

Aprendizagem multirrótulo para a rotulação de imagens

Daniel de Moraes Navarro, Ronaldo C. Prati

Universidade Federal do ABC (UFABC) – Santo André, SP – Brasil

daniel.phn@gmail.com, ronaldo.prati@ufabc.edu.br

Resumo. Neste projeto é proposta uma ferramenta capaz de anotar imagens automaticamente utilizando AM multirrótulo através de um prévio processamento para extração de informações dessas imagens. Diferentemente de AM tradicional, em AM multirrótulo mais de um rótulo pode ser associado a cada novo objeto a ser classificado.

Palavras-chave: Inteligência Artificial, Aprendizado de Máquina, Aprendizado Multirrótulo, Rotulação de Imagens.

1. Contextualização

Embora a evolução das tecnologias e ferramentas de processamento e extração de informações de imagens não tenha se desenvolvido no mesmo ritmo que as outras, tais como as de textos disponíveis há algum tempo, é crescente a importância e o interesse por ferramentas para esse fim no atual cenário da nossa sociedade. Atualmente muitos modelos para recuperação de imagens baseada, por exemplo, em seu conteúdo estão sendo investigados.

A solução comumente adotada, utilizada por muitas organizações para indexar e posteriormente recuperar coleções de imagens é a chamada anotação de imagens, que consiste na anotação manual de rótulos, ou palavras-chave, às imagens digitais através de metadados de acordo com a sua representatividade e o seu conteúdo, e então procurar por essas anotações. Apesar de esse modelo permitir o retorno de arquivos baseado em conceitos semânticos, a anotação manual das imagens é um processo trabalhoso e de alto custo que, na maioria dos casos, não garante o retorno correto das informações que se procura [Jeon et al 2003]. A anotação automática de imagens, por sua vez, é o processo pelo qual um sistema computacional realiza essa associação de rótulos ou palavras-chave para uma imagem digital [Li and Wang 2003]. Nesse caso, uma das abordagens investigadas para rotulação automática de imagens é através de algoritmos de Aprendizagem de Máquina (AM) na identificação de ocorrências de um determinado assunto nas imagens, facilitando a recuperação dessas imagens.

Neste projeto é investigado o uso de algoritmos de AM baseadas em classificação multirrótulo para a anotação de imagens. Em abordagens tradicionais de AM supervisionada, um único rótulo é atribuído a cada novo objeto a ser classificado. Em AM multirrótulo, em contrapartida, mais de um rótulo pode ser atribuído ao novo objeto a ser classificado. Essa atribuição de múltiplos rótulos pode ser tanto um subconjunto dos rótulos disponíveis quanto uma lista ordenada de rótulos.

2. Metodologia

Este projeto propõe a investigação e a anotação automática de imagens utilizando e desenvolvendo métodos de AM e da Mineração de Dados (DM) de maneira a permitir a recuperação de imagens através de parâmetros semânticos através da execução de uma ferramenta aplicada a uma coleção de imagens e que é integrado com o software Waikato Environment for Knowledge Analysis (WEKA), que possui um conjunto de algoritmos de diversas técnicas de DM.

O processamento das imagens é feito de forma que um conjunto de imagens de treinamento previamente anotadas é processado para extrair características. Inicialmente foi escolhido o processamento por histogramas, sendo que outras abordagens estão sendo desenvolvidas. As etapas desse processamento são, sucintamente, a divisão da imagem em segmentos, o processamento de cada um desses segmentos por histogramas de maneira independente e, por fim, os valores de cada um dos histogramas processados são sequencialmente armazenados em um vetor com número de posições equivalente à quantidade de valores gerados pelo processamento dos histogramas. A ferramenta exporta os dados resultantes do processamento diretamente no formato Attribute Relation File Format (ARFF), que é o formato dos arquivos de entrada do WEKA [Witten and Frank. 2005]. Uma vez construído esse conjunto de dados, algoritmos de aprendizado de máquina multirrótulo disponíveis no pacote MULAN (uma extensão do WEKA para AM multirrótulo disponível em <http://mulan.sourceforge.net>) podem ser utilizados para construir modelos de rotulação de novas imagens.

Para validar a proposta foram realizados experimentos utilizando uma base de dados composta por 139 imagens de cenas urbanas de Barcelona (disponível em <http://mlg.ucd.ie/content/view/61>) previamente anotadas com quatro rótulos: Buildings, Flora, People e Sky. Cada imagem possui no mínimo dois e no máximo quatro rótulos, sendo que cada rótulo está presente em pelo menos 60 imagens, gerando um arquivo ARFF que pôde ser avaliado utilizando-se uma técnica de classificação multirrótulo por transformação disponível no pacote MULAN juntamente com os modelos de classificação J48 e kNN, disponíveis no pacote WEKA. As taxas de acerto estimadas usando validação cruzada 10 folds foram de $62,5\% \pm 4,04$ e $64,78\% \pm 8,28$, respectivamente.

3. Conclusões

Neste projeto é utilizado AM multirrótulo para a rotulação de imagens. As imagens são segmentadas e processadas para a extração de características. Classificadores multirrótulos podem ser treinados, de tal maneira que múltiplos rótulos podem ser atribuídos para uma mesma imagem.

4. Referências

- Jeon, J., Lavrenko, V., and Manmatha, R. (2003) "Automatic image annotation and retrieval using cross-media relevance models", In ACM SIGIR'2003), p. 119-126.
- Li, J. and Wang, J. Z. (2003) "Automatic linguistic indexing of pictures by a statistical modeling approach", In IEEE Trans. Pattern Anal. Mach. Intell., 25(9), p. 1075-1088.
- Witten, I. H., Frank, E. (2005) "Data Mining: Practical Machine Learning Tools and Techniques", 2ª edição. New York: Morgan Kaufmann, 524 p.