深蓝学院
shenlanxueyuan.com

语音合成：第四章作业讲评

主讲人 谢启聪

作业

深蓝学院
shenlanxueyuan.com

- 第一部分：实现Tacotron系统中的CBHG模块
- 第二部分：完成模型训练和测试

作业

- 需要实现的CBHG编码器的部分在*models/basic_model.py*中
- 参数配置可以参考Tacotron1[1]

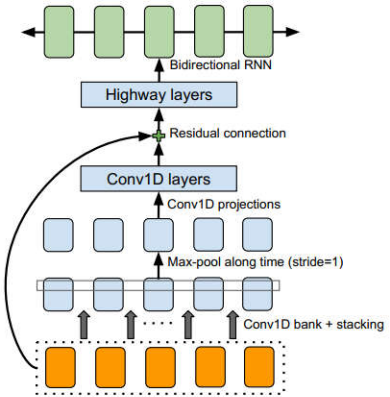


Figure 2: The CBHG (1-D convolution bank + highway network + bidirectional GRU) module adapted from Lee et al. (2016).

Spectral analysis	*pre-emphasis*: 0.97; *frame length*: 50 ms; *frame shift*: 12.5 ms; *window type*: Hann
Character embedding	256-D
Encoder CBHG	*Conv1D bank*: K=16, conv-k-128-ReLU *Max pooling*: stride=1, width=2 *Conv1D projections*: conv-3-128-ReLU → conv-3-128-Linear *Highway net*: 4 layers of FC-128-ReLU *Bidirectional GRU*: 128 cells
Encoder pre-net	FC-256-ReLU → Dropout(0.5) → FC-128-ReLU → Dropout(0.5)
Decoder pre-net	FC-256-ReLU → Dropout(0.5)→ FC-128-ReLU → Dropout(0.5)
Decoder RNN	2-layer residual GRU (256 cells)
Attention RNN	1-layer GRU (256 cells)
Post-processing net CBHG	*Conv1D bank*: K=8, conv-k-128-ReLU *Max pooling*: stride=1, width=2 *Conv1D projections*: conv-3-256-ReLU → conv-3-80-Linear *Highway net*: 4 layers of FC-128-ReLU *Bidirectional GRU*: 128 cells
Reduction factor (r)	2

[1] Wang, Yuxuan, et al. "Tacotron: Towards end-to-end speech synthesis." *arXiv preprint arXiv:1703.10135* (2017).

作业

深蓝学院
shenlanxueyuan.com

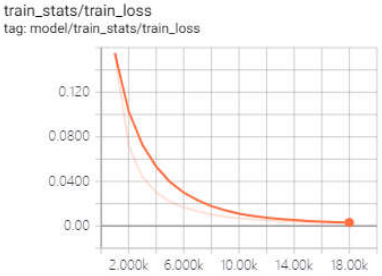
- 第一部分：实现Tacotron系统中的CBHG模块
- **第二部分：完成模型训练和测试**

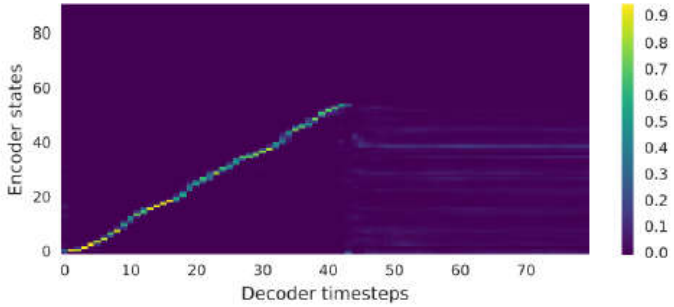
作业

判断模型是否收敛

1） loss曲线

2）attention的对齐





深蓝学院
shenlanxueyuan.com

感谢各位聆听！

Thanks for Listening