Table 8: Results of 5-layer ConvNet

	Val	Test
Tag Classification	0.47	0.4675
Time Bucket Classification	0.3425	0.35

Table 9: Results of 5-layer ConvNet on adaptive histogram equalized dataset

	Val	Test
Tag Classification	0.49	0.465
Time Bucket Classification	0.3425	0.35

5.3. 5-Layer ConvNet

We trained the 5-layer ConvNet for 100,000 iterations with a minibatch size of 20. The ConvNet performed equally well on the normal and adaptive histogram equalized datasets. Table 8 shows results of 5-layer ConvNet on the normal dataset and Table 9 shows its performance on adaptive histogram equalized dataset. It performed better than the 3-layer ConvNet with tags as expected. Comparing its performance on the two datasets, it does equally well on both the datasets.

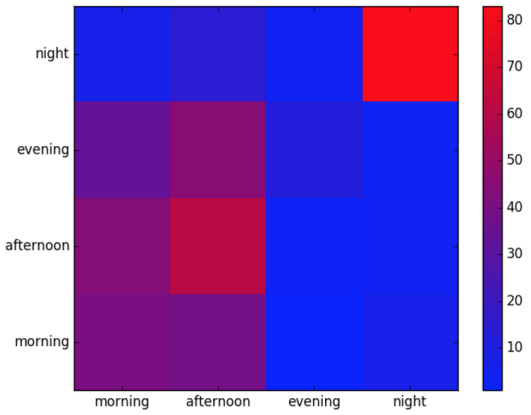

Figure 6: Confusion matrix for 5-layer ConvNet on tag classification with adaptive histogram equalized dataset

Figure 6 the confusion matrix for the 5-layer convnet. It classifies night correctly almost all the time, does well on afternoon, average on morning, and subpar on evening. It confuses afternoon for morning and vice versa

5.4. AlexNet

AlexNet was the most fine-tuned and trained model from the bunch, and is the only model to score test and train accuracies above 50. The learning rate was 0.0012, the batch size was 20 and the training iterations were 20,000. Table 10 shows the results of AlexNet on the tag classification task. AlexNet did not benefit from the adaptive histogram equalization as the accuracies were very close to human accuracy as shown in Table 11.

Table 10: Results of AlexNet

	Val	Test
Tag Classification	0.55	0.6

Table 11: Results of AlexNet on adaptive histogram equalized dataset

	Val	Test
Tag Classification	0.4425	0.42
Time Bucket Classification	0.4	0.42

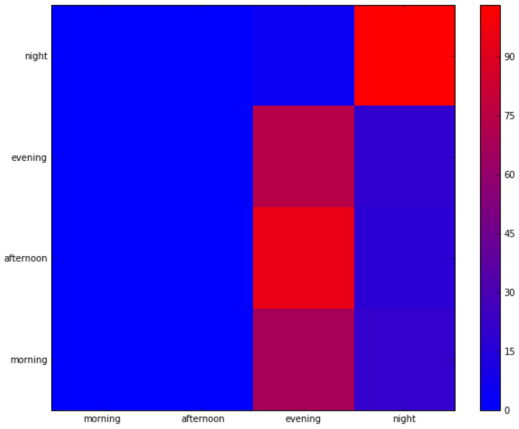

Figure 7: Confusion matrix of AlexNet

Figure 7 the confusion matrix for the AlexNet. Like the 5-layer net, the AlexNet does very well for night. It does well for evening, but does subpar for morning and afternoon. It tends to confuse evening as afternoon and morning.

5.5. VGGNet

VGGnet was the weakest performer of all the models, with accuracies as 0.35 for training set, 0.3 for validation set and 0.33 for test set. This is likely due to the VGG being a more suitable architecture for complex and feature sensitive tasks. In addition, it was extremely tricky to tune the VGG, as the range between under and overfitting was very narrow, so our results may not have reached an optimal level. Table 12 shows the results for VGGNet on adaptive histogram dataset.

5.6. Discussion

All the models perform equal or better as compared to the human accuracy. During the project, we came across