

# Facial Expression Recognition

**Lớp:** CS331.M21

**Giảng viên:** Nguyễn Vinh Tiệp

**Nhóm:** 12



# Thành viên

1

Nguyễn Thị Như Ý

19522555

2

Nguyễn Tấn Tú

19522454

3

Đinh Hoàng Linh Đan

19521309

4

Trần Nguyễn Quỳnh Anh

19521217



# MỤC LỤC

1	Giới thiệu
2	Dataset
3	Model
4	Thách thức bài toán và hướng phát triển
5	Demo



# Giới thiệu

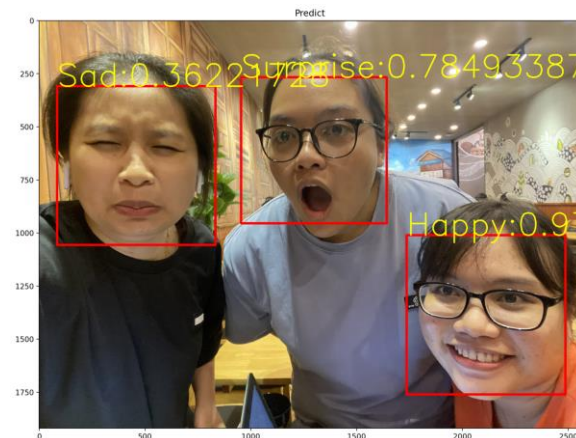
## » Giới thiệu bài toán

- ▶ Nhận diện biểu cảm biểu cảm là một trong những bản năng của con người
- ▶ Con người thể hiện cảm xúc thông qua các biểu cảm khuôn mặt, luôn tồn tại một số đặc trưng chung trên khuôn mặt bất kể độ tuổi, vị trí địa lý hay điều kiện sống... Dựa vào đặc trưng này, ta rút trích ra các đặc điểm quan trọng của cảm xúc, mô hình hóa, và dạy cho máy tính hiểu được cảm xúc đó.
- ▶ Ta chia biểu cảm khuôn mặt vào bảy loại sắc thái chính: tức giận(angry), sợ hãi(fear), ngạc nhiên(surprised), buồn(sad), chán ghét(disgusted), hạnh phúc(happy), trung lập(neutral),



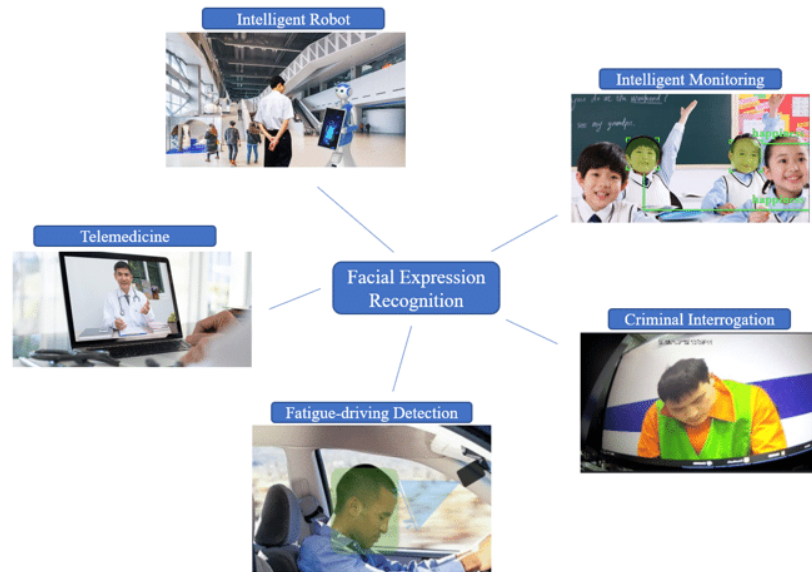
## » Giới thiệu bài toán

- ▶ **Input:** Bức ảnh chứa mặt người
  - Chụp chính diện
  - Ánh sáng tốt
  - Không vật cản lớn
- ▶ **Output:** Dự đoán biểu cảm và accuracy của gương mặt có trong ảnh



# » Ứng dụng

- ▶ **Trong lĩnh vực tiếp thị:** phân tích phản hồi của khách hàng
- ▶ **Trong dịch vụ khách hàng:** phân tích mức độ hài lòng của khách hàng.
- ▶ **Trong trò chơi điện tử và trò chơi thực tế ảo:** kiểm tra trải nghiệm người dùng.
- ▶ **Trong y tế:** Giúp bác sĩ và chuyên gia y tế đánh giá sức khỏe bệnh nhân.
- ▶ Xác định những người có **hành vi đáng ngờ** trong đám đông từ đó có thể ngăn chặn tội phạm, khủng bố.



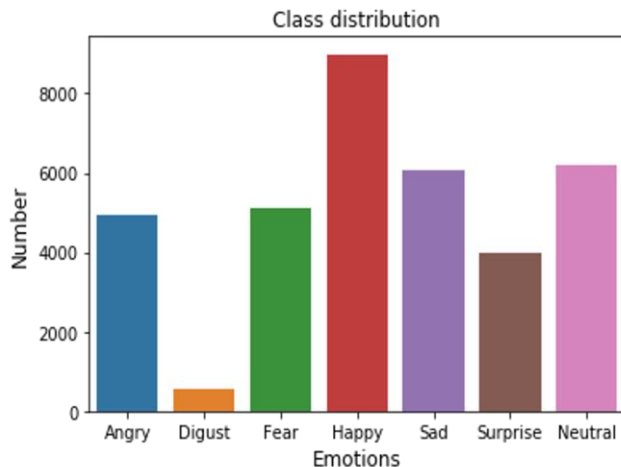


# Dataset



## » Dataset

- ▶ **Dataset:** FER-2013 gồm 35887 ảnh xám kích cỡ 48x48 chia thành 2 tập training và validation theo tỉ lệ lần lượt là 0.8 và 0.2
- ▶ Có 7 lớp biểu cảm
- ▶ **Dữ liệu test:** 883 ảnh chia đúng tỉ lệ với tỉ lệ của dataset



	emotion	number
0	Angry	4953
1	Disgust	547
2	Fear	5121
3	Happy	8989
4	Sad	6077
5	Surprise	4002
6	Neutral	6198

9
1
9
17
11
7
11

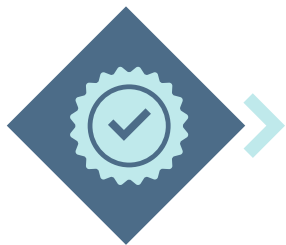


**MODEL**

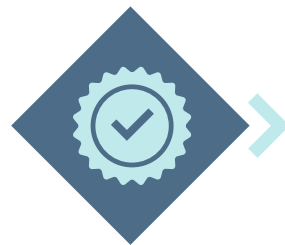
## » Các nghiên cứu trước

Model	Accuracy
VGG19	70.80
EfficientNet_b2b	70.80
Googlenet	71.97
Resnet34	72.42
Inception_v3	72.72
Bam_Resnet50	73.14
Densenet121	73.16
Resnet152	73.22
Cbam_Resnet50	73.39
ResMaskingNet	74.14
ResMaskingNet + 6	76.82

## » Model sử dụng



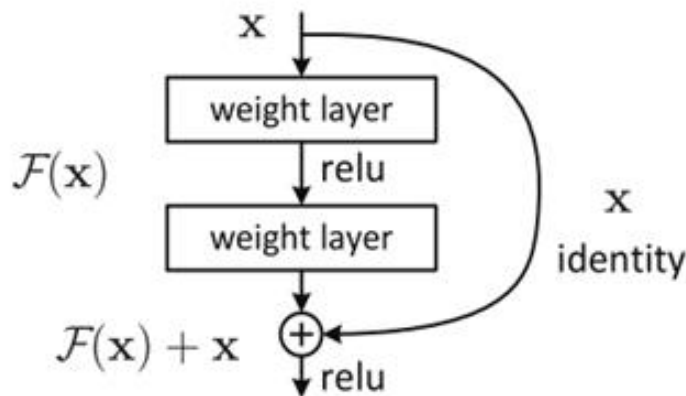
**ResNet-50**



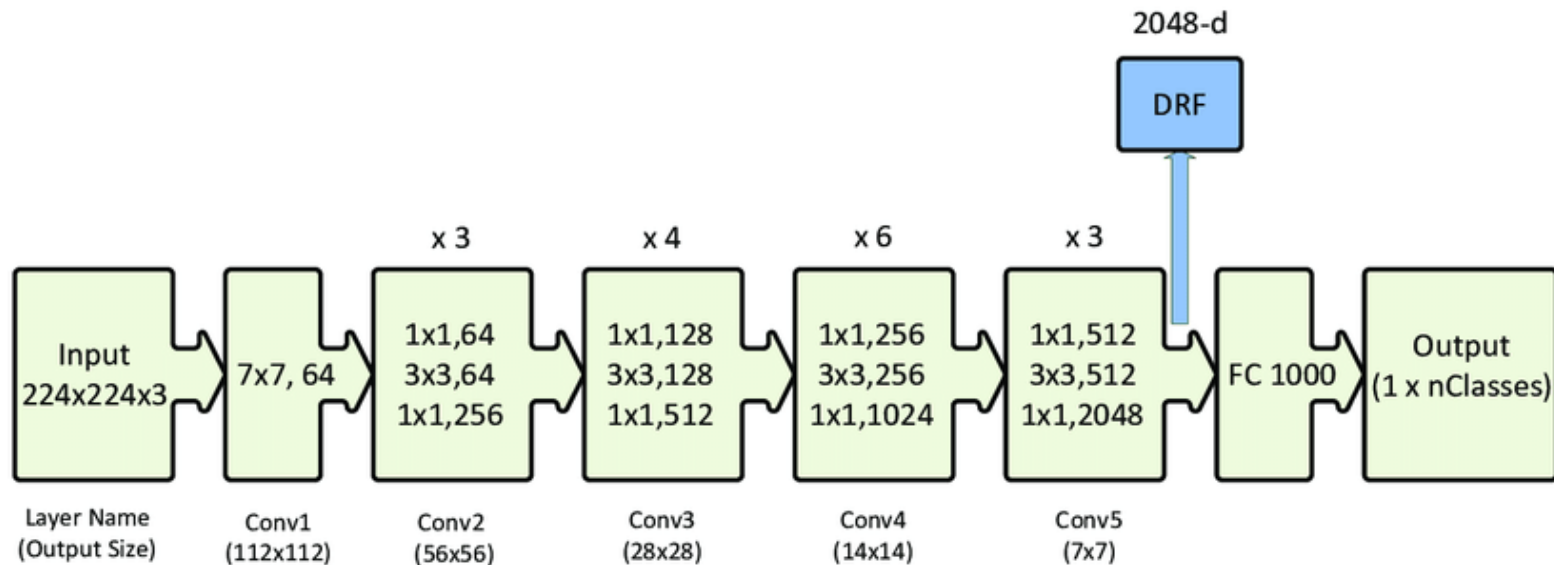
**VGG16**

## » ResNet

- Ý tưởng chính của ResNet là sử dụng kết nối tắt (Skip connection) đồng nhất để xuyên qua một hay nhiều lớp. Các kết nối tắt giúp giữ thông tin không bị mất bằng cách kết nối từ layer sớm trước đó tới layer phía sau và bỏ qua một vài layers trung gian.



## » Kiến trúc ResNet-50

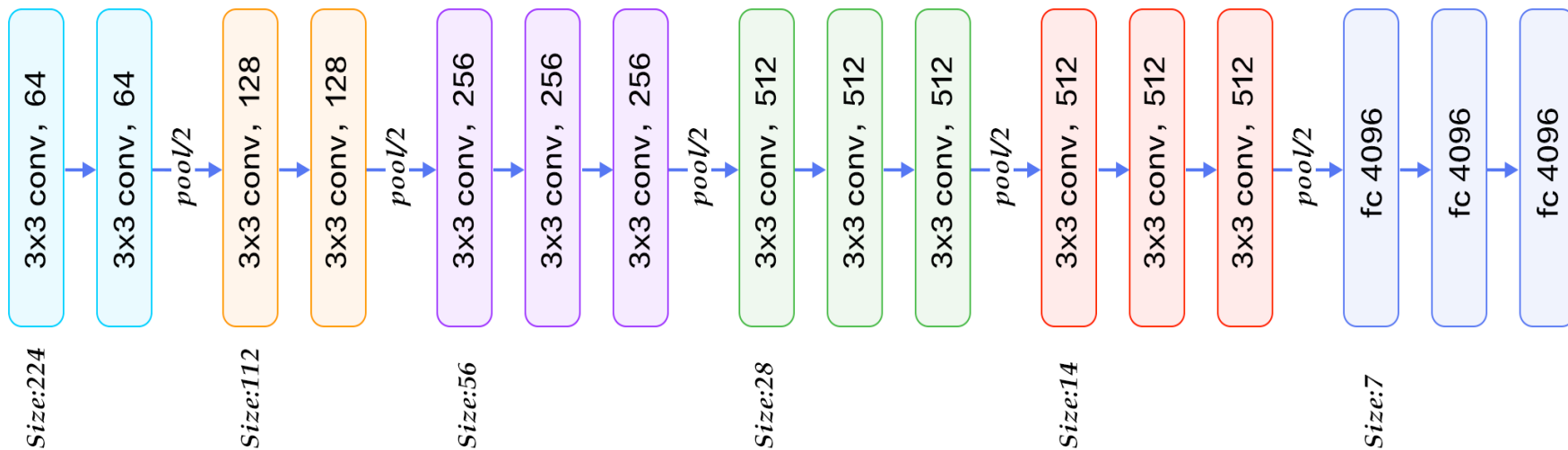


## » Transfer Learning ResNet50

- ▶ Dùng pre-trained model ResNet50
- ▶ Thêm các lớp sau:
  - ◆ GlobalAveragePooling2D()
  - ◆ Dropout(0.2)
  - ◆ Dense (7, activation='softmax')
- ▶ Sử dụng learning rate thấp : 0.0005.
- ▶ Parameters : 23,602,055



# Kiến trúc VGG16



Kiến trúc VGG16 bao gồm 16 layer : 13 Convolution layer đều có Kernel 3x3, sau mỗi convolution layer là maxpooling downsize xuống 0.5, và 3 fully connected layer.  
Số parameters : khoảng 138 triệu.



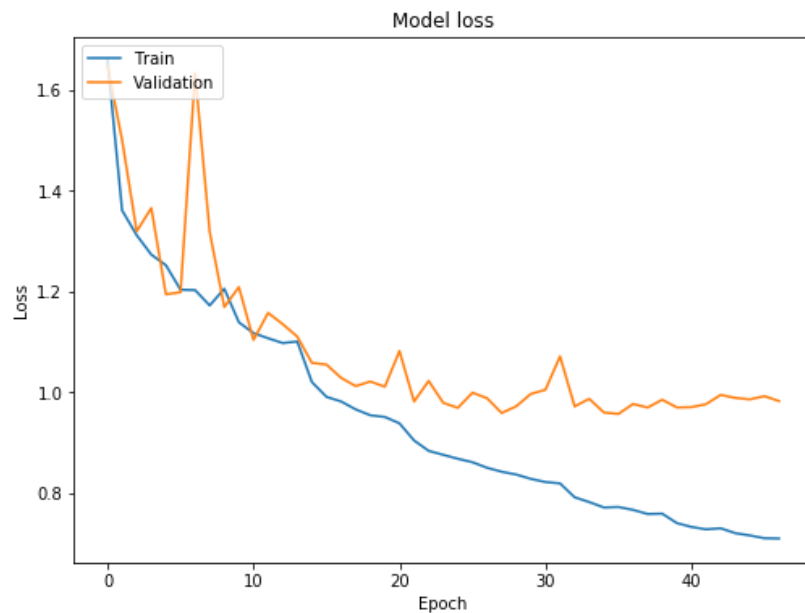
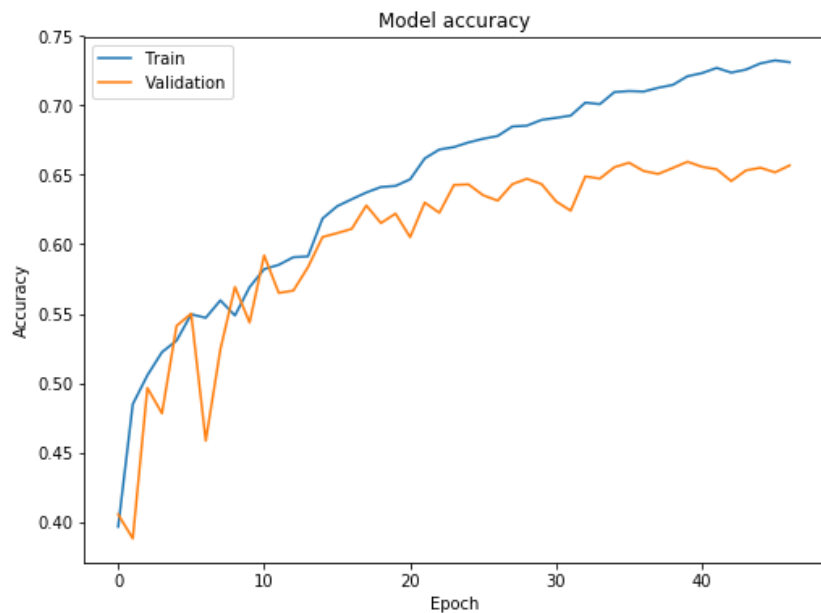
# » Transfer Learning VGG16

- ▶ Dùng pre-trained model VGG16
- ▶ Thêm các lớp sau:
  - ◆ GlobalAveragePooling2D()
  - ◆ Dropout(0.2)
  - ◆ Dense(7, activation='softmax')
- ▶ Sử dụng learning rate thấp : 0.0005.
- ▶ Parameters : 14,718,279

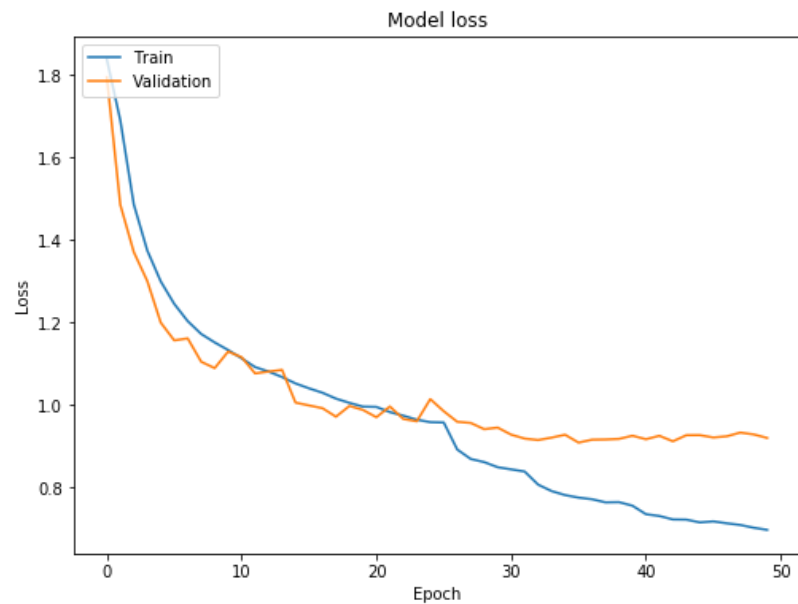
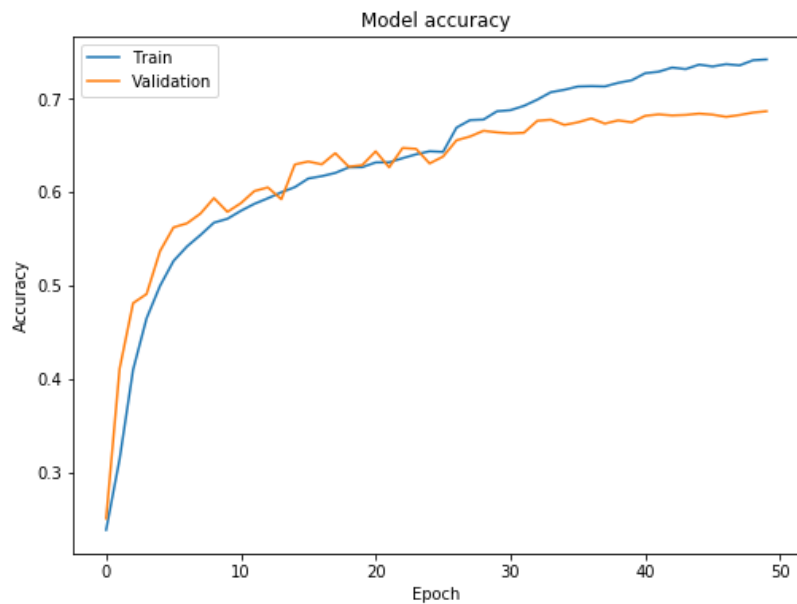


**Kết quả**

# » ResNet50

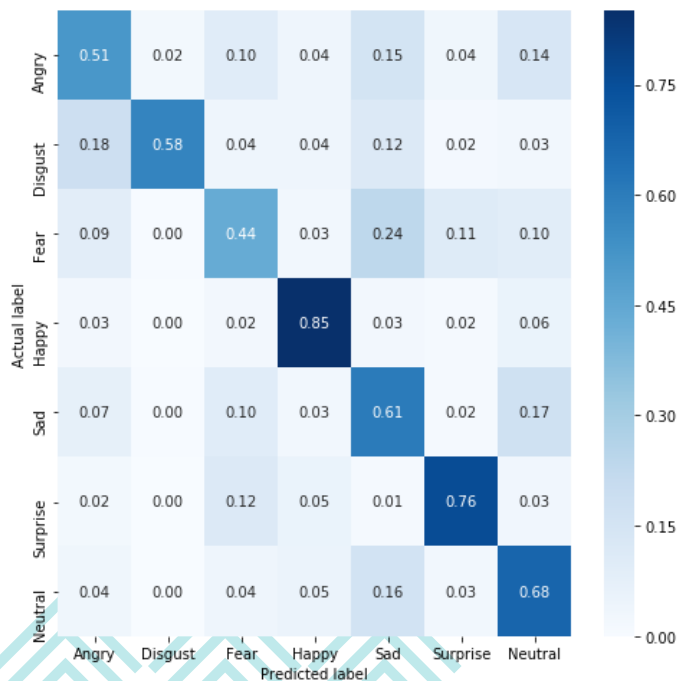


# » VGG16

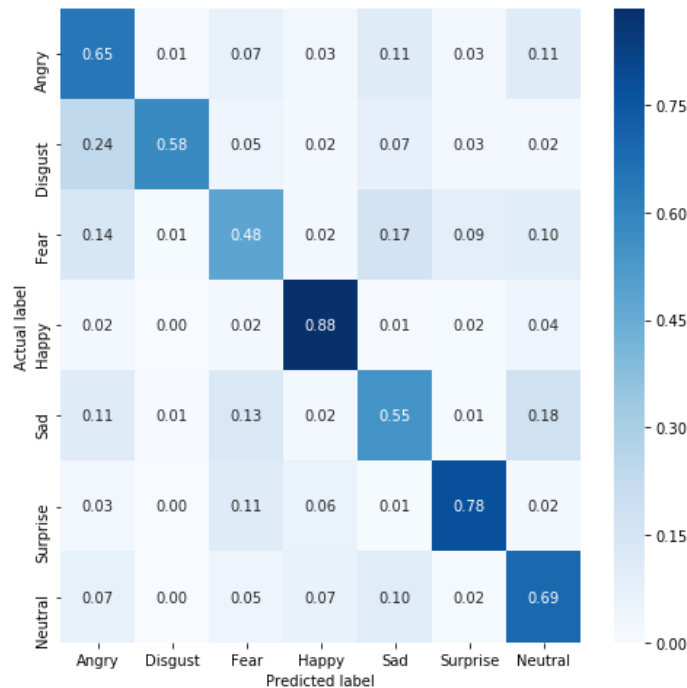


# » So sánh

► Val\_accuracy : ResNet50 (0.6594) , VGG16 (0.6863)



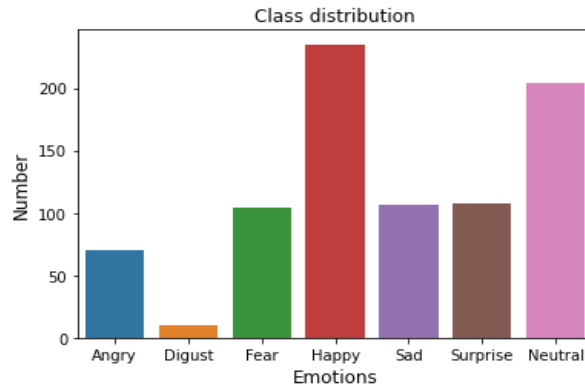
ResNet50



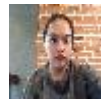
VGG16

## » Thử nghiệm trên tập private test của nhóm

- ▶ Gồm 838 ảnh theo tỉ lệ như tập train.
- ▶ Kết quả accuracy khá kém : ResNet50 (0.23) và VGG16 (0.37)
- ▶ Do các ảnh được chụp bằng điện thoại hiện đại nên có độ phân giải cao (chiều cao và chiều rộng từ 2000 trở lên). Trong khi input của model là ảnh 48x48.
- ▶ Có nhiều ảnh giống nhau và biểu cảm của khuôn mặt người chụp chưa được chuẩn xác.



2640x1980 -> 48x48





# Thách thức của bài toán

## » Dataset

- Dataset được sử dụng để training có gương mặt là người châu Âu, còn tập test có gương mặt châu Á nên kết quả dự đoán cho ra kết quả không cao

Disgust

Angry

Sad

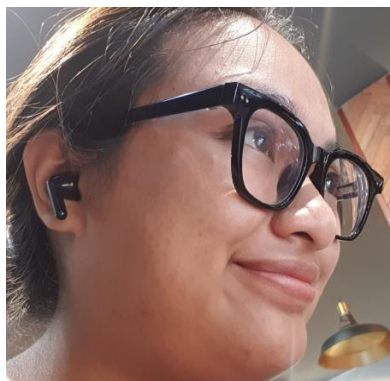
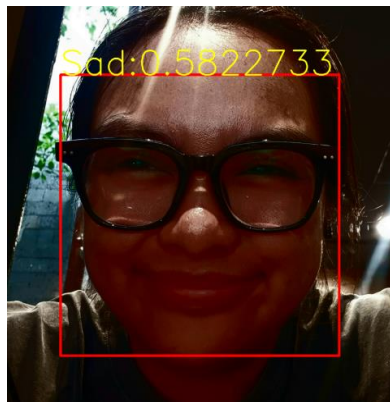
Surprise





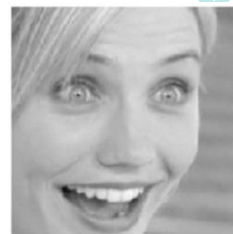
## » Các yếu tố ảnh hưởng

- ▶ Điều kiện ánh sáng
- ▶ Độ phân giải
- ▶ Tư thế của khuôn mặt
- ▶ Những vật làm che khuất khuôn mặt (râu, ria mép, kính râm, khẩu trang, mũ lưỡi trai, ...)



## » Kết quả dự đoán

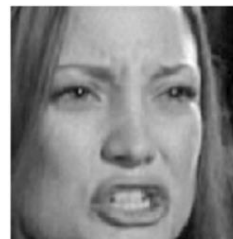
- ▶ Kết quả thử nghiệm thực tế cho thấy mô hình khá nhạy khi nhận biết cảm xúc **Happy**, khá kém với cảm xúc **Disgust**.
- ▶ **Fear** với **Angry**, **Neutral** với **Sad** có biểu cảm khá giống nhau nên model thường nhầm lẫn. Kết quả là cho ra độ chính xác không cao.
- ▶ Việc hầu hết các mô hình được công bố với tập dữ liệu **FER-2013** đều chỉ đạt độ chính xác thấp (**dưới 80%**), điều này có thể cho thấy bộ dữ liệu này có những yếu tố mất cân bằng hoặc nhiều khi gán nhãn dữ liệu.



Happy



Disgusted



Angry



Fear



Neutral



Sad



## » **Hướng phát triển**

- ▶ Thêm dataset có gương mặt châu Á để phù hợp với điều kiện thử nghiệm thực tế
- ▶ Tăng số lượng mẫu của các biểu cảm để hệ thống đánh giá tốt hơn và chính xác hơn
- ▶ Thử nghiệm với nhiều mô hình mạng học sâu khác hơn.
- ▶ Cải thiện web API
- ▶ Hướng đến phát hiện biểu cảm khuôn mặt thông qua video.



## » Tham khảo

- ▶ [Nghiên cứu nhận dạng biểu cảm khuôn mặt bằng phương pháp học sâu sử dụng kiến trúc ResNet - Tài liệu, ebook, giáo trình, hướng dẫn \(timtailieu.vn\)](#)
- ▶ [FER2013 | Kaggle](#)
- ▶ [Simple CNN on fer2013 dataset | Kaggle](#)
- ▶ [facial-emotion-recognition/emotion\\_recognizer.ipynb at main · esra-polat/facial-emotion-recognition · GitHub](#)
- ▶ [37-151.pdf \(vap.ac.vn\)](#)
- ▶ [1804.08348.pdf \(arxiv.org\)](#)



**DEMO**

# Facial Expression Recognition - VGG16

Facial Emotion Recognition VGG16 - a Hugging Face Space

[https://huggingface.co/spaces/linhdan412/Facial\\_Emotion\\_Regconition](https://huggingface.co/spaces/linhdan412/Facial_Emotion_Regconition)

# Facial Expression Recognition - ResNet50

Facial Emotion Recognition ResNet50 - a Hugging Face Space

[https://huggingface.co/spaces/linhdan412/Facial\\_Emotion\\_Regconition\\_ResNet50](https://huggingface.co/spaces/linhdan412/Facial_Emotion_Regconition_ResNet50)

**Cảm ơn thầy và các bạn đã  
theo dõi**

