# Report

**Github Link**
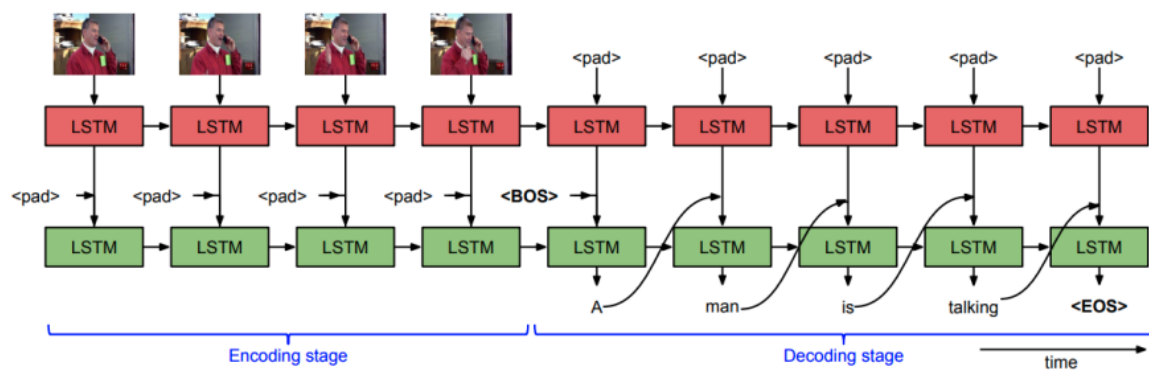
**https://github.com/ruicho/HW2/tree/main/HW2/HW2_1**

**Dataset Tree**

```
├──────MLDS_hw2_1_data
        ├────── bleu_eval.py
        ├────── sample_output_testset.txt
        ├────── testing_data/
             ├────── feat/ #100 files, .npy
             ├────── video/ #100 files, .avi
                     └────── id.txt
        ├────── testing_label.json
        ├────── training_data/
             ├────── feat/ #1450 files, .npy
             ├────── video/ #1450 files, .avi
             │       └────── id.txt
        └────── training_label.json
```

## Network Model

The network structure is similar as the one used in the paper:

Parameteres:

| LSTM dimension | 256 |
|---|---|
| Validation batch size | 100 |
| Number of frames | 80 |
| Maximal caption length | 50 |
| Forget bias (in the red blocks) | 1.0 |
| Forget bias (in the green blocks) | 1.0 |
| Dropout probability | 0.5 |
| Learning rate | 0.5 |
| Number of epochs | 100 |
| Batch size | 256 |

Data directories:

| Data directory | /HW2/HW2_1/MLDS |
|---|---|
| Test data directory | /HW2/HW2_1/MLDS/testing_data |
| The saved model directory | saved_model/ |

Schedule sampling is used in this model, and the sampling probability is 0.6.

The final bleu score for this model:

```
(base) PS C:\Users\10452\Downloads\seq2seq_video2text-master\HW2\HW2_1> python .\bleu_eval.py .\final_output.txt
Average bleu score is 0.701456473412245
(base) PS C:\Users\10452\Downloads\seq2seq_video2text-master\HW2\HW2_1>
```

We can see the bleu score is about 0.7

For running the shell:

Please run:

hw2_1.sh _ _ _

• _ _ _ here represent the data directory, the test data directory, and the name of the output(.txt)

How to run the training:
Run: python model_train.py

How to get the bleu score:
Go to the data set file.
Run: python bleu_eval_origin.py    HW2/HW2_1 /final_output.txt

How to run the testing:

Run python model_test.py