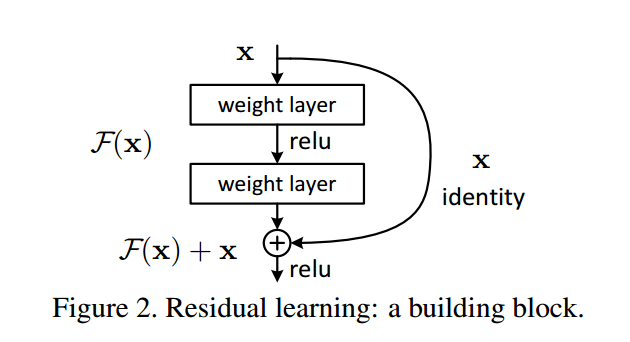
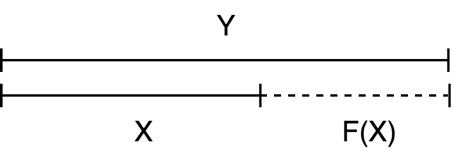
Deep Residual Learning for Image Recognition

Jo MinKi

1. Introduction  
   There was a weird and critical problem in neural network field. Everybody think that deeper network can solve more complex problem. But in reality, when the network gets deeper, the performance (validation error or testing error) of it gets worse. On top of that, when the depth of network goes further than 50 layers, even training error going to down. Which means that network lose the learning ability when the depth goes too far. This phenomena called ‘degradation problem’.
2. Main concept of this is that instead of target function , train the network to fit to the residual function . For example, if the target function is identity function , then training for instead of . This concept could be treated as a shortcut connection, which means connect one or more layers without any parameter.   
   Basic concept of ResNet is based on VGG Net, so the filer has 3x3 fixed size filter. And to make the network gets deeper, the researchers doubled the number of filter when the network has a pooling.
3. Deep Residual Learning   
   (I don’t get this part really much so I’m just stating their logic as they claim without any doubt. But my predecessor’s thinking is that making some plausible excuses is needed for researchers even though themselves don’t agree with that. So I think we can just take this claim like ‘Ah~ that could be the case’.)  
     
   

Question.

1. Do we know the ? If we don’t, where are we heading for?  
   1. No. we are not knowing the exact function that we should head for. In practically, we are just adding a skip connection on the network so we get the ­­ as a result. So don’t take the explanations on the 3.1 too much seriously.
2. Are they saying lots of layers need the identity mapping function?