

第一周总结

第一部分：工作简要描述

第二部分：各项工作实际完成情况及摘录

2.1 阅读官方资源网站，及官方提供的比赛说明文档等内容

2.2 做好相应先导知识学习(简要)

2.2.1 进行简易语音合成实验，使用TTS工具进行简要语音合成工作

2.2.2 进行简易语音合成实验，使用vocaloid等工具进行语音转换工作(此处尝试失败)

2.2.3 进行PyTorch简易学习

2.3 第三部分：论文集中相应资料的整理整合工作

2.4：根据比赛提供的相应指标选取3篇论文作为后续算法复现的参考资源

第五部分：对下一阶段工作的展望

第六部分：结语

第一周总结

学号：16340074

姓名：何自强

专业：软件工程(电子政务方向)

内容：第一周工作总结及摘录

时间段：2019年11月14日-2019年11月20日

备注：所有工作数据，笔记，摘录等均放在<https://github.com/understated-1/->该仓库上，随时可查看，溯源。因为原先使用的github内容较杂，所以新注册了一个github账号

第一部分：工作简要描述

- 阅读官网资源网站，及官方提供的比赛说明文档等内容
- 做好相应先导知识学习(简要)
- 论文集中相关资料的整理整合工作
- 根据比赛提供的相应指标选取3篇论文作为后续算法复现的参考资源

第二部分：各项工作实际完成情况及摘录

2.1 阅读官方资源网站，及官方提供的比赛说明文档等内容

官方资源网站为：<https://www.asvspoof.org/>

官方比赛说明文档及后续赛事结果的详细结果为: <https://arxiv.org/vc/arxiv/papers/1904/1904.05441v1.pdf>

对给部分官方内容进行简要的总结如下:

ASVspoof比赛旨在为**ASV(automatic speaker verification)**系统研究提供应对日益先进的语音欺诈技术的方法, 检验标准, 并进行相应数据的整合工作。

在**ASVspoof2013**比赛中, 赛事主要关注点在于倡导研究者之间通力合作, 研究初步的检测指标, 进行相应数据集的创设工作

在**ASVspoof2015**比赛中, 赛事关注于两类语音篡改(合成)技术引起的欺诈案例, 即语音合成技术(**SS, speech synthesis**)与语音转换技术(**VC, voice conversion**)

在**ASVspoof2017**比赛中, 赛事主要关注点转移至重放攻击(**RA, replay attack**)部分, 并进行相应检测指标, 检测方法的探索

在**ASVspoof2019**比赛中, 赛事主办方将前几届比赛的内容进行一次综合, 将**SS, VC, RA**均纳入考虑范畴之内, 将**SS**与**VC**归类于**LA(logical access)**部分, **RA**归类于**PA(physical access)**部分, 对这两个子工作均创设相应数据集, 并制定赛事的评测指标(**t-DCF**), 集思广益, 以求新方法, 新方向的出现

数据集部分从黎师兄处已下载, 由于过大, 并未考虑将其放在github上

参考部分:

[ASV spoof 2019: Future Horizons in Spoofed and Fake Audio Detection](#)

IIIT-H Spoofing Countermeasures for Automatic Speaker Verification Spoofing and Countermeasures Challenge 2019

2.2 做好相应先导知识学习(简要)

大致简要(非常粗陋)地进行以下工作:

2.2.1 进行简易语音合成实验, 使用**TTS**工具进行简要语音合成工作

因为觉得需要进行相应检测工作, 先要了解敌方的技术, 所以尝试进行相应语音合成工作

利用**espeak**工具进行简要的**TTS**转换, 转换内容

- 字符串: 这里使用台词“just a joker”
- 中短篇长的短文: 《基督山伯爵》(英文版)第一章内容(txt文本共18kb)

实验利用**espeak**工作成功进行了上述两个**TTS**任务, 并对合成数据进行简要数据描述, 详细报告已放至github上, 地址

<https://github.com/understated-1/-/blob/master/> 进度计划/2019年11月14日-2019年11月20日/语音合成实验/TTS/实验记录.pdf

2.2.2 进行简易语音合成实验，使用vocaloid等工具进行语音转换工作(此处尝试失败)

尝试利用TTS工具后，进行语音转换工作的尝试，在知乎，stackoverflow等论坛上进行搜索，确定了下面几种方案：

- 使用utau工具进行语音转换工作
- 使用vocaloid工具进行语音转换工作

第一项方案，进行到一半的时候，发现utau与win10存在兼容性问题，且版本过于旧，安装，运行工作均无法正常进行，最后放弃。

第二项方案，经过较长时间摸索安装成功，但在音源及相应转换工作进行卡住，由于事件有限，此处就没有再进行下去

2.2.3 进行PyTorch简易学习

查阅相应论文，发现均出现大量利用神经网络进行相应检测分类工作，之前有简易学习过Tensorflow，但在部分地方卡住，在同学的介绍下，对另一个框架PyTorch进行简易的学习。

使用教材为：王海玲，刘江峰译制的《PyTorch深度学习》，结合PyTorch文档

具体实操案例：

- 简易实现线性回归模型
- 参照教程相应内容，模仿着实现简单的全连接神经网络的模型建立，数据导入，训练，评测工作，并运行简易的CNN给出的样例代码

笔记内容见github：

<https://github.com/understated-1/-/blob/master>/进度计划/2019年11月14日-2019年11月20日/PyTorch简要学习/PyTorch文档阅读.md

这里简要对PyTorch文档的Tensor部分文档进行学习

2.3 第三部分：论文集中相应资料的整理整合工作

对提供的论文集，利用关键词搜索，进行ASVspoof相关论文进行第一次初筛，然后根据涉及的欺诈技术进行分类为LA及PA，该文档已提交至微信群内，github上也有相应整合后文件。

<https://github.com/understated-1/-/tree/master>/进度计划/2019年11月14日-2019年11月20日/论文整合/整理

筛选流程与方法比较粗糙，若有错漏，后期再修改

2.4: 根据比赛提供的相应指标选取3篇论文作为后续算法复现的参考资源

选取以下三篇论文，后期可能会做出修改：

- 论文1:
 - 标题: Residual Neural Networks for Audio Spoofing Detection
 - 作者: Moustafa Alzantot, Ziqi Wang, Mani B. Srivastava
 - 选取理由: 学习相应神经网络知识, 及相应建模知识, 且作者已于github开源相应代码, 可参考学习
- 论文2:
 - 标题: IIIT-H Spoofing Countermeasures for Automatic Speaker Verification Spoofing and Countermeasures Challenge 2019
 - 作者: K N R K Raju Alluri, Anil Kumar Vuppala
 - 选取理由: 粗略看abstract感觉有之前做的提取MDCT系数用于检测相应篡改的案例的影子, 而且可以加深对相应内容的理解
- 论文3:
 - 暂未定

第三篇论文还并未确定, 在后续的总结报告中会给出

第五部分: 对下一阶段工作的展望

进行以下工作:

- 对论文1进行较为详尽的阅读, 理解
- 对给出的开源代码进行阅读, 理解, 运行, 检测, 检验
- 尝试对开源代码进行修改, 或自己进行相应内容的实现工作
- 以上部分均做出较为详尽的阅读报告, 实验报告

第六部分: 结语

以上为第一周内容的学习总结记录, 若有不当的地方, 还望老师指出, 所有工作均放置于github上, 可随时查阅进度。

