Speech Lab 期末總結Report

B03902013吳克駿  
B03902060江昶翰

1.Paper Reading

* **Sequence to Sequence – Video to Text**<http://www.cs.utexas.edu/~ml/papers/venugopalan.iccv15.pdf>
* **Audio Set: An ontology and human-labeled dataset for audio events**<https://research.google.com/pubs/pub45857.html>
* **Bidirectional Long-Short Term Memory for Video Description**<https://arxiv.org/pdf/1606.04631.pdf>
* **Describing Videos by Exploiting Temporal Structure**<https://arxiv.org/pdf/1502.08029.pdf>
* **Attention-Based Multimodal Fusion for Video Description**<https://www.merl.com/publications/docs/TR2017-008.pdf>

2.實作 (Project Title : 影片敘述之自動生成 -Video caption generation)

* **Project1: ASTMIC**
* **CNN-practice: Image Sentiment Classification**    
  <https://sunprinces.github.io/ML-Assignment3/index.html>
* **RNN-practice: RNN with MNIST**  
  https://github.com/c1mone/Tensorflow-101/blob/master/notebooks/10\_MNIST\_Reccurent\_Neural\_Network.ipynb
* **LSTM S2S-Video2Text: Seq2seq & Attention**  
  <https://docs.google.com/presentation/d/1OtD_BD6_Ljvr3aqLjHnnNX_h55BirD3cxhExq9wySmI/edit#slide=id.g1efeb48205_0_0>
* **Add attention layer in project**
* **Try to add audio in project**