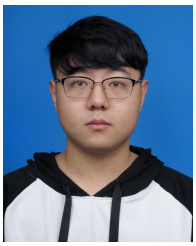

刘博超

政治面貌：中共党员
电　　话：17863928239
电子邮件：liubochao@iie.ac.cn
个人主页：刘博超
研究方向：人工智能安全、计算机视觉、模型安全

教育背景

- 中国科学院大学 · 网络空间安全学院 · 人工智能安全 · 博士（2020 年 9 月-至今）（保送）
- 山东大学 · 信息科学与工程学院 · 电子信息工程 · 本科（2016 年 9 月-2020 年 6 月）

科研方向

- **差分隐私学习**
  主要研究内容为：给定一个隐私数据集或者一个训练好的隐私模型，如何训练一个差分隐私保护的模型用于发布。针对隐私数据集问题，设计了一种判别-生成式双流蒸馏的方法训练一个隐私保护的学生模型。针对隐私模型问题，设计了一种差分隐私的无数据蒸馏方法训练一个隐私保护的学生模型。
- **生成式模型**
  主要研究内容为：训练一个隐私保护的生成式模型用于不同的下游任务。设计了一种基于哈密顿动力学的生成方案和一种基于扩散模型的生成方案。

论文列表

1. **Bochao Liu**, Pengju Wang, Shiming Ge. Learning Differentially Private Diffusion Models via Stochastic Adversarial Distillation. In ECCV, 2024. (**CCF-B，CV 三大顶会之一**)
2. **Bochao Liu**, Pengju Wang, Shikun Li, Dan Zeng and Shiming Ge. Model Conversion via Differentially Private Data-Free Distillation. In IJCAI, 2023. (**CCF-A**)
3. Shiming Ge, **Bochao Liu** and Pengju Wang. Learning Privacy-Preserving Student Networks via Discriminative-Generative Distillation. IEEE TIP, 2023. (**CCF-A，SCI-Q1**, 学生一作)
4. **Bochao Liu**, Jianghu Lu, Pengju Wang and Shiming Ge. Privacy-Preserving Student Learning with Differentially Private Data-Free Distillation. In IEEE MMSP, 2022. (**最佳学生论文提名奖**)
5. Pengju Wang, **Bochao Liu**, Shiming Ge. Fusion of Current and Historical Knowledge for Personalized Federated Learning. In IJCNN, 2024. (**CCF-C**)
6. 王鹏举, 卢江虎, **刘博超**, 葛仕明. 资源受限场景中的联邦学习技术综述. 信息安全学报, 2022. (**CCF-B**)
7. Haolin Liu, Chenyu Li, **Bochao Liu**, Pengju Wang, Shiming Ge*, Weiping Wang. Differentially Private Learning with Grouped Gradient Clipping. In ACM MM Asia, 2021. (**CCF-C**)
8. **Bochao Liu**, Shiming Ge, Pengju Wang and Liansheng Zhuang. Private Gradient Estimation is Useful for Generative Modeling. In ACM MM, 2024. (**CCF-A**, submitted)
9. **Bochao Liu**, Shiming Ge, and Tongliang Liu. Privacy-Preserving Model Transcription with Differentially Private Synthetic Distillation. IEEE TPAMI. (**CCF-A，SCI-Q1**, submitted)

授权专利

- 王伟平; 葛仕明; **刘博超**; 李晨钰。一种多方参与数据不共享的网络模型训练方法。
- 葛仕明; 刘浩林; **刘博超**; 王伟平。一种基于少量公共数据的隐私模型训练方法及装置。

实习经历

- 百度 · AI 安全组 · 2023.12-2024.5

  负责图像隐水印算法的扩容和鲁棒性研究。设计了一种基于 DCT 的图像隐水印方案；成功破解提取了一言的水印；改进了 RivaGAN（32bit—>64bit）；设计了一种基于 DiT 的水印方案。

- 华为 · 数据存储产品线 · 2024.6-至今

  负责存储服务器中的防窃取相关事项，准备设计一个病毒数据采集平台用于自动收集测试病毒数据。

项目实践

- **行业大数据安全受控共享核心系统研发（北京市科学技术委员会）-核心算法研发人员**

  该项目面向数据安全领域研发共享数据资产地图、数据分级分类管理系统和数据安全受控共享系统。本人主要负责数据安全受控共享系统中数据分布极端不均匀情况下的算法研究。

- **信息缺失情况下资源动态需求的因素分析（国家重点研发计划）-核心算法研究人员**

  该项目主要研究多维度表征与多模态数据补全技术，构建动态需求分析模型。本人主要负责针对文本数据的预测和补全算法研究。

- **Deepfake 检测竞赛（DFDC 2020）**

  该比赛的主要任务是训练一个 AI 模型去检测出那些视频是经过后期修改的。主办方 Facebook 提供了 470G 的训练数据和 400 段视频的验证集。本人参与了该竞赛获得了世界排名前 14% 的成绩。

专业技能

- 熟悉图像生成模型: GAN，DDPM，DiT 等;
- 熟悉基础模型: ResNet、Transformer 等;
- 熟悉大模型微调: LoRA，Dreambooth 等;
- 了解视频生成模型: SVD，Sora，Mora 等;
- 编程语言: Python、C++、Shell 等;
- 深度学习框架: PyTorch、TensorFlow。

获奖情况

- 中国科学院大学优秀党员（2023）
- 山东大学优秀学生奖学金（2016-2020）
- 山东大学三好学生、优秀毕业生（2020）
- 全国大学生数学建模竞赛省级三等奖（2018）
- 全国大学生数学竞赛国家级一等奖（2018）
- 社会实践省级三等奖（2018）
- 全国大学生数学竞赛国家级一等奖（2017）

学生活动

- 山东大学电子信息与物联网党支部组织委员
- 山东大学羽毛球协会副部长

兴趣爱好

- 羽毛球、乒乓球、健身