**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**KHOA KHOA HỌC MÁY TÍNH**

Logo, company name

Description automatically generated

**BÁO CÁO ĐỒ ÁN MÔN HỌC**

**THỊ GIÁC MÁY TÍNH NÂNG CAO**

**CS331.N12.KHCL**

|  |  |
| --- | --- |
| **SINH VIÊN THỰC HIỆN:** | **PHẠM TRẦN ANH TIÊN - 20522012**  **LƯƠNG LÝ CÔNG THỊNH - 20521960** |

**TP. HỒ CHÍ MINH, 11/2022**

[1. GIỚI THIỆU BÀI TOÁN 3](#_Toc120864433)

[1.1. Lí do chọn đề tài: 3](#_Toc120864434)

[1.2. Input và output: 3](#_Toc120864435)

[1.2.1. Input 3](#_Toc120864436)

[1.2.2. Output 4](#_Toc120864437)

[1.3. Dataset 4](#_Toc120864438)

[1.4. Các dạng khuôn mặt 4](#_Toc120864439)

[1.4.1. Heart shape 4](#_Toc120864440)

[1.4.2. Oblong shape 5](#_Toc120864441)

[1.4.3. Round shape 6](#_Toc120864442)

[1.4.4. Square shape 6](#_Toc120864443)

[1.4.5. Oval shape 7](#_Toc120864444)

[2. GIỚI THIỆU VỀ MÔ HÌNH 7](#_Toc120864445)

[2.1. VGG16 8](#_Toc120864446)

[2.2. Resnet50 8](#_Toc120864447)

[3. CÁC HƯỚNG TIẾP CẬN 8](#_Toc120864448)

[3.1. Raw: sử dụng ảnh gốc không qua xử lý 8](#_Toc120864449)

[3.1.1. SVM và KNN 8](#_Toc120864450)

[3.1.2. VGG16 và ResNet50 10](#_Toc120864451)

[3.2. Crop face: 11](#_Toc120864452)

[3.2.1. SVM và KNN 11](#_Toc120864453)

[3.2.2. VGG16 và ResNet50 13](#_Toc120864454)

[3.2.3. Deep feature 14](#_Toc120864455)

[3.3. Landmark: 16](#_Toc120864456)

[3.4. Tổng hợp kết quả: 18](#_Toc120864457)

[4. NHẬN XÉT 18](#_Toc120864458)

[5. Thách thức bài toán và cải tiến 19](#_Toc120864459)

[5.1. Thách thức 19](#_Toc120864460)

[5.1.1. Bộ dữ liệu 19](#_Toc120864461)

[5.1.2. Các yếu tố ảnh hưởng 19](#_Toc120864462)

[5.2. Hướng cải thiện cho bài toán: 19](#_Toc120864463)

[6. THAM KHẢO 20](#_Toc120864464)

FACE SHAPE CLASSIFICATION

# GIỚI THIỆU BÀI TOÁN

## Lí do chọn đề tài:

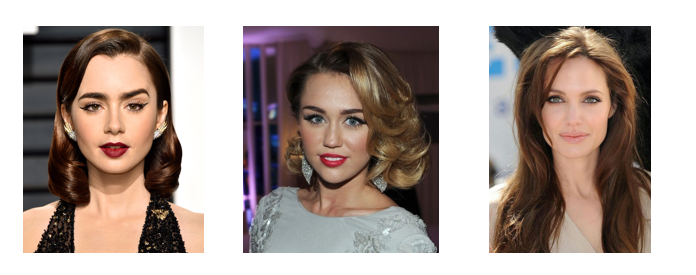
Ngày nay, trong việc tư vấn lựa chọn kiểu tóc hay kiểu gọng kính để phù hợp với khuôn mặt một cách riêng tư, online và tự động hoàn toàn thì việc đầu tiên ta cần làm là xác định hình dạng khuôn mặt. Do đó, chúng ta có thể thực hiện giải pháp bằng xây dựng mô hình máy học để thực hiện việc này với độ chính xác cao.

## Input và output:

### Input

Input là một bức ảnh mặt người

* Chụp chính diện
* Không có vật cản che mặt
* Ảnh sắc nét
* Điều kiện ánh sáng tốt



Hình 1.Ví dụ về input của bài toán

### Output

A picture containing text, person, indoor, computer

Description automatically generatedOutput là hình dạng khuôn mặt có trong ảnh

Hình 2. Ví dụ về output của bài toán

## Dataset

Bộ dữ liệu bao gồm 5000 ảnh và được chia thành 2 tập train và test với tỉ lệ 0.8 và 0.2. Bộ dữ liệu bao gồm 5 lớp: heart, oblong, oval, round và square

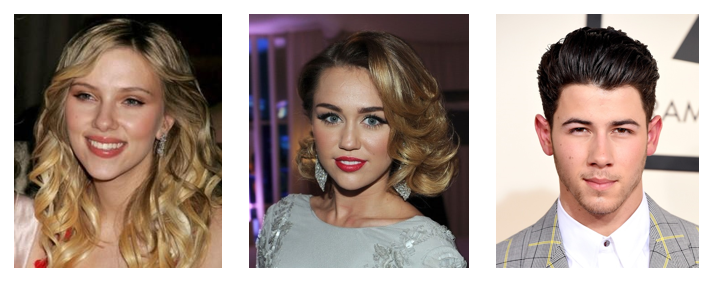
## Các dạng khuôn mặt

Theo trang web [14], định nghĩa các dạng khuôn mặt như sau:

### Heart shape

Đặc điểm mặc trái tim:

* Trán sẽ rộng hơn gò má (hoặc xấp xỉ)
* Chiều dài đường viền hàm của bạn cũng sẽ nhỏ hơn so với trán.
* Trán và gò má sẽ có chiều rộng lớn hơn đường viền hàm
* Cằm nhọn, tương tự như đầu dưới của trái tim.



Hình 3. Ví dụ về mặt heart shape

### Oblong shape

Đặc điểm mặt thuông dài:

* Chiều dài của khuôn mặt sẽ là đặc điểm nổi bật nhất.
* Về chiều rộng, đường viền hàm, trán và gò má sẽ có số đo gần bằng nhau.



Hình 4. Ví dụ về mặt oblong shape

### Round shape

Đặc điểm mặt hình tròn:

* Chiều dài khuôn mặt gần bằng chiều rộng của gò má.
* Chiều rộng đường viền hàm và chiều rộng trán gần bằng nhau.
* Chiều rộng xương gò má và chiều dài khuôn mặt lớn hơn chiều rộng đường viền hàm và trán.



Hình 5. Ví dụ về mặt round shape

### Square shape

Đặc điểm mặt vuông:

* Mặt vuông là khuôn mặt có hàm rộng bằng hoặc hơn chiều dài ngang gò má,
* Phần trán rộng
* Cằm khá góc cạnh.

A collage of people

Description automatically generated with low confidence

Hình 6. Ví dụ về mặt square shape

### Oval shape

Đặc điểm mặt trái xoan:

* Khuôn mặt sẽ dài hơn đáng kể so với chiều rộng gò má
* Đường viền hàm nhìn chung sẽ nhỏ hơn chiều rộng của trán
* Chiểu rộng gò má thường sẽ rộng hơn trán (có thể bằng)

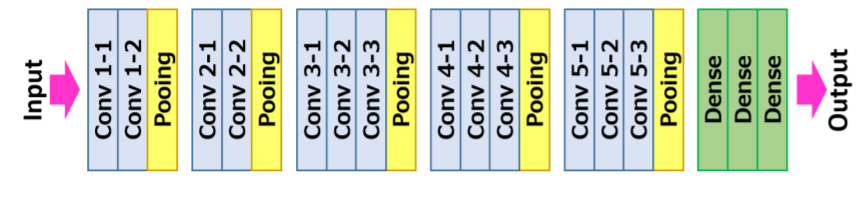


Hình 7. Ví dụ về mặt oval shape

# GIỚI THIỆU VỀ MÔ HÌNH

## VGG16

VGG16 là một mạng CNN phát hiện và phân loại đối tượng xuất hiện vào năm 2014. Model sau khi train bằng VGG16 có thể đạt độ chính xác 92.7% top-5 test trong bộ dữ liệu ImageNet gồm 14 triệu hình ảnh thuộc 1000 lớp khác nhau.. Nó là một trong những thuật toán phổ biến để phân loại hình ảnh và dễ sử dụng với transfer learning.



Hình 8. Minh họa về cấu trúc của model VGG16

## Resnet50

ResNet (Residual Network) được giới thiệu đến công chúng vào năm 2015 và đã giành được vị trí thứ 1 trong cuộc thi ILSVRC 2015 với tỉ lệ lỗi chỉ 3.57%. Và các biến thể có tên là ResNet và theo sau là số lớp nhất định.

ResNet được phát triển để giải quyết vanishing gradient. Vanishing gradient là vấn đề ta gặp phải khi huấn luyện các mạng nơ ron nhân tạo với các phương pháp dựa trên gradient và backpropagation do việc đạo hàm qua nhiều lớp khiến cho việc cật nhật trọng số bé theo cấp theo cấp số nhân và dẫn tới mạng học rất chậm hoặc không học được gì.

Ý tưởng là sử dụng các “kết nối tắt” nhằm bỏ qua một hoặc nhiều lớp thì các lớp ở phía trên có được thông tin trực tiếp hơn từ các lớp dưới nên sẽ điều chỉnh trọng số hiệu quả hơn.

# CÁC HƯỚNG TIẾP CẬN

Các model deep learning sẽ được sử dụng thông qua transfer learing và nhóm chỉ tùy chỉnh và train lại phần classification để phù hợp với yêu cầu bài toán, giữ nguyên tham số của phần feature extraction. Ngoài ra nhóm cũng sử dụng SVM và KNN cho bài toán đang đề cập, 2 model SVM và KNN để sử dụng các tham số mặc định.

## Raw: sử dụng ảnh gốc không qua xử lý

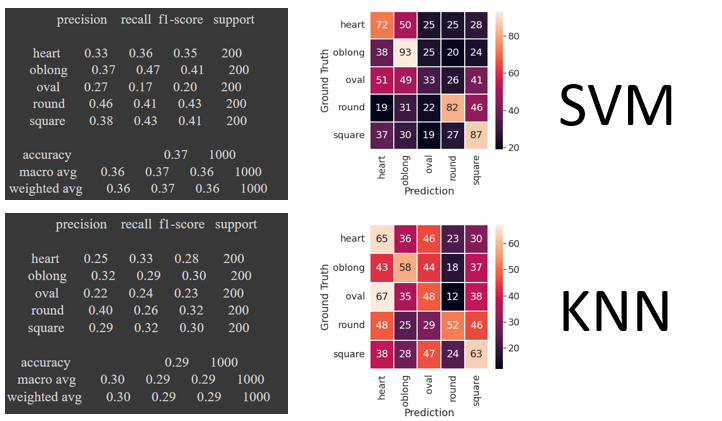
### SVM và KNN

Đối với 2 model SVM và KNN, ảnh đầu vào sẽ được resize về cùng một kích thước là (128,128) và thành ảnh xám. Sau đó sẽ được chuyển thành một véc tơ đặc trưng để model xử lý.

A picture containing timeline

Description automatically generated

Hình 9. Minh họa phương pháp tiếp cận Raw cho 2 mô hình SVM và KNN



Hình 10. Kết quả 2 model SVM và KNN ở phương pháp Raw

### VGG16 và ResNet50

Với 2 model VGG16 và ResNet50, ảnh đầu vào sẽ được resize về cùng kích thước phù hợp với đầu vào của model và đưa vào model để xử lý.

A picture containing graphical user interface

Description automatically generated

Hình 11. Minh họa phương pháp tiếp cận Raw cho 2 mô hình VGG16 và ResNet50

A picture containing chart

Description automatically generated

Hình 12. Kết quả 2 model VGG16 và ResNet50 ở phương pháp Raw

## Crop face:

### SVM và KNN

Một phương pháp dò tìm khuôn mặt sẽ được sử dụng để tìm kiếm khuôn mặt trên ảnh input, từ đó cắt khuôn mặt đã tìm thấy trên ảnh, chuyển thành ảnh xám và biến đổi thành véc tơ đặc trưng để model xử lý.

Graphical user interface

Description automatically generated

Hình 13. Minh họa phương pháp tiếp cận CropFace cho 2 model SVM và KNN

Graphical user interface, application

Description automatically generated

Hình 14. Kết quả 2 model SVM và KNN ở phương pháp CropFace

### VGG16 và ResNet50

Tìm kiếm khuôn mặt trên ảnh input, từ đó cắt khuôn mặt đã tìm thấy trên ảnh và thay đổi kích thước phù hợp để model xử lý

Diagram

Description automatically generated with low confidence Hình 15. Minh họa phương pháp tiếp cận CropFace cho 2 model VGG16 và ResNet50

Graphical user interface

Description automatically generated with medium confidenceHình 16. Kết quả 2 model SVM và KNN ở phương pháp CropFace

### Deep feature

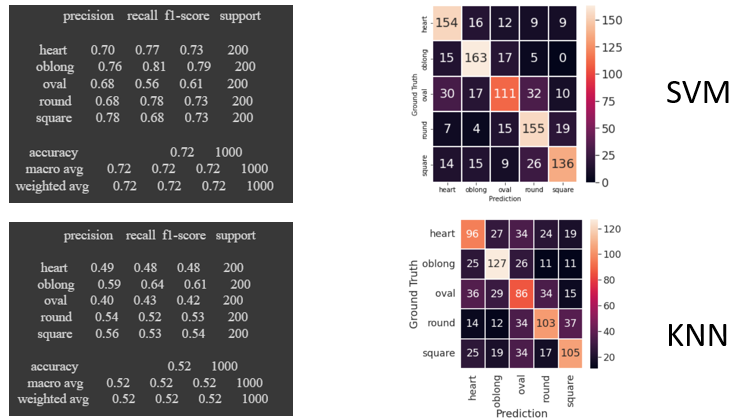
Sử dụng model VGG16 (không có phần classification) để trích xuất đặc trưng và coi output như là một véc tơ đặc trưng, sử dụng hai model SVM và KNN để xử lý các véc tơ đặc trưng đó

Graphical user interface, timeline

Description automatically generated

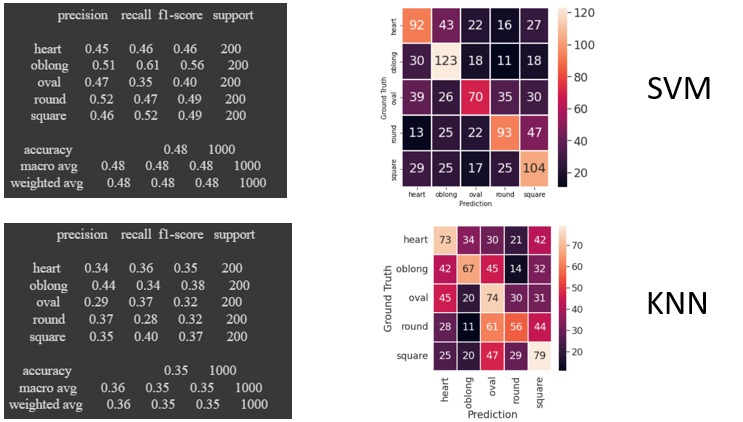
Hình 17. Minh họa từng bước xử lý

* + - 1. **VGG16**



Hình 18. Kết quả sử dụng VGG16 deep feature

* + - 1. **ResNet50**



Hình 19. Kết quả sử dụng ResNet50 deep feature

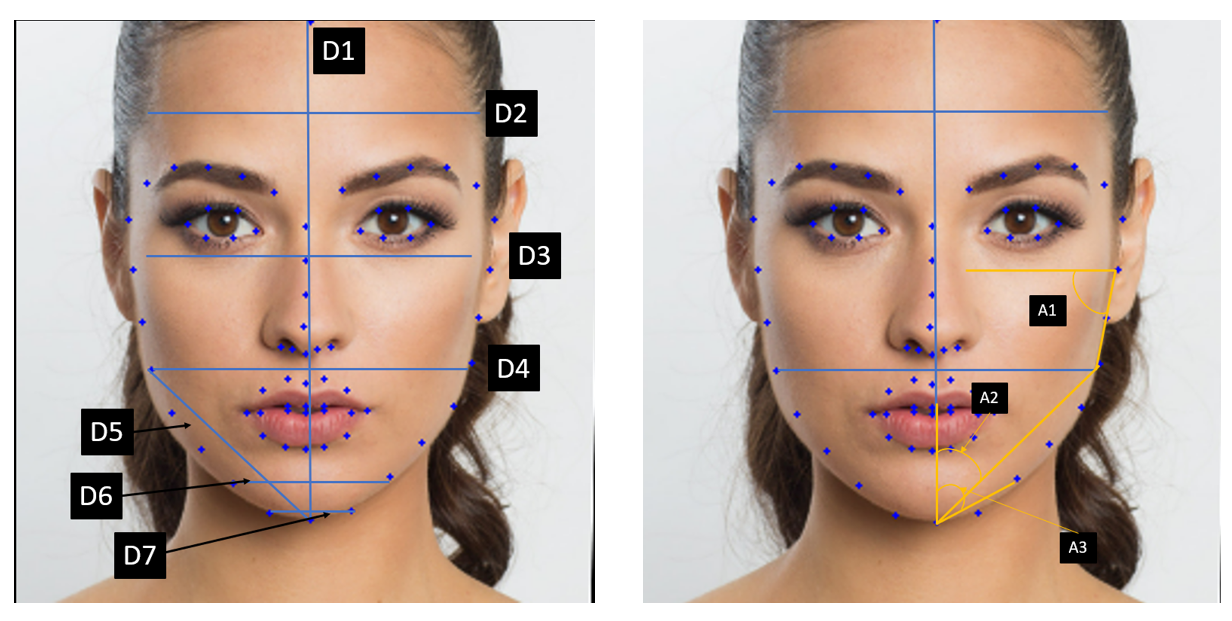
## Landmark:

Từ các hình ảnh đầu vào tiến hành crop face lấy ra vùng chứa khuôn mặt, sau đó sử dụng thư viện mediapipe để lấy được các điểm landmark trên khuôn mặt, từ các điểm landmark(điểm) thực hiện việc tính thông số các đặc điểm khuôn mặt qua quá xử lí và train bằng các thuật toán máy học đơn giản rồi sau đó đưa ra label dự đoán.

Graphical user interface, application

Description automatically generated

Hình 20. Minh họa phương pháp tiến hành với Lankmark



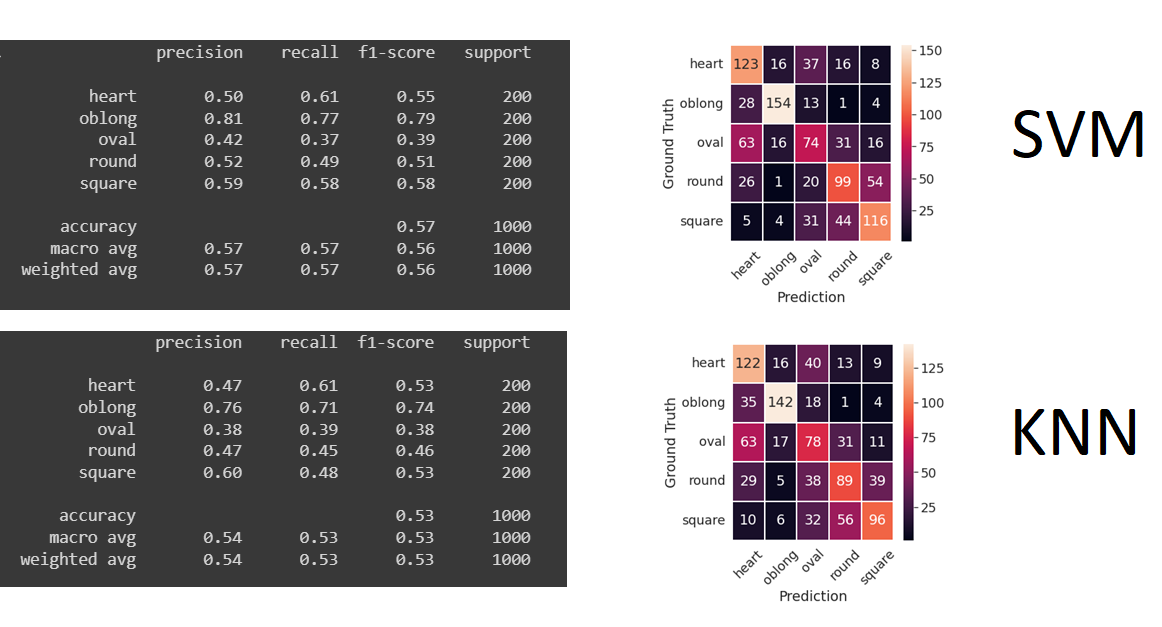
Hình 21. Minh họa phương pháp Landmark

Sử dụng các điểm landmark để tính toán các đặc điểm trên khuôn mặt:

* D1 - Face length: độ dài khuôn mặt tính từ cằm cho đến chân tóc vùng trán-
* D2 - Forehead width: độ rộng vùng trán tình từ đuôi 2 bên chân mày
* D3 - Cheekbone width: độ rộng vùng má
* D4 - Jaw width: độ rộng 2 bên hàm
* D5 - Jawline length: độ dài quai hàm tính từ đầu quai hàm cho đến vùng cằm
* D6 - Chin width 1: độ rộng cằm khúc trên
* D7 - Chin width 2: độ rộng cằm khúc dưới
* Và 3 chỉ số góc A1, A2, A3

Sau đó tiến hành tạo thêm các tỉ lệ từ các đặc điểm độ dài phía trên:

|  |  |
| --- | --- |
| * R1 = D2 / D3 * R2 = D3 / D1 * R3 = D2 / D1 * R4 = D2 / D4 * R5 = D6 / D4 | * R6 = D5 / D6 * R7 = D6 / D3 * R8 = D5 / D2 * R9 = D5 / D4 * R10 = D7 / D6 |





Hình 22. Kết quả 3 model SVM, KNN, Decision tree ở phương pháp Landmark

## Tổng hợp kết quả:

Bảng 1. Tổng hợp kết quả

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method** | **Accuracy** | **Precision** | **Recall** | **F1-score** |
| Raw: SVM | 0.37 | 0.36 | 0.37 | 0.36 |
| Raw: KNN | 0.29 | 0.30 | 0.29 | 0.29 |
| Raw: VGG16 | 0.48 | 0.49 | 0.48 | 0.47 |
| Raw: ResNet | 0.41 | 0.42 | 0.40 | 0.39 |
| Crop: SVM | 0.60 | 0.60 | 0.60 | 0.60 |
| Crop: KNN | 0.46 | 0.46 | 0.45 | 0.46 |
| Crop: VGG16 | 0.52 | 0.56 | 0.52 | 0.51 |
| **Crop: VGG16+SVM** | **0.72** | **0.72** | **0.72** | **0.72** |
| Crop: VGG16+KNN | 0.52 | 0.52 | 0.52 | 0.52 |
| Crop: ResNet | 0.63 | 0.64 | 0.63 | 0.62 |
| Crop: ResNet+SVM | 0.48 | 0.48 | 0.48 | 0.48 |
| Crop: ResNet+KNN | 0.35 | 0.36 | 0.35 | 0.35 |
| LM: SVM | 0.57 | 0.57 | 0.57 | 0.56 |
| LM: KNN | 0.53 | 0.54 | 0.53 | 0.53 |
| LM: Decision Tree | 0.45 | 0.46 | 0.45 | 0.45 |

# NHẬN XÉT

* Trong các cách tiếp cận thì cách crop face nhìn chung có kết quả tốt nhất.
* Trong đó SVM sử dụng VGG16 deep feature vector cho kết quả tốt nhất.
* Tùy theo cách tiếp cận và model thì oval thường bị nhầm lẫn với round và heart, round thì thường bị nhầm lẫn với oval và square. Heart đôi số trường hợp sẽ bị nhầm với oblong và oval.
* Khi sử dụng deep learning model thì ResNet50 cho kết quả tốt hơn so với VGG16 nhưng khi sử dụng các deep learning model để trích xuất feature vector thì VGG16 lại cho kết quả tốt hơn.

# Thách thức bài toán và cải tiến

## Thách thức

### Bộ dữ liệu

Đối với các bài toán trong lĩnh vực máy học và học sâu, dữ liệu luôn là một phần quan trọng của bài toán. Dữ liệu trong các bài toán đòi hòi và nhiều và đa dạng để các model có thể hoạt động tốt.

Tuy nhiên bộ dữ liệu mà nhóm đang sử dụng có các hạn chế như: bộ dữ liệu chỉ bao gồm nữ, chỉ có mặt người phương Tây, một số khuôn mặt còn bị lặp lại và bộ dữ liệu chỉ có 5 dạng khuôn mặt. Vì thiếu sự đa dạng trong dữ liệu nên khi thực nghiệm kết quả sẽ không được tốt.

### Các yếu tố ảnh hưởng

* Ánh sáng: điều kiện ánh sáng cũng có thể gây ra thách thức đáng kể trong việc nhận dạng hình dạng khuôn mặt. Sự chiếu sáng làm thay đổi diện mạo khuôn mặt một cách đáng kể, nếu độ sáng khác nhau, cùng một người chụp bằng một máy ảnh kết quả thu được có thể khác nhau.
* Độ phân giải: Hình ảnh có độ phân giải thấp không cung cấp nhiều thông tin vì hầu hết đã bị mất đi. Đây cũng là một thách thức lớn trong quá trình nhận diện biểu cảm khuôn mặt.
* Tư thế của khuôn mặt: Tư thế của một khuôn mặt thay đổi khi chuyển động của đầu và góc nhìn của người đó thay đổi. Các chuyển động của đầu gây ra những thay đổi về hình dạng khuôn mặt.
* Những vật làm che khuất khuôn mặt như râunlàm che khuất miệng, kính râm che khuất mắt, khẩu trang,… Đây có thể được coi là một trong những thách thức quan trọng nhất của hệ thống nhận dạng khuôn mặt.

## Hướng cải thiện cho bài toán:

* Tăng cường thêm dữ liệu bao gồm: tăng cường mẫu, tăng dự đa dạng(giới tính, dân tộc, hình dạng khuôn mặt, …)
* Thử nghiệm nhiều mô hình
* Tìm hiểu và thực hiện thêm nhiều đặc trưng

# THAM KHẢO

[1] katebloom, Determine Your Face Shape, katebloom, <https://www.katebloom.co.uk/blog/entry/determine-your-face-shape>

[2] adonistio, in Face shape classification using Inception v3. arXiv:1911.07916

[3] tensorflow, <https://www.tensorflow.org/api_docs/python/tf/keras/applications/resnet50>

[4] tensorflow, <https://www.tensorflow.org/api_docs/python/tf/keras/applications/vgg16>

[5] sklearn, <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>

[6] sklearn, <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html>

[7] mediapipe, Mediapipe Face Mesh, <https://google.github.io/mediapipe/solutions/face_mesh.html>

[8] mediapipe, Mediapipe Face Detection, https://google.github.io/mediapipe/solutions/face\_mesh.html

[9] Datum Codes, Custom Face Mesh Mediapipe tutorial | Landmarks, Youtube, https://www.youtube.com/watch?v=LGPBRH6Hqw8

[10] Niten Lama, Face Shape Dataset, Kaggle, [https://www.kaggle.com/datasets/niten19/face-shape-dataset](https://www.kaggle.com/datasets/niten19/face-shape-dataset%20)

[11] HM Edit@rs, How to Determine Your Face Shape: The Ultimate Guide, https://hairmotive.com/how-to-determine-face-shape/

[12] Adrian Rosebrock, Face Alignment with OpenCV and Python, https://pyimagesearch.com/2017/05/22/face-alignment-with-opencv-and-python/

[13] Mehta Ashwinee, Taha Mahamed. Human Face Shape Classification with Machine Learning, DOI:[10.13140/RG.2.2.28282.57282](http://dx.doi.org/10.13140/RG.2.2.28282.57282)

[14] Hair Motive,  [HM Edit@rs](mailto:HM Edit@rs), How to Determine Your Face Shape: The Ultimate Guide, <https://hairmotive.com/how-to-determine-face-shape/>