning object classes with unsupervised segmentation," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, vol. 1, Oct 2005, pp. 756–763 Vol. 1.
- [7] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, June 2005, pp. 886–893.
- [8] A. C. Lorena and A. C. P. L. F. De Carvalho, "Uma introdução às support vector machines," *Revista de Informática e Teoria Aplicada - RITA*, vol. 14, no. 2, pp. 43–67, 2007.
- [9] O. Documentation, "Introduction to support vector machines — opencv 2.4.13.0 documentation," http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html, 2016, acesso 10 de Julho de 2016.
- [10] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing and Management*, vol. 45, no. 4, pp. 427–437, 2009.