

NGHIÊN CỨU THIẾT KẾ THUẬT TOÁN NHẬN DẠNG KHUÔN MẶT DỰA TRÊN MẠNG NƠ RON

AN IMPLEMENTATION OF THE BACKPROPAGATION ALGORITHM FOR IMAGES RECOGNITION

Đậu Trọng Hiền

DH Sư Phạm Kỹ Thuật TP. HCM

TÓM TẮT

Thuật toán lan truyền ngược trong mạng nơ ron là một phương pháp hiệu quả dành cho máy học khi tập dữ liệu đầu vào phức tạp chẳng hạn như hình ảnh. Bài báo này đưa ra một thuật toán nhận dạng ảnh dựa trên thuật toán lan truyền ngược trong mạng nơ ron. Để tăng tốc độ quá trình nhận dạng, ảnh màu được chuyển đổi sang ảnh xám. Quá trình huấn luyện được thực hiện bởi tập ảnh từ rõ đến nhiễu. Điều này giúp mạng thích nghi với vấn đề nhiễu của ảnh trong thực tế.

ABSTRACT

A neural network learning algorithm called Backpropagation is among the most effective approaches to machine learning when the data includes complex sensory input such as images. This document provides an implementation of the Backpropagation algorithm to recognize the images. To increase the speed of recognition color images are scaled to gray images. Note that the training process did not consist of a single call to a training function. Instead, the network was trained several times on various input ideal and noisy images. In this case training a network on different sets of noisy images forced the network to learn how to deal with noise, a common problem in the real world.

I. GIỚI THIỆU

Hệ thống nhận dạng dùng máy móc thay thế cho con người mang lại rất nhiều lợi ích. Riêng về vấn đề nhận dạng mặt người thì hiệu quả kinh tế rất cao. Một máy xét hộ chiếu có thể xử lý nhiều hộ chiếu hơn con người trong cùng một thời gian [1]. Các loại máy nhận dạng giúp ta tiết kiệm thời gian, tiền bạc và loại bỏ những công việc lặp lại nhàm chán. Bài báo này đưa ra giải thuật nhận dạng khuôn mặt người dựa vào mạng nơ ron kết hợp với tiền xử lý ảnh để tăng độ chính xác cũng như tốc độ nhận dạng.

II. THIẾT KẾ THUẬT TOÁN NHẬN DẠNG

Đọc ảnh

Ảnh là tập hợp các ma trận điểm ảnh. Việc đọc ảnh vào chương trình thực chất là việc nạp ảnh vào biến của chương trình để xử lý. Ảnh màu gồm 3 vector ma trận ảnh [3],[4] R (Red: màu đỏ), G (Green: màu xanh lá cây), B (Blue: màu xanh dương), 3 màu cơ bản này

tạo nên tất cả các màu khác theo tỉ lệ khác nhau. Dù là ảnh xám hay ảnh màu thì việc nhận dạng vẫn không thay đổi, khuôn mặt người cần nhận dạng vẫn đảm bảo chính xác. Để tăng tốc độ nhận dạng ta chuyển ảnh màu về ảnh xám là ma trận 2 chiều các điểm ảnh có kích thước mxn.

Chuyển đổi ảnh màu thành ảnh xám

Mục đích của việc chuyển đổi này là biến ảnh màu với ma trận 3 chiều sang ma trận 2 chiều dễ dàng để xử lý nhận dạng. Thuật toán chuyển đổi ảnh màu thành ảnh xám bằng cách chuyển đổi hình thức tổng trọng số của các phần tử R, G, and B :

$$0.2989 * R + 0.5870 * G + 0.1140 * B$$



Ảnh màu ban đầu Ảnh xám
Hình 2.2: Đột ảnh màu thành ảnh xám

Tuy nhiên, ảnh xám vẫn còn chứa rất nhiều yếu tố dư thừa là hình nền của ảnh như cây cối, trời, quần áo,..những yếu tố này không thật sự cần thiết trong quá trình nhận dạng khuôn mặt. Vì thế, cần phải xử lý cắt bỏ các phần dư thừa đó trước khi chuyển sang ảnh xám. Điều quan trọng nữa là cần phải làm cho các hình có kích thước giống nhau để dễ dàng xử lý các điểm ảnh.

Chuẩn hóa kích thước ảnh

Việc chuẩn hóa là nhằm đưa các ảnh đầu vào về cùng một kích thước, nghĩa là có cùng một ma trận điểm ảnh, điều này sẽ tạo một bước thuận lợi cho việc huấn luyện mạng.



30x30 pixel 20x20 pixel **Hình 2.3: ảnh sau khi chuẩn hóa**

Phân lớp:

Trong đề tài này tác giả sử dụng mạng nơ ron để phân lớp ảnh. Việc thiết kế mạng sẽ được đề cập đến trong mục 3.

III. THIẾT KẾ MẠNG NƠRON

Mạng MLP[2] là mạng thường được sử dụng để giải quyết bài toán nhận dạng. Với ba lớp nơron mạng có thể xấp xỉ bất cứ ánh xạ nào trong khi số lượng lớp càng lớn thì độ phức tạp của mạng càng cao cũng như các phép tính mà máy tính phải thực hiện nhiều dẫn đến tốc độ chậm hơn trong khi độ chính xác trong nhận dạng không cải thiện nhiều. Vì vậy trong đề tài này tác giả sử dụng mạng MLP ba lớp: lớp vào, lớp ẩn, lớp ra.

IV. KẾT QUẢ THỰC NGHIỆM

Người thực hiện đề tài chạy thử nghiệm 300 mẫu ảnh trong đó 100 mẫu dùng để huấn luyện mạng, 200 mẫu còn lại dùng để đánh giá độ chính xác của thuật toán nhận dạng.

Thuật toán được chạy thực nghiệm trên máy tính có cấu hình RAM DDR2 1GB, CPU 1.6G 2 nhân.

Chạy thực nghiệm ở độ phân giải ảnh 30x30pixels, 20x20 pixels, 10x10 pixels:

Độ phân giải của ảnh	Thời gian huấn luyện (s)	Tốc độ nhận dạng trung bình 1 khuôn mặt (ms)	Số lượng ảnh nhận dạng đúng trong tập mẫu	Số lượng ảnh nhận dạng đúng trong tập kiểm tra	Độ chính xác trên tập mẫu (%)	Độ chính xác trên tập kiểm tra (%)
30x30 pixels	76.24	5	100	197	100	98.50
20x20 pixels	13.08	4.8	100	192	100	96.00
10x10 pixels	1.8436	0.0045	100	187	100	93.5000

Trong đó tốc độ nhận dạng bao gồm cả phần tiền xử lý ảnh và phần nhận dạng của mạng.

V. KẾT LUẬN

Thuật toán nhận dạng cho phép nhận dạng khuôn mặt với độ chính xác cao với góc nhìn nghiêng không quá 10 độ. Tốc độ của thuật toán là khá nhanh. Có thể nhận dạng tương đối các ảnh nhiễu. Thuật toán nhận dạng khuôn mặt dựa trên mạng nơ ron đem lại kết quả khá cao. Nhận dạng khuôn mặt hay nhận dạng ảnh nói chung là cả một qui trình xử lý thông tin khá phức tạp. Từ bước thu nhận ảnh cho đến bước phân lớp cho một mẫu ta phải đi qua khá nhiều qui trình xử lý và sử dụng nhiều kỹ thuật xử lý ảnh phức tạp[5],[6]. Với kỹ thuật nhận dạng bằng mạng nơron, chúng ta có thể giảm bớt được một số công đoạn xử lý ảnh và đạt được tốc độ nhanh hơn. Hiện nay, các hệ thống đa xử lý (xử lý song song) cũng như các vi mạch tổ hợp phát triển mạnh vì vậy việc tính toán mạng nơron sẽ thuận lợi hơn nhiều. Mạng nơron nhận dạng khuôn mặt như đã thiết kế có khả năng phát triển để nhận dạng và tìm kiếm tội phạm trong cơ sở dữ liệu ảnh lớn hoặc trong những đoạn băng ghi hình do camera quan sát tại hiện trường ghi lại.

TÀI LIỆU THAM KHẢO

- Anil K. Jain, Fundamentals of Digital Image Processing. Prentice Hall, 1991.
- Bernd Jähne, Digital Image Processing, Springer Verlag 1995
- Learning the Uncorrelated, Independent, and Discriminating Color Spaces for Face Recognition Chengjun Liu; Information Forensics and Security, IEEE Transactions on Volume 3, Issue 2, June 2008 Page(s):213 – 222.
- Maximum Confidence Hidden Markov Modeling for Face Recognition Jen-Tzung Chien; Chin-Pin Liao; Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 30, Issue 4, April 2008 Page(s): 606 – 616.
- Nguyễn Đình Thúc, Mạng Nơron, Phương Pháp và Ứng Dụng, NXB Giáo Dục 2000.
- Rein-Lien Hsu, “Face Detection and Modeling for Recognition,” Ph.D. Thesis, Department of Computer Science & Engineering, Michigan State University, USA, 2002.