Analysis of Cat-Dog Classification Models

Tugay Talha İçen

October, 2023

0

1 Introduction

In this report, I present the performance metrics of several VGG model architectures for a cat-dog classification task, including the effect of Image Data Augmentation (IDA) and different learning rates. Each model is assessed based on its training time and classification accuracy. The models vary based on the number of blocks they contain, ranging from one to seven, and whether they include dropout layers or IDA.

2 Models and Training

The models consist of convolutional neural network architectures based on VGG, with blocks varying from one to seven. Each block contains a 2D convolution layer followed by a max pooling layer. The dropout models additionally have dropout layers included after each max pooling operation for regularization. Some models have also incorporated IDA. Finally, I also varied the learning rate in some models to observe its effect on performance and learning time.

3 Results

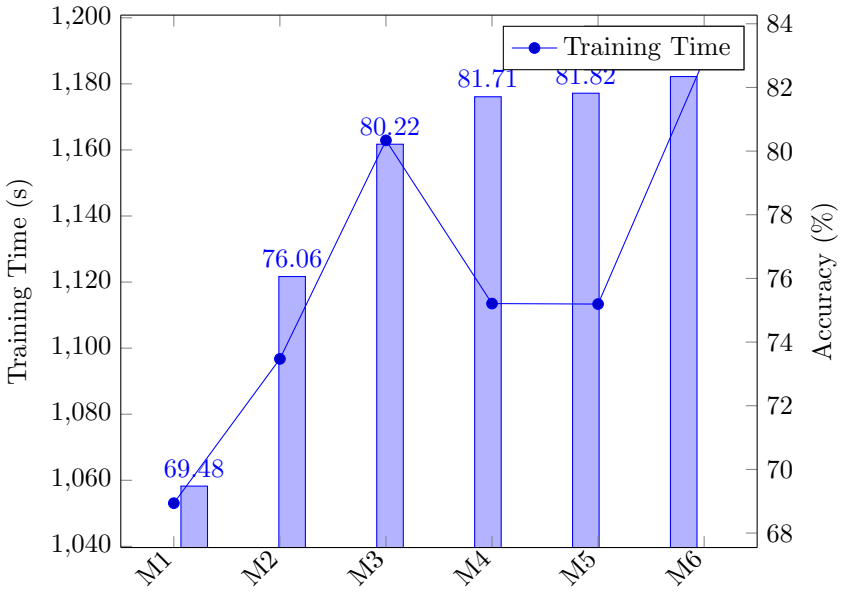
M1: 1 blocks & 4 layers VGG model **M2**: 2 block & 5 layers VGG model
M3: 3 block & 6 layers VGG model **M4**: 4 block & 7 layers VGG model
M5: 5 block & 8 layers VGG model **M6**: 7 block & 10 layers VGG model
Dropout 1: M1 model with dropout regularization.
Dropout 2: M2 model with dropout regularization.
Dropout 3: M3 model with dropout regularization.
Dropout 4: M4 model with dropout regularization.
Dropout 5: M5 model with dropout regularization.
Dropout 6: M6 model with dropout regularization.
IDA Model: M3 model with Image Data Augmentation.

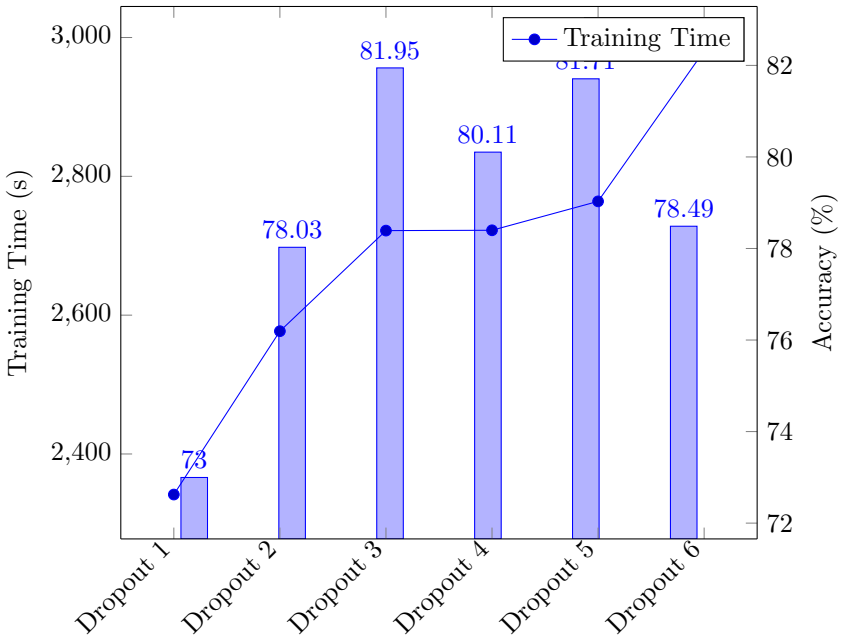
Dropout IDA Model: Dropout 3 model with Image Data Augmentation.
Dropout (LR=X): Dropout 3 model with learning rate X.
Batch Size: X (X= Batch size): M6 model with batch size X.
20 epochs are used for normal models, 50 epochs are used for dropout models and 100 epochs are used for IDA models and learning rate test models.

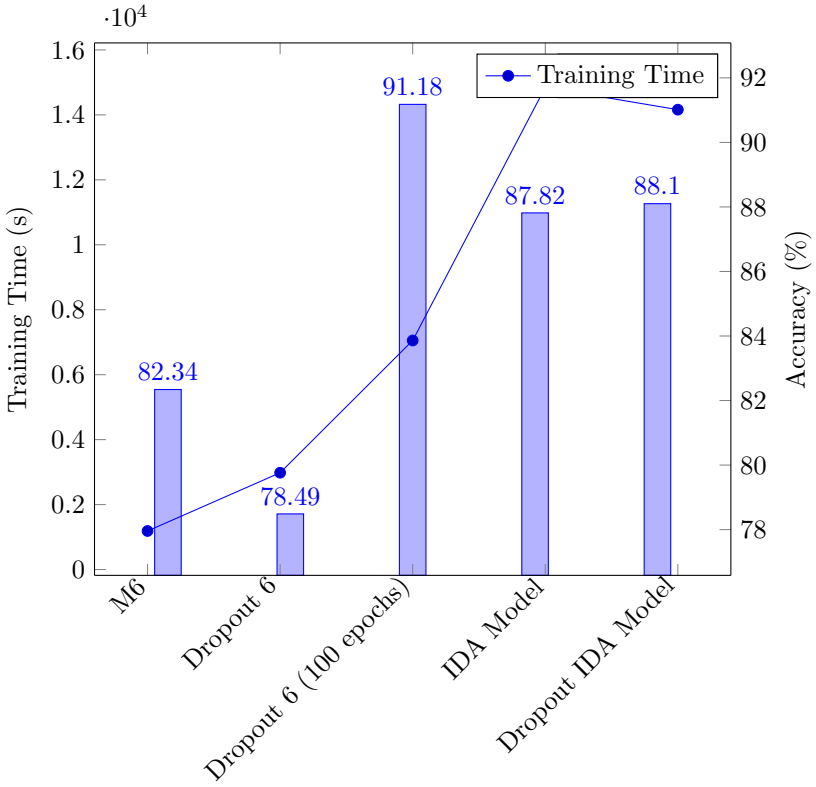
The following table lists the training time and accuracy for each model:

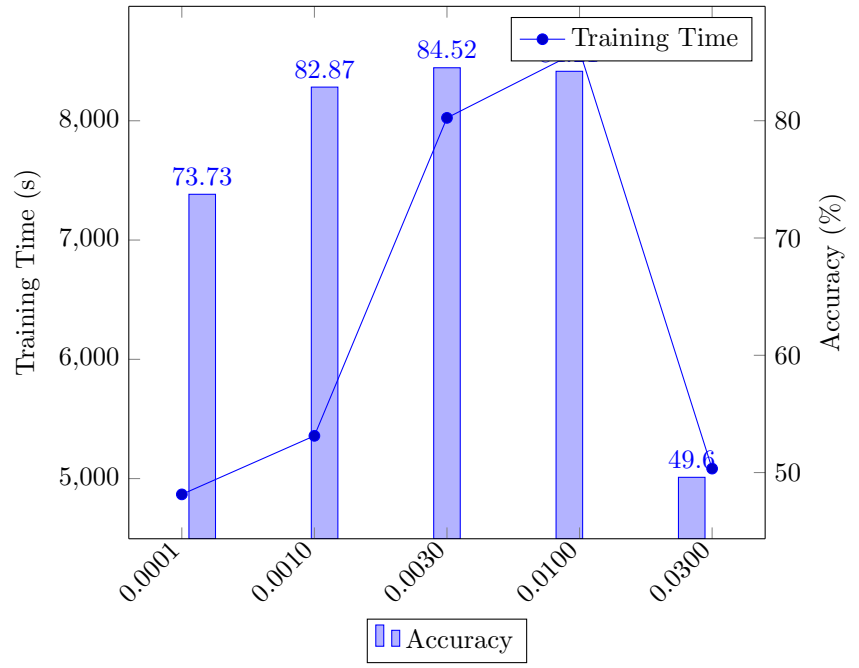
Model	Training Time (seconds)	Accuracy (%)
M1	1053.07	69.475
M2	1096.75	76.059
M3	1162.90	80.216
M4	1113.49	81.707
M5	1113.34	81.818
M6	1187.38	82.342
Dropout 1	2341.55	72.997
Dropout 2	2576.91	78.026
Dropout 3	2721.81	81.945
Dropout 4	2722.27	80.105
Dropout 5	2763.94	81.707
Dropout 6	2980.78	78.486
Dropout 6 (100 epochs)	7053.15	91.179
IDA Model	14849.57	87.815
Dropout IDA Model	14160.34	88.101
Dropout (LR=0.0001)	4867.64	73.727
Dropout (LR=0.0010)	5358.49	82.865
Dropout (LR=0.0030)	8024.23	84.515
Dropout (LR=0.0100)	8587.45	84.214
Dropout (LR=0.0300)	5083.72	49.595
Batch Size: 8	1285.326	0.947
Batch Size: 16	1041.039	0.926
Batch Size: 32	920.357	0.921
Batch Size: 64	873.642	0.904
Batch Size: 128	861.372	0.893
Batch Size: 256	850.574	0.876

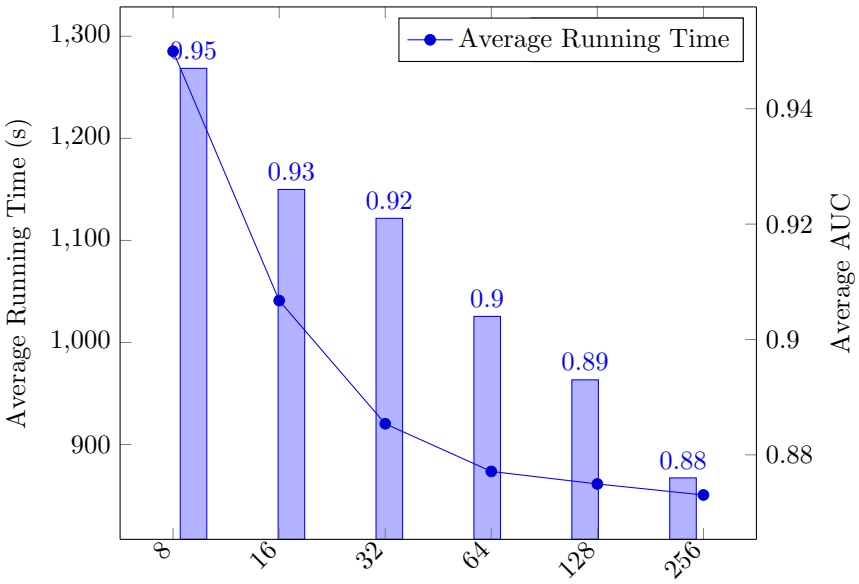
Table 1: Training time and accuracy for each model, including the results from varying batch sizes.


Training Time
Training Time (s)
Accuracy (%)
1,200
1,180
1,160
1,140
1,120
1,100
1,080
1,060
1,040
84
82
80
78
76
74
72
70
68
69.48
76.06
80.22
81.71
81.82
M1
M2
M3
M4
M5
M6


Training Time
Training Time (s)
Accuracy (%)
3,000
2,800
2,600
2,400
82
80
78
76
74
72
73
78.03
81.95
80.11
78.49
Dropout 1
Dropout 2
Dropout 3
Dropout 4
Dropout 5
Dropout 6


$\cdot 10^4$
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Training Time (s)
92
90
88
86
84
82
80
78
Accuracy (%)
Training Time
82.34
78.49
91.18
87.82
88.1
M6
Dropout 6
Dropout 6 (100 epochs)
IDA Model
Dropout IDA Model





4 Conclusion

This report presents results from a cat-dog classification task using several VGG-based models, including variations in the number of blocks, dropout regulariza-

tion, Image Data Augmentation (IDA), different learning rates, and the effects of different batch sizes.
The results demonstrate how these factors affect training time and model accuracy. In general, models with dropout layers and IDA have higher training times due to the increased complexity and the additional data generated for training. The highest accuracy was achieved by the dropout 6 model with 100 epochs. In addition to that by looking at the tests on the 3-layer model, we can say that the image data augmentation models significantly increase the performance, and if I applied it on the dropout 6 model, we could achieve much better results(It didn't due to length of training times). Variations in batch sizes also influenced training times and accuracies; notably, smaller batch sizes, although more computationally intensive, resulted in higher Average AUC scores. About the learning rates, their performance can be better interpreted by looking at the graphs inside the results folder and when we look at that graphs the the fastest and most consistent learning rate for our model is 0.01 lower lrs than that are slower and 0.03 is not learning successfully.