Table 12: Results of VGGNet

	Train	Val	Test
Tag Classification	0.35	0.3	0.33

several noticable insights which can effectively explain the difficulty of the task and the behavior of the models.

5.6.1 Errors in EXIF data

Since the data was collected from Flickr, which is a public image portal, it contained some disparities with the provided and most likely "photo taken" time.

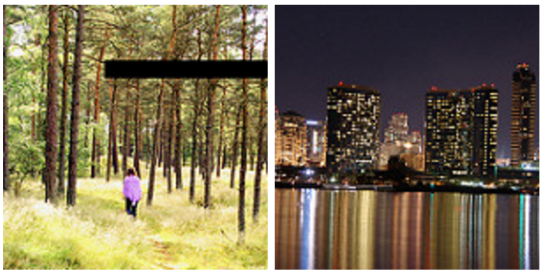

Figure 8: Example images with EXIF errors

In Figure 8, the left image seems to be taken during the day, but the provided time was 00:00:00, which is midnight. The right image seems to have been taken at night but the provided time was 11:37:19, which is around noon time. The cause of these errors cannot be properly deduced , as they may vary from photo to photo. One of the possibilities is that the time obtained from the device could be in a different time zone or may have a miscalibrated system time. The user could have as well misreported or edited the time taken at upload time or afterwards. Furthermore, such errors cannot be manually corrected, even on a small scale, as we don't know the actual time when the image was taken.

5.6.2 Ambiguity in Categories

This task is really difficult even for humans due to the ambiguity between categories. For example, it is very easy to confuse morning and evening images as both have a very similar sky color palette, brightness and contrast due to how similar sunrises and sundowns are. Figure 9 show the ambiguity between images from morning and evening. Furthermore, some of the photographs such as in Figure 10 were verging on the border of two categories, such as evening and night or morning and afternoon. Since the convolutional neural network needs to pick one, it becomes a significantly hard choice to make.

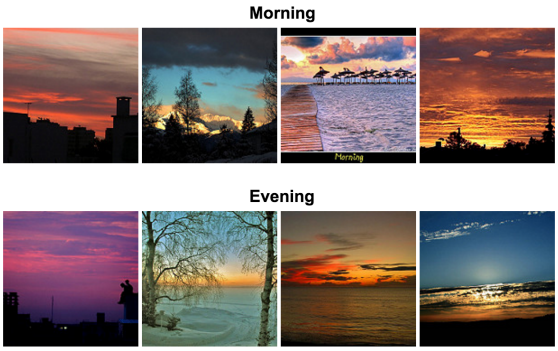

Figure 9: Example images showing ambiguity in categories

Figure 10: Images on the border of evening and night

5.6.3 Edited photographs

Figure 11: Examples of edited photographs in the dataset

Many of the photographs on Flickr are preprocessed in some way for aesthetic purposes, which causes a loss of original information and a distortion between perceived and actual time or category. This distortion in data makes the task of predicting the correct time trickier and might induce a bias towards a particular bucket or label. The images in Figure 11 were taken in the evening but the photographers seem to have toned down the brightness of the image, causing the network to misclassify the images as having been taken at night.