

南开大学计算机学院 HLT 实验室招收推免 硕士研究生测试项目

语音识别深度学习系统

随着深度学习技术的飞速发展，语音识别已经从经典的 HMM-GMM 模型发展到了基于深度学习技术的端到端模型，性能也取得了巨大的提高。本项目目标是深入理解 transformer 模型、复现 Speech Transformer、与基线系统对比、提高 Speech Transformer 的可能的提高的方法的探索或思考。

1. 准备知识:

machine learning, Speech signal processing, speech recognition foundation, language modeling technology, pytorch, torchaudio, librosa, etc.

2. 数据:

Aishell1: <https://www.openslr.org/33/>

3. 参考论文

SPEECH-TRANSFORMER: A NO-RECURRENCE SEQUENCE-TO-SEQUENCE MODEL
FOR SPEECH RECOGNITION



Speech-Transformer_A_No-Recuri

4. 目标:

1. 基本目标, 复现 Speech Transformer。允许搜索和参考前人代码, 建议使用 pytorch 实现。
2. 实验报告: 包括:
 - [1] Understanding of Transformer model (Attention is All You Need)
 - [2] Model building using Aishell1-trainset
 - [3] Model testing using Aishell1-testset
 - [4] Compare with baseline models
 - [5] Improvement (for example, using Conformer model, add language model, etc.)
 - [6] Ablation analysis
 - [7] Future work

5. 提交:

1. 代码仓库 (Gitee 链接)
2. 实验报告
3. 20 分钟工作报告视频链接 (哔哩哔哩)
4. 时间: 7 月 25 号或提前

6. HLT 实验室简介

<https://qinyong.vrdcc.com/>

7. 联系人

秦勇, 13132593367(mobile), kelvinqinyong (wechat)