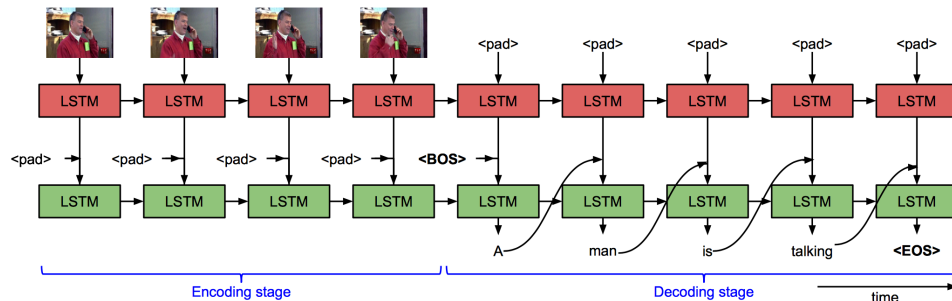


Hw2

- Model description



使用助教提供的 features，每個 video 為 80 個 frames 各 4096 維，中間兩層 lstm 128 維，為 encoder、decoder，sequence 長度為 frames 數 + caption 的字數，字數部分 training 時設定為固定長度 20 個字，少的補 `<pad>`，sequence 長度皆為 80+20，encoder 的 input sequence 為 80 個 frames 的 features，而之後給 padding，decoder 的 lstm input 為合併 encoder 的 output 與 前一次 time step 的正確 word 做 embedding 後的 vector，而 decoder 後段才開始會輸出 word sequences，

- Attention mechanism

- How do you implement attention mechanism?

使用 `tf.contrib.seq2seq.LuongAttention` 宣告 attention mechanism，並且使用 `tf.contrib.seq2seq.AttentionWrapper`，將 attention mechanism 與 decoder 的 lstm cell 包在一起。

- Compare and analyze the results of models with and without attention mechanism.

在這個 dataset 中，效果並不顯著，只能稍微提高準確度。

- How to improve your performance

- Write down the method that makes you outstanding

若在基本的 seq2seq encoder decoder 架構中在加入 attention mechanism，inference 中加入 beamsearch 去存取所有可能的 captions，會優化許多。

- Describe the model or technique
- Why do you use it

- Experimental results and settings (1%)