

INF1771 - INTELIGÊNCIA ARTIFICIAL

TRABALHO 3 – APRENDIZADO DE MÁQUINA

Descrição:

O trabalho 3 consiste em utilizar 4 algoritmos de aprendizado supervisionado de máquina (Árvores de Decisão, K-Nearest Neighbor, Support Vector Machines e Redes Neurais) para reconhecer o estado emocional de pessoas com base em características visuais extraídas da face das pessoas.

As imagens que serão utilizadas para o treinamento e validação dos métodos de aprendizado supervisionado fazem parte do banco de dados de faces conhecido como “Radboud Faces Database” (www.socsci.ru.nl:8180/RaFD2/RaFD?p=main). Este banco de dados consiste em um conjunto de imagens da face de 67 pessoas exibindo 8 expressões emocionais (raiva, nojo, medo, felicidade, tristeza, surpresa, desprezo, e neutro). A figura 1 mostra um exemplo de pessoa expressando essas 8 emoções.

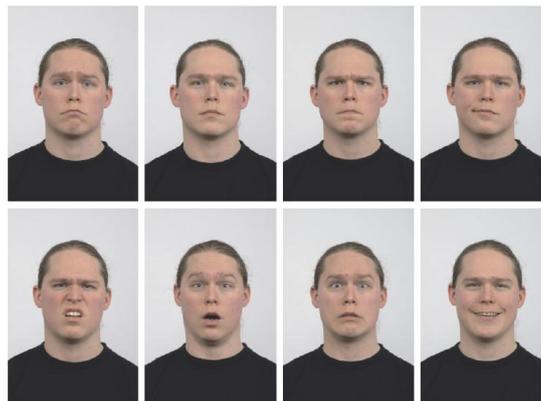


Figura 1. Exemplo de pessoa expressando as 8 emoções.

O trabalho consiste em criar um programa para extrair características relevantes das imagens de forma a gerar uma base de treinamento e uma base de validação. Normalmente gera-se um único conjunto de dados e depois se divide ele em dois conjuntos (treinamento e validação). Cada imagem da base de dados será um exemplo de treinamento. Em seguida deve-se treinar 4 classificadores diferentes com o conjunto de dados de treinamento e depois fazer classificação do conjunto de dados de validação para verificar se o classificador aprendeu a identificar corretamente as emoções expressadas a partir das características visuais extraídas da face das pessoas. Os algoritmos de aprendizado supervisionado que devem ser utilizados são:

- Árvores de Decisão.
- K-Nearest Neighbor (KNN)
- Support Vector Machine (SVM)
- Rede Neural (usando backpropagation)

O objetivo do trabalho é observar como cada um dos classificadores se comporta com este conjunto de dados e qual apresenta o melhor resultado. No final do trabalho, deve ser entregue um relatório descrevendo a etapa de seleção e extração das características das imagens; e também o resultado dos experimentos realizados para cada um dos algoritmos, incluindo a taxa de acertos, tempo gasto no processo de treinamento e tempo gasto no processo de classificação.

Programa Base

O programa base fornecido demonstra como alguns atributos podem ser extraídos das imagens de treinamento. Basicamente o que o programa faz é identificar a posição da face na imagem e, dentro da face, localizar a posição da boca e dos olhos. Em seguida, algumas características estruturais da boca e olhos são extraídas. As características extraídas pelo programa base são:

- 1) Largura do retângulo englobante do maior objeto segmentado encontrado na região da boca.
- 2) Altura do retângulo englobante do maior objeto segmentado encontrado na região da boca.
- 3) Perímetro do maior objeto segmentado encontrado na região da boca.
- 4) Área do maior objeto segmentado encontrado na região da boca.
- 5) 1° momento invariante de Hu do maior objeto segmentado encontrado na região da boca.
- 6) 2° momento invariante de Hu do maior objeto segmentado encontrado na região da boca.
- 7) 3° momento invariante de Hu do maior objeto segmentado encontrado na região da boca.
- 8) 4° momento invariante de Hu do maior objeto segmentado encontrado na região da boca.
- 9) 5° momento invariante de Hu do maior objeto segmentado encontrado na região da boca.
- 10) 6° momento invariante de Hu do maior objeto segmentado encontrado na região da boca.
- 11) 7° momento invariante de Hu do maior objeto segmentado encontrado na região da boca.
- 12) Largura do retângulo englobante do maior objeto segmentado encontrado na região do olho direito.
- 13) Altura do retângulo englobante do maior objeto segmentado encontrado na região do olho direito.
- 14) Perímetro do maior objeto segmentado encontrado na região do olho direito.
- 15) Área do maior objeto segmentado encontrado na região do olho direito.
- 16) 1° momento invariante de Hu do maior objeto segmentado encontrado na região do olho direito.

- 17) 2º momento invariante de Hu do maior objeto segmentado encontrado na região do olho direito.
- 18) 3º momento invariante de Hu do maior objeto segmentado encontrado na região do olho direito.
- 19) 4º momento invariante de Hu do maior objeto segmentado encontrado na região do olho direito.
- 20) 5º momento invariante de Hu do maior objeto segmentado encontrado na região do olho direito.
- 21) 6º momento invariante de Hu do maior objeto segmentado encontrado na região do olho direito.
- 22) 7º momento invariante de Hu do maior objeto segmentado encontrado na região do olho direito.
- 23) Largura do retângulo englobante do maior objeto segmentado encontrado na região do olho esquerdo.
- 24) Altura do retângulo englobante do maior objeto segmentado encontrado na região do olho esquerdo.
- 25) Perímetro do maior objeto segmentado encontrado na região do olho esquerdo.
- 26) Área do maior objeto segmentado encontrado na região do olho esquerdo.
- 27) 1º momento invariante de Hu do maior objeto segmentado encontrado na região do olho esquerdo.
- 28) 2º momento invariante de Hu do maior objeto segmentado encontrado na região do olho esquerdo.
- 29) 3º momento invariante de Hu do maior objeto segmentado encontrado na região do olho esquerdo.
- 30) 4º momento invariante de Hu do maior objeto segmentado encontrado na região do olho esquerdo.
- 31) 5º momento invariante de Hu do maior objeto segmentado encontrado na região do olho esquerdo.
- 32) 6º momento invariante de Hu do maior objeto segmentado encontrado na região do olho esquerdo.
- 33) 7º momento invariante de Hu do maior objeto segmentado encontrado na região do olho esquerdo.

Com esse conjunto de atributos é possível conseguir uma taxa de precisão de aproximadamente 70% (testando com SVM). É bem provável que alguns dos atributos não sejam relevantes e estejam piorando a classificação ou que seja necessário adicionar mais atributos para melhorar ainda mais a taxa de reconhecimento.

O programa base está disponível no seguinte link:

http://edirlei.eternix.com.br/aulas/ia_2011_2/Trabalho3ProgramaBase.zip

O programa base foi desenvolvido utilizando a biblioteca de visão computacional OpenCV (<http://opencv.willowgarage.com/wiki/>). Mais detalhes sobre o funcionamento do programa base e como utilizar o OpenCV serão apresentados durante a aula.

Informações Adicionais:

- Não é necessário implementar todos os classificadores. É permitida a utilização de bibliotecas externas, como por exemplo, a LibSVM (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>).
- Não é necessário desenvolver um programa totalmente integrado. Você pode ter um pequeno programa para cada classificador.
- Os programas podem ser implementados em qualquer linguagem (C, C++, C#, Java...).
- Você deve decidir quais atributos serão utilizados pelos classificadores. A classificação pode piorar ou melhorar dependendo do conjunto de atributos utilizado.
- Os atributos extraídos pelo programa base são apenas exemplos de atributos que podem ser extraídos. Você deve escolher quais atributos serão utilizados e fazer testes para verificar qual conjunto de atributos apresenta melhores resultados.
- Para se conseguir melhores resultados será necessário extrair mais características das imagens de treinamento. O programa base extrai apenas algumas características estruturais, mas existem muitas outras informações úteis que podem ser extraídas computacionalmente das imagens.

Forma de Avaliação:

Será avaliada a qualidade do relatório final e a correta utilização dos classificadores. Neste trabalho, o mais importante será o relatório final contendo os resultados dos testes. Entretanto, também será necessário apresentar os programas para comprovar os dados do relatório.

Bônus:

Os trabalhos que conseguirem uma taxa de reconhecimento superior a 90% receberam 3.0 pontos extras na nota. Podendo tirar até 13 no trabalho.

Data de Entrega:

05/12

Forma de Entrega:

O programa e o relatório devem ser apresentados na aula do dia 05/12 (segunda) e enviados até o dia 05/12 para o email edirlei.slima@gmail.com. Como estamos no final do semestre, não serão aceitos trabalhos enviados depois desta data.