

Introduction

This experiment is aim to improve the performance of ASR system in noise environment

Background

DNN have really good performance when training data is clean while it cannot perform well in noise environment.

Instead of research on improvement of AM structure, we figure out a new training method to make the ASR system do better. This training method derives on Adaboost which is to assemble several weak classifiers into a relatively strong classifier. In this work, the classifier is DNN. We train several small DNN and ensemble them to see whether they can have better performance than the big DNN whose node amount is the same as node amount of them all. This several classifiers is trained iteratively whose training data is based on last trained DNN's training data which is not learned well and we keep the dataset size (num of utterance) to be the same.

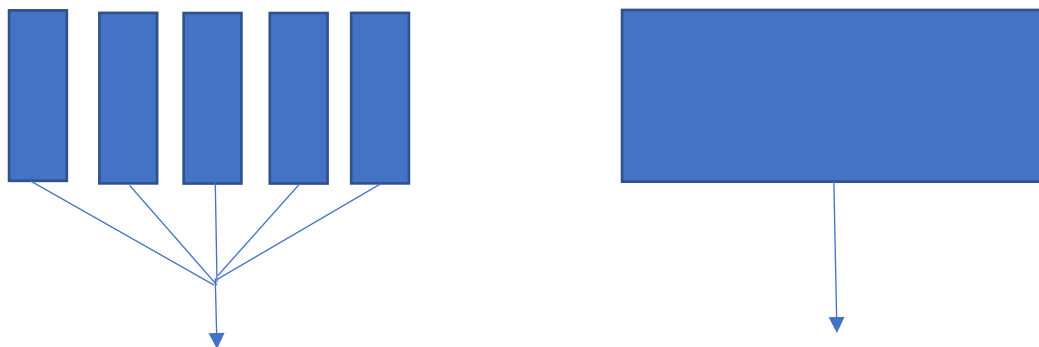
Procedure

- Train a DNN (256*6)
- ```
for (i = 0 ; i < 5 ; i++)
```
- ```
{
```
- Decode the training data to get the WER of each utterance
 - Normalize the probability of each utterance in WER , resample the data and the amount of utterance of new dataset is the same as the old one.
 - Train a new DNN use the newly resampled dataset.
- ```
}
```

Assemble the 5 DNN(256\*6) into a big one.

Train a big DNN(1280\*6)

Compare the performance of the assembled model and the big model.



### Experiment details

Training data: WSJ /train\_si284\_sp\_hires(+1 noise)

Test data: WSJ/test\_eval92 & dev93\_sp\_hires(+1 noise)

### Structure details

Model\_1 (1280 \* 6)

|                          |       |
|--------------------------|-------|
| decode_bd_tgpr_dev93     | 10.37 |
| decode_bd_tgpr_dev93_fg  | 9.21  |
| decode_bd_tgpr_eval92    | 6.50  |
| decode_bd_tgpr_eval92_fg | 5.46  |
| decode_tg_dev93          | 11.85 |
| decode_tg_eval92         | 8.22  |
| decode_tgpr_dev93        | 12.67 |
| decode_tgpr_eval92       | 8.86  |

Model\_2.1/2.2/2.3/2.4/2.5

|                   | 2.1(256*6) | 2.2(256*6) | 2.3(256*6) | 2.4(256*6) | 2.5(256*6) |
|-------------------|------------|------------|------------|------------|------------|
| bd_tgpr_dev93     | 12.74      |            |            |            |            |
| bd_tgpr_dev93_fg  | 11.23      |            |            |            |            |
| bd_tgpr_eval92    | 8.65       |            |            |            |            |
| bd_tgpr_eval92_fg | 7.27       |            |            |            |            |
| tg_dev93          | 14.12      |            |            |            |            |
| tg_eval92         | 10.07      |            |            |            |            |
| tgpr_dev93        | 14.63      |            |            |            |            |
| tgpr_eval92       | 10.85      |            |            |            |            |