

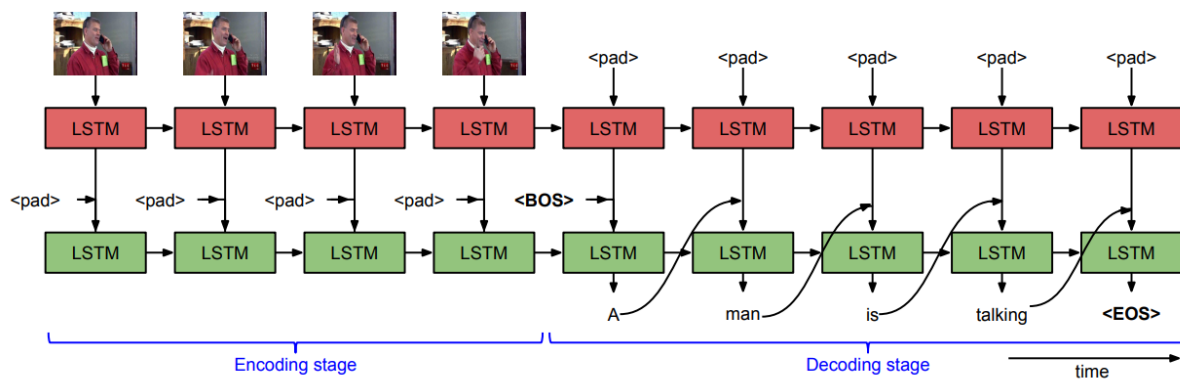
- **Dataset tree:**

```

|—— 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:



- **Parameters:**

The parameters used in this project:

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	0.001
Number of epochs	100
Batch size	256

The data directories I used for this project are shown below:

Data directory	'/home/lingyis/DLhw2/MLDS_hw2_1_data'
Test data directory	'/home/lingyis/DLhw2/MLDS_hw2_1_data/testing_data'
The saved model directory	'save_model/'

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

The final bleu score for this model:

```
(8810) [lingyis@node0097 ~]$ cd DLhw2/MLDS_hw2_1_data/
(8810) [lingyis@node0097 MLDS_hw2_1_data]$ python bleu_eval_origin.py /home/lingyis/DLhw2/final_output.txt
Average bleu score is 0.7076193854480062
(8810) [lingyis@node0097 MLDS_hw2_1_data]$
```

We can see the bleu score is about 0.7.

For running the shell:

Please run:

Hw2_seq2seq.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 home/lingyis/DLhw2/final_output.txt`

How to run the testing:

Run `python model_test.py`