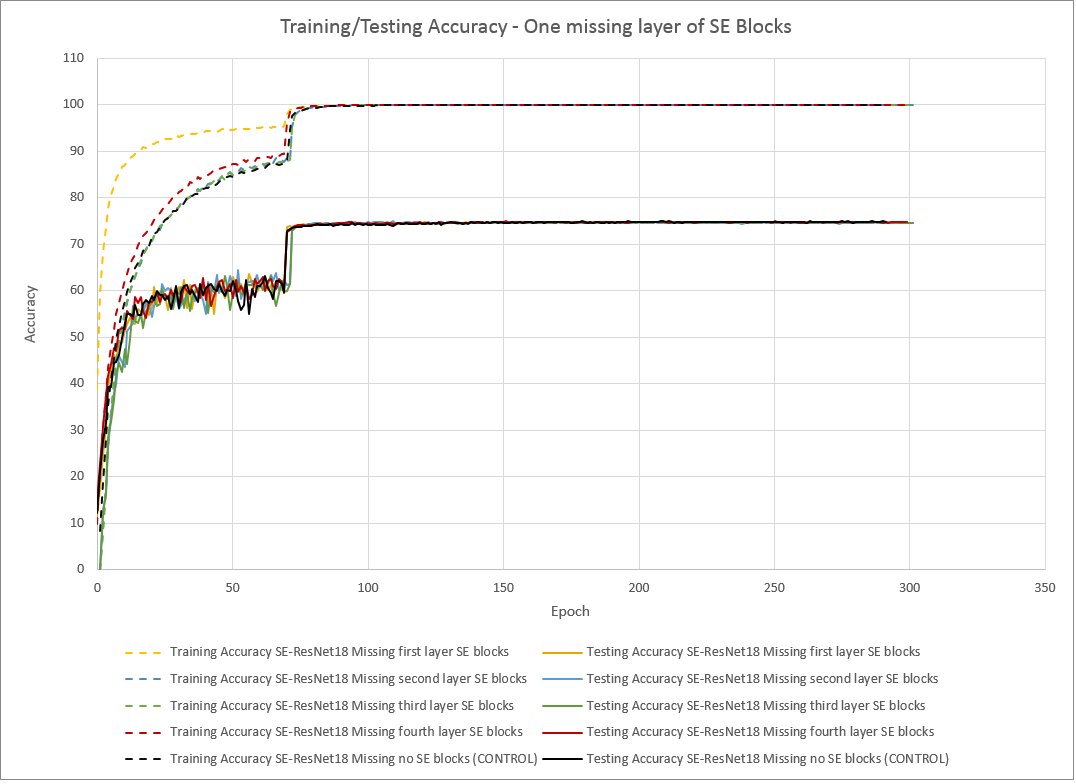
Where are SE blocks most effective in an SE-ResNet network?

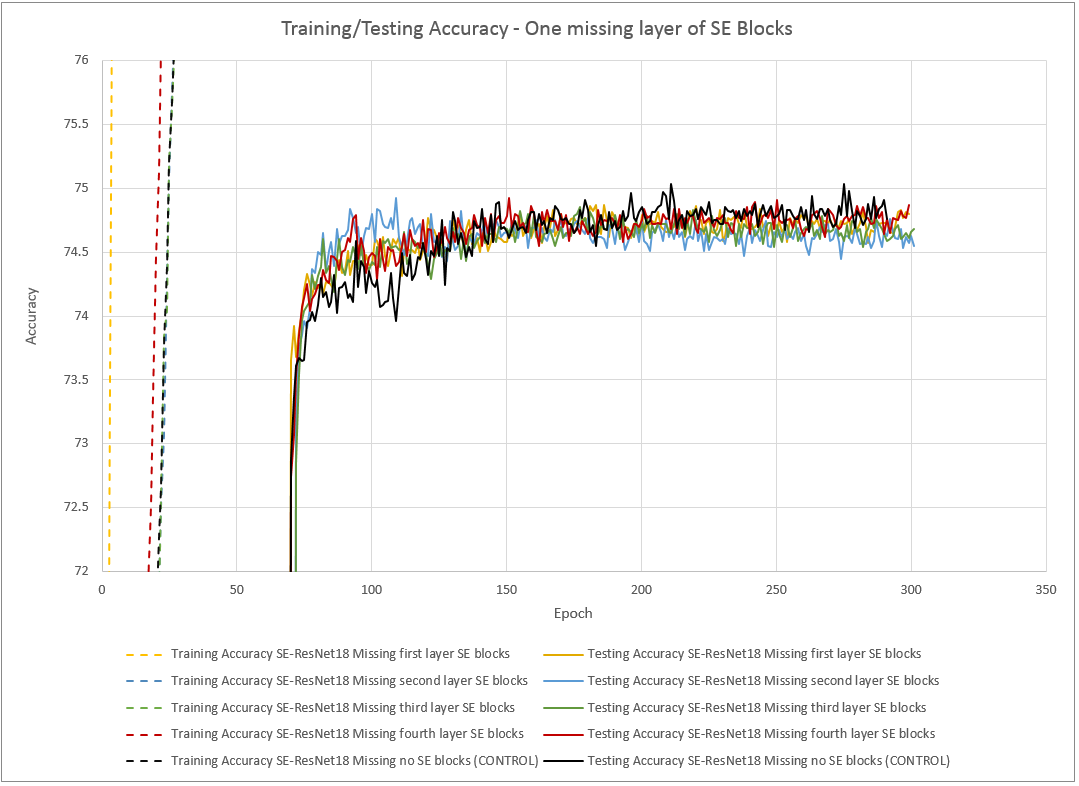
An SE-ResNet18 was used for all of these tests, due to the decreased training time.

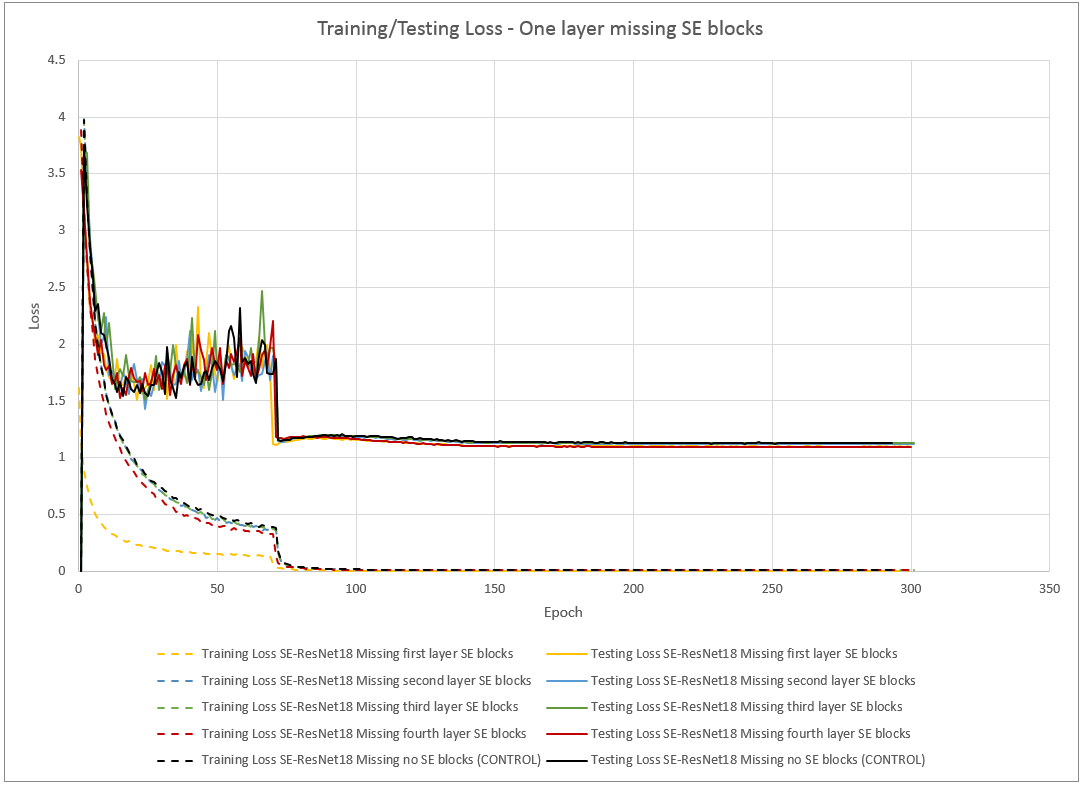
The SE-ResNet18 is initialized each time with one layer’s SE blocks removed. The network is then trained, and its loss and accuracy plots are graphed below.

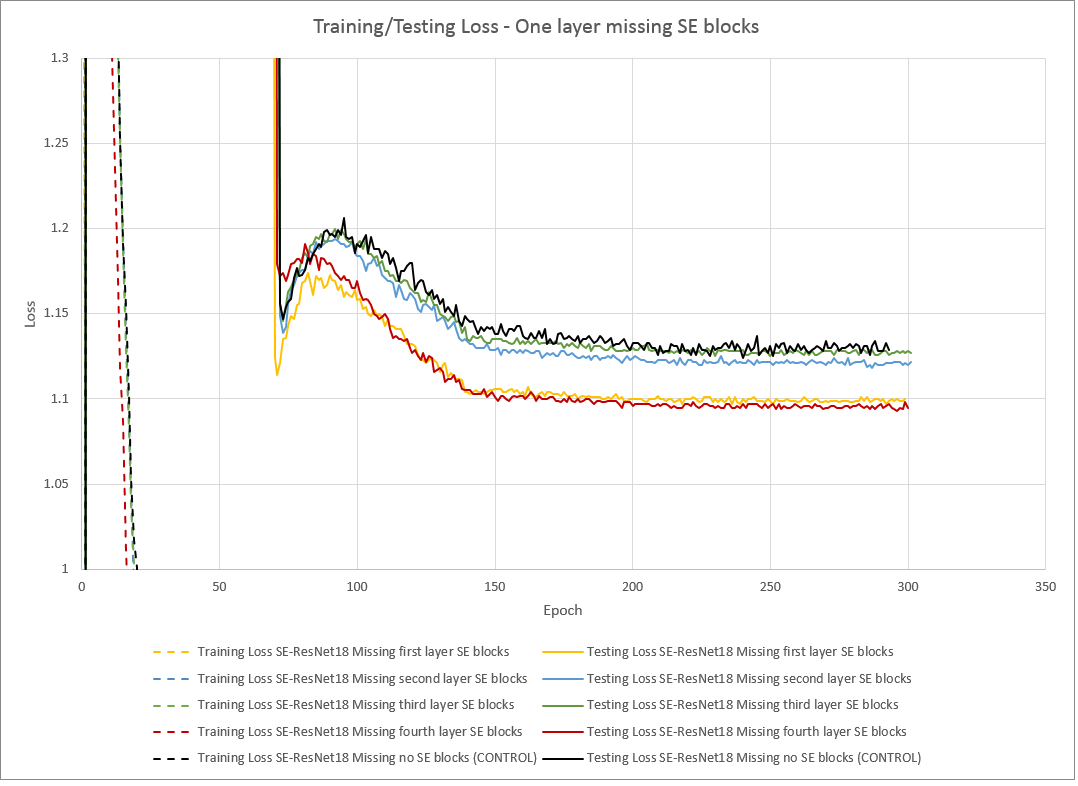
From the graphs it is clear that removing all SE blocks from any single layer in the network has no meaningful effect on the test accuracy, but removing the SE blocks from the first layer cause the network to start overfitting very quickly.

Despite the immeasurable effect on the test accuracy, there are distinct effects on the test loss by removing all SE blocks from individual layers. The removal of either the first or fourth layer of SE blocks decreases the test loss by as much as .2 after training, and removing the second or third layer of SE blocks seem to decrease the loss also, but the effect is negligible, and possibly not significant.



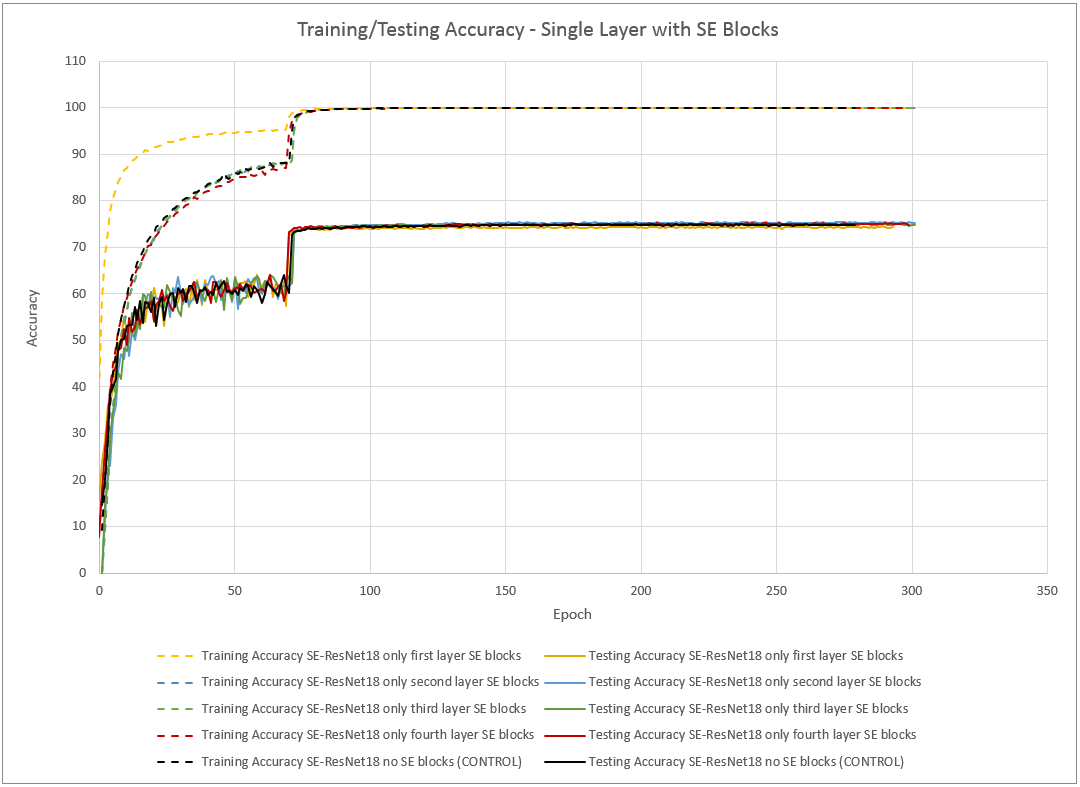


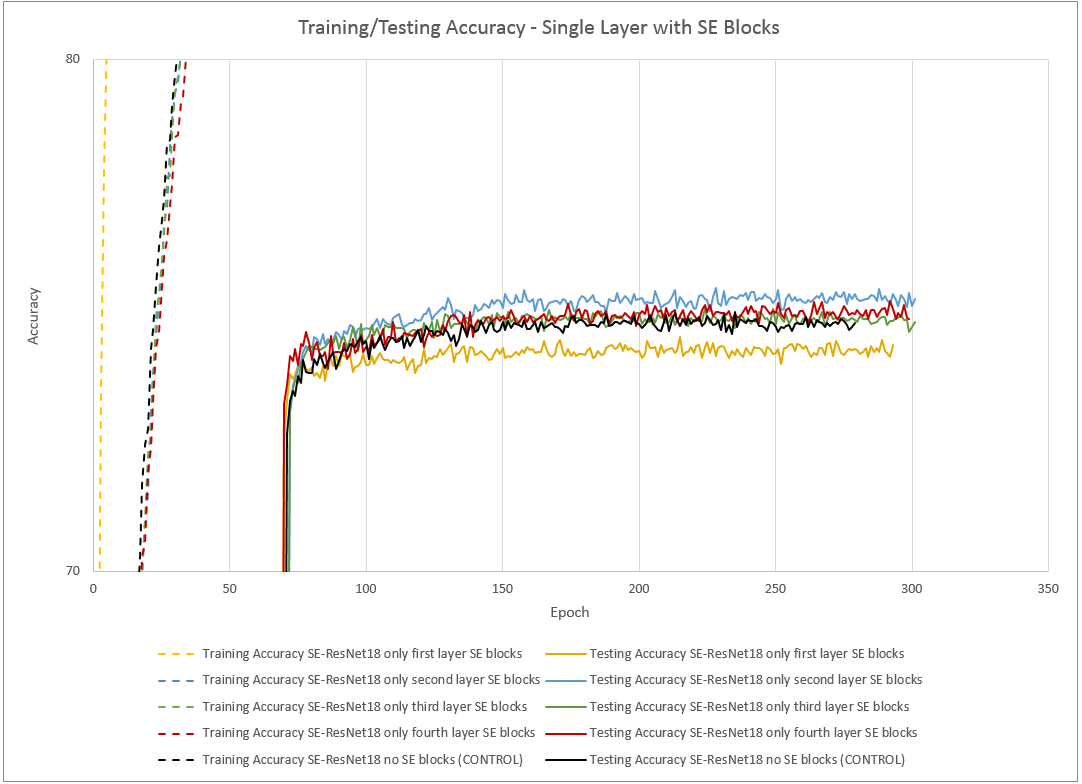


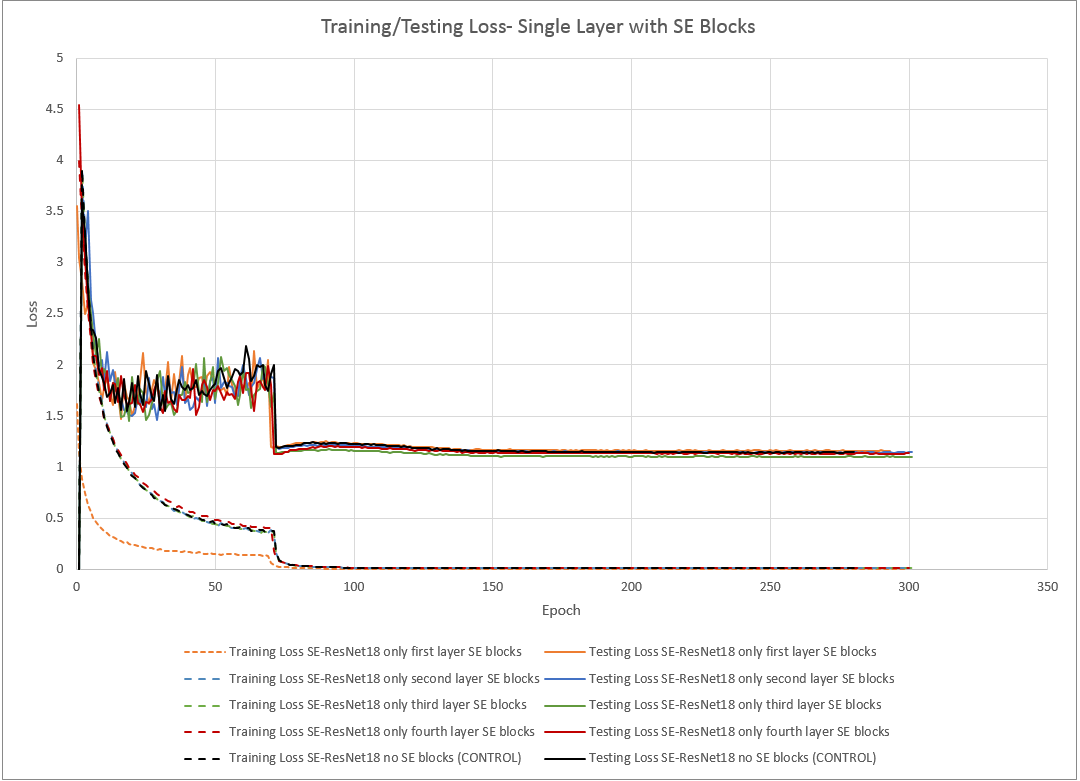


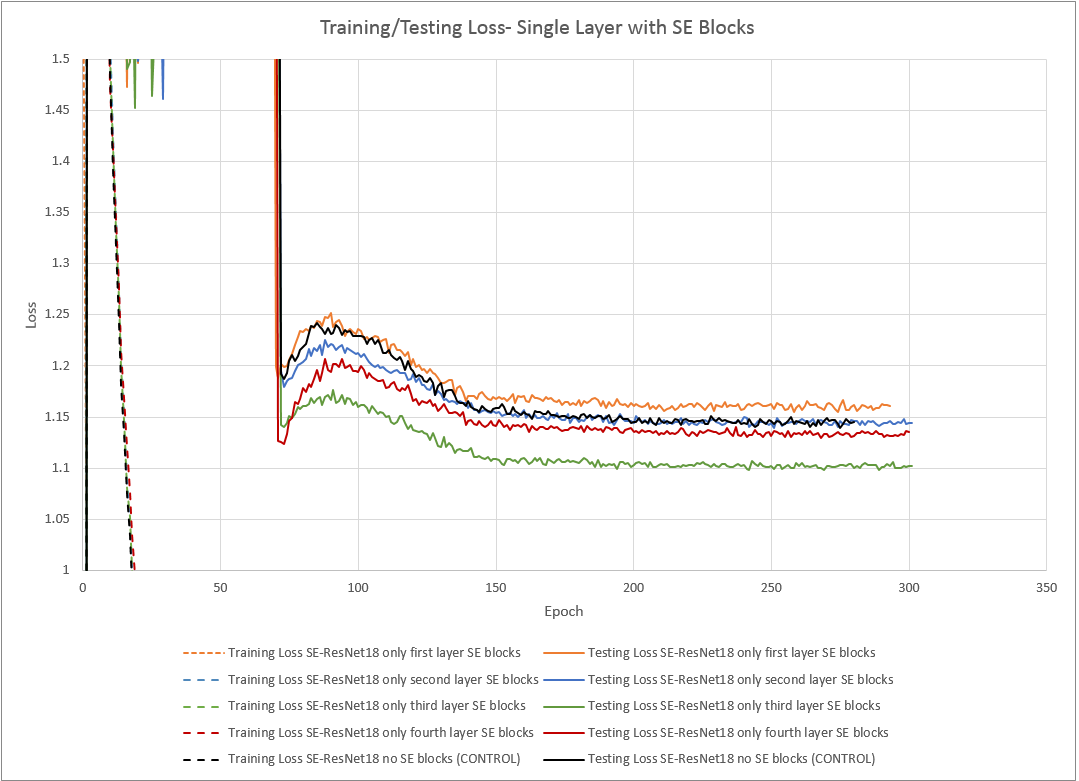
On the other hand, by adding SE blocks to an individual layer (starting with a plain ResNet18), there are much more profound effects.

Adding SE blocks only on the first layer significantly decreases test accuracy and increases test loss compared to the plain ResNet18. Adding SE blocks only to the second layer significantly increases the test accuracy. Adding SE blocks on the third layer only significantly decreases the test loss, without much increase in test accuracy, which would imply overfitting.









Broad Conclusions:

* Both experiments suggest that SE blocks on the first layer of a ResNet will decrease its classification accuracy.
* Both experiments suggest that SE blocks in the second layer of a ResNet increase accuracy the most.
* Both experiments suggest SE blocks in the third layer of a ResNet have little effect on the accuracy or loss.
* The removal of SE blocks in layer 4 decreased the loss with only a small increase in test accuracy, suggesting that SE blocks in layer 4 mostly help prevent overfitting, while contributing only a little to the classification.