Facial Expression Recognition

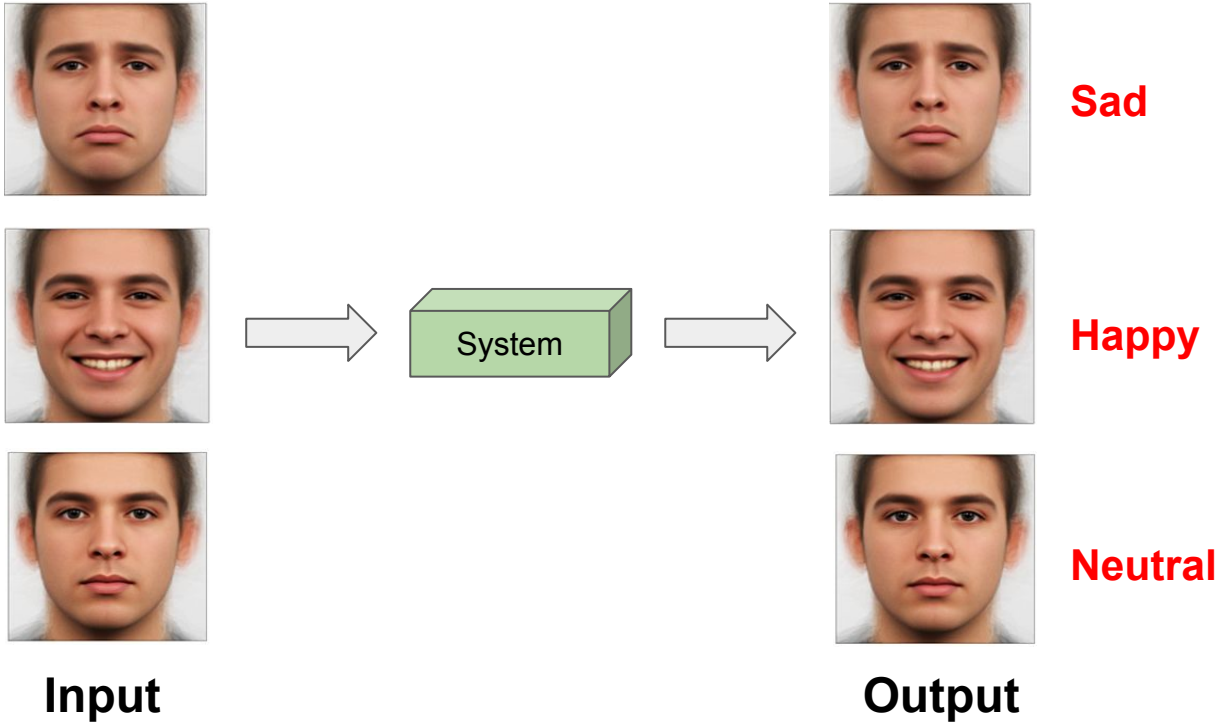
Supervisor: Đinh Viết Sang

Our team:

- Nguyễn Minh Dũng
- Lê Công Pha
- Nguyễn Duy Nhất
- Võ Minh Tâm

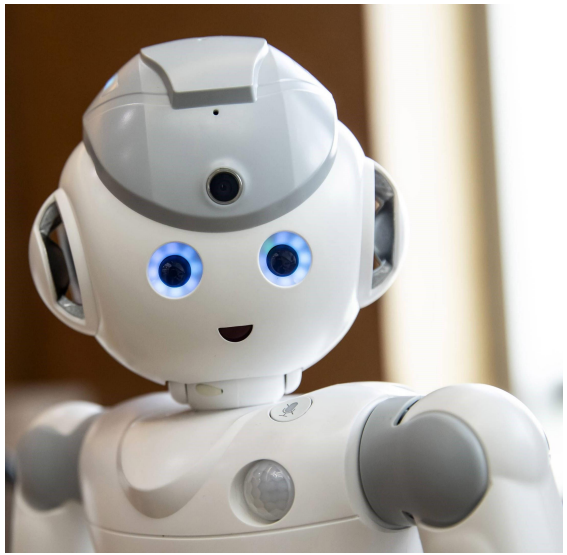
Contents

Introduction - The problem



Introduction - Applications

Customer services



Assistant Robot

Introduction - Applications

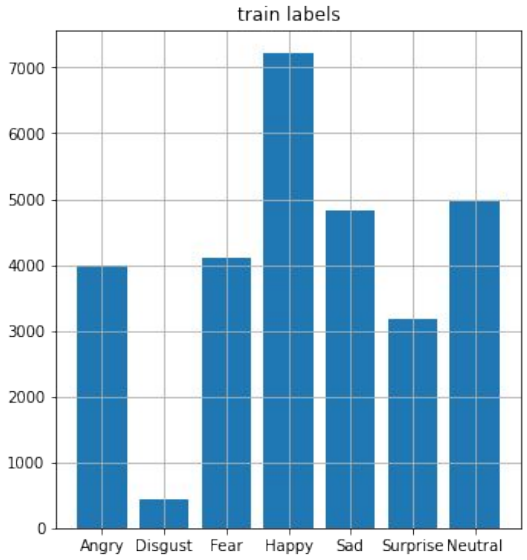
Autopilot

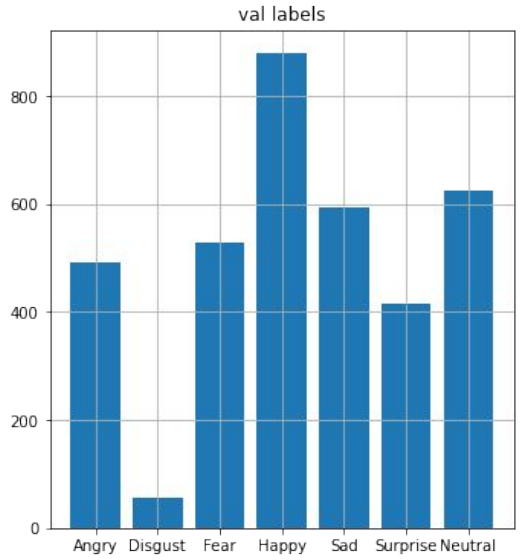
Surveillance

Dataset

Name	**Facial Expression Recognition 2013 (FER-2013)**
Originators	Pierre-Luc Carrier and Aaron Courville
Source	Challenges in Representation Learning: Facial Expression Recognition Challenge
Num of samples	35887 (train:val:test 8:1:1)
Num of classes	7 (0: Angry, 1: Disgust, 2: Fear, 3: Happy, 4: Sad, 5: Surprise, 6: Neutral)
Size of each sample	48x48x1

Dataset

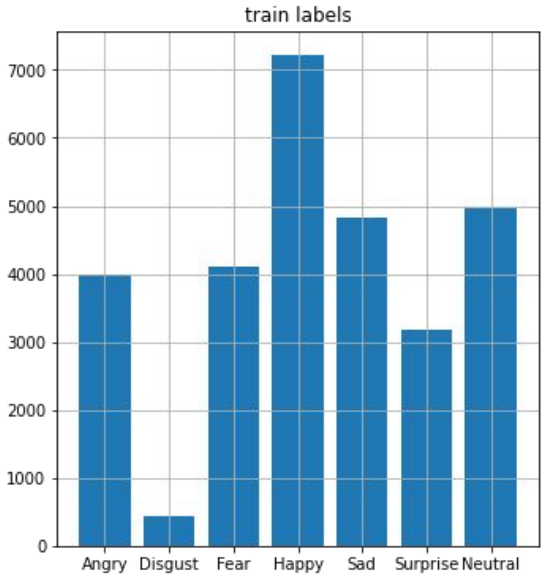




https://datarepository.wolframcloud.com/resources/FER-2013

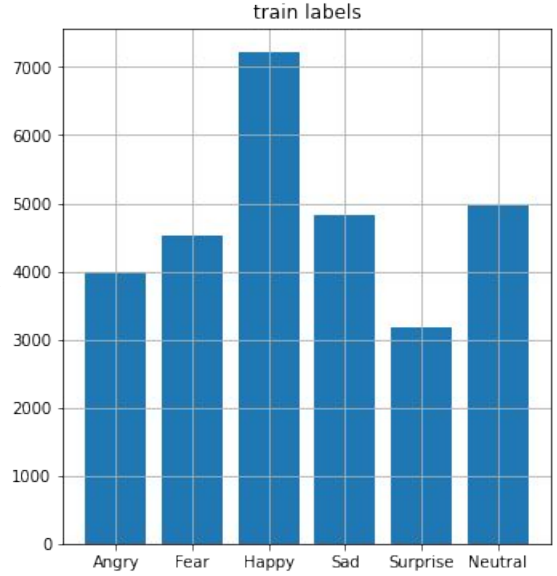
Prior experiments

- **Imbalanced data**

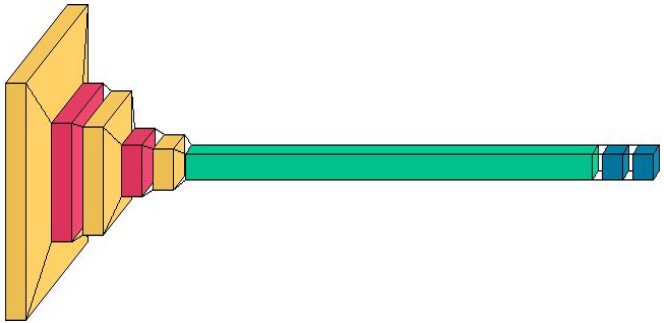


Disgust, Fear

Fear

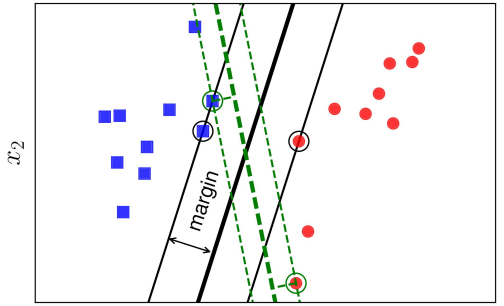


Prior experiments



Shallow CNN

+



SVM [1]

Accuracy: 0.5932

[1] https://machinelearningcoban.com/assets/19_svm/svm5.png

Our Approach - Imbalanced data

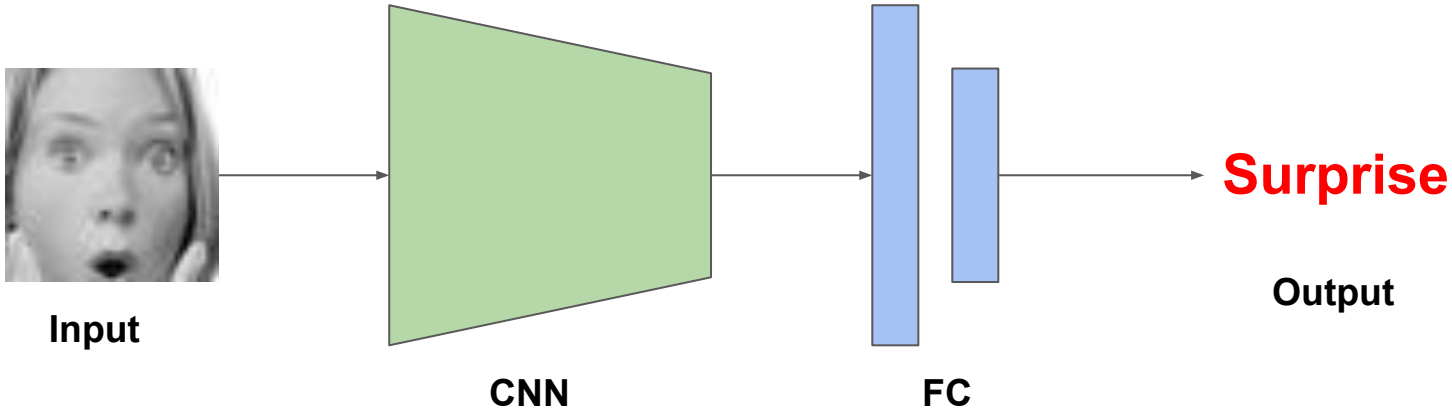
Balance method	Model	Val Accuracy
RandomOverSampling	ResNet50	0.5943
SMOTE	ResNet50	0.5982
AugDisgust + Class weight	ResNet50	0.5690

Some faces generated by SMOTE

Accuracy of RandomOverSampling and SMOTE: approximately the same

--> **We choose RandomOverSampling method to balance data**

Our Approach - Improve accuracy



ResNet50
VGGFace
VGGFace2
EfficientNet B3/B4

Our Approach

Based on approaches in kilean notebook*: fine-tuning VGGFace2

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
vggface_resnet50 (Functional (None, 1, 1, 2048)        23561152
_________________________________________________________________
flatten (Flatten)            (None, 2048)              0
_________________________________________________________________
dropout (Dropout)            (None, 2048)              0
_________________________________________________________________
dense (Dense)                (None, 2048)              4196352
_________________________________________________________________
dropout_1 (Dropout)          (None, 2048)              0
_________________________________________________________________
dense_1 (Dense)              (None, 1024)              2098176
_________________________________________________________________
classifer (Dense)            (None, 7)                 7175
=================================================================
Total params: 29,862,855
Trainable params: 6,301,703
Non-trainable params: 23,561,152
_________________________________________________________________
```

#	△pub	Team Name	Team Members	Score
1	—	RBM		0.71161
2	—	Unsupervised		0.69267
3	—	Maxim Milakov		0.68821

Private Leaderboard of FER

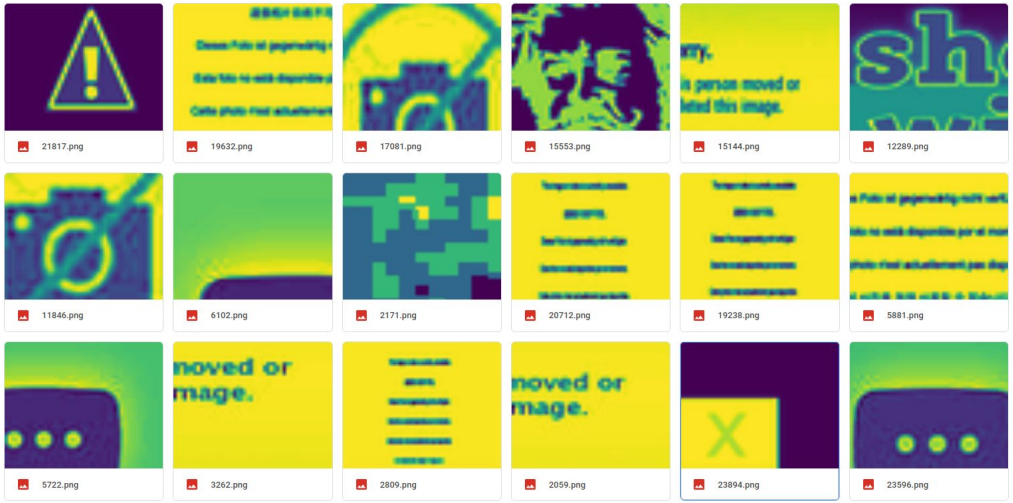
Validation	0.6761
Test	0.6981

Kilean notebook result

*https://www.kaggle.com/kilean/emotion-detection-accuracy70 in Challenges in representation learning: facial expression recognition challenge

Our Approach - Cleaned data

- Removed 86 images not a face
- Fine-tuning VGGFace2 with RanOS, ADASYN

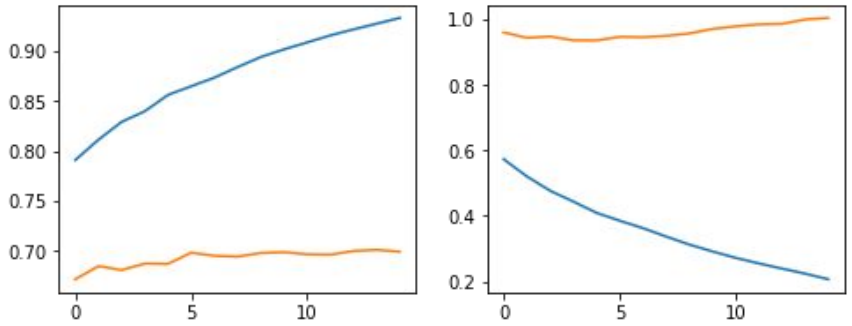


Mô hình	Sử dụng	Dataset	Val Acc
VGGFace2	RandomOverSampler	cleaned FER-2013	0.7031
VGGFace2	ADASYN	cleaned FER-2013	0.7051

--> **Cleaned data is better**

Our Approach - Cleaned data

- Fine-tuning VGGFace1, EfficientNetB3, EfficientNetB4



accuracy(left) and loss(right). lr=1e-4, ep=15

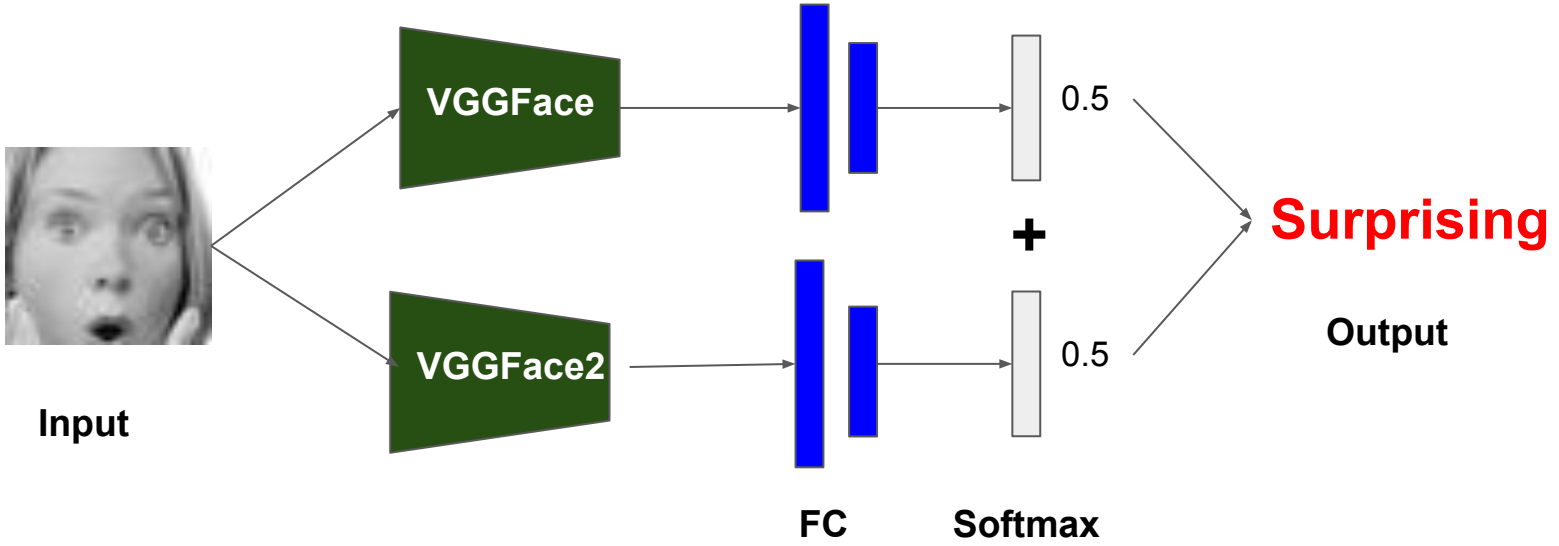
Layer (type)	Output Shape	Param #
vggface_vgg16 (Functional)	(None, 7, 7, 512)	14714688
global_average_pooling2d_3 (	(None, 512)	0
batch_normalization_3 (Batch	(None, 512)	2048
dropout_7 (Dropout)	(None, 512)	0
dense_6 (Dense)	(None, 512)	262656
dropout_8 (Dropout)	(None, 512)	0
dense_7 (Dense)	(None, 128)	65664
classifer (Dense)	(None, 7)	903

Total params: 15,045,959
Trainable params: 330,247
Non-trainable params: 14,715,712

Mô hình	Sử dụng	Dataset	Val Acc
VGGFace1	RandomOverSampler	cleaned FER-2013	0.7062
EfficientNetB4	RandomOverSampler	cleaned FER-2013	0.5234
EfficientNetB3	RAndomOverSampler	cleaned FER-2013	0.5017

--> Acc on VGGFace1 is better than EfficientNet

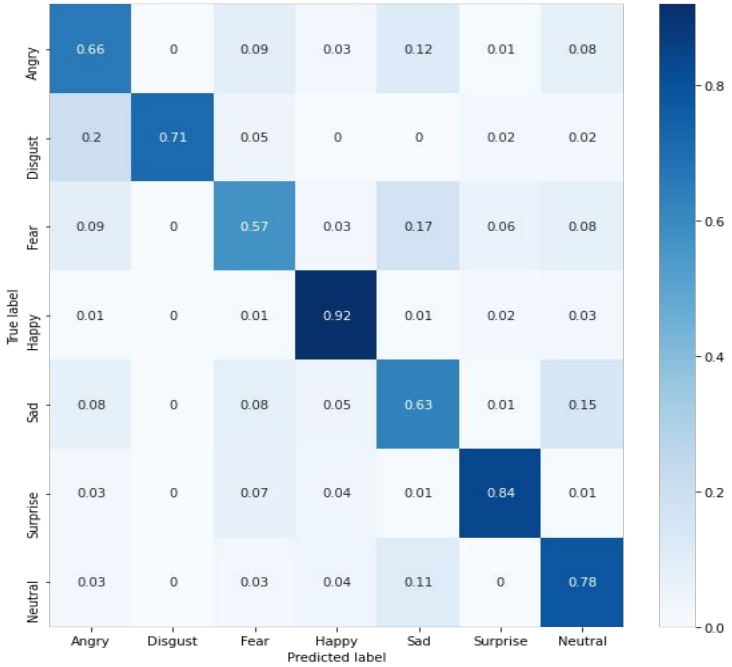
Our Approach - Ensemble learning



Mô hình	Sử dụng	Dataset	Val Acc
VGGFace1+2	RandomOverSampler	cleaned FER-2013	**0.7283**

--> **Ensemble VGGFace with VGGFace2 has the best accuracy on validation set**

Our Approach - Ensemble learning



Methods	Acc on Test
kilean's notebook	0.6761
first leaderboard	0.7116
Our model	**0.74.7**

RANK	MODEL	ACCURACY	EXTRA TRAINING DATA	PAPER	CODE	RESULT	YEAR
1	Ensemble ResMaskingNet with 6 other CNNs	76.82	✓	Challenges in Representation Learning: A report on three machine learning contests			2020
2	Residual Masking Network	74.14	✓	Challenges in Representation Learning: A report on three machine learning contests			2020
3	VGG	72.7	×	Facial Expression Recognition using Convolutional Neural Networks: State of the Art			2016
4	Res-Net	72.4	×	Facial Expression Recognition using Convolutional Neural Networks: State of the Art			2016
5	Inception	71.6	×	Facial Expression Recognition using Convolutional Neural Networks: State of the Art			2016
6	DeepEmotion	70.02	×	Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network			2019

Some result

predict: Angry
label: Angry

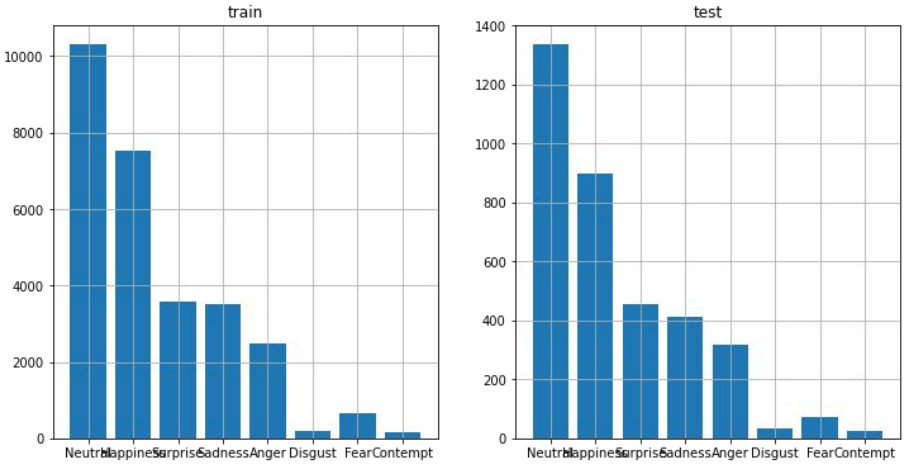
predict: Fear
label: Fear

predict: Angry
label: **Fear**

predict: Angry
label: **Sad**

FER+ dataset

- The FER+ annotations provide a set of new labels (labeled by 10 crowd-sourced taggers) for the standard Emotion FER dataset.
- Class: **neutral, happiness, surprise, sadness, anger, disgust, fear, contempt**, unknown, NF

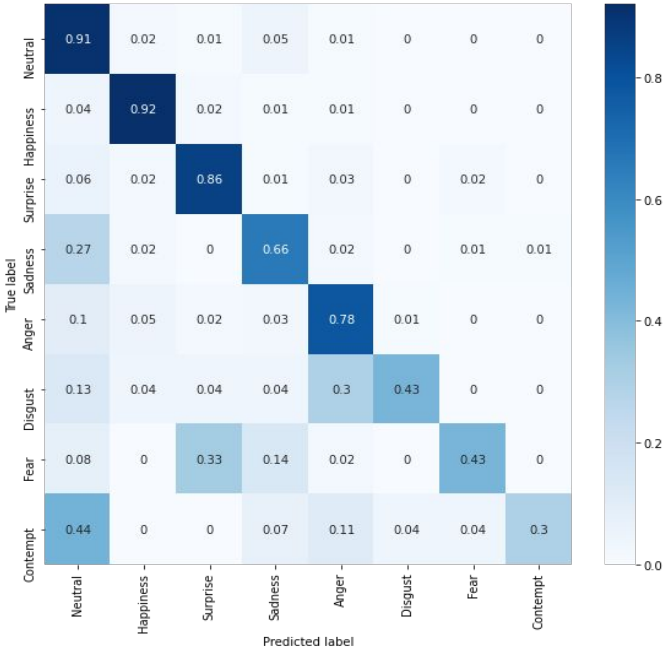


Phân bố dữ liệu FER+

Label of FER (top), new label FER+ (bottom)

https://github.com/microsoft/FERPlus

Ensemble learning on FER+



Method	Acc on Test
Our model	**0.8432**
PSR(VGG16)	0.8975

TN	Mô hình	Sử dụng	Dataset	Val Acc
1	VGGFace1	RandomOverSampler	cleaned FER-2013	0.8477
2	VGGFace2	RandomOverSampler	cleaned FER-2013	0.8365
3	VGGFace1+2	RandomOverSampler	cleaned FER-2013	**0.857**

BẢNG 4.4: Kết quả thực nghiệm trên bộ dữ liệu FER+

Henrique Siqueira, et al, Efficient Facial Feature Learning with Wide Ensemble-Based Convolutional Neural Networks. AAAI 2020

Conclusions

- Use: **RandomOverSampling**, **ADASYN** (Adaptive Synthetic), **Augment+Class weight**, **SMOTE** (Synthetic Minority Over-sampling Technique).
- Use pre-trained models: **VGGFace**, **VGGFace2**, **EfficientNet**, **ResNet50.**

  => VGGFace, VGGFace2, ensemble better.
- A **demo** in Python using Flask framework.

Thank you

Task assignment

		4-Jan	5-Jan	6-Jan	7-Jan	8-Jan	9-Jan	10-Jan	11-Jan	12-Jan	13-Jan	14-Jan	15-Jan	16-Jan	17-Jan	18-Jan
Analyze	Data augmentation	All														
	Pretrained CNN models		Nhat, Dung, Tam													
Act	Training on FER2013				Pha , Nhat, Tam											
	Demo										All					
	Report								All							