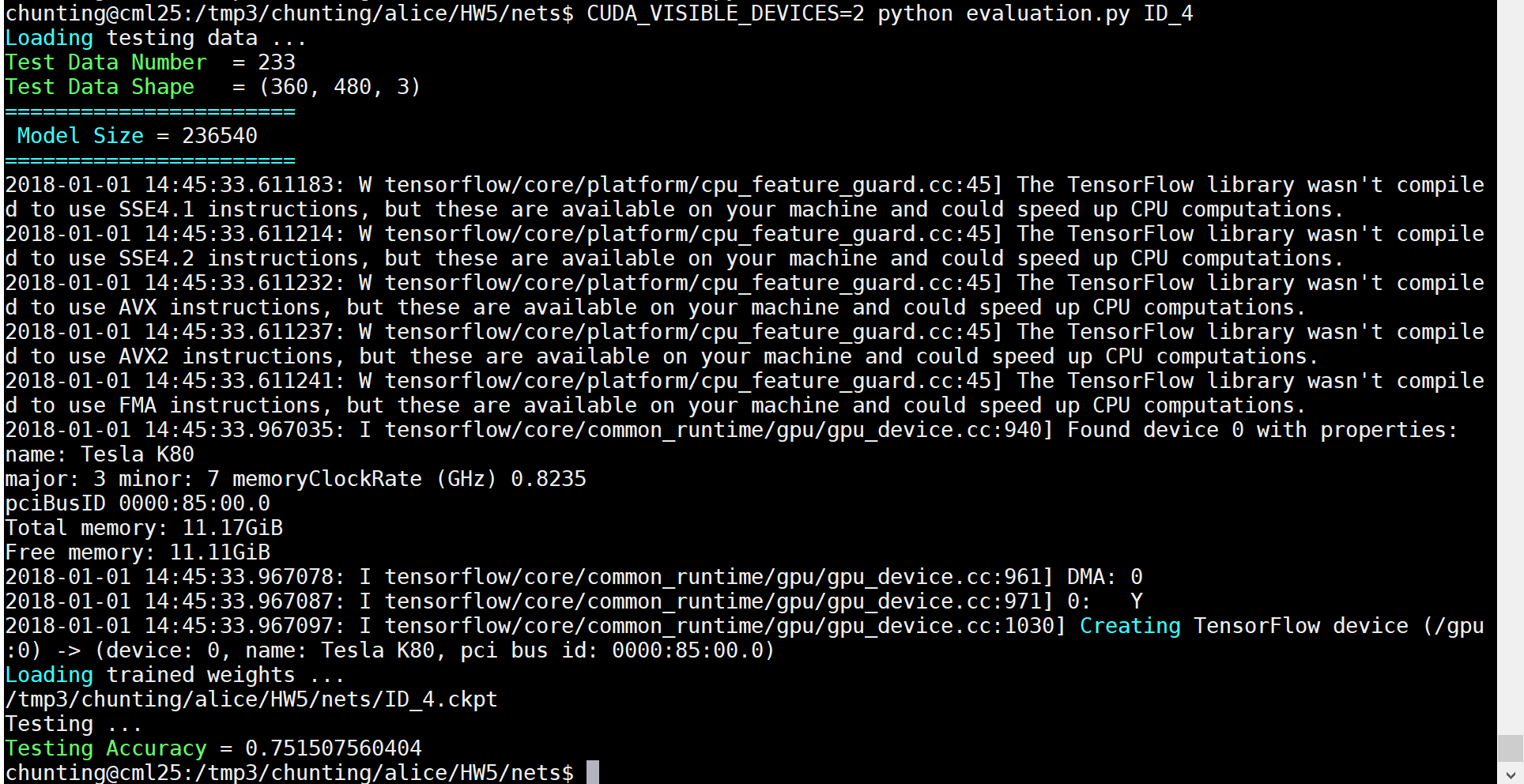
Lab5 Report A061054 蔡昀芸

# Model Design



這次model主要用了八層dilated\_conv空洞捲積及一層捲積，其中發現rate在training過程中會影響很多，我從第一層到第八層分別用的rate是1, 1, 2, 2, 2, 6, 6, 12，rate越小model能學到的會較局部，越大的話涵蓋的範圍越廣可以看到的資訊越多，所以由小的Rate在延伸到大的會比較妥當，此外前四層我都有使用tf,cond做dropout的動作，而後面沒有，因為dropout太多資訊流失，使用的learning rate一樣是sgd，原本有試著換成adam但發現效果並沒有sgd好，雖然她會收斂的比sgd快速但是optimized的時候並沒有找到最佳。結果顯示，model size = 236540，總共跑了epoch 400次可以達到準確率0.7515

Layer 1 : dilated\_conv(rate=1) , batch\_norm, relu, dropout

Layer 2 : dilated\_conv(rate=1) , batch\_norm, relu, dropout

Layer 3 : dilated\_conv(rate=2) , batch\_norm, relu, dropout

Layer 4 : dilated\_conv(rate=2) , batch\_norm, relu, dropout

Layer 5 : dilated\_conv(rate=2) , batch\_norm, relu, dropout

Layer 6 : dilated\_conv(rate=6) , batch\_norm, relu,

Layer 7 : dilated\_conv(rate=6) , batch\_norm, relu

Layer 8 : dilated\_conv(rate=12) , batch\_norm, relu

Layer 9 : conv, batch\_norm, relu

# Question and Analysis

中間嘗試了很多種model，原本是想用Segnet的架構去寫，加上encoder跟decoder，但是發現層數太多train不太起來，另外也有overfitting的問題，test基本上準確率上升很快也能達到89,90，但evaluation卻還是很低，後來有試試看改成沙漏型的model，接很多個conv一起然後調整stride的大小讓每一層從大變小，在從小變大，但這樣channel太多model size會到兩百多萬，最後用了較單純的架構，model size二十幾萬筆之前寫的架構都少很多，收斂的速度中等，optimize大概在87,88這兩個值之間。