

Detecção de imagens: uma abordagem sobre reconhecimento facial.

Iury T. Monte, Ulisses C. Silva

Departamento de Computação – Universidade Federal Rural de Pernambuco
(UFRPE)
Rua Dom Manuel de Medeiros, s/n - Dois Irmãos, Recife - PE, 52171-900 Brasil
{iury.tavares, ulisses.chaves, andre.camara} @ufrpe.br

Abstract. *The present work deals with an analysis, through machine learning, of images, based on photographs, from the UCI Machine Learning Repository, of faces. In this analysis we used three algorithms, KNN, SVM, K-means and compared their performances in the detection of the images. The results obtained with the KNN, SVM and K-means algorithms were considered from medians to promising, with a caveat for SVM that presented accuracy above 70%.*

Resumo. *O presente trabalho trata-se de uma análise, por meio de aprendizagem de máquina, de imagens, tendo como base fotografias, da UCI Machine Learning Repository, de rostos. Nesta análise utilizamos três algoritmos, KNN, SVM, K-means e comparamos seus desempenhos na detecção das imagens. Os resultados obtidos com os algoritmos KNN, SVM e K-means foram considerados de medianos a promissores, com ressalva para o SVM que apresentou acurácia acima de 70%.*

Palavras-chaves. *Inteligência Artificial, Machine Learning, Dectecção de Imagens.*

1. Introdução

Hoje em dia é muito frequente as situações em que as pessoas precisam confirmar a autenticidade de sua identidade, como em eventos, compras via internet e transações bancárias. Uma das formas mais utilizadas para tal tarefa são as senhas pessoais que são facilmente burláveis e que qualquer indivíduo pode conseguir a senha de outra pessoa conseguindo, desse modo, passar pelo bloqueio de segurança imposto pela mesma.

Em primeira análise, o reconhecimento facial pode ser uma das formas mais eficientes para evitar que fraudes de identidade aconteçam, uma vez que o reconhecimento facial é baseado em uma técnica biométrica em que os softwares codificam nossos rostos. Para que o mapeamento de um rosto seja

feito, precisamos utilizar características do rosto de uma, como tamanho do queixo e a distância entre os olhos, conhecidos como pontos nodais.

Nota-se que, para isso, teremos que utilizar algoritmos de aprendizagem de máquina precisos, robustos e a prova de falsificação. A fim de que problemas de autenticidade de identidade, falsas compras e furtos bancários possam ser evitados. Logo, neste artigo iremos abordar a importância de algoritmos de aprendizagem de máquina para a detecção facial, falando um pouco sobre filtros de imagens, sistemas de cores usados para a detecção de uma face e assinatura facial, a fim de desenvolvermos uma solução para essa problemática.

2. Materiais

Nesse trabalho foi utilizado o banco de dados UCI Machine Learning Repository que contém 640 imagens em escala de cinza no formato PNG de 20 pessoas. Existem 32 imagens por assunto, uma por expressão ou configuração facial diferente: centro-luz, com óculos, feliz, luz a esquerda, sem óculos, normal, luz direta, sonolento e surpreso. Desse conjunto de dados, utilizamos 58 imagens das quais 75% foi utilizado para treino e 25% para teste.

3. Métodos

A aprendizagem de máquina é um subcampo da Engenharia e da ciência da computação que evoluiu do estudo de reconhecimento de padrões e da teoria do aprendizado computacional em inteligência artificial, explorando a construção de algoritmos que podem aprender de seus erros e fazer previsões sobre dados.

Consideramos utilizar neste projeto três algoritmos de aprendizagem (dois supervisionados e um não supervisionado): KNN, SVM, KMEANS. Além disso, nós iremos utilizar um método inventado em 2005 chamado Histogramas de Gradientes Orientados (histogram of oriented gradients) — ou apenas HOG para encontrar um rosto em uma imagem.

3.1 Histogramas de Gradientes Orientados (HOG)

A ideia principal do descritor Histograma de Gradientes Orientados é que a aparência e forma de objetos em uma imagem podem ser descritos através da distribuição dos gradientes de intensidade dos pixels ou pelas direções das bordas. O processo para gerar o descritor pode ser dividido em quatro etapas: cálculo do gradiente em cada pixel, agrupamento dos pixels em células, agrupamento das células em blocos e obtenção do descritor. Primeiro utiliza-se máscaras derivada discreta pontual tanto no eixo vertical como horizontal para o cálculo do gradiente de cada pixel. O passo seguinte é responsável por agrupar os pixels de uma determinada região, criando-se o que se chama de célula. Após a segunda etapa, os blocos são criados através do agrupamento de células de uma certa região. Na etapa final, cria-se o descritor, que nada mais é do que uma lista dos histogramas de todas as células de todos os blocos. A atenuação do problema das variações locais de iluminação ou de contraste entre o primeiro plano e o plano de fundo se dá através da normalização de cada histograma de acordo com seus próprios valores.

3.2 K-Nearest Neighbors (KNN)

É um algoritmo de aprendizagem supervisionada. Caracterizado por ser um classificador onde o aprendizado é baseado no quão similar é um dado (vetor) do outro, ou seja, todas as instâncias correspondem a pontos em um espaço n -dimensional e a vizinhança é definida por uma função de similaridade (quanto menor a distância, maior a similaridade), tendo isso em vista a classe de um novo exemplo é definida a partir dos vizinhos mais próximos. Um ponto importante é que a vizinhança é definida comumente pela distância Euclidiana, mas também pode ser usado outras funções para esse cálculo.

3.3 Máquina de Vetores Suporte (SVM)

Máquinas de Vetores Suporte (SVMs, do inglês Support Vector Machines) é um algoritmo de aprendizado de máquina (AM) supervisionado que pode ser utilizado tanto para classificação, quanto para regressão (Soares, 2008). Esta técnica originalmente desenvolvida para classificação binária, busca a construção de um hiperplano como superfície de decisão, de forma que a

separação entre os exemplos seja máxima. SVMs lineares são bons para conjunto de dados linearmente separáveis. No entanto, existem muitos casos em que não é possível dividir o conjunto de treinamento linearmente por um hiperplano. Nesses casos, podemos mapear os dados para um espaço de dimensão mais alta, no qual os dados passam a ser linearmente separáveis. Rótulos ou classes são os fenômenos que desejamos realizar uma previsão. Para SVMs de classificação estes rótulos possuem valores discretos $(1, \dots, n)$. Nos casos que estes valores sejam contínuos temos o SVM para regressão. Para dados que não podem ser separados por um hiperplano, temos o SVM não linear. A técnica adotada aqui é aumentar a dimensionalidade do espaço amostral dos dados, isto é, caso os dados estejam em R^d passamos os dados para o espaço R^x , em que $x > d$, onde os dados possam ser separados por um hiperplano. Para isso é necessário uma função kernel (Soares, 2008).

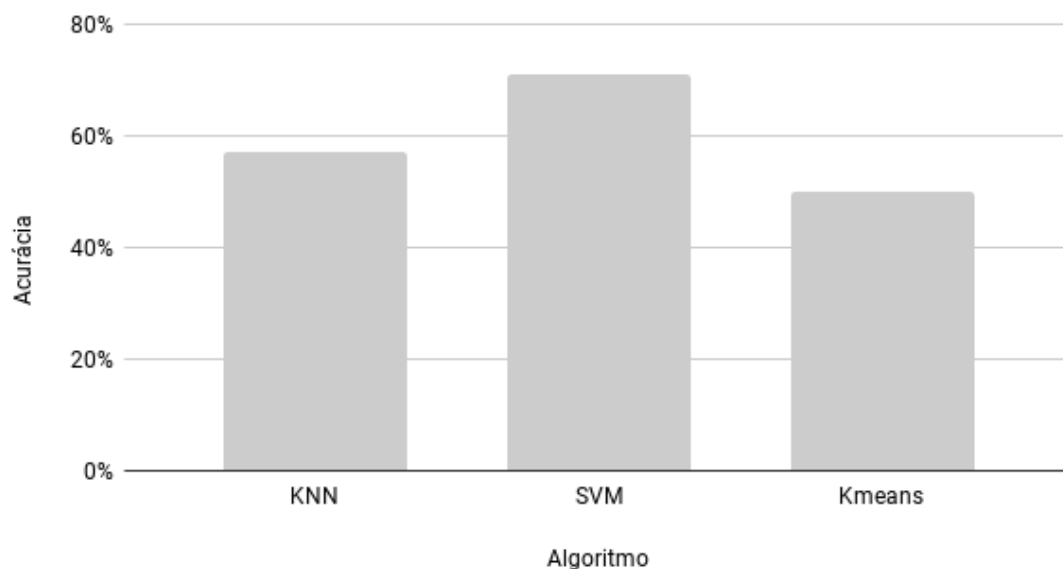
3.4 K-means

Agrupamento k-means é um método de Clustering que objetiva particionar n observações dentre k grupos onde cada observação pertence ao grupo mais próximo da média. Desta maneira, o algoritmo automaticamente vai fornecer uma classificação automática sem a necessidade de nenhuma supervisão humana, ou seja, sem nenhuma pré-classificação existente. O usuário deve fornecer ao algoritmo a quantidade de classes que ele deseja. Este número de classes que deve ser passada para o algoritmo é chamado de k e é daí que vem a primeira letra do algoritmo: K-Means. Para gerar as classes e classificar as ocorrências, o algoritmo faz uma comparação entre cada valor de cada linha por meio da distância. Geralmente utiliza-se a distância euclidiana para calcular o quão 'longe' uma ocorrência está da outra. A maneira de calcular esta distância vai depender da quantidade de atributos da tabela fornecida. Após o cálculo das distâncias o algoritmo calcula centróides para cada uma das classes. Conforme o algoritmo vai iterando, o valor de cada centróide é refinado pela média dos valores de cada atributo de cada ocorrência que pertence a este centróide. Com isso, o algoritmo gera k centróides e coloca as ocorrências da tabela de acordo com sua distância dos centróides.

4. Experimentos e Resultados

Para avaliar a eficiência dos algoritmos K-means, KNN, SVM foram verificadas as suas acurácias através de suas respectivas bibliotecas. Os resultados entre os algoritmos foram parecidos, porém o que se apresentou mais promissor foi o SVM. O algoritmo K-means foi o que teve menor acurácia, apresentando uma taxa de 50% de acerto, sendo um algoritmo de aprendizado não supervisionado de agrupamento. No treino, as duas classes foram divididas por ele e no teste, foi comparado o resultado de um 'predict' do K-Means com as 'labels' usadas nos algoritmos supervisionados. No KNN, utilizamos $K = 5$ (podendo variar com uma base de dados maior e alterando o valor de k), que é definido por padrão da própria biblioteca, com isso o algoritmo conseguiu apresentar acurácia de 57,14%. O algoritmo SVM, se mostrou mais eficiente comparado ao KNN e Kmeans, apresentando acurácia de 71,43%, nele optamos por dividir o conjunto de treinamento linearmente, mapeando os dados em duas classes, o que é rosto e o que não é.

Histograma de Acurácia



5. Conclusão

Após a realização de diversos testes com os algoritmos KNN, SVM e Kmeans ficou comprovado que, para nosso banco de dados, o SVM se mostrou mais

promissor. Ele apresenta uma simples implementação com a maior taxa de acerto. Cabe ressaltar que os algoritmos de reconhecimento facial usualmente demandam de uma grande capacidade de processamento, por trabalharem com a extração de características baseadas nas imagens digitais como um todo. Este fato tem um impacto direto no que tange a precisão e índice de acerto dos algoritmos de reconhecimento facial: quanto mais preciso for o software de reconhecimento, maior a capacidade de processamento necessária para a execução do mesmo.

Referências

ANTONELLO, Leonardo Leite¹ Ricardo. IDENTIFICAÇÃO AUTOMÁTICA DE PLACA DE VEÍCULOS ATRAVÉS DE PROCESSAMENTO DE IMAGEM E VISÃO COMPUTACIONAL. Disponível em: <<http://professor.luzerna.ifc.edu.br/ricardo-antonello/wp-content/uploads/sites/8/2017/02/Livro-Introdu%C3%A7%C3%A3o-a-Vis%C3%A3o-Computacional-com-Python-e-OpenCV-3.pdf>>. Acesso em: 18/07/2019.

GOMIDE, Renato de Sousa; VIEIRA, Marcus Fraga. SEGMENTAÇÃO DA PELE NO ESPAÇO DE COR RGB PARA UTILIZAÇÃO EM AMBIENTES VIRTUAIS. Sociedade Brasileira para o Progresso da Ciência. Disponível em: <<http://www.sbpnet.org.br/livro/63ra/conpeex/mestrado/trabalhos-mestrado/mestrado-renato-sousa.pdf>>. Acesso em: 18/07/2019.

UFOP. Laboratório Mobilis. Detecção de Pedestres Usando o OpenCV. Disponível em: <<http://www.decom.ufop.br/imobilis/deteccao-de-pedestres-usando-o-opencv/>>. Acesso em: 18/07/2019.

UCI. Machine Learning Repository. CMU Face Images Data Set. Disponível em: <<http://archive.ics.uci.edu/ml/datasets/cmu+face+images>>. Acesso em: 18/07/2019.

Learn OpenCV. Histogram of Oriented Gradients. Disponível em: <<https://www.learnopencv.com/histogram-of-oriented-gradients/>>. Acesso em: 18/07/2019.