

Deep Facial Expression Recognition: A Survey

Shan Li and Weihong Deng

Febin Sebastian Elayanithottathil

Introduction

- Facial expression is one of the most powerful, natural and universal signals for human beings to convey their emotional states and intentions.
- Overfitting caused by a lack of sufficient training data and expression-unrelated variations, such as illumination, head pose and identity bias
- Static image FER and Dynamic sequence FER
- Practical importance:
 - Sociable robotics, medical treatment, driver fatigue surveillance, human computer interaction

Facial Expression Database

Database	Samples	Subject	Condit.	Elicit.	Expression distribution	Access
CK+ [33]	593 image sequences	123	Lab	P & S	6 basic expressions plus contempt and neutral	http://www.consortium.ri.cmu.edu/ckagree/
MMI [34], [35]	740 images and 2,900 videos	25	Lab	P	6 basic expressions plus neutral	https://mmifacedb.eu/
JAFPE [36]	213 images	10	Lab	P	6 basic expressions plus neutral	http://www.kasrl.org/jaffe.html
TFD [37]	112,234 images	N/A	Lab	P	6 basic expressions plus neutral	josh@mplab.ucsd.edu
FER-2013 [21]	35,887 images	N/A	Web	P & S	6 basic expressions plus neutral	https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge
AFEW 7.0 [24]	1,809 videos	N/A	Movie	P & S	6 basic expressions plus neutral	https://sites.google.com/site/emotiwchallenge/
SFEW 2.0 [22]	1,766 images	N/A	Movie	P & S	6 basic expressions plus neutral	https://cs.anu.edu.au/few/emotiw2015.html
Multi-PIE [38]	755,370 images	337	Lab	P	Smile, surprised, squint, disgust, scream and neutral	http://www.flintbox.com/public/project/4742/
BU-3DFE [39]	2,500 images	100	Lab	P	6 basic expressions plus neutral	http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html
Oulu-CASIA [40]	2,880 image sequences	80	Lab	P	6 basic expressions	http://www.cse.oulu.fi/CMV/Downloads/Oulu-CASIA
RaFD [41]	1,608 images	67	Lab	P	6 basic expressions plus contempt and neutral	http://www.socsci.ru.nl:8180/RaFD2/RaFD
KDEF [42]	4,900 images	70	Lab	P	6 basic expressions plus neutral	http://www.emotionlab.se/kdef/
EmotioNet [43]	1,000,000 images	N/A	Web	P & S	23 basic expressions or compound expressions	http://cbsl.ece.ohio-state.edu/dbform_emotionet.html
RAF-DB [44], [45]	29672 images	N/A	Web	P & S	6 basic expressions plus neutral and 12 compound expressions	http://www.whdeng.cn/RAF/model1.html
AffectNet [46]	450,000 images (labeled)	N/A	Web	P & S	6 basic expressions plus neutral	http://mohammadmahoor.com/databases-codes/
ExpW [47]	91,793 images	N/A	Web	P & S	6 basic expressions plus neutral	http://mmlab.ie.cuhk.edu.hk/projects/socialrelation/index.html

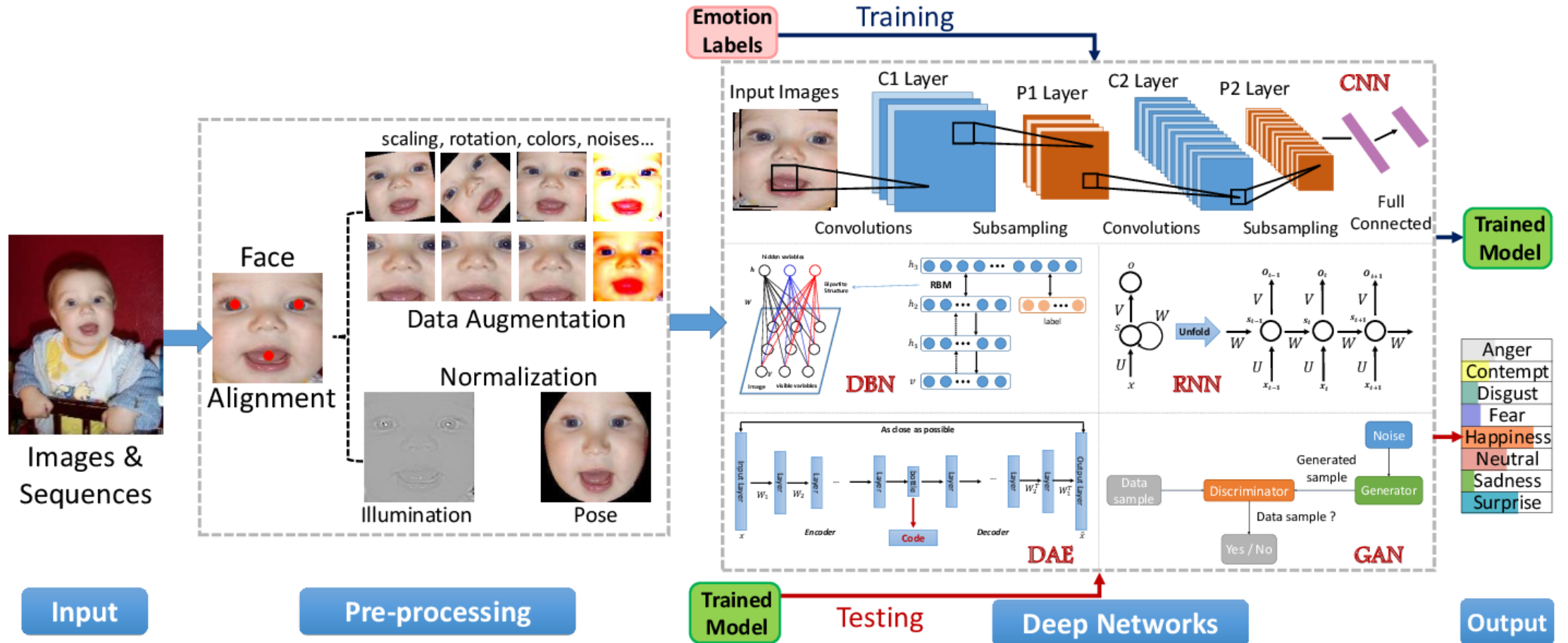
FER2013

- Introduced during the ICML 2013 Challenge
- Registered and resized to 48X48 pixels
- 28,709 training images, 3,589 validation images, 3,589 test images
- 7 expression labels (anger, disgust, fear, happiness, sadness, surprise and neutral)

Deep Facial Expression Recognition

- Pre-processing
 - Face alignment
 - Data augmentation
 - Illumination and normalization
 - Pose normalization
- Deep feature learning
 - Convolution neural network(CNN)
 - Deep belief network(DBN)
 - Deep autoencoder
 - Recurrent neural network (RNN)
 - Generative Adversarial Network (GAN)
- Deep feature classification

General pipeline of deep facial expression recognition systems.

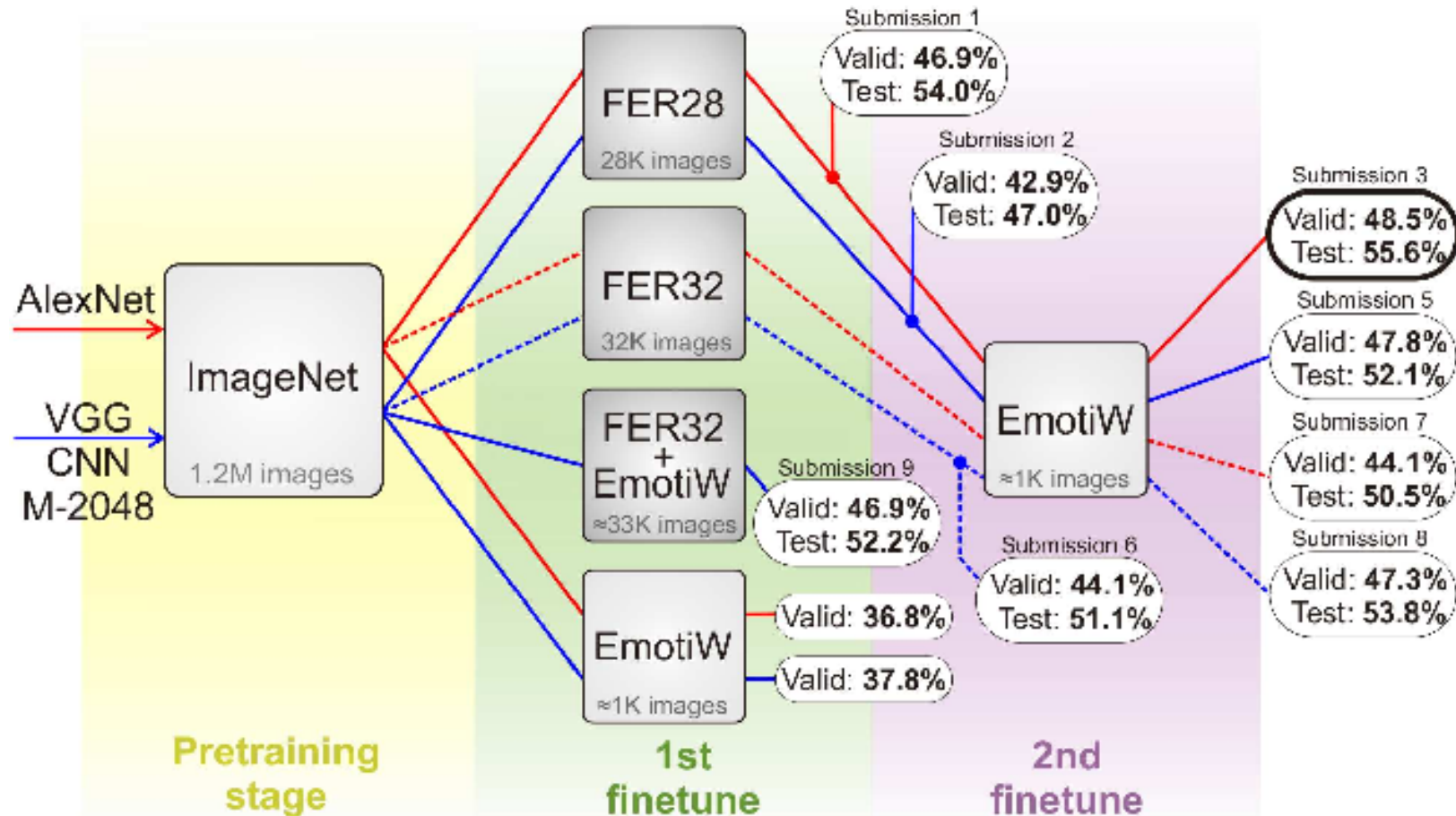


Deep FER network for static images

Two key issues:

- lack of plentiful diverse training data
- Expression unrelated variations, such as illumination, head pose and identity.
- Pre-train and fine-tune
- Diverse input
- Multitaks networks
- Cascaded network
- Network ensemble

Two-stage finetuning strategy



Deep FER network for dynamic images sequences

- Frame aggregation
 - Aggregate the network output for frames in each sequence to improve the performance
- Expression Intensity network
 - Takes training samples with different intensities as input.

Challenges and Opportunities

- Facial expression dataset
 - Lack of training data in terms of both quality and quantity
 - People of different age ranges, cultures and genders display and interpret facial expression in different ways
 - Occlusion and multipose
- Multimodal affect recognition
- Dataset bias and imbalanced distribution
- FER on infrared data
- FER on 3D static and dynamic data

Experiment

- Data Set: FER2013
- Tensorflow, keras
- Tensorflow js
- Colab
- Optimizer : adam
- Loss function : categorical_crossentropy
- Activation function: softmax
- Epochs: 110

Model architecture

```
IMAGE_SIZE = (48, 48)
INPUT_SHAPE = IMAGE_SIZE + (1,)

input_tensor = Input(shape=INPUT_SHAPE)
model = MobileNet(input_tensor=input_tensor, alpha=1.0,
                  include_top=False, dropout = 0.5, weights=None)

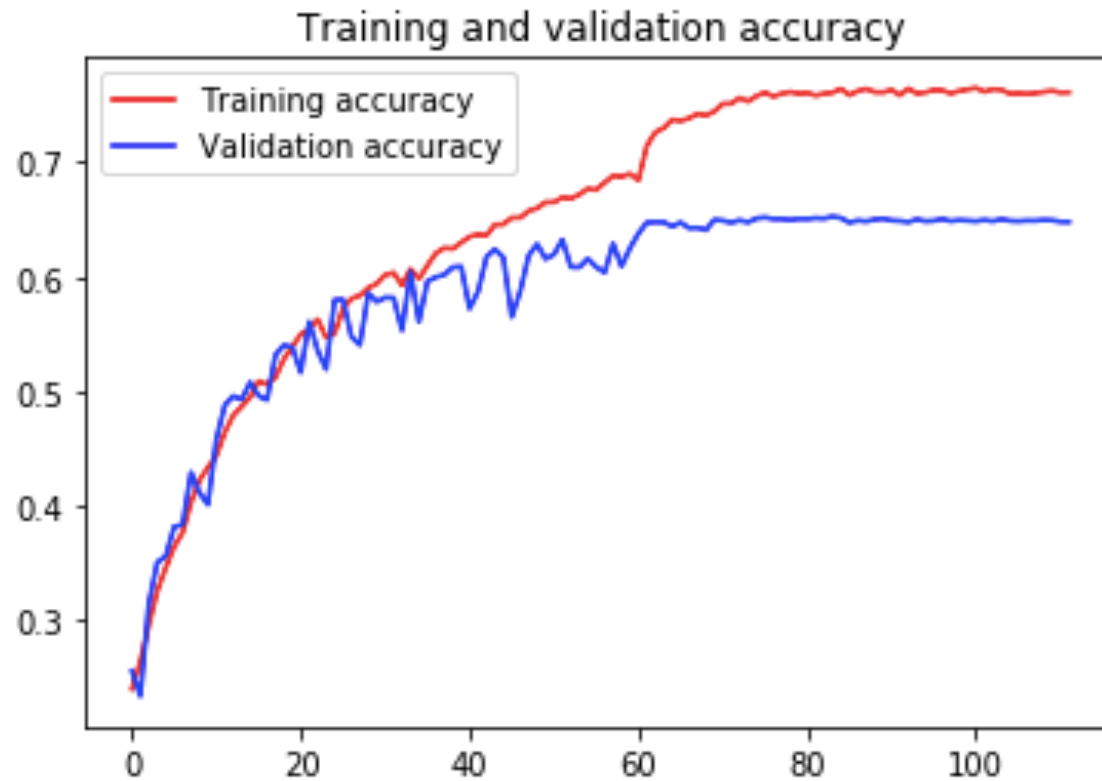
output = tf.keras.layers.Reshape((1024,))(model.output)
output = tf.keras.layers.Dense(7, activation='softmax')(output)
model = tf.keras.Model(model.input, output)
```

Model Summary

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 48, 48, 1)]	0
conv1_pad (ZeroPadding2D)	(None, 49, 49, 1)	0
conv1 (Conv2D)	(None, 24, 24, 32)	288
conv1_bn (BatchNormalization)	(None, 24, 24, 32)	128
conv1_relu (ReLU)	(None, 24, 24, 32)	0
conv_dw_1 (DepthwiseConv2D)	(None, 24, 24, 32)	288
conv_dw_1_bn (BatchNormaliza	(None, 24, 24, 32)	128
conv_dw_1_relu (ReLU)	(None, 24, 24, 32)	0

conv_pw_12_bn (BatchNormaliz	(None, 1, 1, 1024)	4096
conv_pw_12_relu (ReLU)	(None, 1, 1, 1024)	0
conv_dw_13 (DepthwiseConv2D)	(None, 1, 1, 1024)	9216
conv_dw_13_bn (BatchNormaliz	(None, 1, 1, 1024)	4096
conv_dw_13_relu (ReLU)	(None, 1, 1, 1024)	0
conv_pw_13 (Conv2D)	(None, 1, 1, 1024)	1048576
conv_pw_13_bn (BatchNormaliz	(None, 1, 1, 1024)	4096
conv_pw_13_relu (ReLU)	(None, 1, 1, 1024)	0
reshape (Reshape)	(None, 1024)	0
dense (Dense)	(None, 7)	7175
=====		
Total params: 3,235,463		
Trainable params: 3,213,575		
Non-trainable params: 21,888		

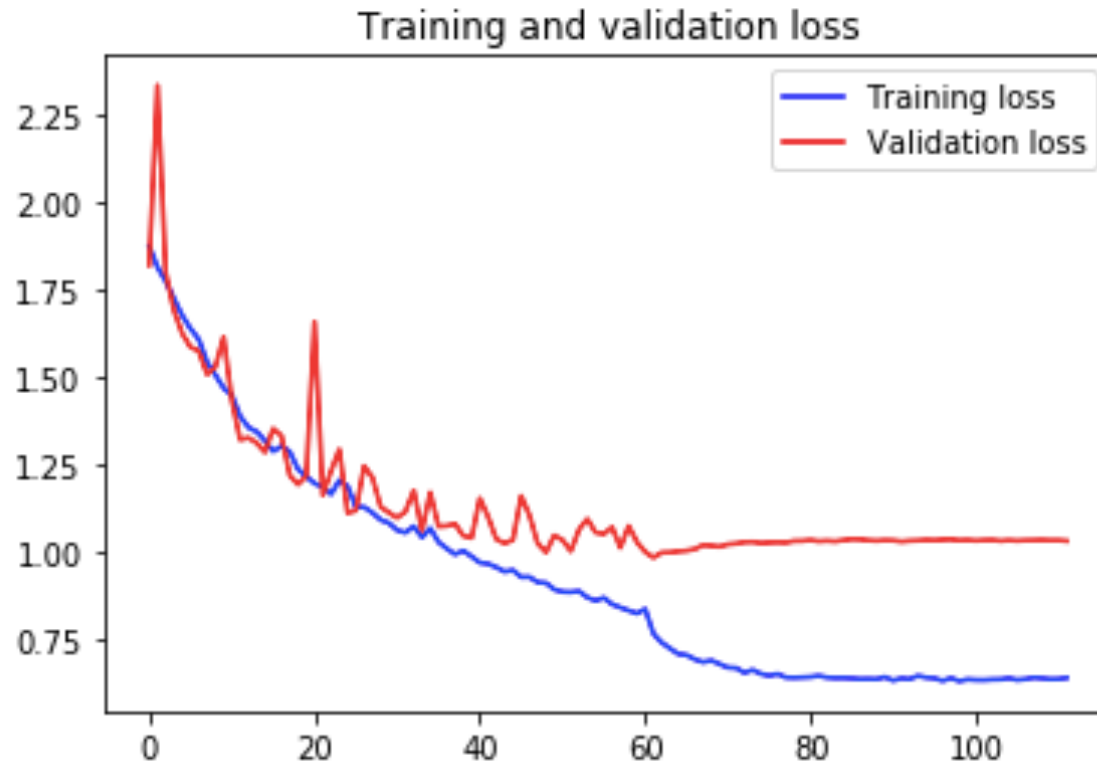
Accuracy



52 minutes

acc: 0.7614
val_acc: 0.6481

Training and Validation Loss



loss: 0.6401
val_loss: 1.0315





Demo

References



[REFERENCE PAPPER](#)



[COLAB](#)



[GITHUB](#)



Thanks!
Any questions ?