Layer-wise gradients in a branched ResNet34 epoch:0 epoch:25 10^{-2} epoch:50 epoch:75 epoch:100 epoch:125 epoch:150 epoch:175 10^{-3} Average gradient 01 10-5 - 10^{-6} -25 10 20 5 15 30 0 Convolutional layer