**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

A blue logo with a black background

Description automatically generated with low confidence

ĐỒ ÁN MÔN HỌC

ĐỀ TÀI

**AGE AND GENDER RECOGNITION**

Môn học: **Nhận dạng – CS338**

Lớp: **CS338.N21**

Giảng viên hướng dẫn: **Đỗ Văn Tiến**

Sinh viên thực hiện:

|  |  |  |
| --- | --- | --- |
| **STT** | **Họ Tên** | **MSSV** |
| **1** | **Bùi khánh Duy** | **19521418** |
| **2** | **Nguyễn Đình Sang** | **19522120** |

*Thành phố Hồ Chí Minh, tháng 6 năm 2023*

Contents

[**1.** **Tổng quan** 1](#_Toc139968636)

[**2.** **Phương pháp** 1](#_Toc139968637)

[**3.** **Quy trình thực hiện** 1](#_Toc139968638)

[**3.1.** **Bộ dữ liệu (dataset)** 1](#_Toc139968639)

[**3.2.** **Xây dựng mô hình CNN** 3](#_Toc139968640)

[**3.3.** **Huấn luyện và kiểm thử mô hình** 5](#_Toc139968641)

[**4.** **Kết luận** 7](#_Toc139968642)

# **Tổng quan**

Ước tính tuổi và giới tính từ một hình ảnh khuôn mặt duy nhất là một nhiệm vụ thiết yếu trong lĩnh vực tương tác giữa người và máy tính và thị giác máy tính, điều đó mang nhiều giá trị ứng dụng thực tế.

Nhiều thuộc tính khuôn mặt từ hình ảnh khuôn mặt một người được biết đến là rất hữu ích cho việc mô tả đặc điểm của cá nhân đó. Trong đó tuổi tác và giới tính là một trong những đặc tính quan trọng được coi là một thuộc tính vốn có và là một yếu tố sinh học quan trọng đặc trưng, đóng một vai trò cơ bản trong sự tương tác xã hội giữa người với người. Vì thế ước tính tuổi và giới tính tự động từ hình ảnh khuôn mặt là một nội dung nghiên cứu quan trọng trong lĩnh vực trí tuệ nhân tạo.

1. **Phương pháp**

Trước đây, các công trình nghiên cứu dự đoán tuổi và giới tính thường liên quan đến các kĩ thuật dựa trên các phép đo khác nhau lên các đặc điểm trên khuôn mặt như mắt, mũi, khoảng cách từ cằm đến trán,... Những phương pháp trên được gọi là các phương pháp nhân trắc học.

Các phương pháp ban đầu dựa trên việc trích xuất thủ công các đặc trưng (manual extraction of features) như PCA, LBP, LDA. Sau đó các đặc trưng được trích xuất này được đưa vào các mô hình máy học cổ điển như SVM, decision tree, logistic regression. Nhược điểm rõ ràng của phương pháp này là khó có thể đo lường được các chỉ số nhân trắc học và khó có thể khái quát hóa vì những người ở những độ tuổi khác nhau vẫn có thể có các số đo nhân trắc học giống nhau.

Gần đây, việc sử dụng mạng thần kinh tích chập (convolutional neural network – CNN) để dự đoán tuổi và giới tính được áp dụng rộng rãi vì CNN khá mạnh mẽ và cho ra kết quả nổi bật khi thử nghiệm trên hình ảnh khuôn mặt với độ che khuất, độ nghiêng và ánh sáng thay đổi. Từ đó cho thấy sự mạnh mẽ trong việc trích xuất đặc trưng (feature extraction) của CNN.

# **Quy trình thực hiện**

* 1. **Bộ dữ liệu (dataset)**

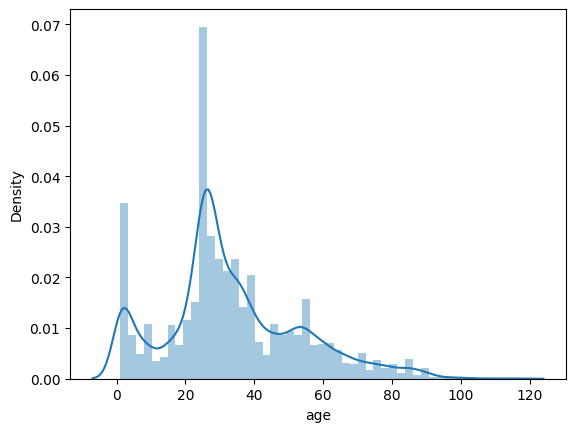
Trong lần thực hiện này, nhóm sử dụng UTKFace dataset với hơn 20000 tấm ảnh khuôn mặt người ở nhiều độ tuổi (khoảng từ 0 đến 116 tuổi), giới tính và sắc tộc. Dataset này có 23708 ảnh, các hình ảnh có nhiều sự khác biệt về biểu cảm khuôn mặt, độ sáng, góc nghiêng, độ phân giải và độ che phủ.

Các hình ảnh trong dataset phần đầu được đặt tên theo cú pháp <tuổi>\_<giới tính: 0-nam|1-nữ>\_<chủng tộc: 0-da trắng|1-da đen|2-châu Á|3-Ấn Độ|4-khác>\_etc, khi đó chúng ta sẽ dễ dàng phân loại trong quá trình sử dụng.

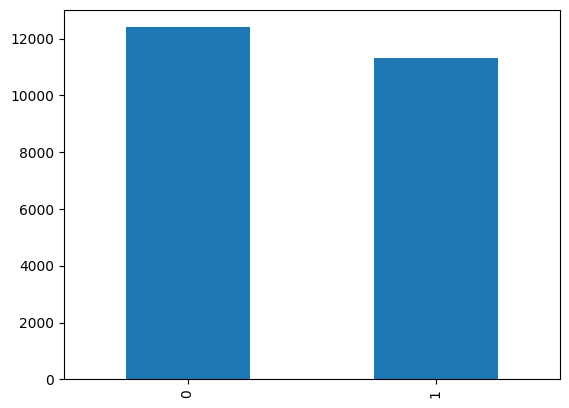
A collage of a group of people's faces

Description automatically generated

*Một số ảnh có trong bộ dữ liệu*



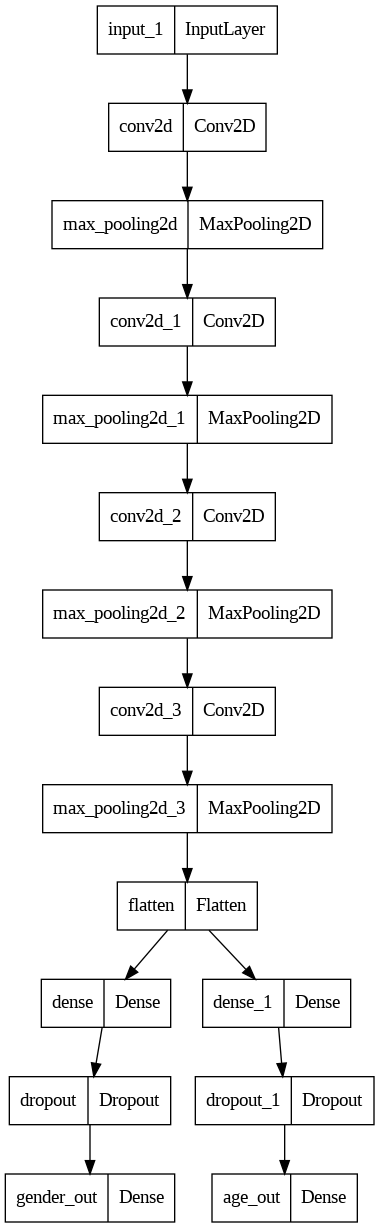
*Phân bố độ tuổi có trong dataset*



*Phân bố giới tính có trong dataset*

* 1. **Xây dựng mô hình CNN**

Mô hình bao gồm một loạt các khối tích chập, theo sau là các lớp FC (fully connected- kết nối đầy đủ) để phân loại và hồi quy. Một bức ảnh RBG được đưa vào mô hình với kích thước là 128 x 128 x 3. Mọi khối tích chập đều là một chồng các lớp tích chập (convolutional layer) với kích thước bộ lọc (filter) là 3 x 3 theo sau là kích hoạt phi tuyến tính (RELU), gộp tối đa (Max pooling) là 2 x 2.



*Cấu trúc của CNN*

A screenshot of a computer program

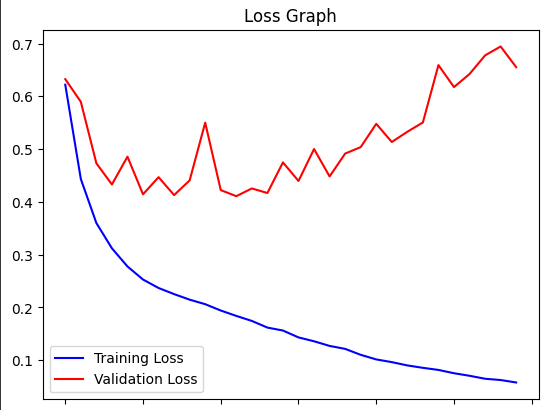
Description automatically generated

A black background with white text

Description automatically generated

## **Huấn luyện và kiểm thử mô hình**

Vì kết nối mạng không ổn định nên nhóm chỉ huấn luyện mô hình với 30epochs, ta thấy loss của mô hình vẫn còn khá cao sau khi huấn luyện, khoảng 4.07.



*Đồ thị loss của phần dự đoán giới tính*

*A graph with red lines and blue lines

Description automatically generated*

*Đồ thị loss của phần dự đoán tuổi*

Sau một vài thử nghiệm, thì ta thấy được mô hình có tỉ lệ dự đoán chính xác giới tính khá cao, tuy nhiên độ tuổi vẫn còn sai lệch khá nhiều.

A close-up of a person's face

Description automatically generated

Mô hình cần được xem xét lại về cấu trúc mạng CNN và số lượng epoch cần để huấn luyện để tối có kết quả chính xác hơn, cũng như thử nghiệm trên nhiều bộ dữ liệu khác nữa.

# **Kết luận**

Tuy nhiên, việc nhận diện tuổi và giới tính là một vấn đề rắc rối vốn có từ lúc được đặt ra, và hơn một số vấn đề khác trong thị giác máy tính. Lý do cơ bản của sự rắc rối này nằm ở số lượng thông tin cần thiết để chuẩn bị cho các mô hình dự đoán. Mặc dù đã sẽ có rất nhiều hình ảnh đã được chuẩn bị nhưng số lượng được thống kê tên, tuổi và giới tính là vô cùng hạn chế.

Mô hình đã không đạt được tỷ lệ chính xác quá cao vì thế cần phải cải thiện thuật toán trong mô hình.

1. **Tài liệu tham khảo**

[Age and Gender Prediction using Deep CNNs and Transfer Learning](https://arxiv.org/pdf/2110.12633)*, Vikas Sheoran, Shreyansh Joshi and Tanisha R. Bhayani*

[Detection of Gender and Age using Machine Learning](https://www.ijraset.com/research-paper/detection-of-gender-and-age-using-ml), Mr. Aditya Kulkarni, Mr. Parth Joshi, Mr. Shaunak Sindgi, Mr. Shreyas Rakshasbhuvankar, Mr. Vivek Kumar, Prof. Madhavi Dachawar

[Building CNN model, Layer Patterns and Rules!](https://viblo.asia/p/building-cnn-model-layer-patterns-and-rules-Do754qXQKM6)

[Convolutional Neural Network (CNN)](https://www.tensorflow.org/tutorials/images/cnn), *TensorFlow*

[Whole model saving & loading](https://keras.io/api/saving/model_saving_and_loading/), *Keras*