Mohamad Radwan

Part 4. Open-Ended Exploration:

Question: How do hyperparameters interact?

Since the dataset of part 3 is a smaller dataset we proceed to use that one since when training the function loops over the epochs faster.

To have a good comparison, all models are run with EPOCHS = 150.

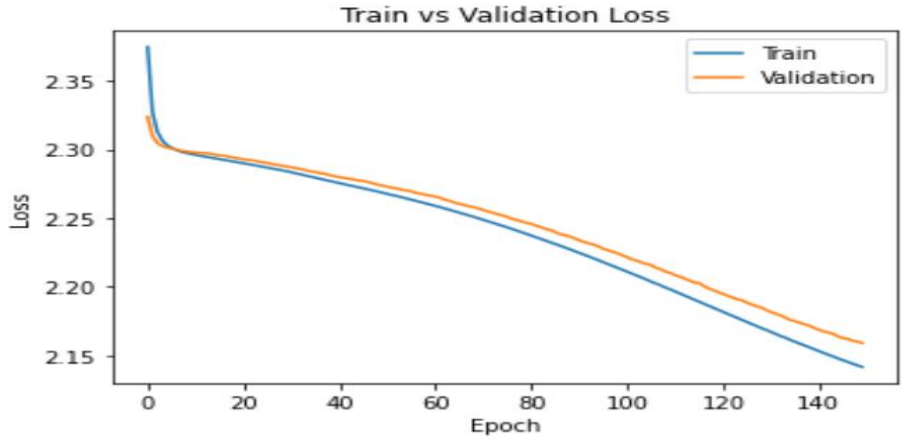
_Demonstrate for at-least two hyperparameters, where each independently increases the performance, and also increase the performance when used together(interpret "performance" as performance on the validation set):

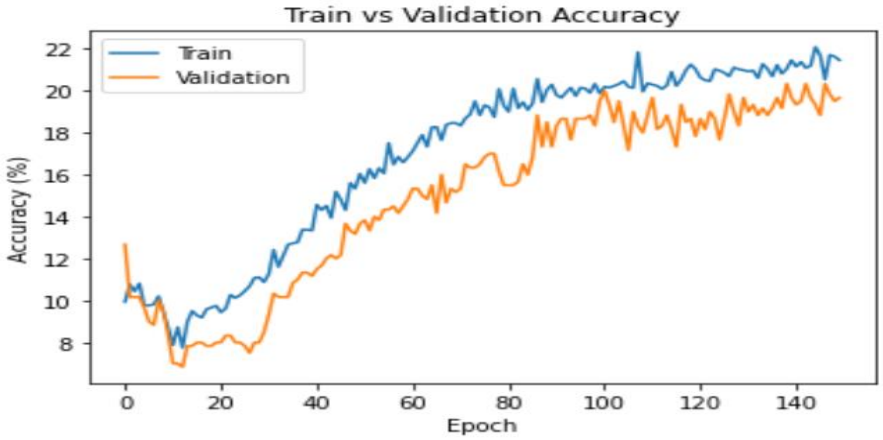
For this, we need a base model and 3 models to compare with.

We vary the hyperparameters BATCH and LEARNING_RATE, keeping Nb_Outputs_ConvolutionLayer1 and Nb_Convolution_Layers constant at 16 and 1 respectively.

Base Model: BATCH = 32, LEARNING_RATE = 5e-3

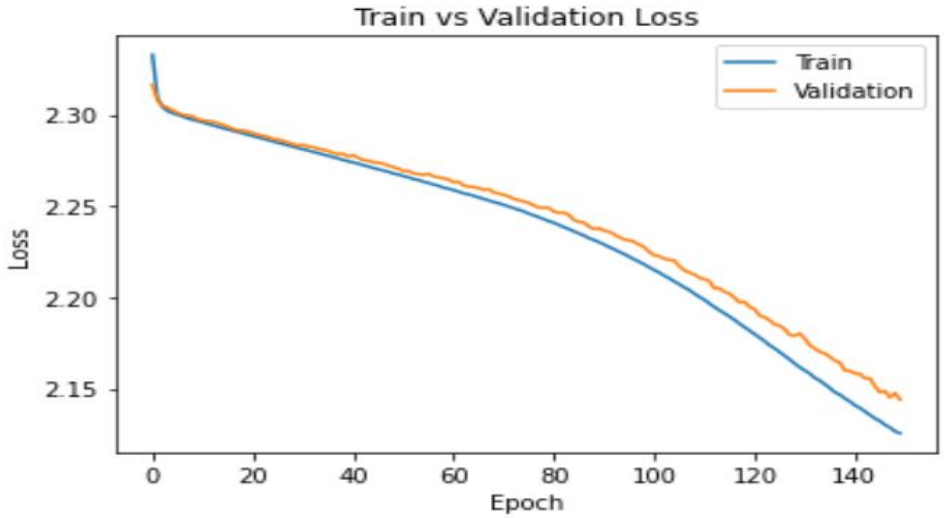
```
Base: Max Validation Accuracy:  20.333 % at epoch  139
```

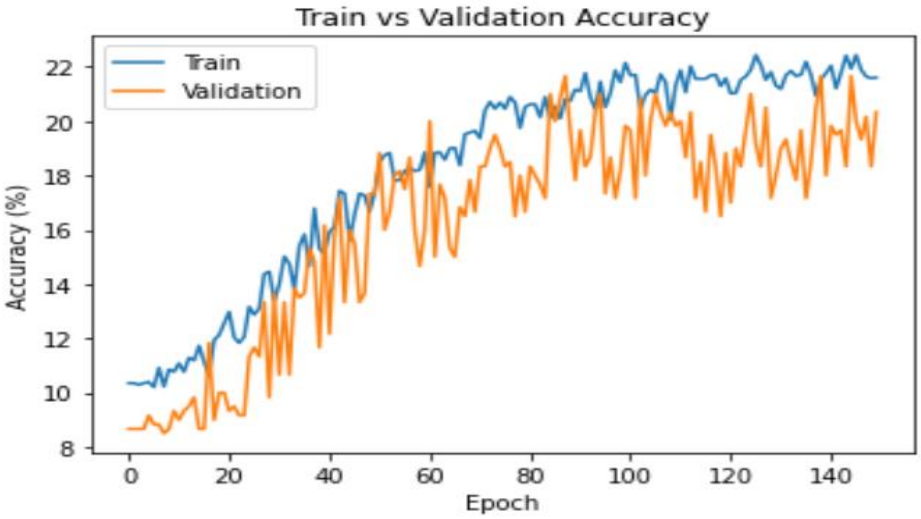




M1: BATCH = 16, LEARNING_RATE = 5e-3

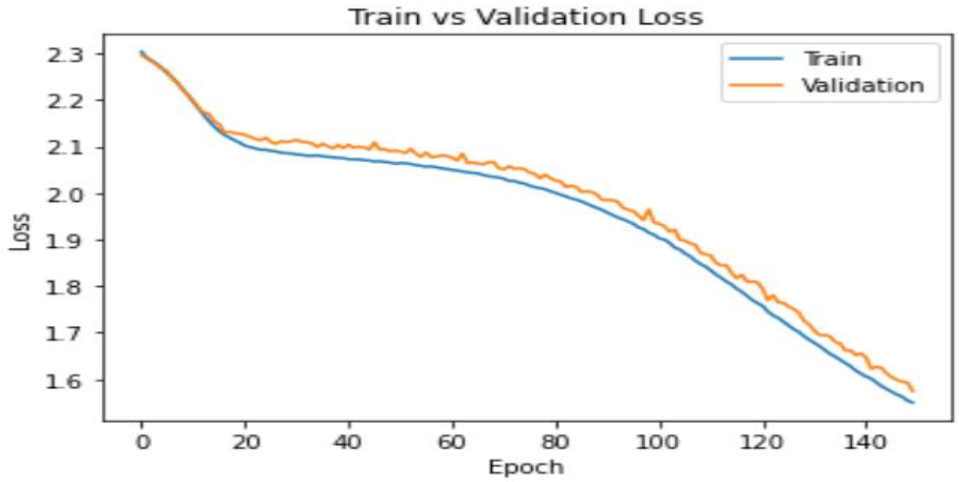
```
M1: Max Validation Accuracy:  21.667 % at epoch  88
```

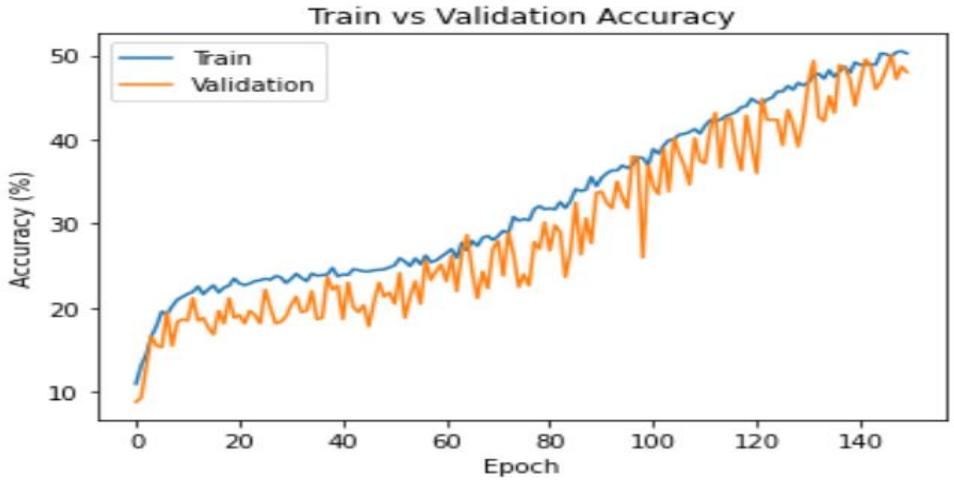




M2: BATCH = 32, LEARNING_RATE = 5e-2

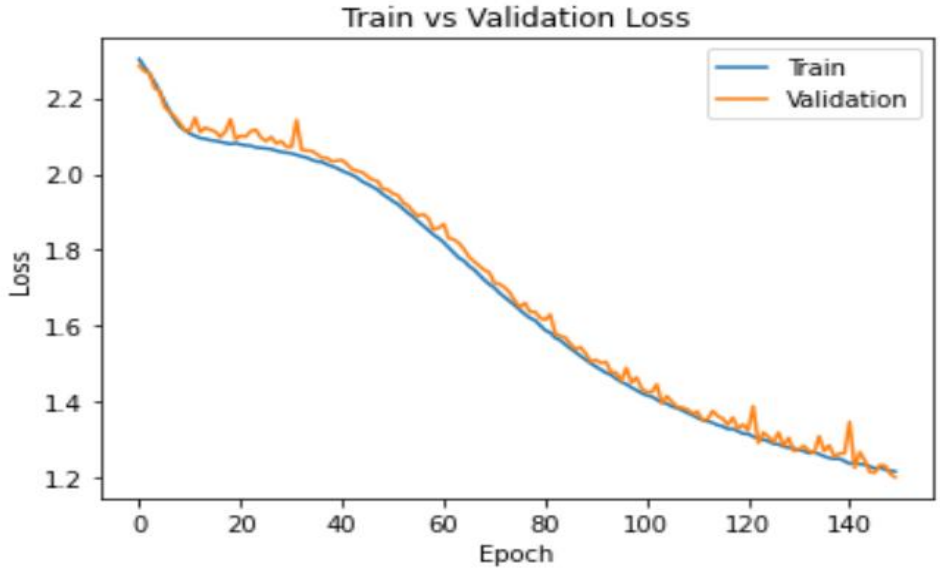
```
M2: Max Validation Accuracy:  50.0 % at epoch  147
```

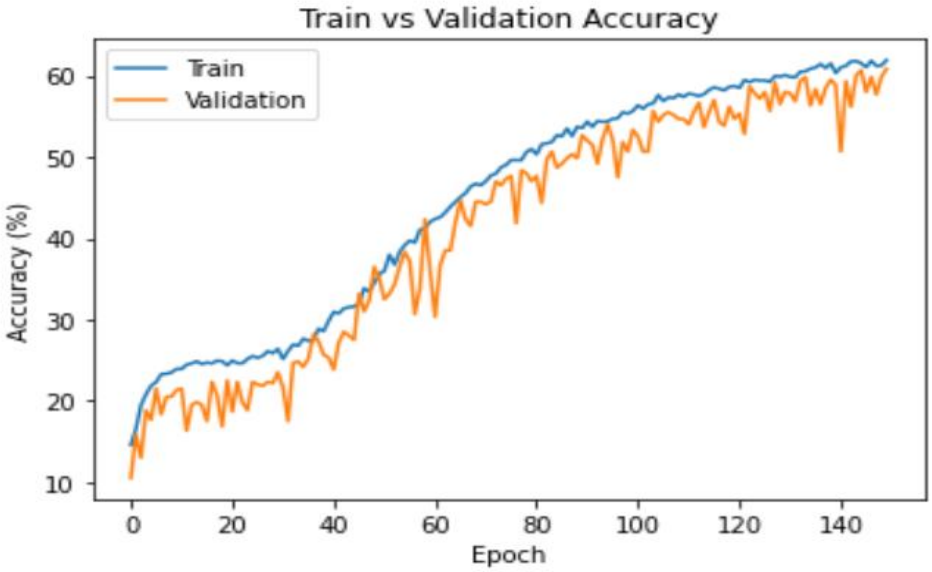




M3: BATCH = 16, LEARNING_RATE = 5e-2

```
M3: Max Validation Accuracy:  60.833 % at epoch  150
```





Summarizing important results:

_Base – M1 Comparison: Decreasing the batch size from 32 to 16 keeping the learning rate at 5e-3 increased the maximum validation accuracy obtained by a small amount.

_Base – M2 Comparison: Increasing the learning rate to 5e-2 keeping the batch size at 32 increased the maximum validation accuracy to 50.0 % (compared to the base model which had 20.333%)

_Base – M3 Comparison: Decreasing the batch size to 16 and increasing the learning rate to 5e-2 increased the maximum validation accuracy to 60.833 %

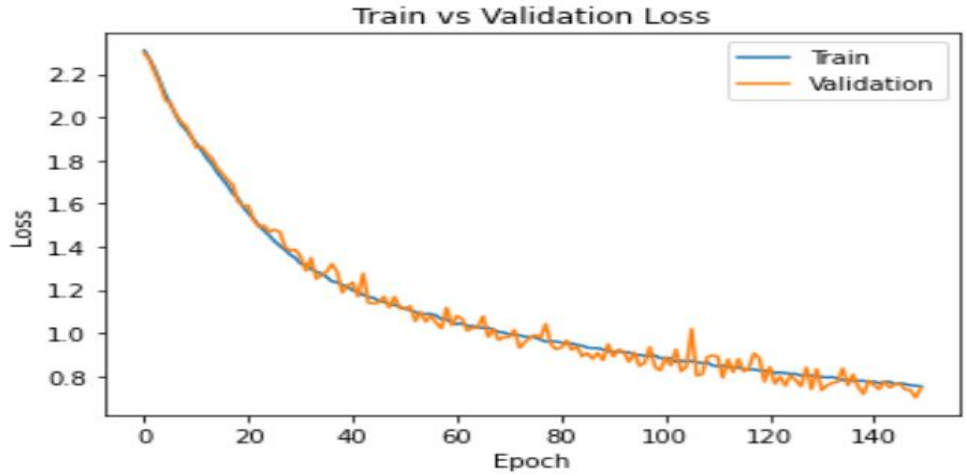
_Demonstrate for at-least two hyperparameters, where each independently increases the performance, but does not increase the performance or produce worse performance when used together.

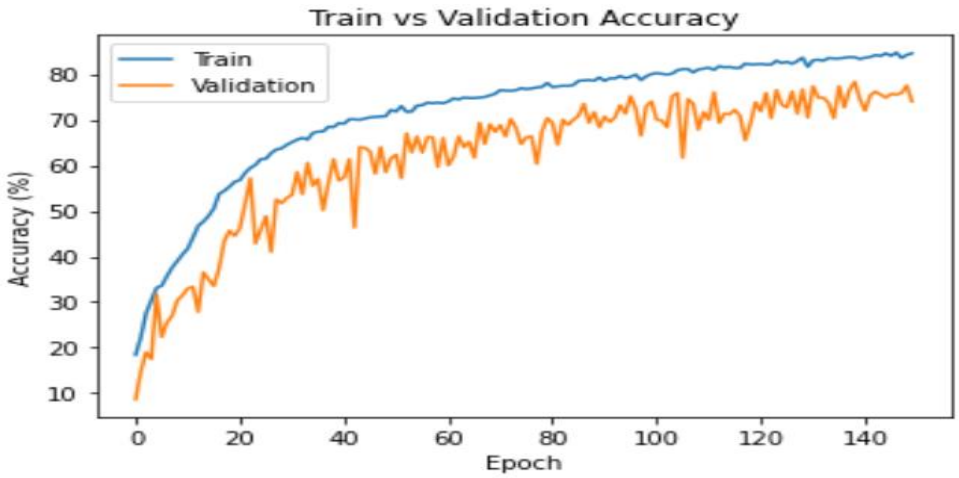
For this, we need a base model and 3 models to compare with.

We vary the hyperparameters Nb_Outputs_ConvolutionLayer1 and Nb_Convolution_Layers, keeping BATCH and LEARNING_RATE constant at 16 and 5e-2 respectively. (These values of BATCH and LEARNING_RATE got the highest max validation accuracy across the 4 previous models)

Base Model: Nb_Outputs_ConvolutionLayer1 = 128, Nb_Convolution_Layers = 2

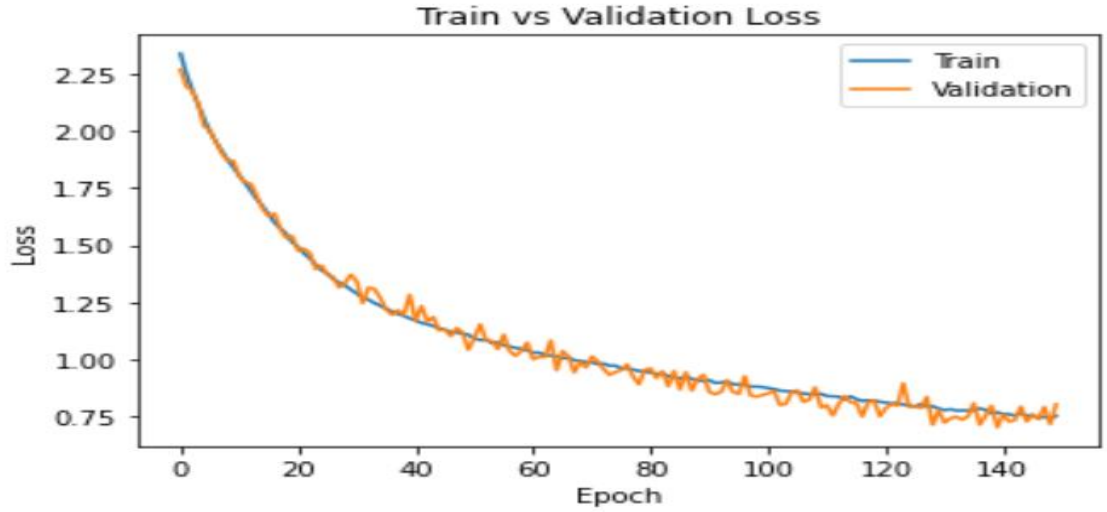
```
Base2: Max Validation Accuracy:  78.333 % at epoch  139
```

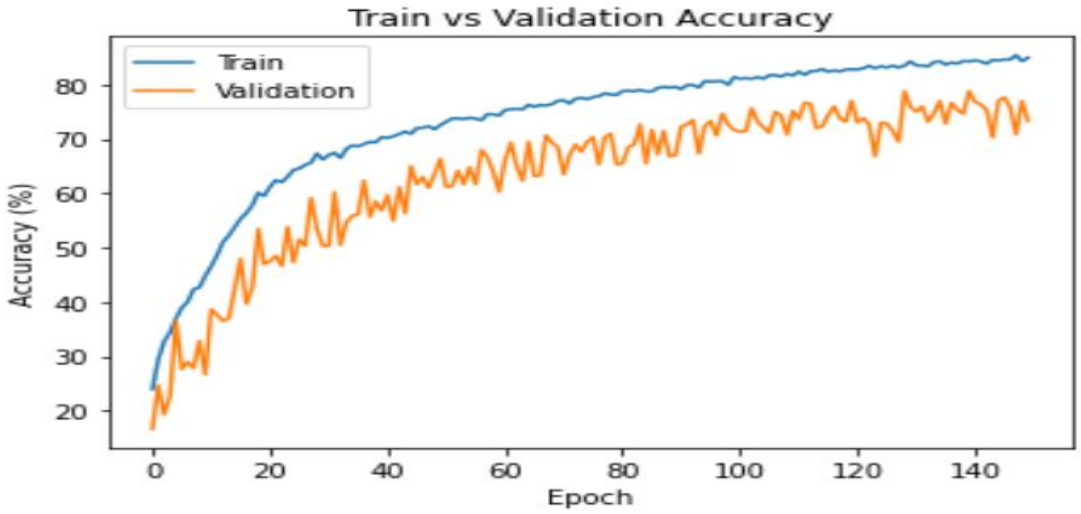




M4: Nb_Outputs_ConvolutionLayer1 = 256, Nb_Convolution_Layers = 2

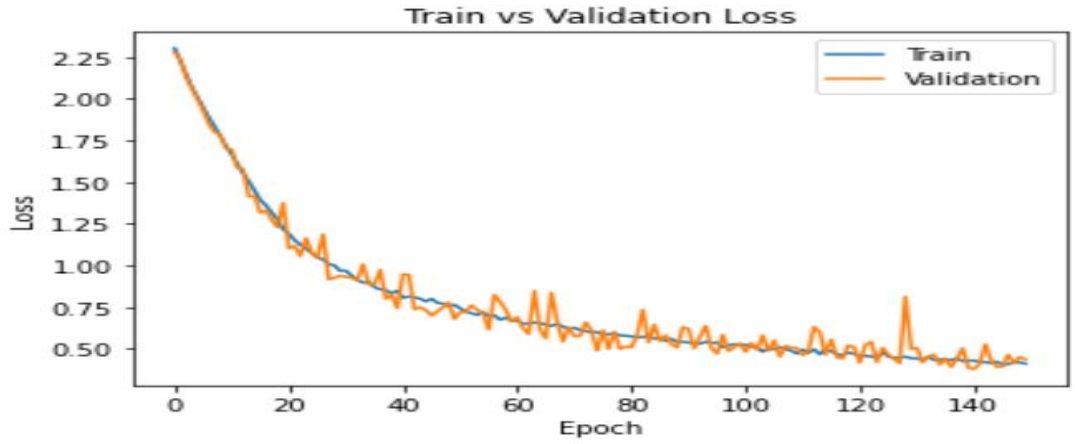
```
M4: Max Validation Accuracy:  78.833 % at epoch  129
```

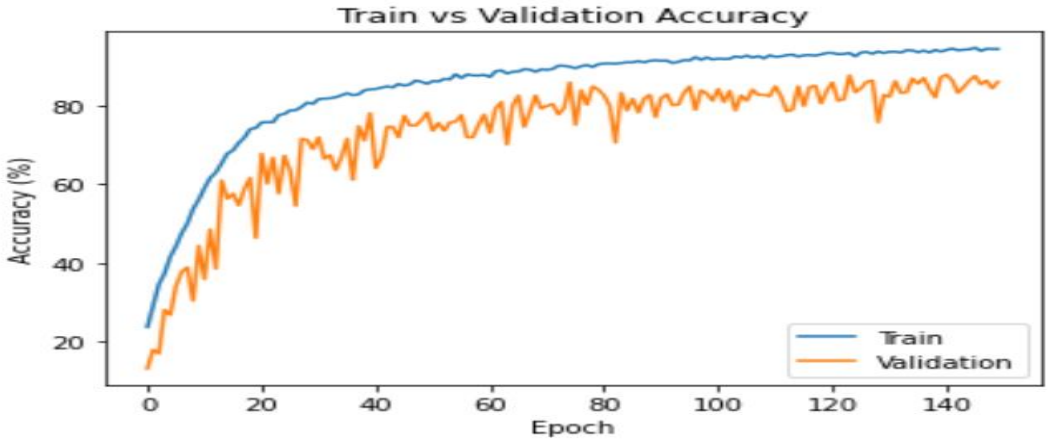




M5: Nb_Outputs_ConvolutionLayer1 = 128, Nb_Convolution_Layers = 3

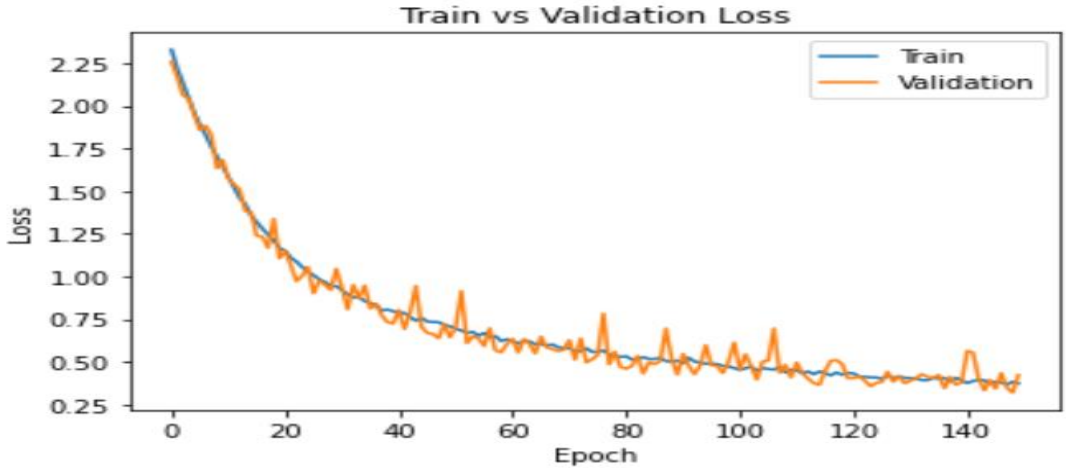
```
M5: Max Validation Accuracy:  88.0 % at epoch  141
```

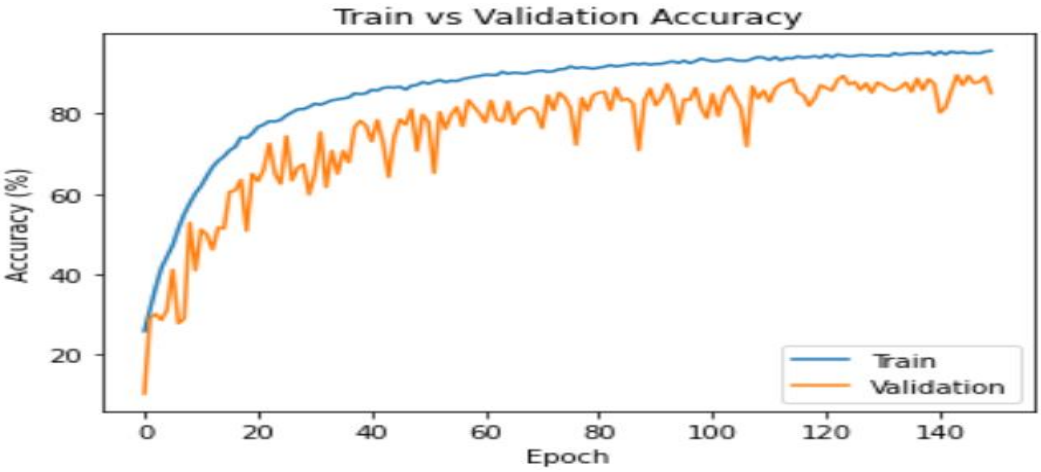




M6: Nb_Outputs_ConvolutionLayer1 = 256, Nb_Convolution_Layers = 3

```
M6: Max Validation Accuracy:  89.5 % at epoch  144
```





Summarizing important results:

_Base – M4 Comparison: Increasing the number of outputs of convolution layer 1 from 128 to 256 keeping the number of convolution layers at 2 increased the maximum validation accuracy obtained by a small amount.

_Base – M5 Comparison: Increasing the number of convolution layers to 3 keeping the number of outputs of convolution layer 1 at 128 increased the maximum validation accuracy to 88.0 % (compared to the base model which had 78.333%)

_Base – M6 Comparison: Increasing the number of outputs of convolution layer 1 to 256 and increasing the number of convolution layers to 3 increased the maximum validation accuracy to 89.5 %

_M5 – M6 Comparison: Increasing the number of outputs of convolution layer 1 from 128 to 256 keeping the number of convolution layers at 3 increased the maximum validation accuracy obtained by a negligible amount so we can consider that increasing the number of outputs of convolution layer 1 from 128 to 256 did not have much effect on the maximum validation accuracy obtained (as

compared to the Base – M5 comparison where the maximum validation accuracy was increased by about 10% just through increasing the number of convolution layers from 2 to 3.

Final Model: M6 is chosen as the final model having the highest max validation accuracy (89.5 %) across all 8 models. M6 model hyperparameters were arrived at by tuning the hyperparameters one at a time as demonstrated in the aforementioned tests. The testing accuracy achieved with this model is 82.667%.