

Projeto de uma Rede Neural Artificial para Reconhecimento Facial

Matheus Augusto Wisniewski

Professora: Lucia Valeria Ramos de Arruda

Disciplina: RNA - Redes Neurais Artificiais

Universidade Tecnológica Federal do Paraná

Curitiba, Março de 2016

Resumo: Neste trabalho são apresentados os passos para a implementação de uma Rede Neural para o Reconhecimento Facial simples, além da justificativa da escolha de cada um deles.

Palavras-chave: Redes Neurais, Reconhecimento Facial

1 Introdução

O reconhecimento facial é uma tarefa tão comum aos seres humanos que um indivíduo nem percebe o inacreditável número de vezes que é realizado todos os dias, como ilustrado na Figura 1.

Embora a pesquisa em reconhecimento facial automático venha sendo realizada desde a década de 1960, este tema capturou a atenção da comunidade científica apenas perto do ano 2000. Muitas técnicas de análise e modelagem de rosto têm progredido significativamente desde essa data [1].

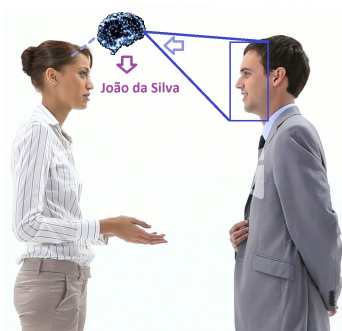


Figura 1: Reconhecimento facial no dia a dia

2 Objetivo

O reconhecimento facial é um método não intrusivo, e imagens de rosto são a característica biométrica mais comumente utilizadas para reconhecimento de identidade [2].

Além disso, a habilidade de realizar análise automática e confiável de rostos possibilita um número muito grande de aplicações. Existem utilidades na área de telecomunicações visuais (como a codificação e a melhoria da qualidade das imagens) e na área de segurança e monitoramento (como a interpretação mais inteligente das cenas), para citar algumas [3].

Esses fatores somados ao objetivo de tornar o reconhecimento facial por redes neurais mais simples de ser entendido são as motivações principais desse trabalho.

3 Justificativa

A complexidade dos problemas de visão computacional é agravada pelo fato de que lida-se com enormes blocos de dados. A imagem em escala de cinza típica tem 640x480 pixels, cada um com 8-bits (ou 256 valores de cinza) de intensidade. Portanto, o tamanho de toda a imagem é 640x480x8 o que resulta no total de 2.457.600 bits [4].

Por conta da altíssima quantidade de dados a serem processados, algoritmos de elevada complexidade tornam-se computacionalmente lentos e, consequentemente, deve-se optar por técnicas que

simplifiquem o processamento [4].

Como se trata de um problema de reconhecimento de padrões, o reconhecimento facial consiste de três partes (após a detecção do rosto, como pode ser visto na Figura 2): o alinhamento, a extração de características e a correspondência com o banco de dados. Essas últimas podem ser realizadas com alto desempenho através das redes neurais [5].

Este trabalho visa elucidar as fases posteriores à detecção de rostos, através de redes neurais.

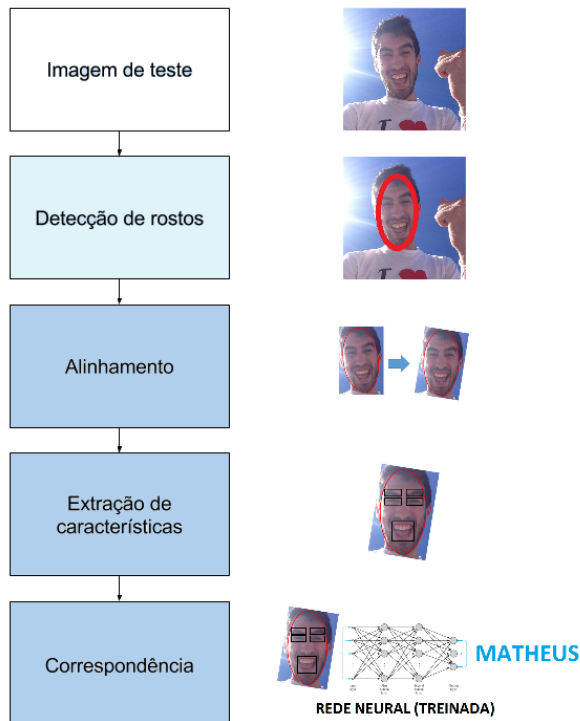


Figura 2: Sistema de reconhecimento facial

4 Descrição do problema

O problema a ser tratado neste trabalho é o reconhecimento facial simples e replicável através de redes neurais.

4.1 Localização do rosto

A primeira etapa da Figura 2 foge do escopo deste estudo, já que para fins de teste pode-se utilizar uma base de dados e imagens de teste com a localização do rosto controlada, como a base de dados de rostos de Yale [7], por exemplo.

Alternativamente, algoritmos populares como o de Viola & Jones [8] também são recomendados para ambientes não-controlados.

4.2 Identificação do rosto

As três etapas que seguem a detecção do rosto na Figura 2 são de responsabilidade das redes neurais implementadas conforme descrito na Seção 5.

5 Construção do sistema neural

A escolha da rede neural foi baseada em um artigo que comparou a eficiência de várias combinações de redes neurais para o reconhecimento facial [9].

5.1 Tipo de Rede

A rede neural foi escolhida com base no estudo comparativo de eficiência de redes neurais feito por Oravec & Pavlovicova [9], como pode ser visto na Tabela 1:

Redes Neurais	% de sucesso
PCA-hloMLP	83,07
blockMLP-RBF	82,29
PCA	81,51
blockSOM-RBF	80,73
blockSOM-MLP	79,95
RBF	78,12
MLP	78,12
MLP-MLP	74,74
blockMLP-MLP	73,70
MLP-RBF	72,40

Tabela 1: Comparativo de redes neurais [9]

Pela facilidade de implementação, aliada com a grande eficiência, a rede neural escolhida foi a **blockMLP-RBF**.

Esta é a combinação de uma rede Perceptron Multicamadas (MLP, do inglês *Multi-Layer Perceptron*) no início, para comprimir a imagem inicial (extraíndo dessa forma as características principais), e uma rede Função de Base Radial (RBF, do inglês *Radial-Base Function*) como classificadora (aquela que define a quem pertence o rosto no final). A Figura 3 ilustra o esquema:

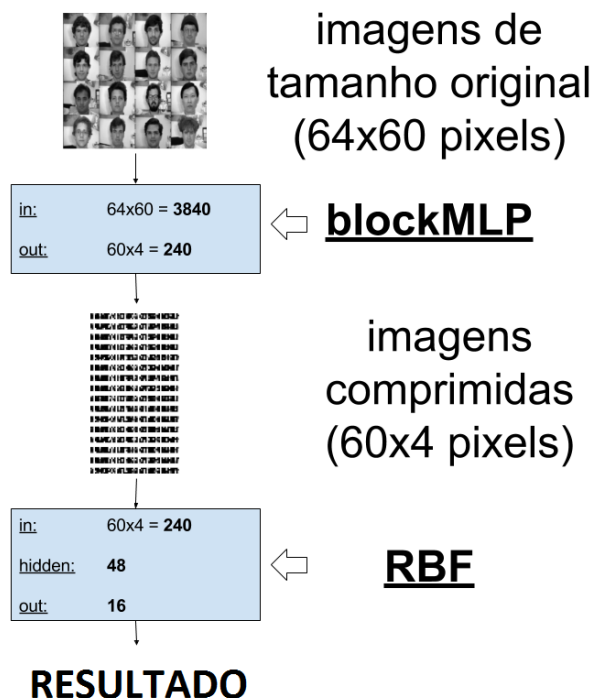


Figura 3: Redes neurais do projeto

5.2 Entradas e saídas

Nos blocos azuis da Figura 3 vemos o número de entradas e saídas das redes.

Na rede blockMLP, o número de neurônios na camada de entrada é o número total de pixels da imagem original (no exemplo dado, 64x60 pixels = 3840), e o número de neurônios de saída é 240 para ter imagens comprimidas de 60x4.

Já a rede RBF possui 240 neurônios na camada de entrada por esse ser o número de neurônios na saída da compressão da rede blockMLP, e o número de neurônios de saída é o número total de possibilidades de identidades (no exemplo dado, eram 16 pessoas possíveis).

5.3 Número de camadas escondidas

Como a rede blockMLP serve simplesmente para a compressão da imagem, ela não possui uma camada escondida (não faz extração de características).

A rede RBF, por sua vez, possui uma camada escondida com 48 neurônios por se tratar de uma rede neural classificadora. Essa camada escondida faz a extração das características necessárias para a identificação das pessoas.

5.4 Função de ativação

A função de ativação escolhida foi a função sigmóide, a função de ativação mais comum na construção de redes neurais [10], como pode-se ver na Equação 1:

$$\Phi(v) = \frac{1}{1 + e^{(-av)}} \quad (1)$$

Enquanto a função de Heavyside (degrau unitário) assume apenas os valores 0 ou 1, a função sigmóide pode assumir todos os valores entre 0 e 1, o que a torna diferenciável. Esse é um ponto importante para o treinamento da rede escolhida.

5.5 Treinamento da rede

O algoritmo escolhido para o treinamento da rede é o de retropropagação do erro, descrito pela Equação 2 (para mais informações de como funciona esse tipo de treinamento, consultar Haykin, 1999 [10]).

$$\Delta w_{ji}(n) = \eta X \delta_j(n) X y_i(n) \quad (2)$$

Onde:

$\Delta w_{ji}(n)$ = Correção do peso

η = Parâmetro de aprendizado

$\delta_j(n)$ = Gradiente local

$y_i(n)$ = Sinal de entrada do neurônio anterior

Como a totalidade da rede pode ser resumida por uma rede RBF com 3 camadas escondidas, o treinamento se dá da seguinte forma:

1. Pelo menos 3 fotos de cada pessoa é colocada na entrada (na entrada da rede blockMLP)
2. O usuário ensina a rede as respostas esperadas para cada uma das fotos dando a saída correta delas (a saída da rede RBF)
3. A partir da última camada e até a primeira, os ajustes dos pesos dos neurônios se propagam a cada iteração até que um limiar de sucesso seja atingido

6 Conclusões

Este trabalho mostra que é suficientemente simples a construção e treinamento de uma rede neural de reconhecimento facial com uma taxa de sucesso acima de 80%.

Se levarmos em conta ainda que o software MATLAB pode ser utilizado para que a implementação da rede e do algoritmo de treinamento não sejam necessários (estes já vêm instalados no software), vemos que qualquer um é capaz de treinar e testar o funcionamento desse tipo de rede neural.

Como sugestão de futuras melhorias, poderia ser feita uma implementação completa dessa rede neural na linguagem C++ para se ter controle total sobre os algoritmos de treinamento, dessa forma podendo medir com mais segurança o desempenho da rede.

Tendo feito esse passo, uma verificação das taxas de sucesso da Tabela 1 também teria um valor muito grande.

Concluindo, o potencial para as redes neurais de reconhecimento facial é enorme, haja visto o número de aplicações, e a sua implementação é altamente acessível.

Referências

- [1] FAGERTUN, J. *Face Recognition*. 209 f. Dissertação (Mestrado) - Programa de Mestrado em informática e modelagem matemática, Universidade Técnica da Dinamarca, 2005. Disponível em: <http://etd.dtu.dk/thesis/185830/imm4014.pdf>. [Acesso em 11 de novembro de 2015].
- [2] YESU, Kolhandai. *Hybrid Features Based Face Recognition Method Using Artificial Neural Network*. National Conference on Emerging Trends and Applications in Computer Science, 2012.
- [3] LINGGARD, R.; MYERS, D.J.; NIGHTINGALE, C. *Neural Networks for Vision, Speech and Natural Language*. Estados Unidos da América: Springer Science & Business Media, 2012.
- [4] BALASURIYA, L. S. *Frontal View Human Face Detection and Recognition*. 109 f. Tese (Graduação) - Department of Statistics and Computer Science, University of Colombo, 2000. Disponível em: http://www.dcs.gla.ac.uk/~sumitha/papers/sumitha_undergraduate_thesis.pdf. [Acesso em 09 de novembro de 2015].
- [5] LE, Thai Hoang. *Applying Artificial Neural Networks for Face Recognition*. Disponível em: <http://www.hindawi.com/journals/aans/2011/673016/> [Acesso em 27 de outubro de 2015].
- [6] LI, S. Z. ; JAIN, A. K. *Handbook of Face Recognition*. Springer, Nova Iorque, NY, Estados Unidos da América, 2004.
- [7] Yale Face Database. Disponível em <http://vision.ucsd.edu/content/yale-face-database>. [Acesso em 15 de março de 2016].
- [8] VIOLA, P.; JONES, M. *Rapid object detection using a boosted cascade of simple features*. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 511-518, dezembro de 2001.
- [9] ORAVEC, Miloš; PAVLOVIČOVÁ, Jarmila. *Face Recognition Methods Based on Feedforward Neural Networks, Principal Component Analysis and Self-Organizing Map*. Radioengineering, vol.16, no.1, Bratislava, Eslováquia, Abril de 2007.
- [10] HAYKIN, Simon. *Neural networks and learning machines*. 3 ed. Estados Unidos da América: Pearson, 1999.