



李志豪

13289817328 zhihaoli@stu.xidian.edu.cn 西安

https://lzhms.github.io/



西安电子科技大学
XIDIAN UNIVERSITY

教育经历

西安电子科技大学

211

双一流

2021.09 - 2025.07

计算机科学与技术 本科 计算机科学与技术学院

前四学期学分绩排名: 1/89 (1%) 2022年度综合排名: 3/417 (0.72%) 2021年度综合排名: 3/70 (4.2% 新工科创新实验班)

课程成绩

数学基础成绩: 高等数学 (I) 95 高等数学 (II) 98 线性代数 97 离散数学 97 概率论与数理统计 95

计算机核心成绩: 数据结构与算法 100 数字电路与逻辑设计 97 模拟电子技术基础 93 计算机通信与网络 93
操作系统原理与设计 94 数据库系统 97 计算机组织与体系结构课程设计 97 微机原理与系统设计 93

英语能力: 四级优秀 (571) 六级通过 (523)

所获荣誉

两次国家励志奖学金 (试点班前 1%) 两次校优秀学生荣誉称号 (学校前 1%) 2022年青春楷模年度人物荣誉称号 (学校前 1%)

华萌奖学金 (学校前 2) 华为“智能基座”产教融合协同育人基地奖学金 (学校前 1.3%)

竞赛奖项

2023年中国大学生服务外包创新创业大赛 全国一等奖 (前 0.3%)

2023年全国大学生数学建模竞赛 全国二等奖 (前 1.53%)

陕西省高等数学竞赛 省级特等奖 (前 3%) 全国大学生数学竞赛 省级一等奖

2022年美国大学生数学建模大赛 国际二等奖

比赛与项目经历

2023 中国大学生服务外包创新创业大赛

所获奖项: 全国一等奖 (前 0.3%)

个人职责: 团队核心负责人, 协调团队成员的分工与合作

主要负责: 模型架构的设计与优化, 通过对训练后的模型预剪枝测试, 确定最佳的剪枝策略并进行迭代行剪枝, 并且设计 Img2Col 算法优化模型推理。



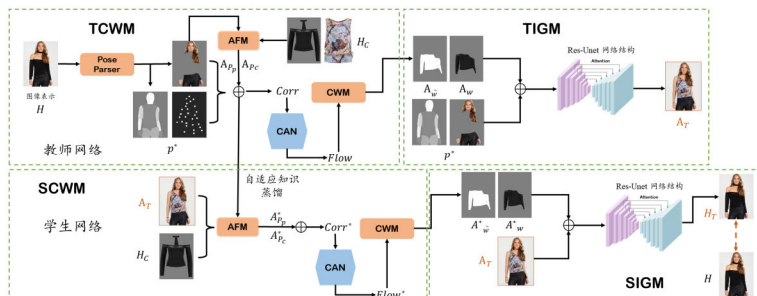
中国大学生服务外包创新创业大赛——虚拟试衣算法

项目概述

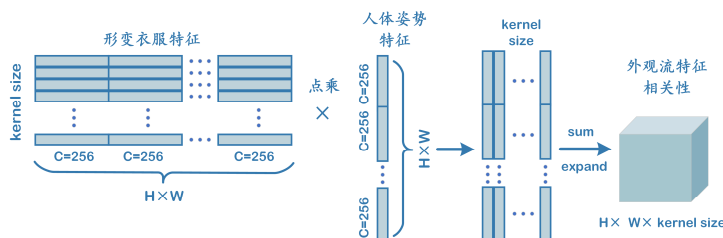
虚拟试衣指根据输入的人物图像与衣服图像, 实现衣服的试穿。本项目要求参赛者设计技术方案流程, 以及完成虚拟试衣模型训练、优化、工程化部署等工作, 最终我们模型 FID 指标达到了 8.906, 非常接近顶刊论文的结果并且工程化落地应用后模型大小从 279MB 轻量化到了 34.95MB, FID 能够维持在 11.503 水平, 压缩比达到 87.47%, 加速比达到 10.896 倍。



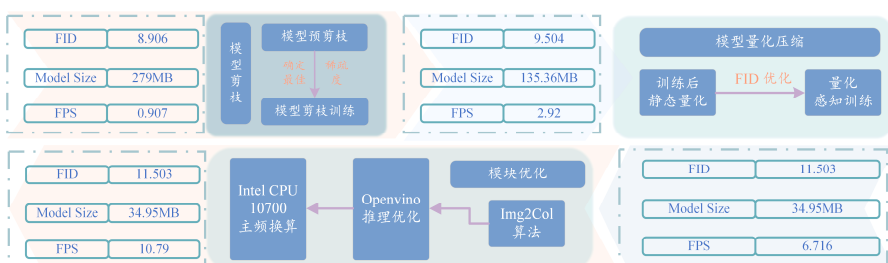
项目地址



Img2Col 算法设计



项目工程化部署总图



编程语言: 擅长 Python, 熟练运用 C、C++、MATLAB、Shell
项目工具: Git、Github、MySQL、Linux
项目习惯: 阅读论文、记录项目日志、科研汇报总结
主流框架: Pytorch, Dassel

研究规划

个人特点: 具有强烈的上进心与自驱力, 以发现新的研究结论为乐趣; 具有很强的学习能力与实践能力, 坚定终身学习的方向
研究兴趣: 大模型微调, 多模态学习, 自然语言处理
个人陈述:

我热衷于大模型的迁移学习技术, 通过调节少量参数使其迅速地适应各种下游任务, 尤其是在多模态领域中怎样使得模型迁移后仍具有较高的鲁棒性以全面拓宽应用领域, 这一难题不断激发着我的研究兴趣。随着ChatGPT等大模型的出现, 微调以及鲁棒性研究能够为人机交互以及通用人工智能的发展开辟新的可能。

我最近的研究工作主要围绕视觉语言模型 (VLMs) 以及提示学习 (Prompt Learning) 优化进行, 利用 OpenAI 发布的预训练模型, 对文本编码器与视觉编码器提取的多层特征进行深入挖掘, 旨在寻找一种匹配学习模式, 使得模型能够将每一级的视觉特征与提示的文本特征联系起来, 以克服噪声标签所带来的较大影响。这种匹配学习模式, 能够将图像特征与文本特征的匹配度分散到多级关系中, 而不仅仅只是最后的深层特征, 通过这种学习模式模型能够重视过程学习, 而非结果学习。

研究经历

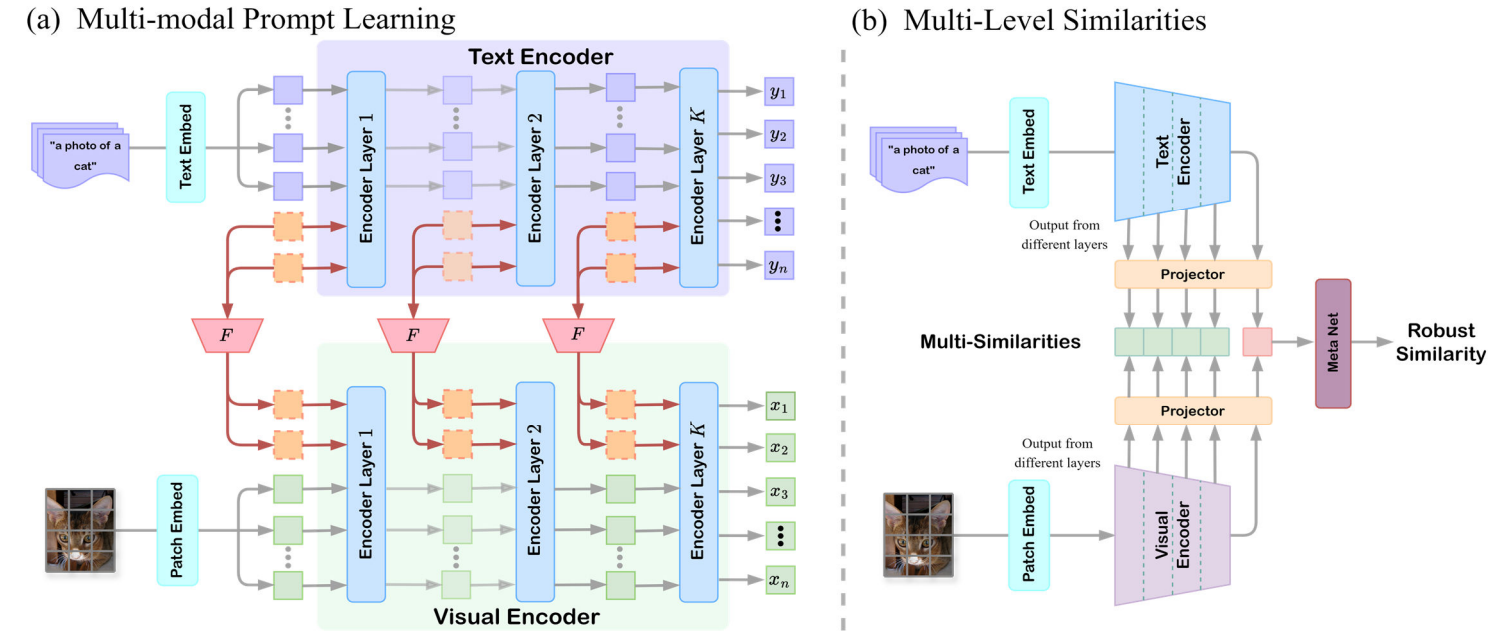
论文复现——视觉语言模型的提示学习 CoOP、CoCoOP 2023年11月 —— 2023年12月
核心观点: CoOP 首次引入可学习的提示以提取任务相关的新知识, CoCoOP 表明应基于每张图像调节提示特征以避免模型对某些类别的偏好作用;
改进思路: 在论文复现的基础上, 为了使得模型迁移到下游任务更加具有鲁棒性, 初步引入广义交叉熵损失 (GCE) 进行改进;
效果对比:

Method	Noise Rate			
	0	12.5	25	50
CoOP	63.13	59.40	54.83	44.80
RCoOP	63.67	61.90	58.37	46.30



项目地址

My Idea——视觉语言模型的鲁棒性多模态学习 2024年1月 —— 至今



核心观点: 提出一种鲁棒性的多模态学习机制, 通过将文本编码器与视觉编码器的各层提取出的特征进行匹配学习计算多级相似度, 并利用 Meta Net 网络学习最佳的相似度分布以得到鲁棒性相似度, 这样能够有效地缓解噪声标签对于模型迁移性能的影响, 使得微调后的模型能够充分利用特征级的匹配关系进行分类, 而不仅仅是标签级别;
效果表现:

Dataset	Method	Noise Rate			
		0	12.5	25	50
Dtd	MaPLe	70.70	62.93	55.43	39.73
	MaPLe+GCE	71.53	69.63	67.77	59.33
	RMaPLe	69.20	66.93	64.87	55.90
	RMaPLe+GCE	68.47	68.83	66.90	61.10



项目地址