



TRƯỜNG ĐẠI HỌC QUỐC TẾ SÀI GÒN
THE SAIGON INTERNATIONAL UNIVERSITY

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUỐC TẾ SÀI GÒN
KHOA CÔNG NGHỆ THÔNG TIN

ĐỒ ÁN MÔN HỌC
NHẬP MÔN THỊ GIÁC MÁY TÍNH

ĐỀ TÀI: NHẬN DIỆN KHUÔN MẶT, GIỚI TÍNH
VÀ DỰ ĐOÁN ĐỘ TUỔI REAL TIME

Ngành: **KHOA HỌC MÁY TÍNH**

Chuyên ngành: **KHOA HỌC MÁY TÍNH**

KỸ THUẬT PHẦN MỀM

Giảng viên hướng dẫn : ThS. Lê Ngọc Thạch

Sinh viên thực hiện :

- Họ và tên : Phạm Nguyễn Hữu Phương

MSSV : 81011801420

- Họ và tên : Phạm Thị Quỳnh

MSSV : 91011801415

TP. Hồ Chí Minh, năm 2020

ĐỒ ÁN MÔN HỌC
NHẬP MÔN THỊ GIÁC MÁY TÍNH

ĐỀ TÀI: NHẬN DIỆN KHUÔN MẶT, GIỚI TÍNH
VÀ DỰ ĐOÁN ĐỘ TUỔI REAL TIME

Ngành: **KHOA HỌC MÁY TÍNH**

Chuyên ngành: **KHOA HỌC MÁY TÍNH**
KỸ THUẬT PHẦN MỀM

Giảng viên hướng dẫn : ThS. Lê Ngọc Thạch

Sinh viên thực hiện :

- Họ và tên : Phạm Nguyễn Hữu Phương
MSSV : 81011801420
- Họ và tên : Phạm Thị Quỳnh
MSSV : 91011801415

TP. Hồ Chí Minh, năm 2020

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Tp. Hồ Chí Minh, ngày tháng năm 2020

Giảng viên hướng dẫn

LỜI CẢM ƠN

Để hoàn thành đồ án môn học Trí tuệ nhân tạo với đề tài “Nhận diện khuôn mặt, giới tính và dự đoán độ tuổi Real time”, trước tiên cho phép em xin gửi lời cảm ơn tới thầy Lê Ngọc Thạch đã giúp đỡ chúng em rất nhiệt tình trong suốt thời gian qua. Hơn nữa, đồ án của chúng em không thể hoàn thành tốt nếu không có sự hướng dẫn tận tình của quý thầy/cô giảng viên khoa Công nghệ thông tin – Trường Đại học Quốc tế Sài Gòn.

Thời gian thực hiện đồ án tuy ngắn, nhưng nhờ sự hướng dẫn của thầy Lê Ngọc Thạch đã tạo cơ hội cho chúng em áp dụng nền tảng môn học Trí tuệ nhân tạo vào công tác nghiên cứu. Trong suốt thời gian hoàn thành đồ án, chúng em đã có cơ hội rèn luyện được các kỹ năng làm việc và nâng cao hiểu biết của mình trong việc thực hiện viết báo cáo và xây dựng chương trình, từ đó nhận thức rõ hơn về tầm quan trọng của kiến thức cũng như các kỹ năng thực hiện đồ án.

Vì vốn kiến thức và kinh nghiệm còn hạn chế nên bài báo cáo và chương trình thực nghiệm không thể tránh khỏi những hạn chế, thiếu sót. Chúng em rất mong muốn nhận được sự góp ý của các thầy/cô để giúp chúng em hoàn thiện hơn về nghiệp vụ của mình để chúng em có cơ sở, nền tảng kiến thức phục vụ cho công tác sau này với hy vọng những đồ án tiếp theo trong chương trình học được hoàn thành tốt hơn.

Một lần nữa cho phép chúng em xin chân thành cảm ơn thầy Lê Ngọc Thạch và các quý thầy/cô giảng viên khoa Công nghệ thông tin – Trường Đại học Quốc tế Sài Gòn đã tạo điều kiện cho sinh viên chúng em có cơ hội được phát triển ý tưởng sáng tạo, tìm tòi, học hỏi và biết cách áp dụng những kiến thức đã học để xây dựng ứng dụng thực tế, đó chính là kết quả của quá trình truyền đạt kiến thức của quý thầy/cô và sự trau dồi kiến thức của bản thân chúng em.

Chúng em xin cảm ơn quý thầy/cô giảng viên trong Khoa đã giúp đỡ chúng em hoàn thành đồ án và báo cáo này.

MỤC LỤC

LỜI CẢM ƠN	i
MỤC LỤC.....	ii
DANH MỤC HÌNH ẢNH	1
CHƯƠNG 1: TỔNG QUAN	2
1.1. Lý do chọn đề tài.....	2
1.2. Các bước hoàn thành dự án	2
1.3. Phương tiện nghiên cứu.....	3
1.4. Cấu trúc đồ án	3
CHƯƠNG 2: CƠ SỞ LÝ THUYẾT	4
2.1. Khái quát về trí tuệ nhân tạo.....	4
2.2. Khái quát về xử lý ảnh	4
2.3. Phát hiện khuôn mặt bằng OpenCV.....	5
2.2. Nhận diện khuôn mặt bằng thuật toán LBPH.....	8
CHƯƠNG 3: KẾT QUẢ THỰC NGHIỆM	15
3.1. Mô hình hoạt động của chương trình:.....	15
3.2. Giao diện chính:	17
3.3. Giao diện đăng ký người dùng (Register):	18
3.4. Giao diện chương trình nhận diện và dự đoán độ tuổi, giới tính.....	18
CHƯƠNG 4: KẾT LUẬN.....	19
4.1. Kết quả đạt được	19
4.2. Vấn đề còn tồn tại	19
4.3. Đánh giá phần mềm	19

DANH MỤC HÌNH ẢNH

Hình 2.1. Đặc trưng Haar-like.....	7
Hình 2.2. Đặc trưng Haar-like khi được đặt lên hình ảnh	7
Hình 2.3. Ví dụ về LBP và độ tương phản cục bộ C.....	9
Hình 2.4. Tập hợp các điểm xung quanh Ptt.....	9
Hình 2.5. Các biến thể của LBP đồng dạng.....	11
Hình 2.6. Bảng thống kê các mẫu của uniform LBP.....	11
Hình 2.7 Mô hình CaffeNet	13
Hình 3.1. Mô hình hoạt động của chương trình	15
Hình 3.2. Mô hình quá trình nhận diện khuôn mặt	16
Hình 3.3. Mô hình quá trình dự đoán độ tuổi, giới tính	17
Hình 3.4. Giao diện chính của chương trình.....	17
Hình 3.5. Giao diện đăng ký	18
Hình 3.6. Giao diện nhận diện và dự đoán	18

CHƯƠNG 1: TỔNG QUAN

1.1. Lý do chọn đề tài

- Ngày nay, ứng dụng của khoa học công nghệ được áp dụng vào hầu hết các lĩnh vực. Với những ứng dụng này, cuộc sống của con người được cải thiện rất nhiều.
- Đặc biệt, trong những năm gần đây, mảng Trí tuệ nhân tạo của công nghệ thông tin đang được nghiên cứu và phát triển rất mạnh mẽ. Rất nhiều công nghệ ứng dụng trí tuệ nhân tạo ra đời và được áp dụng. Trong đó, không thể không kể đến công nghệ nhận diện nói chung và nhận diện khuôn mặt nói riêng trong bảo mật và an ninh.
- Nhận diện khuôn mặt là một trong những cách thức để giám sát an ninh. Trước đây, khi chưa xuất hiện trí tuệ nhân tạo trong nhận diện, việc quản lý giám sát khá khó khăn và không đem lại hiệu quả cao. Tuy nhiên, từ khi trí tuệ nhân tạo được áp dụng, các hệ thống nhận dạng hiện nay có độ tin cậy ngày càng cao.
- Vì vậy, em quyết định lựa chọn đề tài “Nhận diện khuôn mặt, giới tính và dự đoán độ tuổi real time” với mục đích xây dựng một chương trình có khả năng phát hiện khuôn mặt thông qua camera, từ đó sử dụng các thuật toán xử lý, huấn luyện để nhận diện khuôn mặt, giới tính và dự đoán độ tuổi của người sử dụng.

1.2. Các bước hoàn thành dự án

- Thu thập, nghiên cứu tính thiết thực của đề tài: Khảo sát tình hình thực tiễn, tìm hiểu nhu cầu, phương pháp. Tìm hiểu các công cụ, ngôn ngữ lập trình, các kiến thức cơ bản cần thiết, liên quan đến đề tài.
- Nghiên cứu vấn đề: Đọc hiểu tài liệu. Nghiên cứu, sử dụng để nắm rõ vai trò, chức năng của các công cụ, ngôn ngữ lập trình.
- Thiết kế, xây dựng phần mềm: Dựa vào kiến thức, dữ liệu đã có, tiến hành thiết kế, xây dựng phần mềm theo đúng mục đích, định hướng ban đầu.
- Kiểm thử chương trình: Chạy, kiểm thử chương trình. Xem xét, đánh giá hiệu quả của từng chức năng. Sửa lỗi và hoàn thiện chương trình.
- Kết luận: tổng kết, viết báo cáo, rút ra kết luận chung cho đề tài.

1.3. Phương tiện nghiên cứu

- Nghiên cứu cơ sở lý thuyết về trí tuệ nhân tạo, xử lý ảnh thông qua sách báo, Internet.
- Sử dụng các công cụ: Visual Studio Code để lập trình, thiết kế giao diện; SQLite Studio để quản trị cơ sở dữ liệu cho đề tài.

1.4. Cấu trúc đồ án

- Chương 1: Tổng quan: Tóm tắt những lý thuyết, tài liệu có liên quan đến đề tài.
- Chương 2: Cơ sở lý thuyết : Các khái niệm và phương pháp để giải quyết nhiệm vụ của đồ án.
- Chương 3: Kết quả thực nghiệm: Giao diện và đặc điểm chức năng của chương trình
- Chương 4: Kết luận: những kết quả chung khẳng định những kết quả đã đạt được.

CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

2.1. Khái quát về trí tuệ nhân tạo

- Trí tuệ nhân tạo là trí thông minh của máy do con người tạo ra, hướng đến phát triển hệ thống máy tính (gồm cả phần cứng và phần mềm) sao cho nó có khả năng thông minh như loài người.
- Có thể hiểu rằng, trí tuệ nhân tạo ở đây là nói đến khả năng của máy khi thực hiện các công việc mà con người thường phải xử lý; và khi đáng về ứng xử hoặc kết quả thực hiện của máy là tốt hơn hoặc tương đương với con người thì ta gọi đó là máy thông minh hay máy đó có trí thông minh.
- Trí tuệ nhân tạo khác với việc lập trình logic trong các ngôn ngữ lập trình là ở việc ứng dụng các hệ thống học máy (machine learning) để mô phỏng trí tuệ của con người trong các xử lý mà con người làm tốt hơn máy tính. Có 4 quan điểm về AI như sau:
 - + Suy nghĩ như con người.
 - + Hành động như con người.
 - + Suy nghĩ có lý trí.
 - + Hành động có lý trí.
- Trí tuệ nhân tạo xuất hiện, tạo ra sự thay đổi lớn trong khoa học công nghệ của thế giới. Nhờ việc áp dụng trí tuệ nhân tạo, máy móc có thể thay thế con người làm việc, suy nghĩ, tìm hướng giải quyết vấn đề cơ bản.
- Trí tuệ nhân tạo đang và sẽ là xu hướng công nghệ mà con người theo đuổi trong tương lai. Nó là nền tảng cốt lõi của cuộc cách mạng 4.0.

2.2. Khái quát về xử lý ảnh

2.2.1. Xử lý ảnh là gì?

- Xử lý ảnh là đối tượng nghiên cứu của lĩnh vực thị giác máy, là quá trình biến đổi từ một ảnh ban đầu sang một ảnh mới với các đặc tính và tuân theo ý muốn của người sử dụng. Xử lý ảnh có thể gồm quá trình phân tích, phân lớp các đối tượng, làm tăng chất lượng, phân đoạn và tách cạnh, gán nhãn cho vùng hay quá trình biên dịch các thông tin hình ảnh của ảnh.

2.2.2. Một số vấn đề cơ bản trong xử lý ảnh:

- Một số khái niệm cơ bản:
 - + Ảnh là một tập hợp hữu hạn các điểm ảnh kề nhau. Khi được số hóa, nó thường được biểu diễn bằng một ma trận hai chiều, trong đó, mỗi phần tử tương ứng là một điểm ảnh.
 - + Điểm ảnh là một phần tử của ảnh số tại tọa độ (x, y) với độ xám hoặc màu nhất định.
 - + Mức xám của điểm ảnh là cường độ sáng của nó được gán bằng giá trị số tại điểm đó.
- Biểu diễn ảnh
 - + Ảnh trên máy tính là kết quả thu nhận theo các phương pháp số hoá được nhúng trong các thiết bị kỹ thuật khác nhau. Quá trình lưu trữ ảnh nhằm 2 mục đích:
 - Tiết kiệm bộ nhớ
 - Giảm thời gian xử lý
 - + Việc lưu trữ thông tin trong bộ nhớ có ảnh hưởng rất lớn đến việc hiển thị, in ấn và xử lý ảnh được xem như là một tập hợp các điểm với cùng kích thước nếu sử dụng càng nhiều điểm ảnh thì bức ảnh càng đẹp, càng mịn và càng thể hiện rõ hơn chi tiết của ảnh người ta gọi đặc điểm này là độ phân giải.

2.2.3. Mục đích của xử lý ảnh:

- Biến đổi ảnh làm tăng chất lượng ảnh.
- Tự động nhận dạng ảnh, đoán nhận ảnh, đánh giá các nội dung của ảnh.

2.3. Phát hiện khuôn mặt bằng OpenCV

2.3.1. Thư viện mã nguồn mở OpenCV

2.3.1.1. OpenCv là gì?

- OpenCV, viết tắt của Open Source Computer Vision được xem là một trong những thư viện mã nguồn mở hàng đầu cho việc xử lý ảnh theo thời gian thực (real time).
- OpenCV chính thức được ra mắt đầu tiên năm 1999. Nó được mở miễn phí cho cả học thuật và thương mại.
- OpenCV hỗ trợ đa nền tảng gồm Windows, Linux, Mac OS, IOS và Android.
- OpenCv được viết bằng C/C++ và tích hợp OpenCL, nó hỗ trợ ngôn ngữ lập trình C/C++, Python, C# và Java.

- Hiện nay, OpenCv có hơn 2500 thuật toán được tối ưu hóa.

2.1.1.1. Chức năng có trong OpenCv

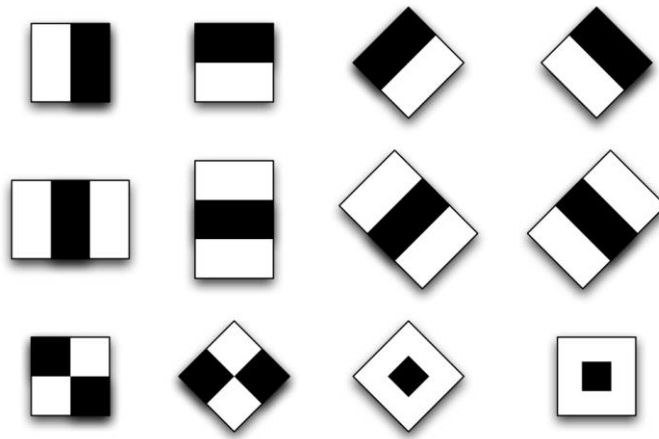
- Các chức năng nổi bật:
 - + Bộ công cụ hỗ trợ 2D và 3D.
 - + Nhận diện khuôn mặt, cử chỉ, chuyển động, đối tượng.
 - + Tương tác giữa con người và máy tính.
 - + Điều khiển robot.
 - + Hỗ trợ thực tế tăng cường.

2.1.1.2. Ứng dụng của OpenCV

- OpenCV hiện có hơn 47000 người dùng và ước tính hơn 14 triệu số lượt tải xuống, thư viện OpenCV được sử dụng rộng rãi trong nhiều lĩnh vực.
 - + Giám sát tự động.
 - + Tìm kiếm, phục hồi, xử lý ảnh.
 - + Nhận dạng khuôn mặt, cử chỉ.
 - + Nhận dạng chữ viết, con số, ký tự.

2.1.2. *Phát hiện khuôn mặt*

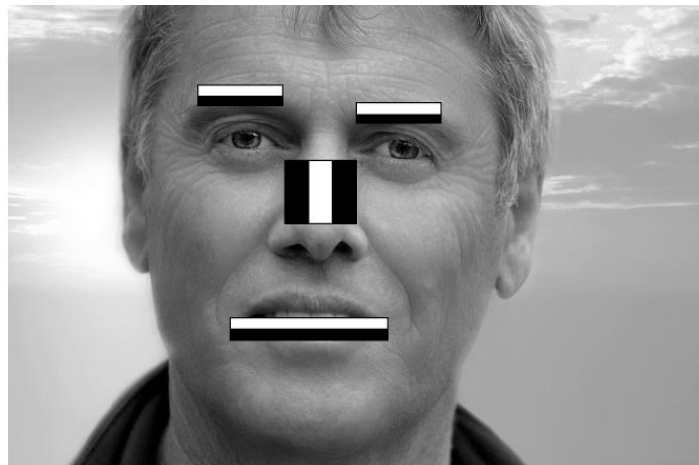
- Phát hiện khuôn mặt là bước đầu tiên của bài toán nhận diện khuôn mặt.
- Với ảnh xám, vùng ảnh khuôn mặt là tập hợp của những điểm ảnh có mối quan hệ khác biệt so với các vùng ảnh khác. Những mối quan hệ này tạo nên các đặc trưng riêng của khuôn mặt. tất cả khuôn mặt của con người đều có chung các đặc điểm như:
 - + Vùng hai mắt sẽ tối hơn (mức xám cao vượt trội hơn) vùng má và vùng trán.
 - + Vùng giữa sống mũi tối hơn vùng hai bên mũi.
 - + ...
- Dựa vào các đặc điểm đó, phương pháp nhận dạng dựa vào đặc trưng Haar-like để phát hiện khuôn mặt xuất hiện.
- Các đặc trưng haar-like (Haar-Like feature) là những hình chữ nhật được phân thành các vùng khác nhau.



Hình 2.1. Đặc trưng Haar-like

- Các đặc trưng trung của Haar-like được tính dựa trên độ chênh lệch giữ tổng các pixel của vùng đen và vùng trắng, công thức tính:

$$f(x) = \text{Tổng vùng đen (các mức xám của pixel)} - \text{Tổng vùng trắng (các mức xám của pixel)}$$
- Để phát hiện khuôn mặt, một cửa sổ con có kích thước cố định sẽ quét lên toàn bộ ảnh đầu vào (ảnh đã được chuyển về ảnh xám). Ứng với từng cửa sổ con quét qua, các ảnh con được đặc trưng Haar-like đặt lên; từ đó tính ra giá trị đặc trưng. Thông qua bộ phân loại, mà hệ thống sẽ xác nhận khung hình đó có phải khuôn mặt hay không.



Hình 2.2. Đặc trưng Haar-like khi được đặt lên hình ảnh

- Trong thư viện OpenCV đã cung cấp sẵn các trình phân loại được huấn luyện sẵn để phát hiện các đối tượng. Để phát hiện khuôn mặt sử dụng trình phân loại được lưu trữ trong tệp : haarcascade_frontalface_default.xml.

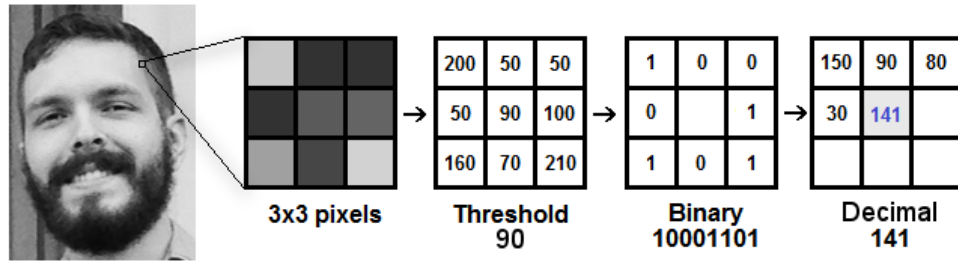
2.2. Nhận diện khuôn mặt bằng thuật toán LBPH

2.2.1. Tạo cơ sở dữ liệu

- Chụp ảnh người đó để nhận diện khuôn mặt. Nó tự động tạo thư mục huấn luyện trong thư mục cơ sở dữ liệu chứa khuôn mặt được nhận dạng.
- Trong khi tạo cơ sở dữ liệu, hình ảnh khuôn mặt phải có các biểu thức khác nhau, đó là lý do tại sao độ trễ 0,38 giây được đưa ra trong mã để tạo tập dữ liệu. Có khoảng 22 tấm hình/khuôn mặt và trích xuất khuôn mặt, chuyển đổi nó thành thang độ xám và lưu nó vào thư mục cơ sở dữ liệu với tên của nó.

2.2.2. Thuật toán LBPH

- LBPH là viết tắt của Local Binary Patterns Histograms.
- Local Binary Pattern hay là mẫu nhị phân địa phương được Ojala trình bày vào năm 1996 như là một cách đo độ tương phản cục bộ của ảnh. Phiên bản đầu tiên của LBP được dùng với 8 điểm ảnh xung quanh và sử dụng giá trị của điểm ảnh ở trung tâm làm ngưỡng. Giá trị LBP được xác định bằng cách nhân các giá trị ngưỡng với trọng số ứng với mỗi điểm ảnh sau đó cộng tổng lại. Hình dưới minh họa cách tính độ tương phản trực giao (C) là hiệu cấp độ xám trung bình của các điểm ảnh lớn hơn hoặc bằng ngưỡng với các điểm ảnh thấp hơn ngưỡng.
- Kể từ khi được đưa ra, theo định nghĩa là bất biến với những thay đổi đơn điệu trong ảnh đen trắng. Để cải tiến phương pháp, bổ sung thêm phương pháp tương phản trực giao địa phương. Hình dưới minh họa cách tính độ tương phản trực giao (C) là ký hiệu cấp độ xám trung bình của các điểm ảnh lớn hơn hoặc bằng ngưỡng với các điểm ảnh thấp hơn ngưỡng. Phân phối hai chiều của mã LBP và độ tương phản cục bộ được lấy làm đặc trưng gọi là LBP/C.

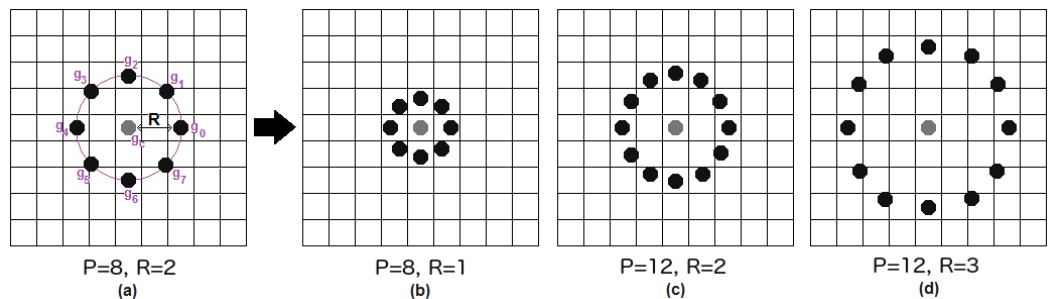


Hình 2.3. Ví dụ về LBP và độ tương phản cục bộ C.

- Dãy LBP được Ojala trình bày vào năm 2002. Định nghĩa một cấu trúc điểm ảnh T là một phân phối đại số của cấp độ xám của $P + 1$ ($P > 0$) điểm ảnh.

$$T = t(g_c, g_0, \dots, g_{P-1})$$

- Với g_c ứng với cấp độ xám của điểm ảnh trung tâm P_{tt} , g_p ($p = 0, \dots, P-1$) tương ứng với P điểm ảnh xung quanh, P điểm ảnh này nằm trên đường tròn bán kính R và tâm là P_{tt} .



Hình 2.4. Tập hợp các điểm xung quanh P_{tt} .

- Không mất thông tin, có thể trừ g_c đi một lượng là g_c

$$T = t(g_c, g_0 - g_c, \dots, g_{P-1} - g_c)$$

- Giả sử sự sai số giữa g_p và g_c là độc lập với g_c , ta có thể nhân tử hóa g_c như sau:

$$T = t(g_c) t(g_0 - g_c, \dots, g_{P-1} - g_c)$$

- $t(g_c)$ biểu thị xu hướng độ sáng tối của cả bức ảnh nên không liên quan đến kết cấu của ảnh cục bộ do đó có thể bỏ qua

$$T \sim t((g_0 - g_c), \dots, (g_{P-1} - g_c))$$

- Mặc dù tính bất biến ngược với độ thay đổi tỷ lệ xám của điểm ảnh, sự khác biệt ảnh hưởng bởi tỷ lệ. Để thu được đặc điểm bất biến với bất kỳ một sự thay đổi nào của ảnh đen trắng (gray scale) chỉ quan tâm đến dấu của độ lệch:

$$T \sim t(s(g_0 - g_c), \dots, s(g_{p-1} - g_c))$$

- Với s là hàm dấu:
$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$
- Trọng số 2^p được dùng cho các hàm dấu, $s(g_p - g_c)$ để chuyển sự khác biệt giữa các điểm ảnh bên cạnh về một giá trị duy nhất.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) * 2^p$$

- Với P pixel thì có 2^P giá trị $LBP_{P,R}$ trong khoảng $[0, 2^P - 1]$ nhưng để đơn giản ta có thể chọn một số giá trị trong 2^P giá trị ký hiệu là $LBP_{P,R}^u$

- **Thuật toán LBP:**

- + Thông tin LBP của pixel tại trung tâm của mỗi khối ảnh sẽ được tính dựa trên thông tin của các pixel lân cận. Có thể tóm tắt các bước tiến hành như sau:
- + Bước 1: Xác định bán kính làm việc.
- + Bước 2: Tính giá trị LBP cho pixel ở trung tâm (x_c, y_c) khối ảnh dựa trên thông tin của các pixel lân cận:
- + Trong đó, (g_p) là giá trị grayscale của các pixel lân cận, (g_c) là giá trị grayscale của các trung tâm và (s) là hàm nhị phân được xác định như sau: $s(z) = 1$ nếu giá trị $z \geq 0$.

- Ví dụ:

47	51	65
62	70	70
80	83	78

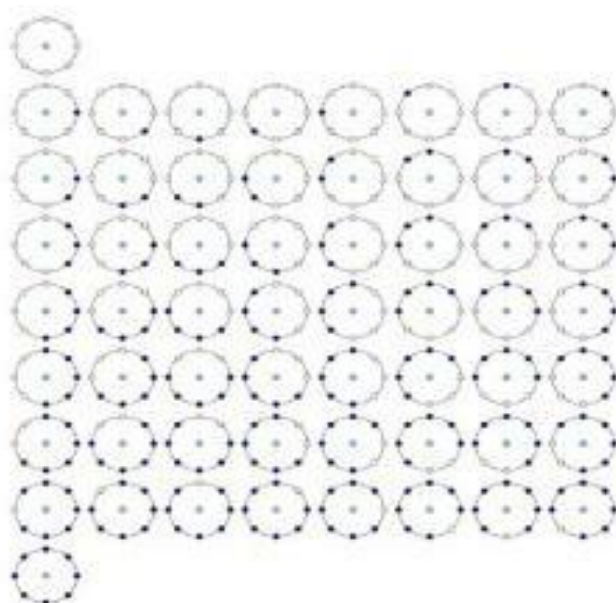
-23	-19	-5
-8	*	0
10	13	8

0	0	0
0	*	1
1	1	1

$$1*2^0 + 1*2^1 + 1*2^2 + 1*2^3 + 0*2^4 + 0*2^5 + 0*2^6 + 0*2^7 = 15.$$

Hình 2.5. Các biến thể của LBP đồng dạng.

- Một mẫu nhị phân được gọi là đồng dạng khi xét chuỗi bit xoay vòng thì có nhiều nhất là 2 lần thay đổi (transitions) từ giá trị bit 0 sang 1 hoặc từ giá trị bit 1 sang 0. Ví dụ: 00000000 có 0 transitions, 01110000 có 2 transitions, 11001111 có 2 transitions nên đây là uniform LBP. 11001001 có 4 transitions, 01010011 có 6 transitions nên không phải là uniform LBP.
- Dựa trên định nghĩa này, bảng ánh xạ cho bán kính làm việc P - neighbours sẽ có $P(P-1) + 3$ nhãn. Có nghĩa là có 59 nhãn trong trường hợp làm việc với 8-neighbour. Hình vẽ sau đây thể hiện 59 nhãn (mẫu) và minh họa về histogram của đặc trưng LBP đồng dạng.



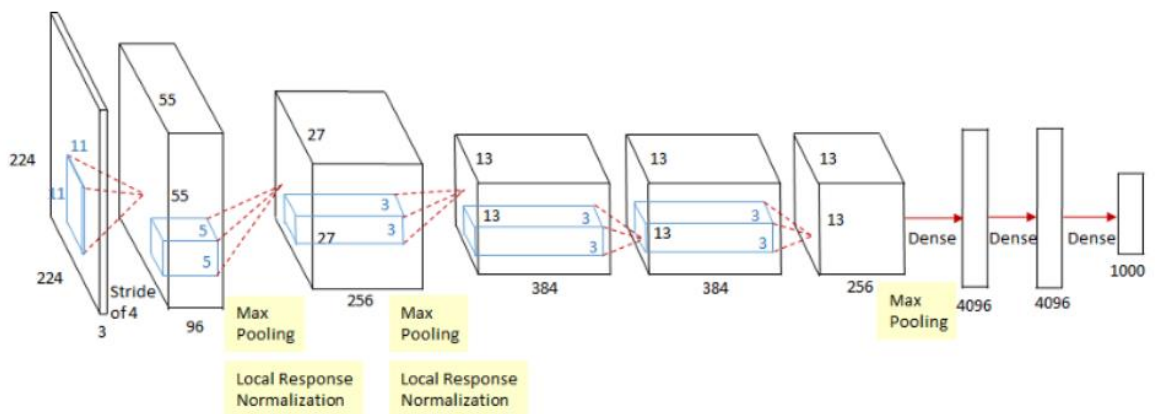
Hình 2.6. Bảng thống kê các mẫu của uniform LBP

- Sau tác vụ thu thập hình ảnh và xử lý trước, thực hiện đào tạo tập dữ liệu. Đối với giai đoạn đào tạo, đào tạo trình nhận dạng được áp dụng để lưu trữ các giá trị biểu đồ của khuôn mặt

2.2.3. Dự đoán giới tính và độ tuổi với CNN

- Để xác định giới tính và độ tuổi, trong đồ án này chúng em sử dụng caffeNet - là một biến thể của AlexNet – một mạng thần kinh tích chập(Convolutional Neural Network – CNN).
- Convolutional Neural Network là một trong những mô hình Deep Learning tiên tiến. Nó giúp cho trong việc xây dựng những hệ thống thông minh với độ chính xác cao như hiện nay. Một số ví dụ như hệ thống xử lý ảnh lớn như Facebook, Google hay Amazon đã đưa vào sản phẩm của mình những chức năng thông minh như nhận diện khuôn mặt người dùng, phát triển xe hơi tự lái hay drone giao hàng tự động.
- Cấu trúc của mạng CNN: Mạng CNN là một tập hợp các lớp Convolution chồng lên nhau và sử dụng các hàm nonlinear activation như ReLU và tanh để kích hoạt các trọng số trong các node. Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo.
- Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo. Trong mô hình mạng truyền ngược (feedforward neural network) thì mỗi neural đầu vào (input node) cho mỗi neural đầu ra trong các lớp tiếp theo.
- Mạng CNN sử dụng 3 ý tưởng cơ bản:
 - + Các trường tiếp nhận cục bộ (local receptive field)
 - + Trọng số chia sẻ (shared weights)
 - + Tổng hợp (pooling).
- Local receptive field: Đầu vào của mạng CNN là một ảnh. Ví dụ như ảnh có kích thước 28×28 thì tương ứng đầu vào là một ma trận có 28×28 và giá trị mỗi điểm ảnh là một ô trong ma trận. Trong mô hình mạng ANN(Artificial Neural network) truyền thống thì ta sẽ kết nối các neuron đầu vào vào tầng ảnh.

- Tuy nhiên trong CNN không làm như vậy mà chỉ kết nối trong một vùng nhỏ của các neuron đầu vào như một filter có kích thước 5×5 tương ứng $(28 - 5 + 1) = 24$ điểm ảnh đầu vào. Mỗi một kết nối sẽ học một trọng số và mỗi neuron ẩn sẽ học một bias. Mỗi một vùng 5×5 đây gọi là một trường tiếp nhận cục bộ.
- Sử dụng gói DNN trong mô hình CNN được nghiên cứu bởi Gil Levi và Tal Hassner để thực hiện dự đoán giới tính và độ tuổi.
- OpenCV cung cấp một lớp có tên Net để tạo ra một mạng lưới thần kinh. Hai nhà khoa học đã tạo ra mô hình CNN dưới dạng mô hình Caffe.
- Mô hình này có hai tệp liên quan:
 - + prototxt: CNN đi vào từ tệp này. Nó xác định các lớp trong mạng thần kinh và các đầu vào, đầu ra và chức năng của mỗi lớp.
 - + .caffemodel: Chứa thông tin của mạng lưới thần kinh được huấn luyện.



Hình 2.7 Mô hình CaffeNet

2.2.3.1. Dự đoán giới tính

- Dự đoán Giới tính như một bài toán phân loại. Lớp đầu ra trong mạng dự đoán giới tính là hai lớp “Male” và “Female”.
- Thông qua hai tệp ‘gender_deploy.prototxt’ và ‘gender_net.caffemodel’.

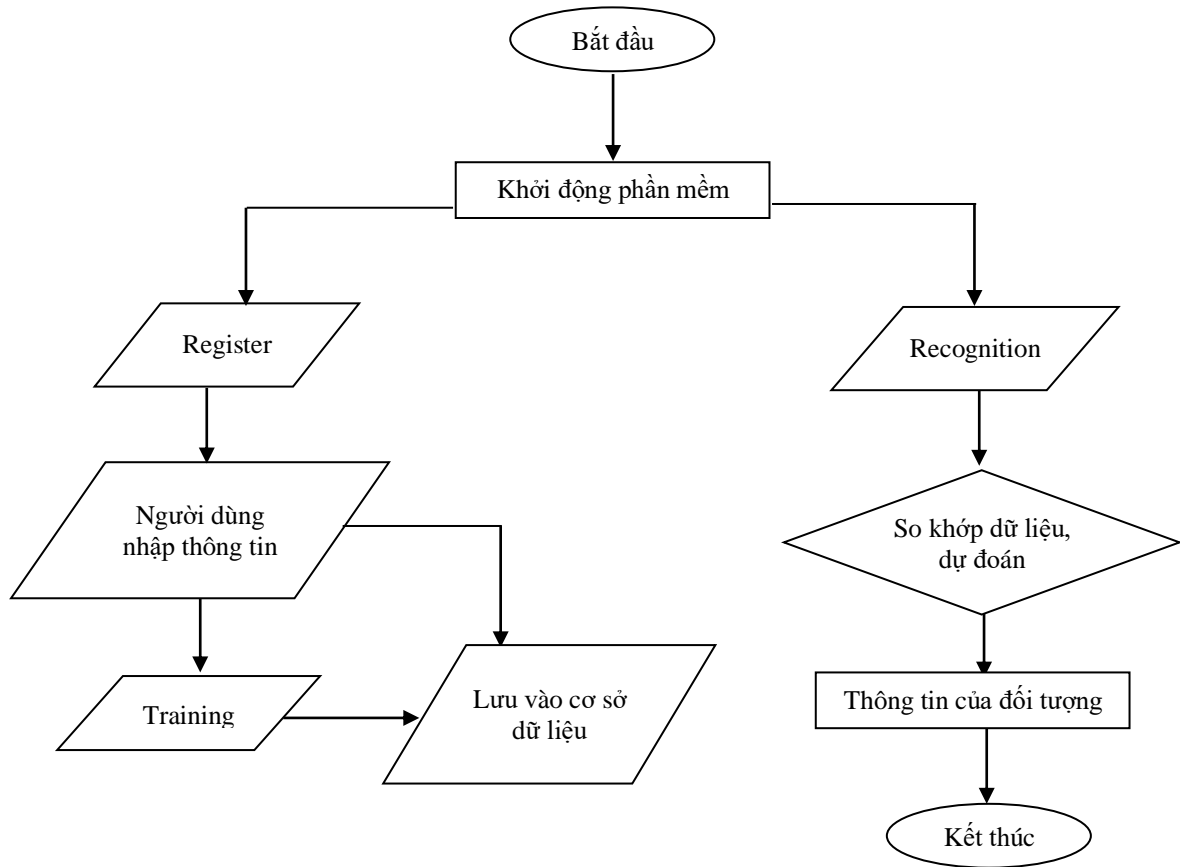
2.2.3.2. Dự đoán độ tuổi

- Thông qua hai tệp ‘age_deploy.prototxt’ và ‘age_net.caffemodel’.

- Tập dữ liệu Đối tượng có 8 lớp được chia thành các nhóm tuổi sau [(0 - 2), (4 - 6), (8 - 12), (15 - 20), (25 - 32), (38 - 43), (48 - 53), (60 - 100)].
- Dự đoán tuổi phụ thuộc vào nhiều yếu tố, những người cùng một độ tuổi nhưng có thể có đặc điểm khác nhau. Do đó, không thể dự đoán tuổi một cách hoàn toàn chính xác.

CHƯƠNG 3: KẾT QUẢ THỰC NGHIỆM

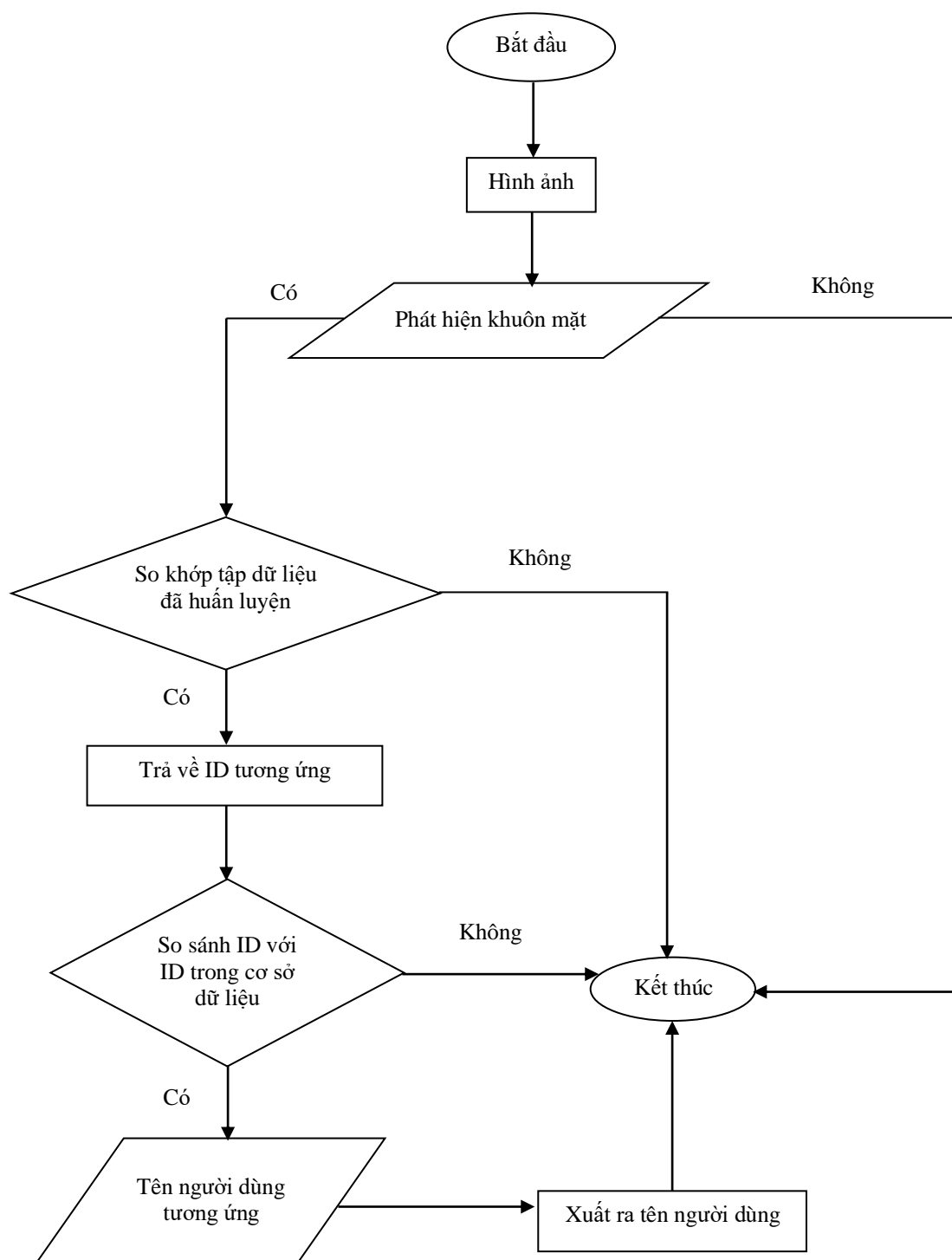
3.1. Mô hình hoạt động của chương trình:



Hình 3.1. Mô hình hoạt động của chương trình

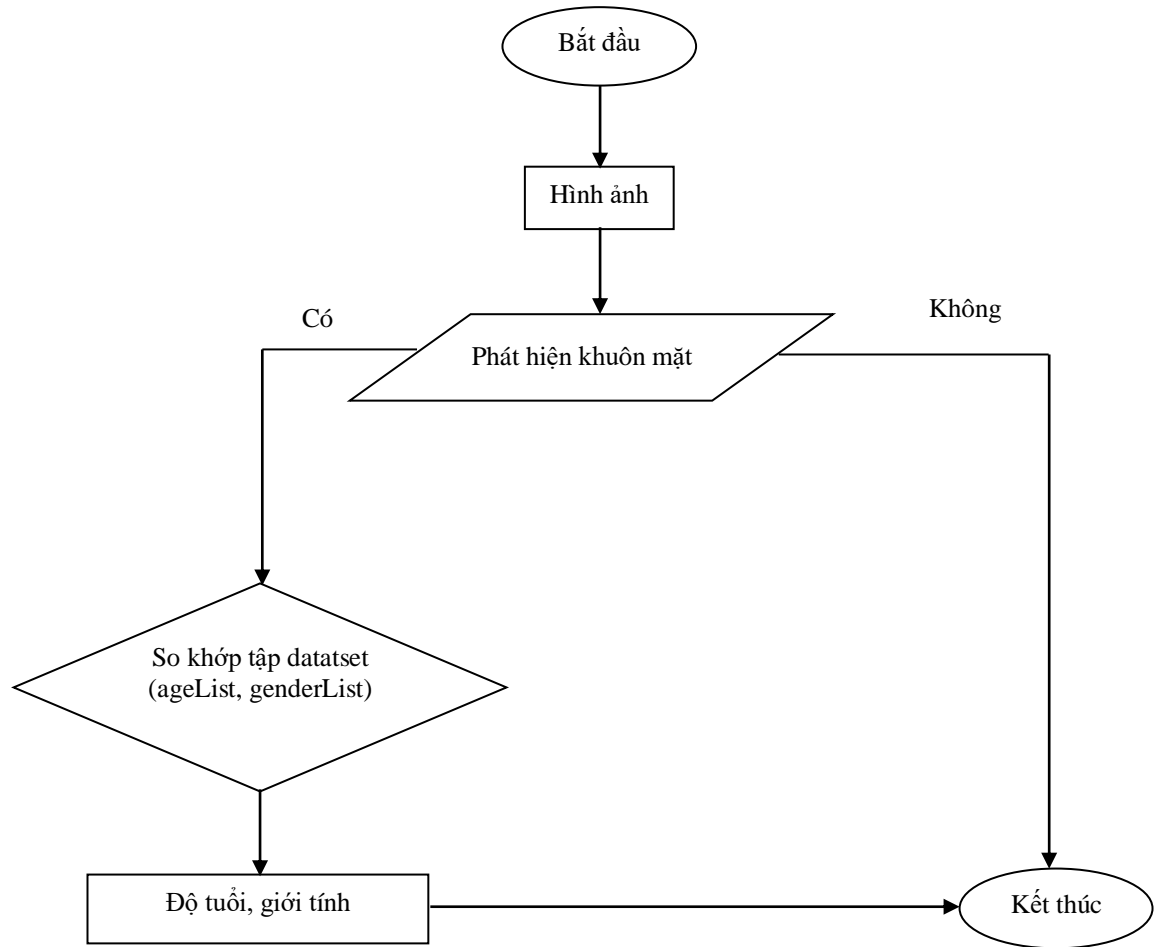
- Khi chương trình được thực thi, giao diện phần mềm hiển thị cho phép người dùng lựa chọn chức năng: Register (đăng ký) và Recognition (nhận diện).
- Với chức năng Register, người dùng sẽ nhập thông tin bao gồm ID và Name. Hai thông tin này sẽ được lưu vào cơ sở dữ liệu. Hình ảnh khuôn mặt được phát hiện sẽ được lưu vào tập huấn luyện của chương trình (training và facebase).
- Chức năng Recognition cho phép nhận diện được khuôn mặt là của người dùng nào trong cơ sở dữ liệu dựa vào dữ liệu đã huấn luyện. Đồng thời dự đoán được độ tuổi và giới tính của người đó thông qua bộ dataset sẵn có.

3.1.1. Mô hình quá trình nhận diện khuôn mặt



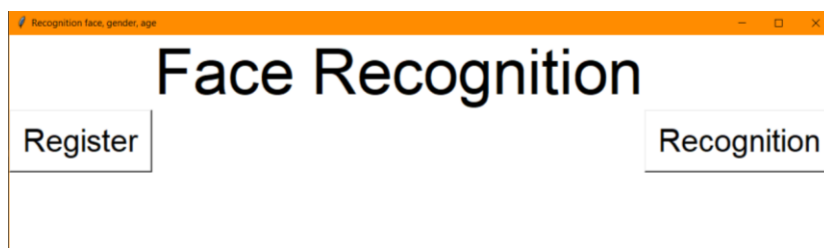
Hình 3.2. Mô hình quá trình nhận diện khuôn mặt

3.1.2. Mô hình quá trình dự đoán độ tuổi, giới tính



Hình 3.3. Mô hình quá trình dự đoán độ tuổi, giới tính

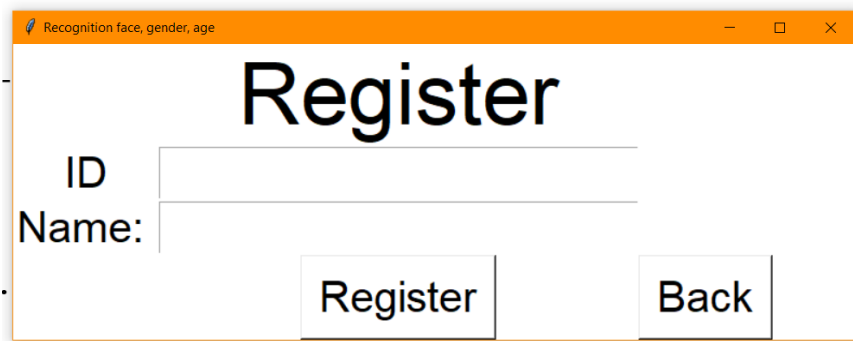
3.2. Giao diện chính:



Hình 3.4. Giao diện chính của chương trình

- Giao diện chính của phương trình được thiết kế rất đơn giản gồm 2 button Register (cho phép người dùng đăng ký) và button Recognition (để nhận diện khuôn mặt).

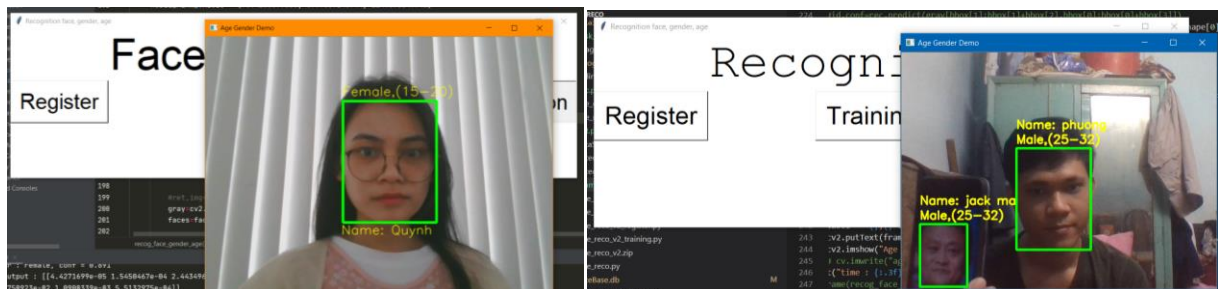
3.3. Giao diện đăng ký người dùng (Register):



Hình 3.5. Giao diện đăng ký

- Để đăng ký, người dùng nhập vào hai thông tin là ID và Name.
- Sau khi nhập đầy đủ thông tin, nhấn button Register để đăng ký. Chương trình sẽ mở camera của thiết bị để xác định vùng có khuôn mặt và chụp lại 61 hình ảnh khuôn mặt để đưa vào tập dữ liệu huấn luyện và thực hiện huấn luyện.
- Nhấn button Back nếu muốn quay lại giao diện chương trình chính.

3.4. Giao diện chương trình nhận diện và dự đoán độ tuổi, giới tính



Hình 3.6. Giao diện nhận diện và dự đoán

- Để thực hiện nhận diện khuôn mặt, giới tính và độ tuổi, người dùng nhấn button Recognition. Một cửa sổ chứa hình ảnh được camera ghi lại xuất hiện. Nếu trong đó chứa hình ảnh khuôn mặt, chương trình thực hiện xác định khuôn mặt đồng thời xác định tên tương ứng và dự đoán độ tuổi, giới tính.

CHƯƠNG 4: KẾT LUẬN

4.1. Kết quả đạt được

- Phát hiện được vùng khuôn mặt có trong hình ảnh.
- Thực hiện nhận diện đúng được khuôn mặt.
- Thực hiện dự đoán độ tuổi, giới tính.

4.2. Vấn đề còn tồn tại

- Chưa phát hiện được khuôn mặt với các góc độ khác nhau, chỉ dừng lại ở phát hiện khuôn mặt ở góc trực diện.
- Dự đoán độ tuổi, giới tính còn chưa chính xác.

4.3. Đánh giá phần mềm

4.2.1. Ưu điểm

- Chương trình dễ sử dụng, thuận tiện cho người sử dụng.

4.2.2. Nhược điểm

- Giao diện chưa được sinh động
- Cơ sở dữ liệu chưa được chặt chẽ, đầy đủ.
- Tính chuyên nghiệp của chương trình chưa cao, có thể xuất hiện một vài lỗi nhỏ, một số tính năng chưa tối ưu.
- Độ chính xác còn thấp.
- Thời gian thực thi chương trình chưa nhanh.

4.2.3. Hướng phát triển

- Thiết kế chương trình và cải tiến giao diện mang tính chuyên nghiệp hơn.
- Khắc phục một số lỗi còn tồn tại.
- Xây dựng chương trình cho phép thực thi trên nhiều hệ điều hành.
- Cải tiến tốc độ chương trình.

TÀI LIỆU THAM KHẢO

- [1] <https://stackoverflow.com>
- [2] <https://tutorialspoint.com>
- [3] <https://youtube.com>