**[CS-8395 Spring 2020]**

**Deep Learning in Medical Image Computing**

**\* Please print and bring it before each class**

Name: Daniel Yan

VUID: yand1

Date: 1/10/2020

Paper Title: Deep Residual Learning

Please summarize the paper using your own words: (<100 words)

The authors begin by describing an existing problem for neural networks: as networks get deeper, they tend to perform more poorly. The authors emphasize that this is not due to overfitting, since the training error gets worse, and also note that this occurs even with the use of batch normalization, so it may not be due to exploding/vanishing gradients either. To solve this problem, the authors propose the idea of residual learning such that the output of layer l is added to the input to layer l before passing the result to layer l + 1. In several tests on existing datasets, the residual network was able to obtain a better result than existing networks by using a larger number of layers.

Question 1 for the paper:

What are some results for residual networks on non-image datasets, and how do those results compare to results for plain networks?

Question 2 for the paper:

Do residual networks still perform better than plain networks when using non-SGD optimizers such as Adam, RMSProp, etc?