附件 6

中山大学大学生创新训练项目

申请书

项目编号 _____

项目名称 基于联邦学习的隐私保护 Deepfake 检测模型研究

项目负责人 孙天一 联系电话 18476892802

所在学院 计算机学院

学号 23336214 专业班级 计算机科学与技术

指导教师 郑培嘉

E-mail sunty27@mail2.sysu.edu.cn

申请日期 12 月 10 日

中山大学 教务部

删除[Iridescent]: 123

删除[Iridescent]: xs

删除[Iridescent]: 01

删除[Iridescent]: 17311113333

删除[Iridescent]: 测试

删除[Iridescent]: 学院

删除[Iridescent]: xs01

删除[Iridescent]: 测试专业

删除[Iridescent]:

删除[Iridescent]: ***

删除[Iridescent]: ****@mail.sysu.edu.cn

删除[Iridescent]: 自动获取

一、 基本情况

项目名称	基于联邦学习的隐私保护 Deepfake 检测模型研究						
所属学科	学科一级类： 计算机科学　　学科二级类： 计算机应用技术						
项目来源	■ A、学生自主选题，来源于自己对课题的长期积累与兴趣 □ B、来源于教师科研项目选题 □ C、学生承担社会、企业委托项目选题 （根据实际情况选择）						
申请金额	（按校级项目资助金额填写）	项目期限	一年期	拟申报项目级别	校级		
负责人	孙天一	性别	男	民族	汉族	出生年月	2005 年 3 月
学号	23336214	联系电话	手机： 18476892802				
指导教师	郑培嘉	联系电话	手机：				
项目简介	本项目旨在结合联邦学习技术，开发一种隐私保护的 Deepfake 检测模型。通过在不共享原始数据的情况下实现协同训练，提升检测精度，并保障用户隐私。项目将设计基于联邦学习的检测算法，应用于教育、媒体等行业，推动隐私计算和 Deepfake 检测技术的发展。						
负责人曾经参	无						

带格式表格[lridescent]

删除[Zheng]: 123

删除[lridescent]: （下拉选择）

删除[lridescent]: （下拉选择）

删除[lridescent]: xs01

删除[lridescent]: xs01

删除[lridescent]: 17311113333

删除[lridescent]: ****

删除[lridescent]: （负责人填写）

删除[lridescent]: （负责人填写）

与科研的情况						
指导教师承担科研课题情况	1. 国家自然科学面上项目，62272498，云计算环境下加密视频分析技术研究，在研，**主持** 2. 国家重点研发计划，2022YFB3103500，多媒体大数据的隐私保护技术，在研，子课题负责人 3. 国家自然科学青年科学基金项目，61502547，云计算下的加密域多媒体水印与模式匹配，结题，**主持** 4. 广东省自然科学基金青年提升项目，2023A1515030087，加密域视频水印技术研究，在研，**主持** 5. 广东省自然科学基金，2022A1515011897，加密域音频水印和语音分类技术研究，在研，**主持**					
指导教师对本项目的支持情况	提供选题研究方向指导；实验室设施及算力支持；以及项目研究学术指导。					
项目组主要成员	姓名	学号	学院	专业班级	联系电话	项目分工
	许才明	23336270	计算机学院	计算机科学与技术	18288425902	数据采集标记
	廖嘉辉	22368031	计算机学院	计算机科学与技术	13702508741	模型搭建
指导教师	姓名	工号	学院/单位	职称	联系电话	电子邮件
	郑培嘉	140207	计算机学院	副教授	13560164255	zhpj@mail.sysu.edu.cn

删除[Iridescent]: xs01

删除[Iridescent]: xs01

删除[Iridescent]: **********

删除[Iridescent]: 111

删除[Iridescent]: 测试学院

删除[Iridescent]: 测试专业

删除[Iridescent]: 测试学院

删除[Iridescent]: ***

删除[Iridescent]: 1******

二、 立项依据（可加页）

1.研究目的

（负责人填写）

随着 Deepfake 技术的快速发展，通过虚假合成人脸或视频伪造身份的技术给社会带来了严重威胁。现有的 Deepfake 检测技术大多依赖于集中化的数据训练，但在数据隐私保护和跨境数据流通方面面临挑战。

本项目旨在结合联邦学习技术，设计一套隐私保护的 Deepfake 检测模型，使参与方能够在不共享原始数据的情况下实现协同训练，提高 Deepfake 检测的精准度。通过研究，项目计划达到以下目标和意义：

- **目标**：开发一种基于联邦学习的 Deepfake 检测模型，提升检测精度，同时保护用户隐私和敏感数据安全。
- **应用价值**：支持教育、媒体、金融等行业在数据隐私保护下进行 Deepfake 内容检测，减少 Deepfake 滥用对社会带来的负面影响。
- **预期意义**：通过创新性方法，为隐私计算与 Deepfake 检测技术提供新的研究路径，助力未来网络安全发展。

2.研究内容

（负责人填写）

1. **研究对象**

- Deepfake 生成内容（包括人脸视频伪造、图像合成）的特征分析。
- 联邦学习框架在隐私保护场景下的应用。

2. **研究范围**

- 数据来源：使用公开 Deepfake 数据集（如 FaceForensics++和 DFDC），并模拟生成多样性数据。
- 模型开发：设计基于联邦学习的 Deepfake 检测算法框架。
- 系统应用：支持多参与方在分布式环境中的数据协同训练。

3. **研究方法**

- **数据预处理**：对公开 Deepfake 数据集进行特征提取和标注。
- **联邦学习算法设计**：结合加密通信与梯度更新机制，设计适合 Deepfake 检测的分布式模型。
- **模型验证与优化**：通过多轮实验评估模型性能，优化检测准确性和鲁棒性。

4. 研究步骤

- 第 1 步：数据集收集与预处理。
- 第 2 步：初步构建联邦学习框架并进行模型实验。
- 第 3 步：优化联邦学习算法，提升检测模型性能。
- 第 4 步：完成模型开发，撰写项目总结报告。

3.国、内外研究现状和发展动态

（负责人填写）

1. 国内研究现状

- 国内在 Deepfake 检测领域的研究多集中于集中式训练方式，隐私保护与分布式学习方面研究较少。
- 联邦学习主要应用于医疗、金融等敏感数据场景，缺乏在 Deepfake 检测方面的探索。

2. 国外研究现状

- 国外研究团队（如 Google AI、Meta AI）已开发出多个 Deepfake 检测工具，关注数据隐私与

算法优化。

- 联邦学习在隐私保护下的数据协同应用逐渐成为研究热点，但在 Deepfake 检测领域尚未有成熟实践。

3. 发展趋势

- 隐私计算技术（如联邦学习）逐渐与检测模型融合，支持隐私保护下的跨域协作。
- 对抗样本防御成为重点，关注模型在复杂攻击场景下的鲁棒性。

4.创新点与项目特色

（负责人填写）

1. 技术创新

- 融合联邦学习与 Deepfake 检测技术，解决传统集中式训练面临的隐私问题。
- 设计通信高效、检测精度高的联邦学习算法框架。

2. 方法创新

- 通过联邦学习机制，保证跨境或分布式数据环境下的隐私安全。
- 引入对抗训练技术，增强模型在复杂环境下的鲁棒性。

3. 应用特色

- 项目成果适用于跨行业数据协作场景，例如教育、企业数据共享平台等。

5.技术路线、拟解决的问题及预期成果

（负责人填写）

1. 技术路线图

- 数据准备：对公开数据集进行预处理，提取 Deepfake 相关特征。
- 算法设计：基于联邦学习框架，设计分布式 Deepfake 检测模型。
- 模型训练与优化：通过实验验证和改进模型鲁棒性。
- 系统开发：集成算法到演示系统中，实现验证与展示。

2. 拟解决的问题

- 如何在保护数据隐私的前提下实现高精度的 Deepfake 检测？
- 如何优化联邦学习通信机制以降低分布式环境下的计算开销？

3. 预期成果

- 研发 1 个基于联邦学习的 Deepfake 检测模型并进行系统原型演示。
- 发表 1 篇大学生学术论文或技术报告。
- 为未来深度学习与隐私保护研究提供可行性参考。

6. 项目研究进度安排

（负责人填写）

时间段	任务内容
第 1-3 月	文献调研与数据集整理，完成 Deepfake 数据预处理
第 4-6 月	构建初步联邦学习框架，设计并验证基础模型
第 7-9 月	优化算法，提高检测精度与通信效率
第 10-12 月	完成原型系统开发与性能测试，撰写项目总结报告

7. 已有基础

（1）与本项目有关的研究积累和已取得的成绩

（负责人填写）

- 团队已完成相关课程（如《人工智能基础》、《网络安全技术》）。
- 指导教师在深度学习与隐私保护方面有一定研究积累，曾指导多项类似项目。

（2）已具备的条件，尚缺少的条件及解决方法

（负责人填写）

1. 已具备的条件

- 计算资源：学校实验室提供高性能计算设备（GPU 集群）。
- 技术资源：团队掌握 TensorFlow、PyTorch 等常用深度学习框架。

2. 尚缺条件及解决策略

- 缺少对分布式框架优化的经验：通过查阅文献或外部指导获取支持。

- 跨域数据共享模拟环境：使用模拟数据或与外部企业合作建立测试环境。

三、 经费预算

开支科目	预算经费（元）	主要用途	阶段下达经费计划（元）	
			前半阶段	后半阶段
预算经费总额	4000.00	购买书籍	2000.00	2000.00
1. 业务费	2000.00	无	1000.00	1000.00
（1）计算、分析、测试费	2000.00	无	1000.00	1000.00
（2）能源动力费	0.00	无	0.00	0.00
（3）会议、差旅费	0.00	无	0.00	0.00
（4）文献检索费	0.00	无	0.00	0.00
（5）论文出版费	0.00	无	0.00	0.00
2. 仪器设备购置费	2000.00	无	1000.00	1000.00
3. 实验装置试制费	0.00	无	0.00	0.00
4. 材料费	0.00	无	0.00	0.00

四、 指导教师意见

导师：***

年　　月　　日

五、 院系推荐意见

盖 章：

年　　月　　日