

# 刘知远

✉ zhiyuanliu@stu.hit.edu.cn · ☎ (+86)-13961568108 · 🌐 https://tpssp.github.io

## 🎓 教育背景

哈尔滨工业大学（威海），软件工程，本科 2022.08 – 至今  
GPA: 3.97/4.0, 排名: 8/231 (Top 3.5%)  
主修课程: 线性代数 (100), 工科数学分析 (99), 大学物理 (98), 算法设计与分析 (98), 概率论与统计 (97)

## 📄 代表性成果

**Dataset Distillation with Neural Characteristic Function: A Minmax Perspective** [1]  
*CVPR 2025 (Rating: 5/5/5)*  
Shaobo Wang, Yicun Yang, Zhiyuan Liu, Chenghao Sun, Xuming Hu, Conghui He, Linfeng Zhang.

## 🔬 科研经历

上海交通大学 EPIC Lab 2024.08 – 至今  
科研实习（高效深度学习），师从张林峰助理教授  
在张林峰教授指导下，参与数据集蒸馏、知识蒸馏、大模型数据集压缩等方向研究。主要包括：

- 首次提出了基于特征函数的数据集蒸馏方法 (NCFM)，以小于 5% 的数据实现了 ImageNet Subsets 数据集的无损压缩。NCFM 将数据集蒸馏重新定义为最大最小化优化问题，利用特征函数 (Characteristic Function) 的频率参数采样策略，最大化分布差异。NCFM 相较于现有 SOTA 方法实现了 300 倍以上的 GPU 内存占用降低和 20 倍以上的速度提升。成果已发表于 CVPR 2025[1]，并取得了 5/5/5 全满分的评审成绩。  
主要贡献：实现核心算法 (CF) 模块开发；全部实验部分；论文后半部分和补充材料撰写。
- 首次将特征函数引入特征匹配型知识蒸馏方法。相比于传统基于点对点 (MSE) 或低阶矩 (MMD) 匹配的方法，CF 能够更全面地匹配教师和学生模型之间中间层特征的分布，提升模型蒸馏性能。
- 目前正在调研大模型的数据集蒸馏方向。

## 📁 项目经历

**Awesome Dataset Reduction Survey 项目** 2024.12 – 2025.01  
项目参与者，参与数据集缩减 (Dataset Reduction) 和数据集蒸馏评测 (DD-Ranking) 项目  
参与构建开源项目 Awesome Dataset Reduction，负责数据集蒸馏 (Dataset Distillation) 和应用 (Application) 领域最新论文的收集、整理与分析。同时为数据集蒸馏评测项目 (DD-Ranking) 提供了 NCFM 的评测方法和评测结果。

**基于 YOLO 算法的行人检测模型实现** 2022.11 – 2023.06  
项目负责人，已顺利结题  
基于 YOLOv5 和 DeepSORT 算法，设计并实现了一套行人检测模型。

## 🏆 荣誉奖项

国家奖学金 (Top 1%) 2023  
华为智能基座奖学金 (Top 1%) 2023  
全国大学生英语竞赛 C 类三等奖 2023  
全国大学生数学建模竞赛省二等奖 2023

## 🔗 语言与技能

英语: CET-4: 657, CET-6: 593  
技能: PyTorch, C++, Java, 较好的与人相处能力 (互评成绩班级排名 1/32)