Neural Network & Fuzzy Logic

Paper Presentation

Hand Gesture Recognition

Using an Adapted CNN with Data Augmentation

Hand Gesture Recognition Challenges

1. Complexity of hand gesture structures
2. Differences in hand sizes and color
3. Differences in hand postures
4. Variation of light and background
5. Variation in gesture
6. Observability of hand

Previous Method (Classical Vision Algorithm)

1. Extract Features
2. Apply Machine Learning Classifier

Deep Learning Algorithm Advantage

1. Ability to extract robust and significant features of input via several nonlinear hidden layers.
2. Merge multiple extracted features efficiently.
3. Prevent overfitting (eg:- dropout).

Convolutional Neural Network

Primarily used for image classification, Objects detections, recognition faces etc

1. 3 Layers
   a. Convolutional layer
   b. Sub sampling or pooling layer
   c. Fully connected layer
2. Convolutional layers are followed by pooling layers
   a. Pooling layers could be maximum or average function
   b. Reduce size of features and decrease compilation time
3. Features extracted are hierarchical
   a. Bottom level layers:- low level features
   b. High level layers:- Abstract information, features useful for classification

API's and Libraries used:

1. Keras API of Tensorflow
2. Sklearn
3. Matplotlib
4. Google.colab

Platform used for processing - Google Colab

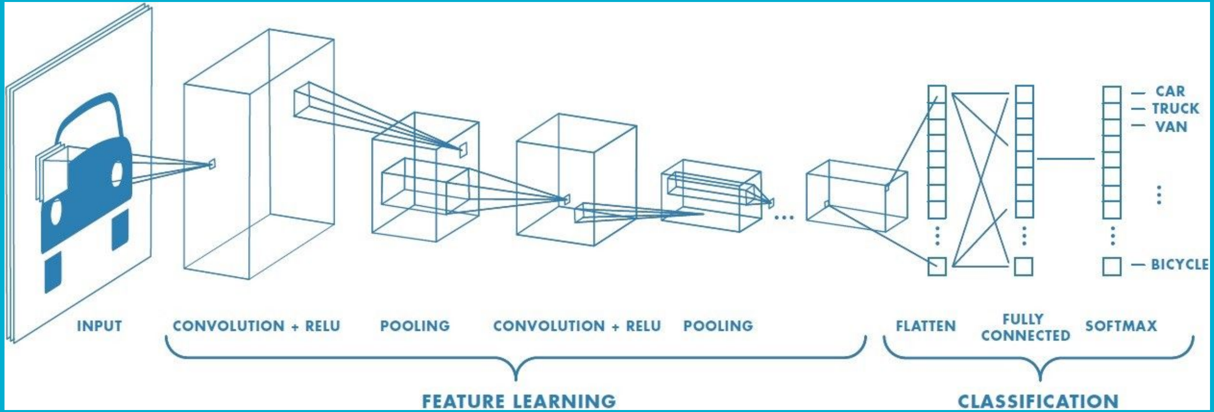
Dataset

1. Number of Classes = 29
    a. 26 alphabets (A - Z)
    b. Space, Del, Nothing
2. 1500 images for each class
3. Varied Illumination
4. Size = 200x200 initially, to be reduced to 128x96
5. Training set images= 39150; Validation set images= 4321
6. Validation ratio = 0.1 from training set
7. Testing set = New 300 images

Image Preprocessing:- For real time classification, images have been converted to grayscale

Baseline CNN Architecture

1. Two Convolutional Layers
2. Two Pooling Layer
3. Two Fully Connected Layers (ReLU)
4. Three Dropout Performance
    a. One dropout each after the two pooling layers
    b. One dropout after the first fully connected layer



Data Augmentation

Transforms base data and increases size of the dataset.

New images are generated from the original usually during training.

Goal is not only to reduce overfitting but also to augment data to improve the classifier.

Augmentations ->[Height and Width shift range = 0.1]

Proposed ADCNN

Performance of baseline CNN is improved by tuning parameters to include kernel initialization and regularization.

*__Data Augmentation__ is done on both Baseline CNN and ADCNN

Network Initialization

Initializing weight correctly can influence how easily the network learn.

1. Uniform ‘He Initialization’ for ReLU layers (**he_uniform**)
2. Uniform ‘Xavier Initialization’ for output softmax (**glorot_uniform**)

Regularization

We use Kernel regularizer L2 for ReLU layers.

Kernel regularizers tries to reduce the weights W (excluding bias).

L2 Regularization

Decreases complexity of model while maintaining the same parameter count.

Penalizes weights with large magnitudes by minimizing L2 norm

Lambda = 0.0001

Evaluation Measures

1. Precision = $\dfrac{TP}{TP + FP}$

2. Recall = $\dfrac{TP}{TP + FN}$

3. F1 Score = $\dfrac{2 * Precision * Recall}{Precision + Recall} \Rightarrow [0.\ 1]|$

4. Accuracy = $\dfrac{TP + TN}{TP + TN + FP + FN}$

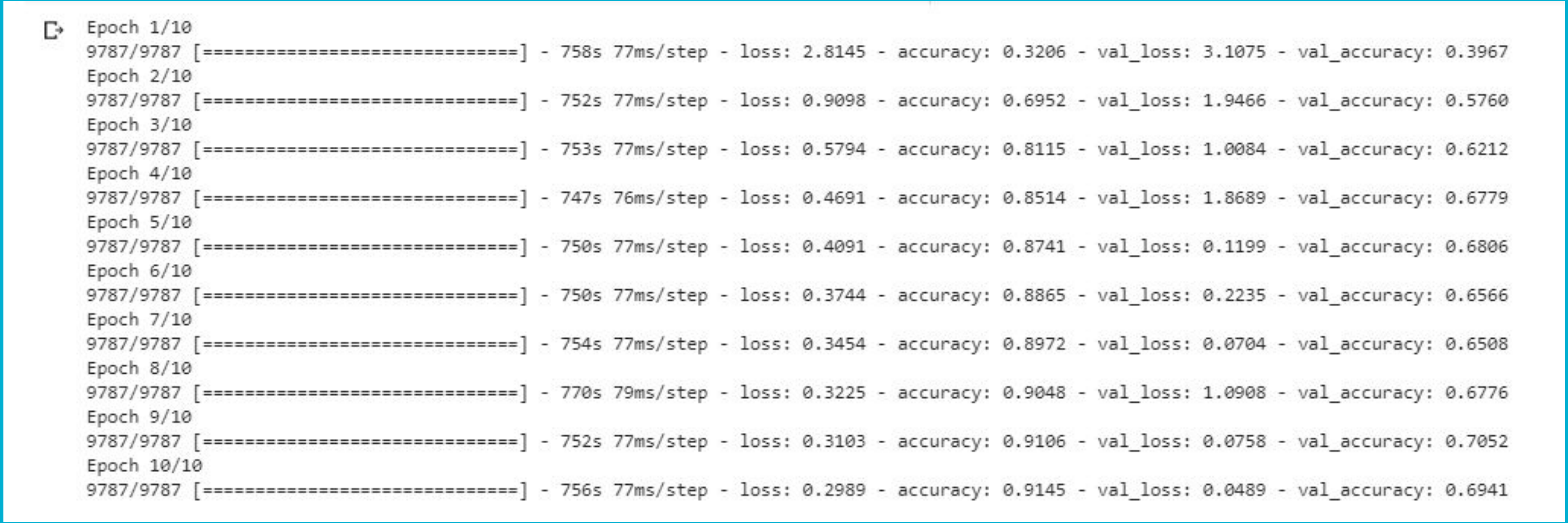
CNN Training Process

1. Combination of Back Propagation with Stochastic Gradient Descent.
2. Cost of function = Categorical Cross Entropy Loss Function
3. Adam Optimizer as optimization function
4. Number of Epochs = 10

For Baseline CNN

Results – For Baseline CNN

For Baseline CNN, Final Validation accuracy = 69.41 %

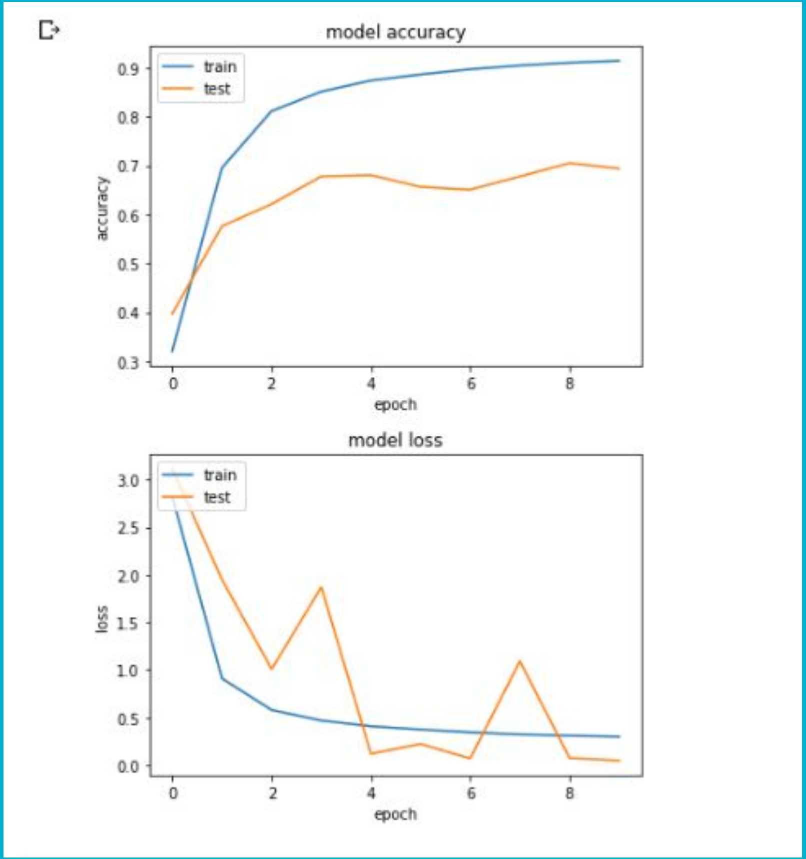


```
Epoch 1/10
9787/9787 [==============================] - 758s 77ms/step - loss: 2.8145 - accuracy: 0.3206 - val_loss: 3.1075 - val_accuracy: 0.3967
Epoch 2/10
9787/9787 [==============================] - 752s 77ms/step - loss: 0.9098 - accuracy: 0.6952 - val_loss: 1.9466 - val_accuracy: 0.5760
Epoch 3/10
9787/9787 [==============================] - 753s 77ms/step - loss: 0.5794 - accuracy: 0.8115 - val_loss: 1.0084 - val_accuracy: 0.6212
Epoch 4/10
9787/9787 [==============================] - 747s 76ms/step - loss: 0.4691 - accuracy: 0.8514 - val_loss: 1.8689 - val_accuracy: 0.6779
Epoch 5/10
9787/9787 [==============================] - 750s 77ms/step - loss: 0.4091 - accuracy: 0.8741 - val_loss: 0.1199 - val_accuracy: 0.6806
Epoch 6/10
9787/9787 [==============================] - 750s 77ms/step - loss: 0.3744 - accuracy: 0.8865 - val_loss: 0.2235 - val_accuracy: 0.6566
Epoch 7/10
9787/9787 [==============================] - 754s 77ms/step - loss: 0.3454 - accuracy: 0.8972 - val_loss: 0.0704 - val_accuracy: 0.6508
Epoch 8/10
9787/9787 [==============================] - 770s 79ms/step - loss: 0.3225 - accuracy: 0.9048 - val_loss: 1.0908 - val_accuracy: 0.6776
Epoch 9/10
9787/9787 [==============================] - 752s 77ms/step - loss: 0.3103 - accuracy: 0.9106 - val_loss: 0.0758 - val_accuracy: 0.7052
Epoch 10/10
9787/9787 [==============================] - 756s 77ms/step - loss: 0.2989 - accuracy: 0.9145 - val_loss: 0.0489 - val_accuracy: 0.6941
```

For Baseline CNN

Results – For Baseline CNN

Loss, accuracy v/s Epoch

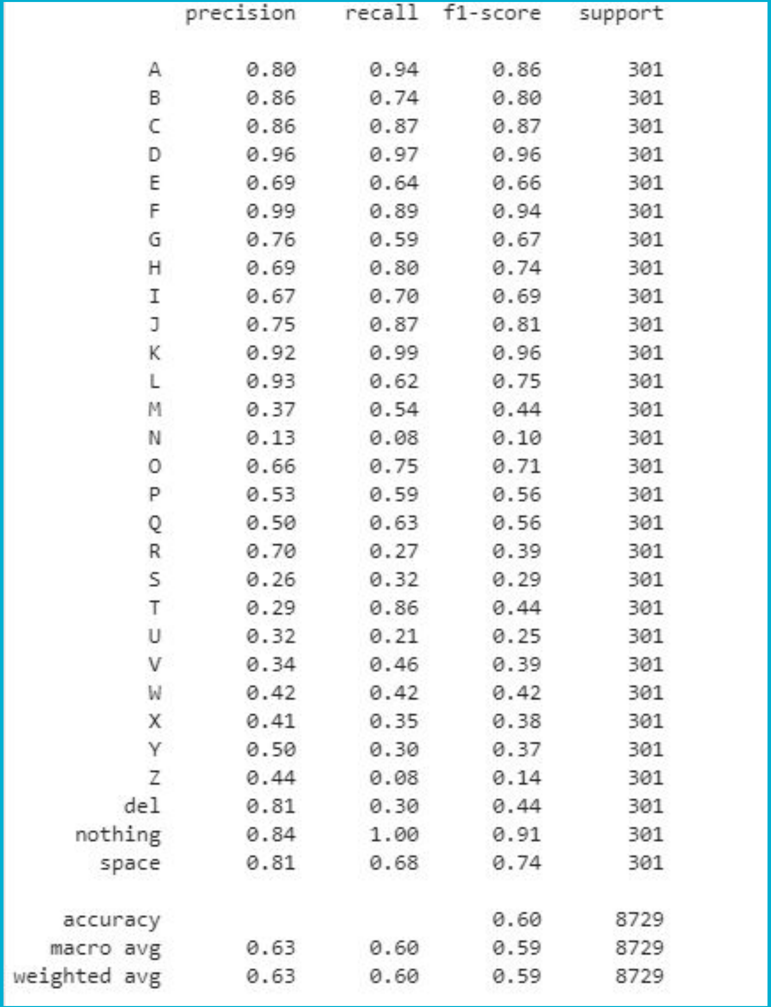


For Baseline CNN

Results -

Precision , recall and F-1 scores

With test accuracy of 60%

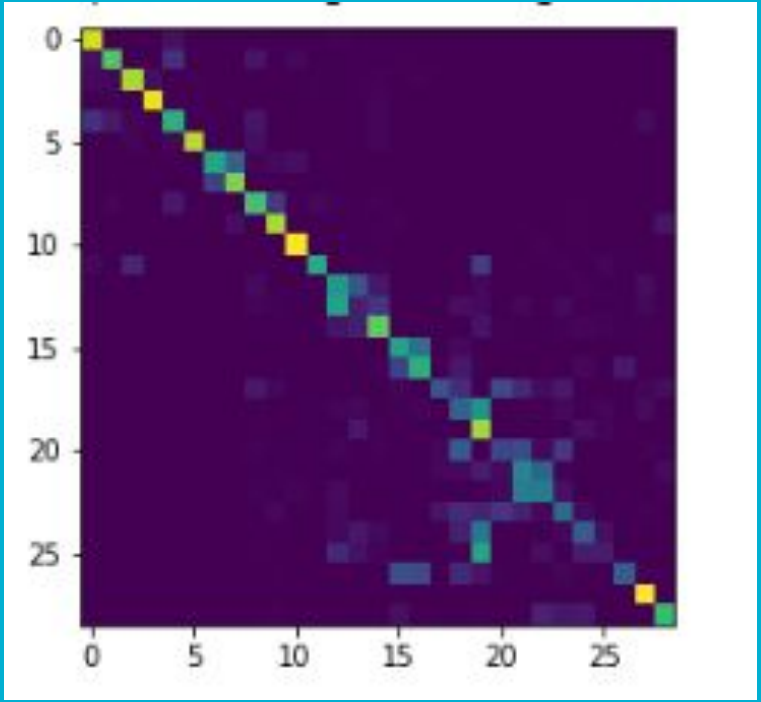


```
              precision    recall  f1-score   support

           A       0.80      0.94      0.86       301
           B       0.86      0.74      0.80       301
           C       0.86      0.87      0.87       301
           D       0.96      0.97      0.96       301
           E       0.69      0.64      0.66       301
           F       0.99      0.89      0.94       301
           G       0.76      0.59      0.67       301
           H       0.69      0.80      0.74       301
           I       0.67      0.70      0.69       301
           J       0.75      0.87      0.81       301
           K       0.92      0.99      0.96       301
           L       0.93      0.62      0.75       301
           M       0.37      0.54      0.44       301
           N       0.13      0.08      0.10       301
           O       0.66      0.75      0.71       301
           P       0.53      0.59      0.56       301
           Q       0.50      0.63      0.56       301
           R       0.70      0.27      0.39       301
           S       0.26      0.32      0.29       301
           T       0.29      0.86      0.44       301
           U       0.32      0.21      0.25       301
           V       0.34      0.46      0.39       301
           W       0.42      0.42      0.42       301
           X       0.41      0.35      0.38       301
           Y       0.50      0.30      0.37       301
           Z       0.44      0.08      0.14       301
         del       0.81      0.30      0.44       301
     nothing       0.84      1.00      0.91       301
       space       0.81      0.68      0.74       301

    accuracy                           0.60      8729
   macro avg       0.63      0.60      0.59      8729
weighted avg       0.63      0.60      0.59      8729
```

For Baseline CNN

Results – For Baseline CNN

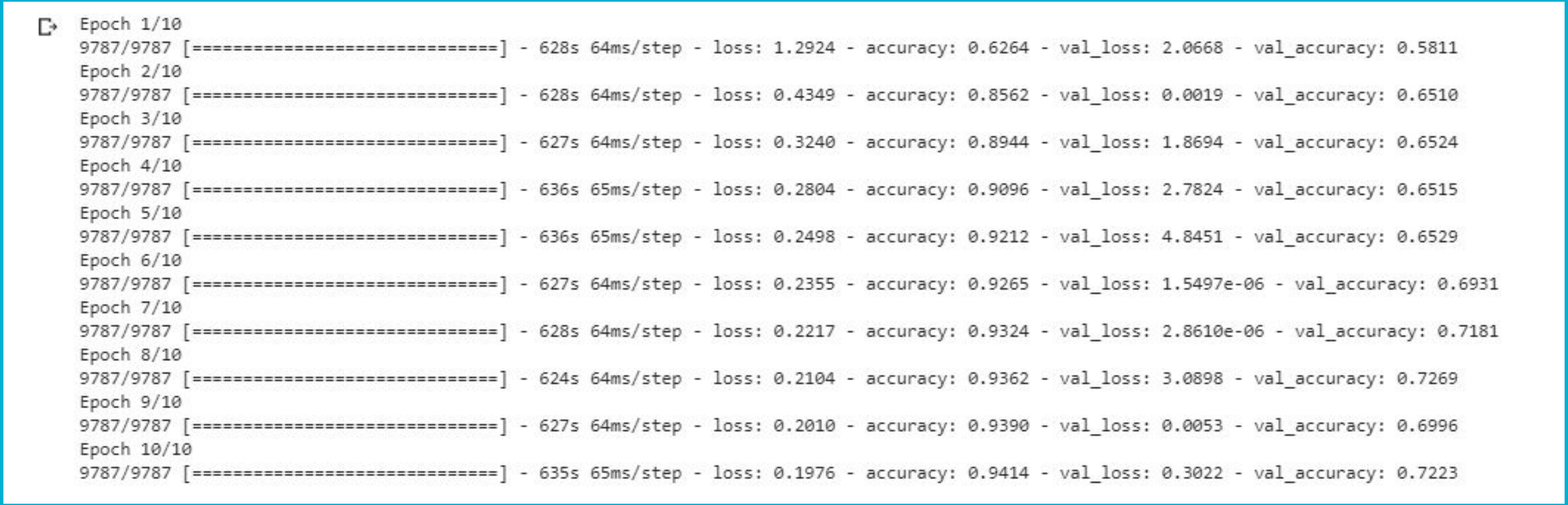
Confusion matrix



ADCNN

Results - ADCNN

For ADCNN, Final Validation accuracy = 72.23 %

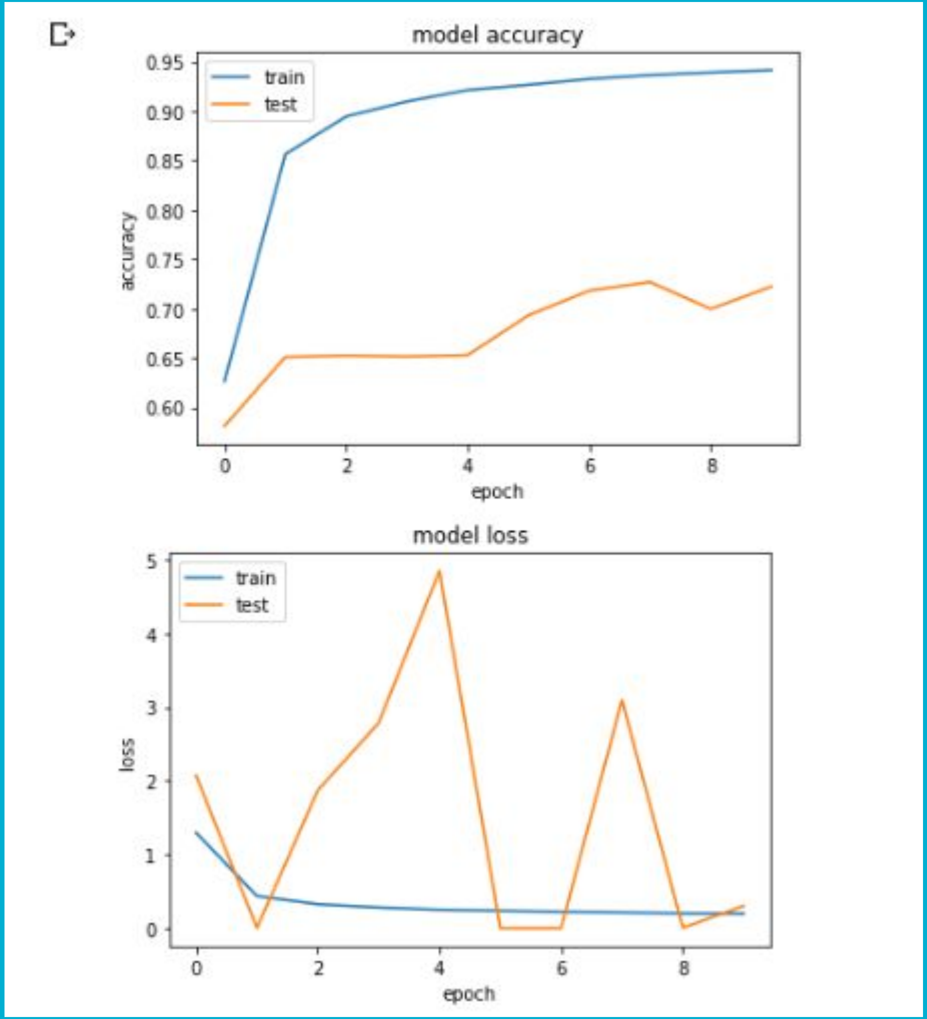


```
Epoch 1/10
9787/9787 [==============================] - 628s 64ms/step - loss: 1.2924 - accuracy: 0.6264 - val_loss: 2.0668 - val_accuracy: 0.5811
Epoch 2/10
9787/9787 [==============================] - 628s 64ms/step - loss: 0.4349 - accuracy: 0.8562 - val_loss: 0.0019 - val_accuracy: 0.6510
Epoch 3/10
9787/9787 [==============================] - 627s 64ms/step - loss: 0.3240 - accuracy: 0.8944 - val_loss: 1.8694 - val_accuracy: 0.6524
Epoch 4/10
9787/9787 [==============================] - 636s 65ms/step - loss: 0.2804 - accuracy: 0.9096 - val_loss: 2.7824 - val_accuracy: 0.6515
Epoch 5/10
9787/9787 [==============================] - 636s 65ms/step - loss: 0.2498 - accuracy: 0.9212 - val_loss: 4.8451 - val_accuracy: 0.6529
Epoch 6/10
9787/9787 [==============================] - 627s 64ms/step - loss: 0.2355 - accuracy: 0.9265 - val_loss: 1.5497e-06 - val_accuracy: 0.6931
Epoch 7/10
9787/9787 [==============================] - 628s 64ms/step - loss: 0.2217 - accuracy: 0.9324 - val_loss: 2.8610e-06 - val_accuracy: 0.7181
Epoch 8/10
9787/9787 [==============================] - 624s 64ms/step - loss: 0.2104 - accuracy: 0.9362 - val_loss: 3.0898 - val_accuracy: 0.7269
Epoch 9/10
9787/9787 [==============================] - 627s 64ms/step - loss: 0.2010 - accuracy: 0.9390 - val_loss: 0.0053 - val_accuracy: 0.6996
Epoch 10/10
9787/9787 [==============================] - 635s 65ms/step - loss: 0.1976 - accuracy: 0.9414 - val_loss: 0.3022 - val_accuracy: 0.7223
```

For ADCNN

Results - ADCNN

Loss, accuracy v/s Epoch

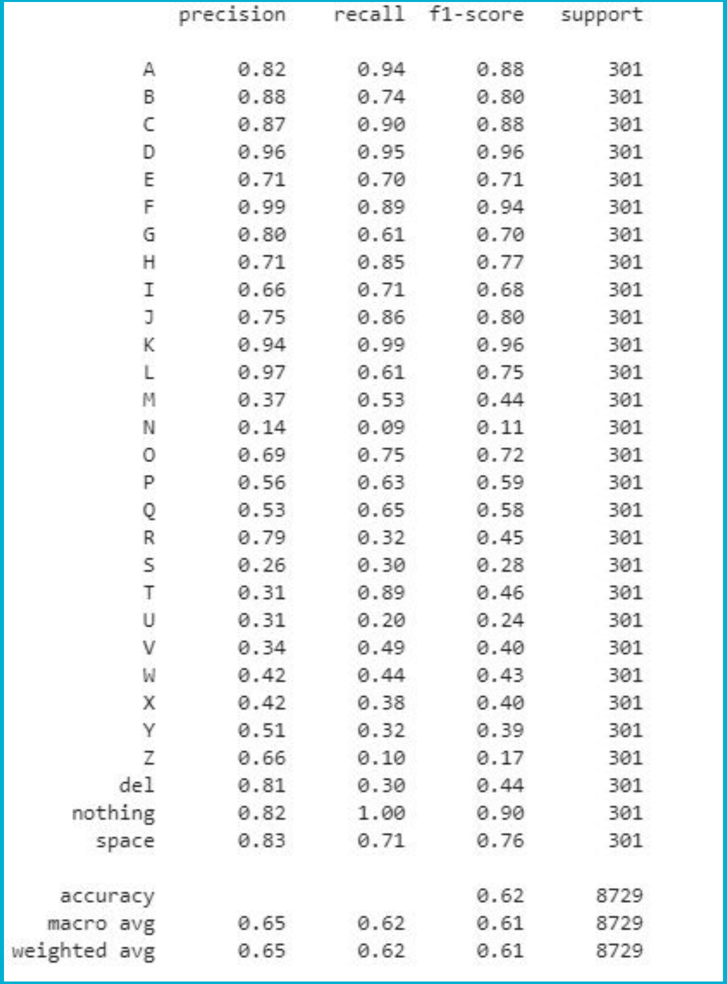


For ADCNN

Results –

Precision , recall and F-1 scores

With test accuracy of 62%

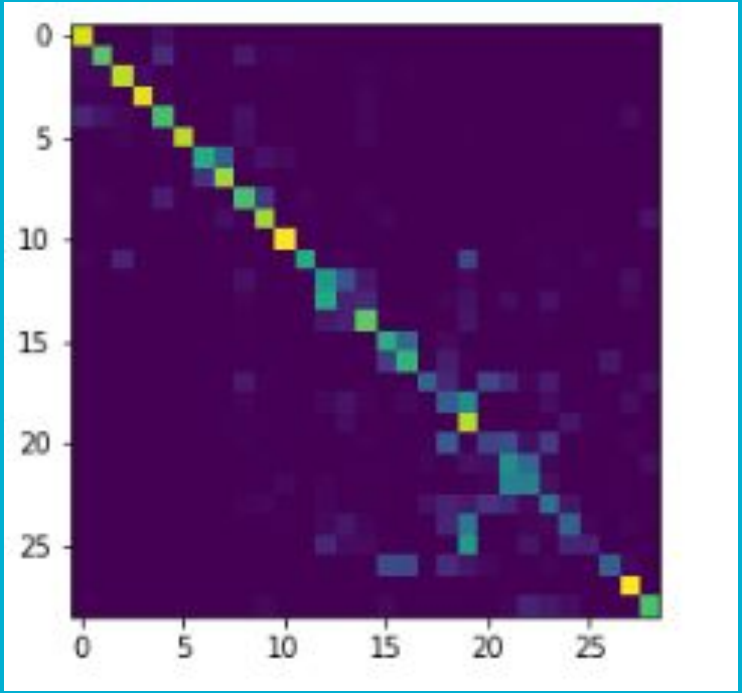


```
              precision    recall  f1-score   support

           A       0.82      0.94      0.88       301
           B       0.88      0.74      0.80       301
           C       0.87      0.90      0.88       301
           D       0.96      0.95      0.96       301
           E       0.71      0.70      0.71       301
           F       0.99      0.89      0.94       301
           G       0.80      0.61      0.70       301
           H       0.71      0.85      0.77       301
           I       0.66      0.71      0.68       301
           J       0.75      0.86      0.80       301
           K       0.94      0.99      0.96       301
           L       0.97      0.61      0.75       301
           M       0.37      0.53      0.44       301
           N       0.14      0.09      0.11       301
           O       0.69      0.75      0.72       301
           P       0.56      0.63      0.59       301
           Q       0.53      0.65      0.58       301
           R       0.79      0.32      0.45       301
           S       0.26      0.30      0.28       301
           T       0.31      0.89      0.46       301
           U       0.31      0.20      0.24       301
           V       0.34      0.49      0.40       301
           W       0.42      0.44      0.43       301
           X       0.42      0.38      0.40       301
           Y       0.51      0.32      0.39       301
           Z       0.66      0.10      0.17       301
         del       0.81      0.30      0.44       301
     nothing       0.82      1.00      0.90       301
       space       0.83      0.71      0.76       301

    accuracy                           0.62      8729
   macro avg       0.65      0.62      0.61      8729
weighted avg       0.65      0.62      0.61      8729
```

For ADCNN

Results - ADCNN

Confusion matrix


0
5
10
15
20
25
0
5
10
15
20
25