Highway Network

Jo MinKi

1. Introduction

Question.

1. The solution part in Section 3, what does the ‘even in the convolutions’ means?
2. Later part of Section 3 and Earlier part of Section 4, I don’t get the whole picture that they want to explain.   
   What does they say??
3. Explanation about the dropout? At the Section 3? So if non-uniform sparse data structure are bad, why it doesn’t matter at the AlexNet? Or they just didn’t realize or ignored that it is matter?
4. What’s the effect of pooling? I thought we do this just decrease the amount of computation, but it seems there’s more than that.
5. Are they using the ReLU? If it is, why they attach the auxiliary classifier around the network?