MA 598 MACHINE LEARNING SEMINAR
SUMMARY 4

ID 5107

4. Deep Residual Learning for Image Recognition

What is their primary result? The authors present a residual learning framework whose purpose is to ease the training of deep networks.

Why is this important? The increasing complexity of neural networks comes at a cost. Adding more layers to a network leads to overfitting, exploding/vanishing gradient, and degradation. The biggest problem seems to be in the way solvers approximate mappings. Since the approximation is done by nonlinear layers, multiple layers have a difficulty approximating identity mappings and lead to the degradation problem.

What are their key ideas?

What are the limitations, either in performance or applicability?

What might be an interesting next step based on this work?

What's the architecture?

How did they train and evaluate it?

Did they implement something?

Date: January 25, 2019.