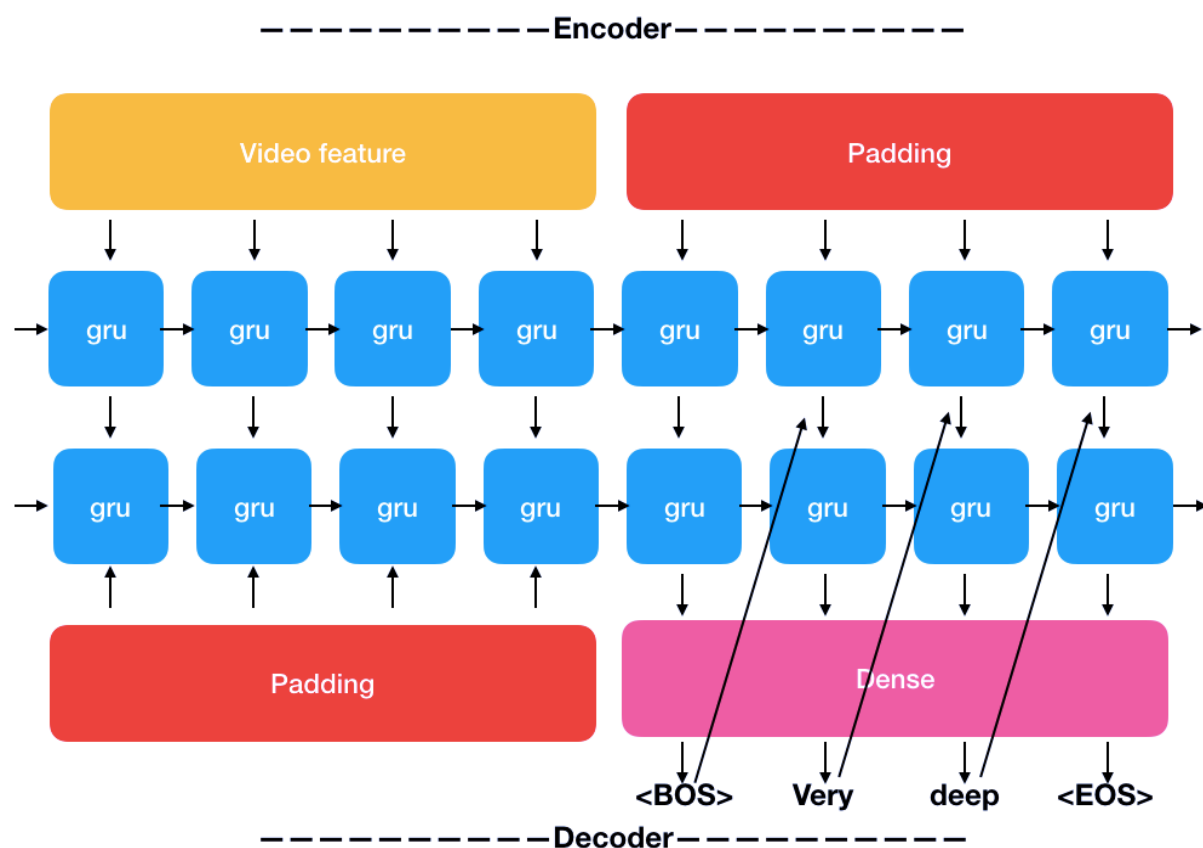


Video Caption Generation

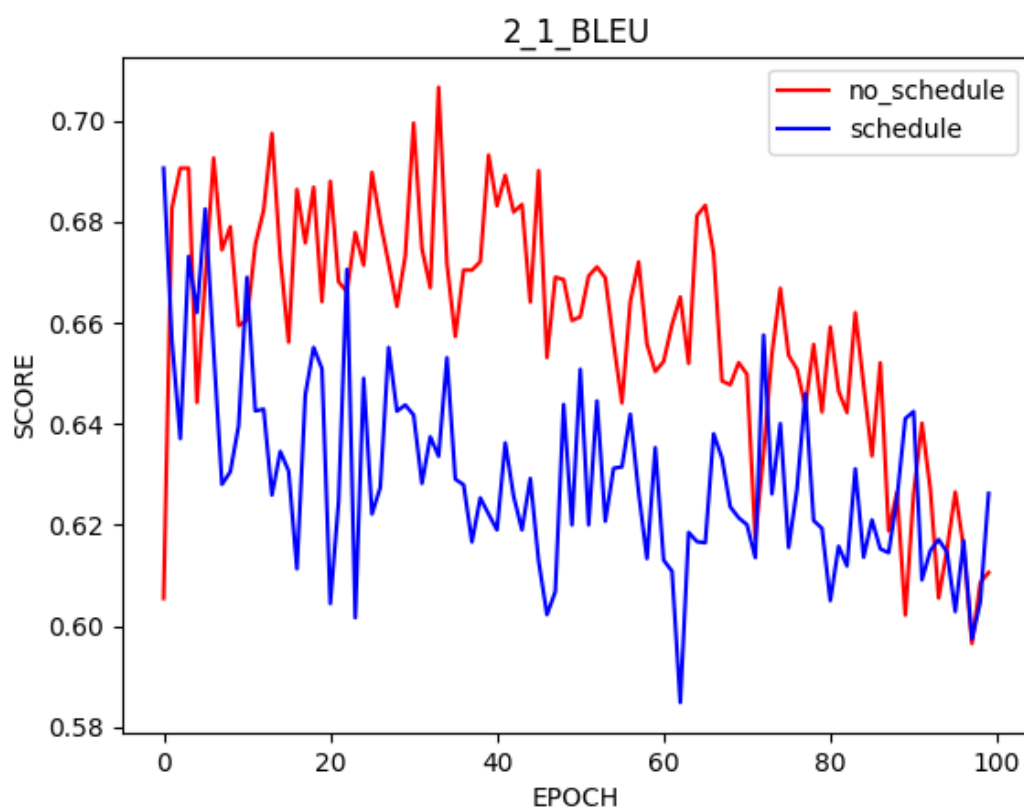
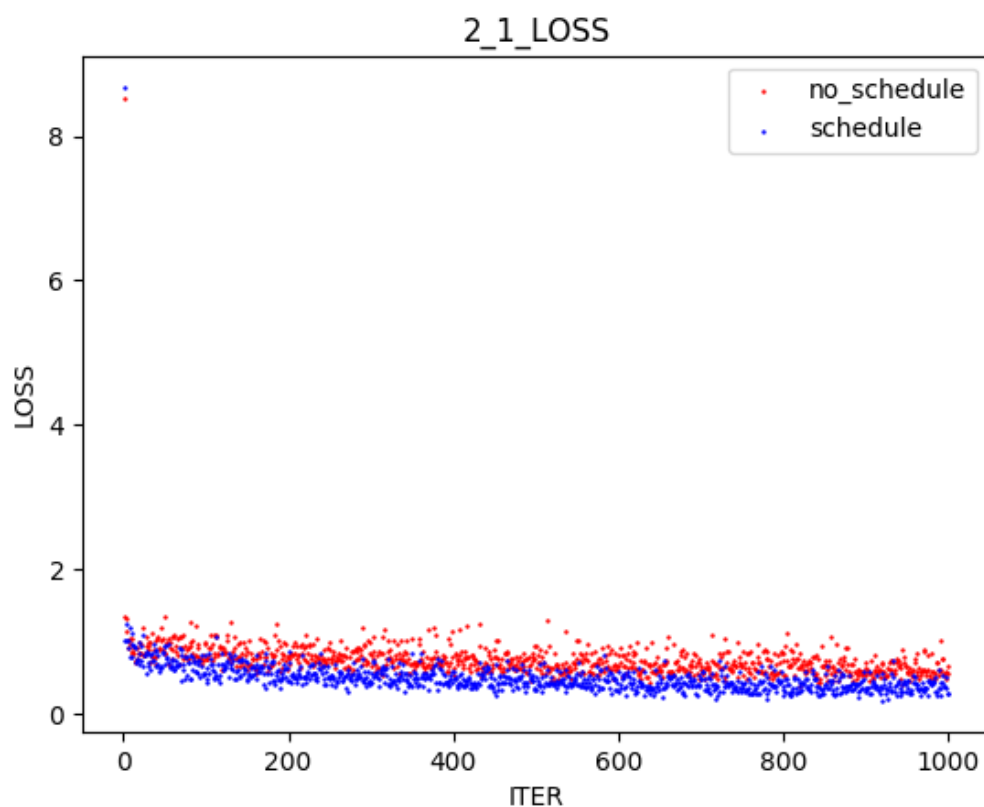
學號：b05902031, 系級：資工二, 姓名：謝議霆, 學號：b05902008, 系級：資工二, 姓名：王行健

Model description



How to improve your performance

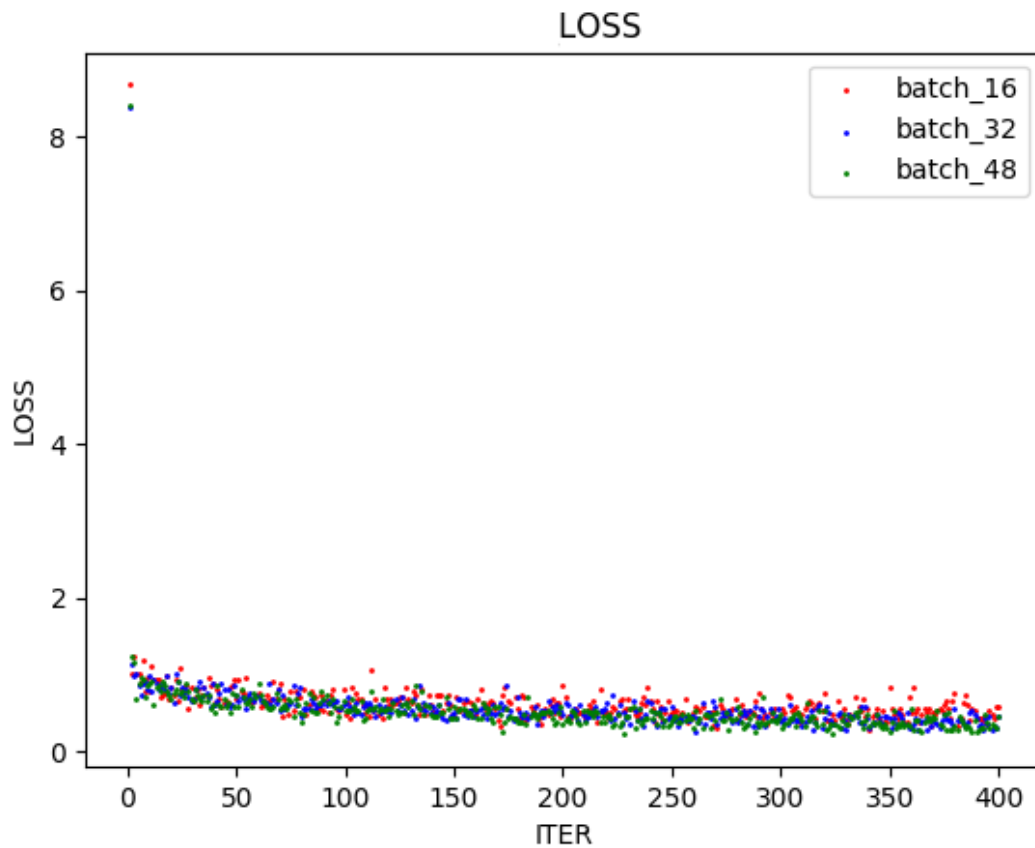
- 我們在最後decode出句子的每一個字的時候，使用schedule sampling，我們適當地在某一次decode的輸入，改成正確答案的前一個詞的embedding，原本是用前一個預測的字的embedding當輸入，頻率是0.5，代表預測的句子有50%是用正確答案的資訊當輸入。
- 因為每次decode出下一個字的時候，我們會用到預測的前一個字的embedding，但是當我們第一個字就預測錯了，那之後也預測不出好結果，所以使用schedule sampling

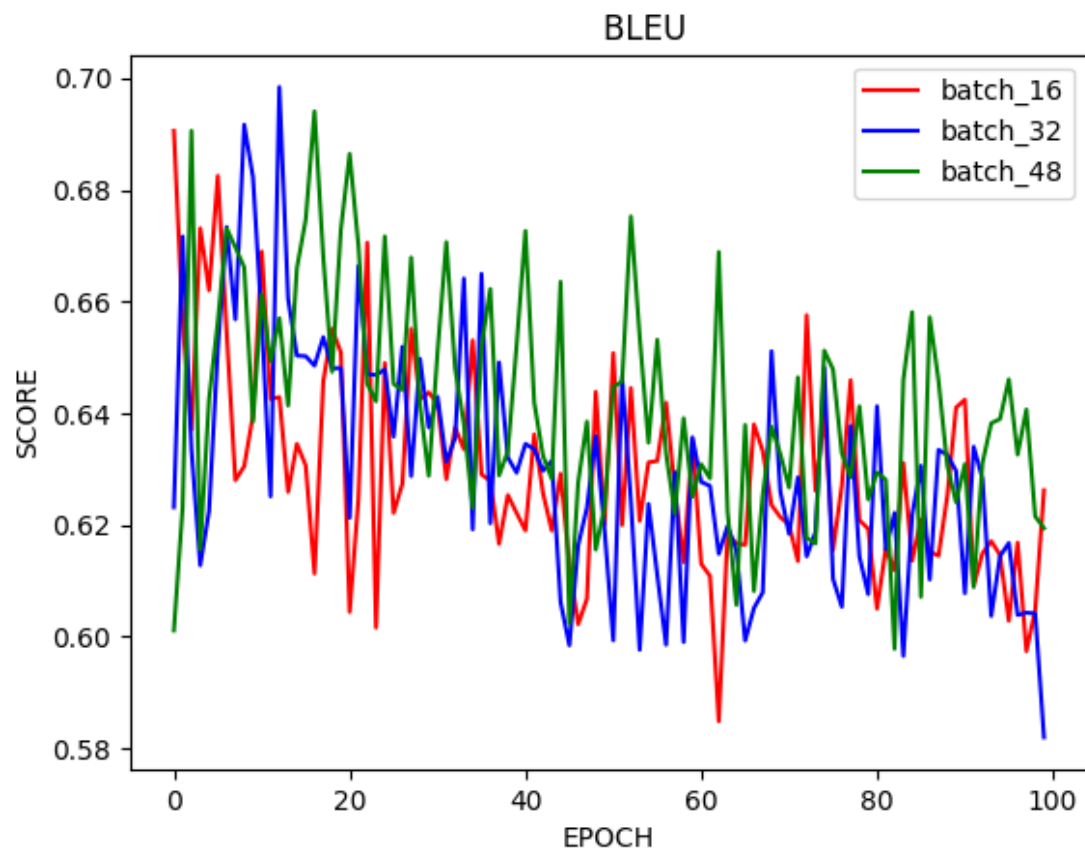


- 從圖上可以看出，有使用schedule sampling的loss都比沒使用的低，而bleu的分數，有使用schedule sampling的都比較高，雖然後面因為overfit的關係整體的bleu趨勢有往下掉，但結論是使用schedule sampling的performance會比較好。

Experimental results and setting

- Embedding size = 512
- Encoder GRU hidden size = 256
- Decoder GRU hidden size = 512
 - 當我們嘗試用更小的hidden size的時候，訓練到後來，其中testing data有一部影片是一隻貓在彈鋼琴，小一點的hidden size會出現"A cat is playing a pizza"，hidden size大一點才會輸出"A cat is playing piano"
- Batch size = 48
- Schedule sampling rate = 0.5
- Optimizer : Adam, learning rate = 0.001
- Epoch = 30
- 最長的句子長度：50
- Different batch size : 將不同的batch size拿來比較，雖然在圖上看起來差不多，但是batch_size 48的loss和bleu score明顯都比較好





分工表

B05902008 : report , 參數調整 + 測試

B05902031 : 架構調整 , 程式編寫