

NGHIÊN CỨU THIẾT KẾ PHẦN MỀM NHẬN DIỆN KHUÔN MẶT ỨNG DỤNG THUẬT TOÁN TRÍ TUỆ NHÂN TẠO HAAR CASCADES

RESEARCH AND DESIGN FACE RECOGNITION SOFTWARE USING HAAR CASCADES ARTIFICIAL INTELLIGENCE ALGORITHM

Hoàng Anh Tuấn, Phạm Trung Thiên*, Đỗ Văn Tĩnh, Nguyễn Ngọc Thế

Khoa Cơ khí, Trường Đại học Kinh tế - Kỹ thuật Công nghiệp

Đến Tòa soạn ngày 14/12/2023, chấp nhận đăng ngày 27/02/2024

Tóm tắt: Nhận diện khuôn mặt (Face detection) là một phần quan trọng của thị giác máy tính trí tuệ nhân tạo, nhận diện khuôn mặt được ứng dụng khá phổ biến cả trong công nghiệp sản xuất lẫn đời sống hàng ngày. Bài báo này nghiên cứu hệ thống nhận diện khuôn mặt sử dụng thuật toán Haar Cascades để nhận dạng đối tượng khuôn mặt trong khung hình theo thời gian thực và định vị khuôn mặt trên khung hình. Hệ thống gồm module huấn luyện nhận dạng khuôn mặt với dữ liệu huấn luyện đầu vào của module và thuật toán trích trọt đặc trưng khuôn mặt. Sau khi có hình ảnh theo thời gian thực từ camera, tiền xử lý dữ liệu và đẩy dữ liệu vào module đã được huấn luyện nhận dạng khuôn mặt để trả ra vị trí của khuôn mặt, có vị trí khuôn mặt trên khung hình sẽ khoanh vùng khuôn mặt để hiển thị kết quả khuôn mặt trên khung hình. Kết quả nghiên cứu cho tỉ lệ chính xác với điều kiện ánh sáng ổn định lên đến 96%, tuy nhiên với điều kiện ánh sáng không ổn định, khoảng cách camera hướng tới đối tượng xa trên 500 mm độ chính xác chỉ đạt 66%.

Từ khóa: Nhận diện khuôn mặt, Haar Cascades, trí tuệ nhân tạo.

Abstract: Face detection is an essential component of computer vision in artificial intelligence, and it finds widespread applications in both industrial manufacturing and daily life. This paper investigates a face detection system using the Haar Cascades algorithm to recognize face objects in real-time frames and locate faces within the frames. The system comprises a face recognition training module with input training data and a feature extraction algorithm for faces. After obtaining real-time images from the camera, data preprocessing is conducted, and the data is fed into the trained face recognition module to output the position of the face. Once the face's position on the frame is determined, the face area is highlighted to display the face detection results on the frame. The research results indicate an accuracy rate of up to 96% under stable lighting conditions. However, under unstable lighting conditions, with the camera focused on distant objects beyond 500mm, the accuracy rate drops to 66%.

Keywords: Face detection, haar cascades, artificial intelligence.

1. ĐẶT VẤN ĐỀ

Bài toán nhận dạng khuôn mặt là một bài toán tồn tại từ lâu và vẫn đang đối mặt với những thách thức khó khăn, chưa đạt được sự hoàn hảo. Tuy nhiên, với sự phát triển đột phá của khoa học máy tính, các công bố của các nhóm nghiên cứu trên thế giới đã giúp bài toán nhận

dạng khuôn mặt được ứng dụng rộng rãi trong nhiều lĩnh vực như y học, quân sự, công nghiệp và cả đời sống. Có nhiều phương pháp khác nhau để cải thiện hiệu suất độ chính xác khi nhận diện với các yếu tố ảnh hưởng như ánh sáng yếu, góc chụp, kích thước hình ảnh và ảnh hưởng của môi trường.

Tích hợp nhận dạng khuôn mặt là một lĩnh vực quan trọng trong nghiên cứu về trí tuệ nhân tạo và thị giác máy tính. Hiện nay, các phương pháp nhận dạng khuôn mặt được phân loại thành ba hướng tiếp cận chính dựa trên các tiêu chí khác nhau. Đầu tiên, có hướng tiếp cận toàn cục (Global) như Eigenfaces-PCA và Fisherfaces-LDA. Thứ hai, là hướng tiếp cận dựa trên các đặc điểm cục bộ (Local Feature Based), ví dụ như LBP và Gabor Wavelets. Cuối cùng, hướng tiếp cận lai (Hybrid) là sự kết hợp của hai hướng tiếp cận trước.

Thuật toán nhận dạng khuôn mặt Haar Cascades là một hướng tiếp cận dựa trên các đặc điểm cục bộ (Local Feature Based). Thuật toán này sử dụng một bộ lọc phức tạp được gọi là Haar-like features để phát hiện các đặc điểm cục bộ trên khuôn mặt, chẳng hạn như cạnh mắt, mũi, miệng, và các đặc trưng quan trọng khác. Sau đó, kết hợp các đặc điểm này để xác định xem có một khuôn mặt trong ảnh hay không. Mặc dù Haar Cascades có khả năng phát hiện khuôn mặt trong thời gian thực và hiệu quả, nhưng nó chủ yếu tập trung vào việc phát hiện các đặc điểm cục bộ trên khuôn mặt, nên nó thuộc vào hướng tiếp cận dựa trên các đặc điểm cục bộ (Local Feature Based) hơn là hướng tiếp cận toàn cục (Global) hoặc hướng tiếp cận lai (Hybrid).

Thuật toán Haar Cascades đã đánh dấu sự phát triển đáng kể trong lĩnh vực nhận dạng khuôn mặt và nhận dạng đối tượng. Một số nghiên cứu nổi bật liên quan đến Haar Cascades bao gồm: Paul Viola và Michael Jones (2001): Hai nhà nghiên cứu này đã công bố bài báo nổi tiếng về việc sử dụng Haar Cascades cho việc nhận dạng khuôn mặt trong lĩnh vực thị giác máy tính. OpenCV Library: OpenCV (Open Source Computer Vision Library) là một thư viện mã nguồn mở rất phổ

biến trong việc xây dựng ứng dụng thị giác máy tính. Thư viện này đã tích hợp Haar Cascades cho việc nhận dạng khuôn mặt và đối tượng. Sự phát triển và cải tiến của OpenCV trong việc sử dụng Haar Cascades đã giúp phổ cập công nghệ này trong cộng đồng phát triển phần mềm và nghiên cứu. Các ứng dụng trong an ninh và thị giác máy tính: Haar Cascades đã được áp dụng rộng rãi trong các ứng dụng thực tế như hệ thống giám sát an ninh, nhận dạng khuôn mặt trên các nền tảng di động, nhận dạng biểu cảm khuôn mặt, và nhiều lĩnh vực khác.

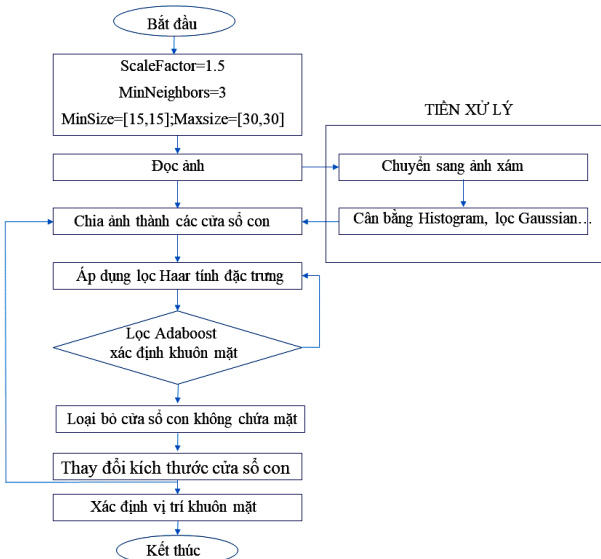
2. XÂY DỰNG HỆ THỐNG

2.1. Nguyên lý thuật toán

Hệ thống sử dụng thuật toán Haar Cascades một phương pháp tiên tiến và nhanh chóng hơn trong xử lý hình ảnh và phát hiện khuôn mặt, sử dụng các đặc điểm hình chữ nhật. Các đặc điểm hình chữ nhật này được sáng tạo để tạo ra các bộ lọc đặc trưng. Mỗi bộ lọc đặc trưng này được thiết kế để phát hiện các đặc điểm riêng biệt trên khuôn mặt, bao gồm các đường viền, đường thẳng và các đặc điểm khác. Các đặc điểm hình chữ nhật này hoạt động bằng cách đo sự chênh lệch về cường độ pixel giữa vùng màu trắng và màu đen với một bộ lọc được tạo để phát hiện các cạnh trong hình ảnh, bộ lọc được thiết kế để phát hiện các đường thẳng trong hình ảnh và bộ lọc với đặc điểm hình chữ nhật có 4 hình vuông, chúng được sử dụng để xác định sự chênh lệch về cường độ giữa các vùng mắt và miệng trên khuôn mặt.

Mỗi bộ lọc Haar tập trung đến một vùng cụ thể trong hình ảnh để phát hiện, và tạo ra hàng nghìn điểm đặc trưng cho hình ảnh. Tuy nhiên, việc tính toán tổng số pixel trong các vùng trắng và đen trên toàn bộ hình ảnh kết hợp sử dụng một phương pháp "ảnh tích phân," cho

phép tính toán trên bốn điểm ảnh thay vì trên toàn bộ hình ảnh. Hệ thống sử dụng hàng nghìn hình ảnh đã được gắn nhãn để chuyển đổi chúng thành một bản đồ tính năng tầng HAAR và đào tạo một bộ phân loại nhận diện khuôn mặt bằng cách sử dụng mô hình học máy.

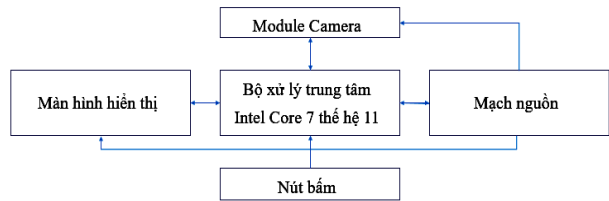


Hình 1. Sơ đồ thuật toán Haar Cascades

Thuật toán Haar Cassades hình 1 hoạt động theo nguyên lý: Khởi tạo hình ảnh đầu vào và các tham số cấu hình. Chia hình ảnh thành các cửa sổ con (sliding windows) với kích thước cố định. Áp dụng bộ lọc Haar cho từng cửa sổ con để tính toán giá trị đặc trưng (feature value) dựa trên sự chênh lệch cường độ pixel giữa các vùng trắng và đen trong cửa sổ con. Duyệt qua các cửa sổ con trên toàn bộ hình ảnh và tính toán giá trị đặc trưng cho mỗi cửa sổ. Sử dụng các bộ phân loại AdaBoost để xác định xem cửa sổ con có chứa khuôn mặt hay không, dựa trên giá trị đặc trưng tính được. Các bộ phân loại này đã được đào tạo trước đó trên hàng nghìn hình ảnh khuôn mặt đã được gắn nhãn. Loại bỏ các cửa sổ con không chứa khuôn mặt dựa trên kết quả từ bước trước. Lặp lại quá trình trên toàn bộ hình ảnh với các cửa sổ con có kích thước khác nhau để

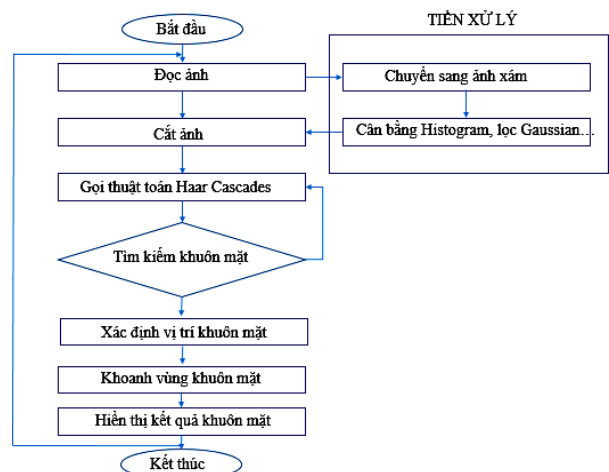
xác định khuôn mặt ở các tỷ lệ khác nhau. Kết hợp các kết quả từ các tỷ lệ khác nhau để đảm bảo rằng khuôn mặt được phát hiện ở mọi kích thước. Trả về vị trí của các khuôn mặt đã được phát hiện trên hình ảnh ban đầu. Kết thúc thuật toán và xuất kết quả.

2.2. Nguyên lý hệ thống



Hình 2. Sơ đồ phần cứng

Phần cứng hệ thống gồm: Camera quét dữ liệu và chụp ảnh theo thời gian thực; vi xử lý sử dụng nền tảng máy tính cấu hình bộ xử lý Intel® Core™ 7 thế hệ thứ 11 với tốc độ lên đến 5 GHz; các thiết bị ngoại vi như nút bấm, màn hình hiển thị, nguồn cấp để hệ thống hoạt động.



Hình 3. Sơ đồ thuật toán nhận diện khoanh vùng khuôn mặt

Hệ thống hoạt động theo nguyên lý: Camera quét dữ liệu theo thời gian thực và lưu lại dữ liệu dạng ảnh. Trong môi trường xử lý ảnh OpenCV, dữ liệu ảnh sau khi chạy qua thuật toán Haar Cascades nhận dạng được vị trí khuôn mặt qua các bước

Bước 1: Hình ảnh được chia thành các phần nhỏ hoặc cửa sổ con.

Bước 2: Nhiều bộ lọc đặc trưng (các bộ dò) được sử dụng để phát hiện các đặc trưng khác nhau trong hình ảnh, và mỗi bộ dò đại diện cho sự kết hợp của các loại đặc trưng khác nhau.

Bước 3: Hình ảnh có độ tin cậy cao nhất sau khi phát hiện khuôn mặt sẽ được xác định và đưa vào bộ tích lũy, trong khi các phần còn lại sẽ bị từ chối. Sau đó, quá trình này được lặp lại cho các khung hình tiếp theo. Đây là cách mà Haar Cascade hoạt động để nhận diện khuôn mặt trong ảnh.

Các vị trí của khuôn mặt sẽ được định vị vẽ đường bao quanh và hiển thị lên màn hình hiển thị.

2.3. Thư viện hỗ trợ

Thư viện openCV (Open source computer vision library) là một thư viện nguồn mở cho máy tính để xử lý hình ảnh, phát triển các ứng dụng đồ họa trong thời gian thực. OpenCV cho phép cải thiện tốc độ của CPU khi thực hiện các hoạt động real time và cung cấp một số lượng lớn các mã xử lý phục vụ cho quy trình của thị giác máy tính hay các learning machine khác.

Thư viện CascadeClassifier các khuôn mặt với nhiều góc nhìn khác nhau, khoảng cách khác nhau, cường độ sáng khác nhau, đặc trưng khuôn mặt khác nhau để đưa vào module huấn luyện với thuật toán Haar Cascades.

Hàm gọi thuật toán Haar Cascades:

```
face_coordinates=trained_face_data.detectMultiScale
```

Trong đó cơ bốn tham số: scaleFactor - Tham số chỉ định kích thước hình ảnh được giảm bao nhiêu ở mỗi tỷ lệ hình ảnh, MinNeighbors

- Tham số chỉ định có bao nhiêu lân cận mỗi hình chữ nhật ứng viên để giữ lại nó. Tham số này sẽ ảnh hưởng đến chất lượng của các khuôn mặt được phát hiện. Giá trị cao hơn dẫn đến ít phát hiện hơn nhưng với chất lượng cao hơn. 3~6 là một giá trị tốt, MinSize - kích thước đối tượng tối thiểu có thể. Các đối tượng nhỏ hơn bị bỏ qua. Tham số này xác định kích thước nhỏ muốn phát hiện. Thông thường, [30, 30] là một khởi đầu tốt để nhận diện khuôn mặt; MaxSize - kích thước đối tượng tối đa có thể. Các đối tượng lớn hơn này bị bỏ qua. Tham số này xác định kích thước lớn như thế nào muốn phát hiện.

3. KẾT QUẢ THẢO LUẬN

Tiến hành thử nghiệm ngẫu nhiên với phần mềm nhận diện khuôn mặt 50 lần với môi trường ánh sáng khác nhau, khoảng cách khác nhau, góc nghiêng nhận diện khác nhau.

Số mẫu	Khoảng cách 0-200mm	Khoảng cách 200-400mm	Khoảng cách 500-1000mm
50	Chính xác 76%	Chính xác 92%	Chính xác 60%

Đánh giá kết quả thử nghiệm với khoảng cách nhận diện khác nhau từ 200 đến 1000 mm cho thấy khoảng cách tối ưu để nhận diện vào khoảng 200 đến 400 mm đạt kết quả nhận diện trên 92%. Với khoảng cách từ 500 đến 1000 m kết quả chỉ đạt 60%. Với khoảng cách từ 0 đến 200 mm kết quả chỉ đạt 76%.

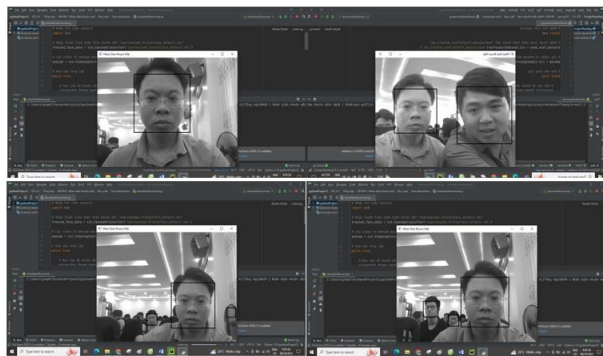
Số mẫu	Ánh sáng tốt	Ánh sáng trung bình	Thiếu sáng
50	Chính xác 96%	Chính xác 70%	Chính xác 66%

Đánh giá kết quả thử nghiệm với cường độ sáng khác nhau kết quả nghiên cứu cho thấy: Với điều kiện ánh sáng tốt, không bị nhiễu sáng cho kết quả chính xác 96%, còn với

trường hợp điều kiện thiếu sáng, hoặc cường độ sáng quá cao kết quả chỉ đạt 66%.

Số mẫu	Nhìn chính diện	Mặt nghiêng 30°	Mặt nghiêng 50°
50	Chính xác 96%	Chính xác 80%	Chính xác 30%

Đánh giá kết quả thử nghiệm với góc nghiêng khuôn mặt nhận diện cho thấy: Với các khuôn mặt nhìn chính diện kết quả phần mềm nhận diện được khuôn mặt đạt 96%, tuy nhiên trường hợp mặt nghiêng quá 30° so với góc chính diện của camera cho kết quả không tốt, chỉ đạt 80%, còn khi mặt nghiêng trên 50° so với góc chính diện của camera kết quả nhận diện dưới 30%.



Hình 4. Kết quả thực nghiệm

Trong bài báo này, nhóm nghiên cứu đưa ra thiết kế bằng mã nguồn mở module phần mềm với sự hỗ trợ của thư viện OpenCV trên nền tảng ngôn ngữ python và sử dụng thuật toán Haar Cascades. Hướng nghiên cứu này giúp xây dựng hệ thống đơn giản hơn, dễ dàng

thêm các tính năng bổ sung mà không làm ảnh hưởng đến các chức năng hiện tại của hệ thống. Cấu hình phần cứng đơn giản dễ tiếp cận và có thể thực nghiệm với số mẫu thử lớn. Trong điều kiện ánh sáng lý tưởng, khoảng cách nhận diện từ 200 đến 400 mm, góc nghiêng khuôn mặt khi nhận diện nhìn thẳng vào camera cho thấy kết quả nhận diện và khoanh vùng khuôn mặt hiển thị đạt 96%. Tốc độ phản hồi của hệ thống 0,5 s đảm bảo ngay tức thì theo thời gian thực.

4. KẾT LUẬN

Bài báo này đã trình bày một nghiên cứu thiết kế hệ thống nhận diện khuôn mặt theo thời gian thực xây dựng trên thuật toán Haar Cascades. Trong điều kiện ánh sáng lý tưởng, khoảng cách nhận diện từ 200 đến 400 mm, góc nghiêng khuôn mặt khi nhận diện nhìn thẳng vào camera cho thấy kết quả nhận diện và khoanh vùng khuôn mặt hiển thị đạt 96%. Tuy nhiên với điều kiện thiếu sáng, góc nhìn nghiêng quá 30° và khoảng cách xa trên 500 m từ camera với khuôn mặt cần nhận diện thì kết quả chỉ đạt 66%. Nghiên cứu này hoàn toàn có thể tiếp tục phát triển tích hợp thêm các thuật toán học sâu CNN, từ cơ sở dữ liệu các khuôn mặt lớp sinh viên, gán nhãn cơ sở dữ liệu các khuôn mặt đối tượng huấn luyện module nhận diện được khuôn mặt đối tượng đã gán nhãn để ứng dụng vào các trường hợp tự động nhận diện và quản lý đối tượng trong thực tế.

TÀI LIỆU THAM KHẢO

- [1] Paul Viola, Michael Jones. *Rapid Object Detection using a Boosted Cascade of Simple Features*. Conference on computer vision and pattern recognition, 2001.
- [2] F. Fleuret and D. Geman. *Coarse-to-fine face detection*. *Int. J. Computer Vision*, 2001.
- [3] Matt Richardson, Shawn Wallace. *Getting start with Raspberry Pi*, Published by O'Reilly Media, United States of America, 2012.
- [4] Medak Teena Ravali, Prof. Ranga Sai Komaragiri. *Image processing platform on RASPBERRY PI for face recognition*, *Global Journal of Advanced Engineering Technologies*, Vol3, Issue4, pg. 441-444, 2014.

- [5] Yanbin Sun, Lun Xie, Zhiliang Wang, and Yi An. *An Embedded System of Face Recognition Based on ARM and HMM*, University of Science and Technology, Beijing, 2007.
- [6] B. Koteswar rao, P. Rama Krishna, MA.Wajeed, *Real Time Embedded Face Recognition using ARM7*", *International Journal of Research in Computer and Communication Technology*, 2013. Kumar, N., Tripathi, H., Gandotra, S. *Optimization of Cylindrical Grinding Process Parameters on C40E Steel Using Taguchi Technique*. *Journal of Engineering Research and Applications*, 100-104, 2015.
- [7] Nguyễn Mạnh Hùng. *Nhận dạng mặt người sử dụng đặc trưng PCA*, Luận văn Thạc sĩ Khoa học máy tính, Học viện Công nghệ Bưu chính Viễn thông, Hà Nội, 2013
- [8] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, "Retinaface: Single-stage dense face localisation in the wild," *ArXiv*, vol. abs/1905.00641, 2019
- [9] Zhang K, Zhang Z, Li Z, et al. "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks" [J]. *IEEE Signal Processing Letters*, 2016, 23(10):1499-1503.
- [10] Xiang J, Zhu G. [IEEE "Joint Face Detection and Facial Expression Recognition with MTCNN[C]// International Conference on Information Science & Control Engin" 4th International Conference on Information Science and Control Engineering (ICISCE) - Changsha (2017.7.21-2017.7.23)] 2017.

Thông tin liên hệ: **Phạm Trung Thiên**

Điện thoại: 0963284444 - Email: ptthien.ck@uneti.edu.vn

Khoa Cơ khí, Trường Đại học Kinh tế - Kỹ thuật Công nghiệp.

-
-