FACIAL EXPRESSION IDENTIFICATION

CHIEN-CHIEH, HU (UIN: 670782410)
DURGASHANKAR M P (UIN: 671192091)
KARTHIK RANGAPPA (UIN: 677923220)
LAAVANYA GANESH (UIN: 654324917)
SAVITHA TAMRAPARNI-VITTAL (UIN: 662405004)
VINEET KUMAR SINGH (UIN: 654489295)

Machine Learning with Python Final Project Report

INTRODUCTION

The major emotions which have been categorized universally are anger, disgust, fear, happiness, sadness, neutral and surprise. These sometimes subtle, yet complex, signals in an expression often contain an abundant amount of information about our state of mind.

Human facial recognition has become a widely-researched topic in recent times. In 2013, International Conference on Machine Learning, Facial emotion recognition challenge was first introduced.

It has become an important aspect to understand how well computers can do a job in recognizing human emotions.

We can measure the effects that content and services have on the audience/users through an easy and low-cost procedure. For example, retailers may use these metrics to evaluate customer interest. Healthcare providers can provide better service by using additional information about patents' emotional state during treatment. Entertainment producers can monitor audience engagement in events to consistently create desired content.

The objective of this project is to classify human emotions into these discrete categories. Presently, there is 75% accuracy for unseen data achieved from the deep learning models. The highest accuracy has been achieved through ensemble of 8 deep nets.

The project idea is to gain understanding from such existing models and infuse more approaches with the CNNs to get better accuracy on unseen data.

Another scope of project lies in tackling images with side shots, partial shots of faces and different rotations and scaling required as when compared to frontal shots of images.

DATASET DESCRIPTION

The Dataset used for this project has been taken from Kaggle.

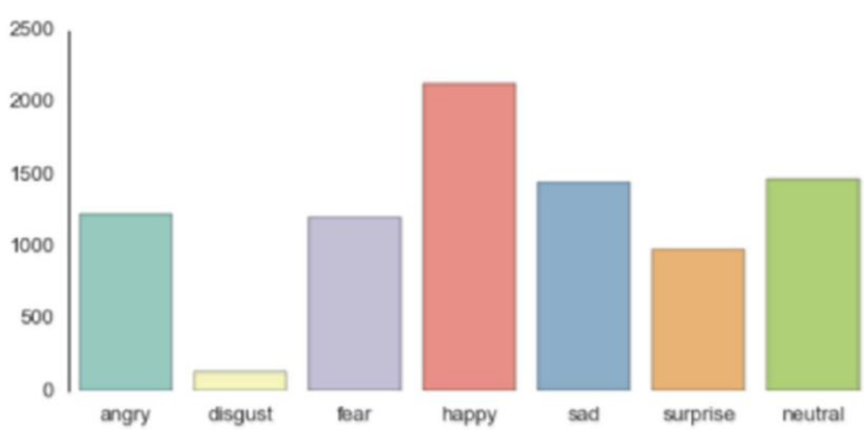
The data consists of 48x48 pixel gray scale images of faces.

The task is to categorize each face based on the emotion shown in the facial expression in to one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise).

The Kaggle dataset has another category 6=Neutral in case the face does not match any of the other emotions.

The training set consists of 28,709 examples. The training dataset contains two columns emo on and pixels. The emotion column contains a numeric code ranging from 0 to 6, inclusive, for the emo on that is present in the image.

The pixel column contains a string surrounded in quotes for each image. The contents of this string a space-separated pixel values in row major order. The test set consists of another 3,589 examples



Overview of Distribution of Categories in the dataset

Emotions	Training Set Count
Angry	4318
Disgust	113
Fear	4097
Happy	7215
Sad	4830
Surprise	3171
Neutral	4965

Happy has the maximum number of labels in the dataset whereas, disgust has a very low number (113).
Hence the dataset looks like a very unbalanced one.

IMAGE PRE-PROCESSING

1. All the images were subtracted from their individual mean.

```
for i in range(X_train.shape[0]):
mean = np.mean(X_train[i, :, :, :])
X_train[i, :, :, :] = (X_train[i, :, :, :] - mean)
for i in range(X_test.shape[0]):
mean = np.mean(X_test[i, :, :, :])
X_test[i, :, :, :] = (X_test[i, :, :, :] - mean
```

2. Normalization of ensure dataset with mean and standard deviation

```
if K.image_dim_ordering() == "th":
for i in range(1):
mean = np.mean(X[:, i, :, :])
std = np.std(X[:, i, :, :])
X_train[:, i, :, :] = (X_train[:, i, :, :] - mean) / std
X_test[:, i, :, :] = (X_test[:, i, :, :] - mean) / std
elif K.image_dim_ordering() == "tf":
for i in range(1):
mean = np.mean(X[:, i, :, :])
std = np.std(X[:, i, :, :])
X_train[:, i, :, :] = (X_train[:, i, :, :] - mean) / std
X_test[:, i, :, :] = (X_test[:, i, :, :] - mean) / std
```

3. Data Augmentation
a. Random Flipping of Images was achieved using ImageDataGenerator
b. 20% of zooming was done on images to focus on more on specific muscle movements on face during each expression. Ex- frown in forehead when angry, curving of lips when happy

etc.
datagen = ImageDataGenerator(zoom_range=0.2, horizontal_flip=True)
The previous model result during first report was built without any data pre-processing. The data preprocessing yielded a significant increase in validation and test accuracy in my model.

MODEL ARCHITECTURE- SEVERAL TRIALS

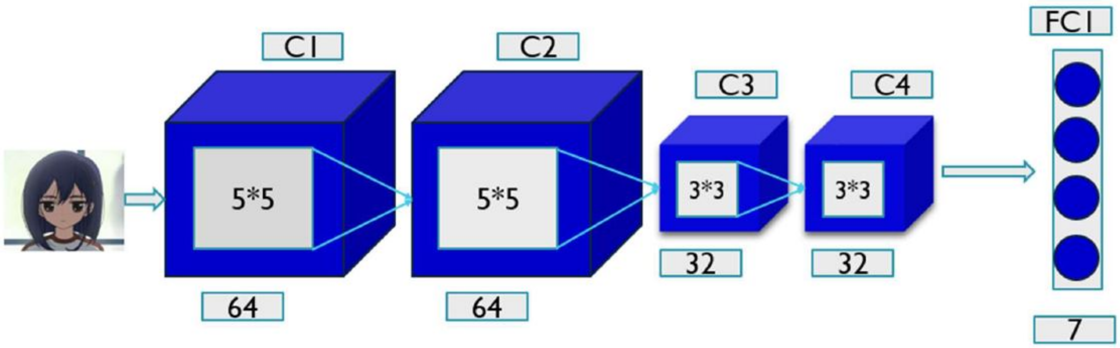
The trials were done various parameters. The following list the parameters for experiments:
1. SVM
2. CNNs with Number of FC layers – 1 & 2
3. Dropout- With 0.2 Dropout or Without Dropout
4. Learning Rate – Constant & Adaptive Scheduler

Please find below the results of 5 such trials and the accuracy from those models based on changing different parameters:

Architecture	Train Acc	Validation Acc	Test Acc	Findings
SVM	99.8%	N/A	48.9%	Very low in testing data (overfitting)
2 CNNs with 1 FC layer	99.6%	59%	59.6%	High training accuracy and low test accuracy indicates strong over-fitting in the model
3 CNNs with 1 FC layer	88.7%	63%	64%	The model with 3 layers definitely reduced overfitting & showed remarkable improvement in test accuracy (5%)
3 CNNs with 2 FC layers	77%	64%	65.4%	The model with one more FC layers was better than the previous one by 1%
4 CNNs with 1 FC layer	70%	65.5%	65.6%	Increasing one more CNN layer increased the model accuracy by 0.2%

FINAL MODEL SELECTION

The input images are of dimension 1*48*48. The model has been built by taking 4 CNN layers with 1 Fully Connected Layer. In all the CNN layers, activation function which has been used is ReLu. The initial two CNN layers have 64 filters of 5*5 dimension. Batch normalization has been used along with 2*2 strides in Max pooling with Zero Padding. The next two layers have 32 filters of 3*3 along with Max Pooling, Zero Padding and Batch Normalization. Additionally, Dropout=0.25 has been used in these two final layers of CNN.
In the fully connected layer, dropout of 0.2, activation function of soft max and cross entropy loss.



RESULTS

The validation and test accuracy on this model is 65% and 66% respectively. According to research study, 65% is human performance in recognizing emotions. This model has been tested and performs equally well as a human being.

On further deep diving into results, there are few insights on the model:

	0	1	2	3	4	5	6
0	282	4	52	19	65	8	61
1	16	20	8	2	5	2	2
2	75	2	201	26	103	63	58
3	19	0	19	760	33	14	34
4	53	0	53	46	314	6	122
5	10	0	47	22	8	318	11
6	44	0	31	44	79	10	418

0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral

Let's take a closer look at predictions for individual emotions.

The matrix gives the counts of emotion predictions and some insights to the performance of the multi-class classification model.

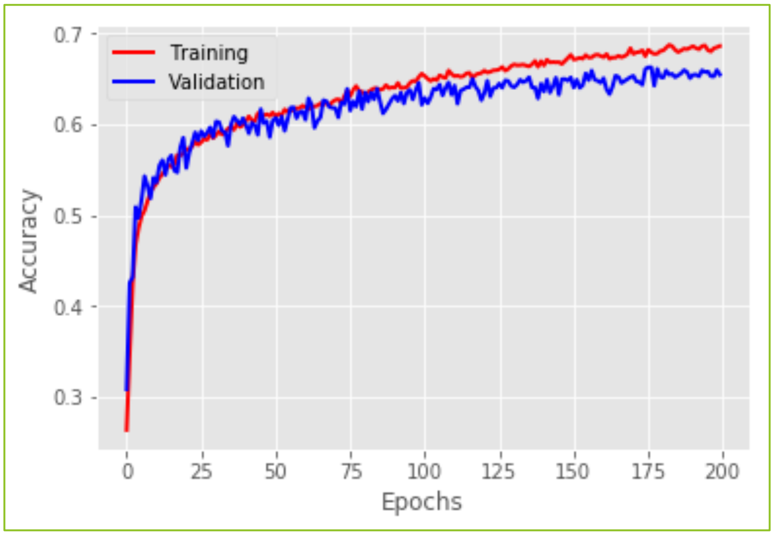
The model performs well on classifying positive emotions resulting in relatively high precision scores for happy and surprised. 760 happy emotions were correctly classified.

Happy has a precision of 76.7% which could be explained by having the most examples (~7000) in the training set.

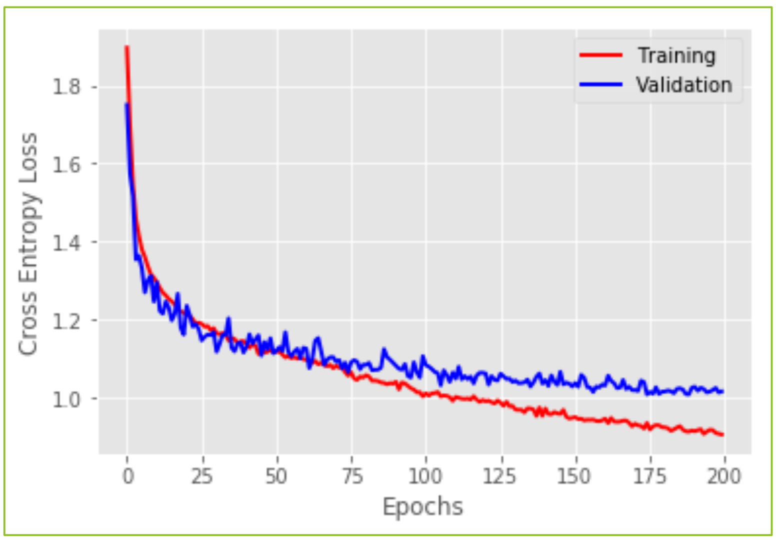
Interestingly, surprise also has high rate of accuracy with 318 faces being classified correctly.

Sad gets misclassified with the next closest negative emotion fear.

The model frequently misclassified angry, fear and neutral as sad. In addition, it is most confused when predicting sad and neutral faces because these two emotions are probably the least expressive (excluding crying faces).

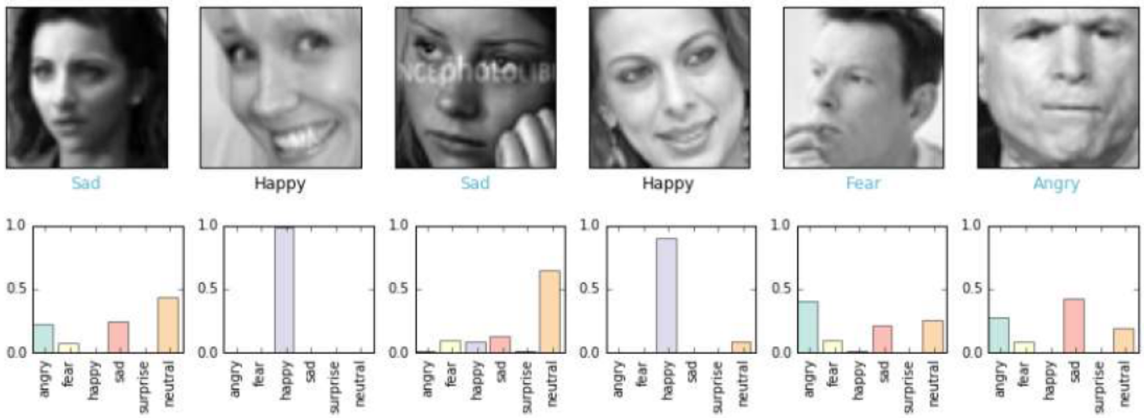


Training Accuracy and Validation improves with each epoch iteration.



Similar trend has been observed for cross entropy loss.

Example of some of the faces which got correctly predicted:



CONCLUSION

The final model accuracy could not reach the state of art accuracy achieved last year which was 75%. The ensemble deep nets would be have been the next future steps to improve the accuracy further. But the model accuracy had improved remarkably alter data augmentation and normalization as compared to first report results where no data pre-processing was done.

Although the accuracy is lower than 75% on closer inspection of mis-classified classes it was observed, emotion is getting labeled to next best emotion category.