

Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação - ICMC

SCC0276 - Aprendizado de Máquina

Projeto 1 - Reconhecimento de Faces

Entrega: 16/05

Este projeto consiste em implementar classificadores de faces em dois conjuntos distintos e avaliar sua performance utilizando métricas estudadas em aula. Também serão utilizadas técnicas de Data Augmentation para enriquecer um dos conjuntos de dados. As bases de dados utilizadas serão:

- The Database of Faces (ORL Faces) [3]: O dataset original possui 10 fotos de 20 pessoas diferentes, totalizando 200 fotos. A orientação da face, bem como presença de acessórios como óculos, muda de foto para foto. A organização é feita por pastas, onde cada pasta representa uma pessoa.
- Pessoas ICMC: Dataset com fotos de professores e funcionários do ICMC. Cada pessoa possui 1 foto, totalizando 20 pessoas e 20 fotos.

Parte 1 - Pré-processamento

Data Augmentation

Algumas vezes é útil "enriquecer" a base de dados, inserindo ruído, realizando rotação, espelhamento, entre outras operações. Esse processo é conhecido como *Data Augmentation* e é muito utilizado para adicionar variabilidade nos dados para, por exemplo, aumentar a robustez de classificadores.

Neste trabalho utilizaremos Data Augmentation para gerar mais casos de exemplo para o dataset "Pessoas ICMC", de forma que cada imagem do conjunto original tenha 9 variações (totalizando 10 imagens por pessoa). Para isso, deverão ser utilizadas as ferramentas vistas em aula, como as utilizadas em [2]. Adicione ruído, rotação, blur, etc, às imagens e expanda o dataset "Pessoas ICMC" para que o mesmo fique com o mesmo número de imagens da versão do dataset "ORL Faces" utilizado.

Extração de características - HOG

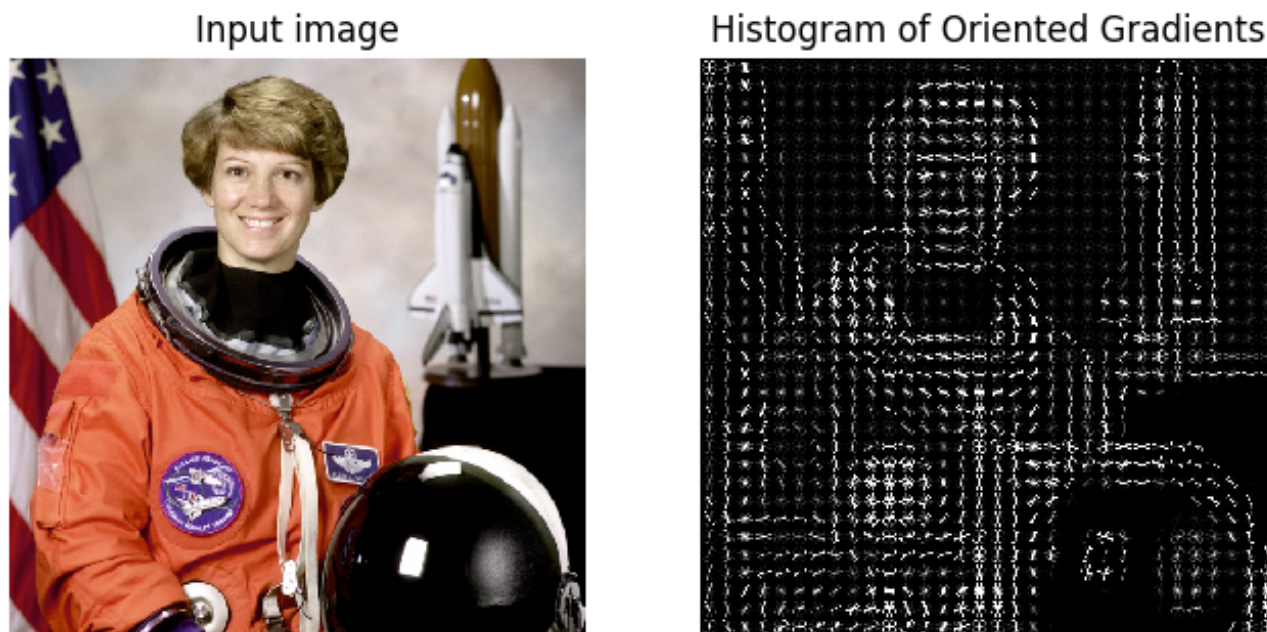


Figura 1: Visualização da técnica Histogram of Oriented Gradients (HOG) [4]

Geralmente, ao se trabalhar com imagens, é comum não realizar a classificação diretamente sobre a imagem original. Primeiramente, é realizada algum tipo de extração de característica na imagem. Essa operação tem como finalidade, entre outros motivos, facilitar o processo de classificação, bem como reduzir a dimensionalidade dos dados. **Para este trabalho, a técnica Histogram of oriented gradients (HOG) [1] deverá ser utilizada para extrair características dos dados ANTES de se realizar classificação.**

Cada imagem de ambos os datasets (no caso do dataset "Pessoas ICMC" o processo de Data Augmentation já deverá ter sido realizado) deverá ter seu HOG calculado e essa será a nova representação da imagem. **Nas próximas etapas do projeto você deve trabalhar com as características HOG extraídas de cada imagem, e NÃO com os dados originais.** O número de características que serão extraídas varia conforme a parametrização do método. Ou seja, a dimensão da representação por HOG de uma imagem vai depender da parametrização utilizada.

Parte 2 - Classificadores

Para classificar os datasets, já pré-processados, dois tipos de classificadores deverão ser utilizados:

- **K-nearest Neighbors (KNN)** [7]: Utilizar KNN para $k = 3$ e $k = 5$ e $k = 7$ nos dois conjuntos.
- **Multilayer Perceptron (MLP) com termo Momentum** [8]: Utilizar uma rede neural do tipo MLP com 1 e 2 camadas intermediárias (variando seus tamanhos) e variar os parâmetros de aprendizado e momentum.

Para determinar a configuração que oferece melhor acurácia [5] para cada tipo de classificador, utilize 10-fold cross-validation estratificado [10]. **Lembre-se de reportar as acurácias de cada configuração testada, de modo que fique claro todas as configurações exploradas e suas respectivas acurácias.**

Após determinar a configuração que apresenta melhor acurácia (tanto para o KNN quanto para o MLP), calcule a matriz de confusão [6] e reporte a acurácia e precisão dos classificadores para cada dataset.

Agora investigue o efeito da redução de dimensionalidade na qualidade dos classificadores. Para isso, **aplique PCA [9] para reduzir o número de atributos para o número de atributos que corresponde a aproximadamente 50% da variância.** Novamente, calcule a matriz de confusão e reporte a acurácia e precisão dos classificadores para cada dataset com dimensionalidade reduzida.

Relatório

O grupo deverá escrever um relatório breve (2 a 5 páginas) que deverá conter:

- Descrição em alto nível da organização do código
- Quais foram as operações utilizadas para Data Augmentation
- A acurácia de cada configuração testada na Parte 2. O uso de tabelas é bem vindo.
- A matriz de confusão, acurácia e precisão dos classificadores que apresentaram melhor desempenho, para os dois conjuntos de dados
- A matriz de confusão, acurácia e precisão dos classificadores que apresentaram melhor desempenho, agora para os dois conjuntos de dados com dimensionalidade reduzida.
- Uma breve discussão comparando o desempenho do KNN e do MLP, na base de dados com e sem redução de dimensionalidade

Entrega

A entrega deverá ser feita pela ferramenta "Escaneinho" do Tidia. Deverão constar na entrega o código fonte e o relatório.

Referências

- [1] *Histogram of oriented gradients*. URL: https://en.wikipedia.org/wiki/Histogram_of_oriented_gradients.
- [2] *Image manipulation augmentation with skimage*. URL: <https://www.kaggle.com/tomahim/image-manipulation-augmentation-with-skimage>.
- [3] Cambridge University Computer Laboratory. *The Database of Faces*. 1994. URL: <https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>.
- [4] *Scikit-image: Histogram of Oriented Gradients*. URL: https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_hog.html.
- [5] *Sklearn: Accuracy Score*. URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.accuracy_score.html.
- [6] *Sklearn: Confusion Matrix*. URL: https://scikit-learn.org/0.17/modules/generated/sklearn.metrics.confusion_matrix.html.
- [7] *Sklearn: K-Neighbors Classifier*. URL: <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html>.
- [8] *Sklearn: MLP Classifier*. URL: https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html.
- [9] *Sklearn: PCA*. URL: <https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>.
- [10] *Sklearn: Stratified K-fold*. URL: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.StratifiedKFold.html.