

成 都 理 工 大 学

专业硕士毕业论文工作计划表

院、系（所）名称	信息科学与技术
专 业	信息与通信工程
研 究 方 向	信号信息处理
导 师	李 灿 平
研 究 生 姓 名	孙 凌 山
学 位 级 别	硕 士

研 究 生 院 制 表

年 月 日

论文题目	基于卷积神经网络的音频场景分类方法研究
课题类型	应用研究
课题来源	科研项目
<p>选题依据（本课题在国内外的研究动态和对此课题开展研究工作的设想，以及准备在哪些方面有所进展或突破）：</p> <p>通过分析声音使设备能够理解其环境是音频场景分类（Acoustic Scene Classification，下文简称 ASC）研究的主要目标。此前音频场景分类仍然基于将通用分类器（高斯混合模型，支持向量机，隐马尔可夫模型）应用于手动提取的特征，例如梅尔频率倒谱系数。近年来，得益于计算机速度的提升与深度学习的快速发展，人们逐渐意识到，可以尝试用深度学习的自动特征提取的特性来代替以往低效的手工提取。</p> <p>Sawhney 和 Maes 在 1997 年使用递归神经网络与 KNN 实现 ASC，达到了 68% 的整体分类准确率。佩尔顿等人在 2001 年通过人类对音频场景典型事件的认识获得了 70% 的分类准确率。2015 年, Piczak 等人对深度学习中的卷积神经网络是否可有效的应用于音频场景分类这一问题进行了探讨。实验结果表明，得益于卷积神经网络对数据集的利用程度高及高效的类别学习特性。</p> <p>在特征提取部分，本文采用效果较好的 log-mel 频谱。为了取得足够好的分类效果，本文选取 Ubransound8K 音频数据集，目的是通过增加数据集的数据量使分类效果达到最优。卷积神经网络部分将引入目前广受欢迎的 Tensorflow 深度学习任务框架。让计算机进行音频自动分类，并探讨深度学习参数如神经元的数量，隐藏层的数量和引入丢失对学习结果造成的影响。</p> <p>由于使用了经试验验证过有效的音频特征，以及之后阶段对神经网络参数的不断优化，理论上音频分类效率相较于传统分类方法应有程度不小的提升。为测试本系统的优越性，特地引入传统的分类基线系统。与之对比的基线系统，本文采用 MFCC+KNN，以对比观察卷积神经网络在音频分类中的优势。KNN 分类器是机器学习领域使用广泛的分类器，其原理简单、效果好，特别是随着训练样本数据的增多，分类效果会更好。</p> <p style="text-align: right;">签名：</p> <p style="text-align: right;">年 月 日</p>	

调查研究（或实验）及论文写作计划

（ 2017 年 10 月 1 日至 2019 年 5 月 1 日）

序号	毕业论文工作阶段要求	实习或出差地	起止日期	检查方式
1	搜集相关基础资料及文献、技术调研	学校	2017.10-2018.05	现场检查
2	开题设计准备及答辩	学校	2018.06	现场检查
3	进行基线系统测试	学校	2018.07-2018.09	现场检查
4	训练卷积神经网络进行音频分类，调试参数并改进系统。	学校	2018.10-2019.01	现场检查
5	完成论文初稿	学校	2019.03	现场检查
6	修改、定稿、送审	学校	2019.04	现场检查
预计初稿完成日期		2019 年 3 月 1 日	计划定稿日期	2019 年 4 月 1 日
经费、设备、图书要求	<p style="text-align: center;">项目经费：已具备项目经费</p> <p>设备：需要 1 台附带高性能显卡的计算机（Nvidia GTX1070 及以上），已具备图书文献等需要 IEEE 文献库、Springer 文献库、arXiv 文献库，图书馆已提供</p>			

研 究 生 作 开 题 报 告	报告日期： 年 月 日
	考评成绩： 考评小组成员签名：（成员由 3-5 人组成）
	提出修改意见：
	导师意见：
<div style="text-align: right;">导师签名：</div>	
系（部）、教研室意见：	
<div style="text-align: right;">系（部）、教研室主任签名：</div>	
学院意见：	
<div style="text-align: right;">院长签名：</div>	

注：1.此表请务必用 A4 幅面纸张双面打印对折；

2.此表一式四份，研究生院培养科（原件）、院、导师及研究生本人各一份。