

王 晨扬

Updated March 10, 2025

Email: chenyangwang600@gmail.com

Phone: 18749480598

Homepage: <https://chenyang4.github.io/>

Wechat: wxyangphd

研究领域

机器学习；OOD 检测/泛化，数据增强，连续学习

教育经历

哈尔滨工业大学

计算机科学与技术，博士在读
导师: 江俊君教授

黑龙江，哈尔滨
Sep. 2021 – Present

哈尔滨工业大学

计算机科学与技术，工学硕士
学分绩: 86.4, 专业前 15%
导师: 江俊君教授

黑龙江，哈尔滨
Sep. 2019 – Jun. 2021

大连理工大学

计算机科学与技术，工学学士
学分绩: 92.6, 专业前 3%

辽宁，大连
Sep. 2015 – Jun. 2019

相关技能

Coding: Python (pytorch); **English:** CET4 (576), CET6 (535)

论文发表

ReSmooth: Detecting and Utilizing OOD Samples When Training With Data Augmentation

IEEE Transactions on Neural Networks and Learning Systems (中科院一区 TOP)

Chenyang Wang, Junjun Jiang, Xiong Zhou, Xianming Liu.

Keywords: data augmentation, OOD detection, sample selection

Abstract: 数据增强 (DA) 是提高深度神经网络训练效果的常用技术。然而，高多样性的增强策略往往会引入分布外 (OOD) 样本，导致性能下降。为了解决这个问题，我们提出了 ReSmooth 框架。首先，使用高斯混合模型 (GMM) 拟合原始和增强数据的损失分布，将其分为分布内 (ID) 样本和分布外 (OOD) 样本。然后，在新训练中，对 ID 和 OOD 样本应用不同的训练策略，优化对多样化增强数据的利用。此外，我们还将 ReSmooth 框架与负数据增强 (NDA) 策略结合。通过差异化处理这些 OOD 样本，大大改善了模型性能。

Enhancing Consistency and Mitigating Bias: A Data Replay Approach for Incremental Learning

Neural Networks (中科院一区 TOP)

Chenyang Wang, Junjun Jiang, Xingyu Hu, Xianming Liu, Xiangyang Ji

Keywords: incremental learning, data replay, data consistency, classifier bias.

Abstract: 深度网络在进行连续学习时容易出现灾难性遗忘。为解决此问题，现有无数据重放方法从模型中反转样本，但存在与真实数据不一致的问题。我们定量测量数据一致性，分析现有的反转技术，开发一种新损失函数以减少不一致性。具体来说，该损失最小化反转数据和真实数据分布之间的 KL 散度。此外，我们观察到旧类权重范数在学习过程中逐渐减少，并提出正则化项来平衡类权重，使旧类样本更易区分。最终，我们提出了具有一致性增强和去偏分类器的数据重放方法，相比之前的方法性能显著提高。

其他论文

XXX for Improving Federated Domain Generalization

submitted to ICML 2025

Chenyang Wang, Junjun Jiang, Xingyu Hu, Xianming Liu

Contribution: 异质数据分布建模与学习

Learning with Noisy Labels via Sparse Regularization

IEEE/CVF international conference on computer vision 2021

Xiong Zhou, Xianming Liu, **Chenyang Wang**, Junjun Jiang, Xiangyang Ji

Incrementally Adapting Pretrained Model Using Network Prior for Multi-Focus Image Fusion

IEEE Transactions on Image Processing

Xingyu Hu, Junjun Jiang, **Chenyang Wang**, Xianming Liu, Jiayi Ma

竞赛经历

AAAI-2022 安全 AI 挑战者计划第八期：以数据为中心的鲁棒机器学习 (29/3692) 王晨扬

赛事要求: 寻找有效的以数据为中心的技术，通过数据增强、标签细化、制造对抗性数据等方法构建训练集，促进训练更加鲁棒的机器学习模型。

数据方案: 所采用的训练集包含三种不同来源的数据，即原始训练数据、增强数据以及可迁移对抗样本。

研究经历

数据增强算法研究

Aug. 2020 – Jan. 2022

哈尔滨工业大学

主要研究模型训练过程中数据增强引入的数据异质性，主要关注于增强的数据分布建模及学习策略。该工作被收录于 TNNLS。

连续学习算法研究

Feb. 2022 – Feb. 2023

鹏城实验室、哈尔滨工业大学

主要研究如何建模各连续学习阶段的数据分布，并在增量学习新数据分布的过程中，保留在旧数据分布上的性能。该工作被收录于 NN。

OOD 泛化算法研究

Sep. 2023 – May. 2024

启元实验室、哈尔滨工业大学

联邦场景下，OOD 泛化算法需同时考虑训练客户端间的分布异质性及训练客户端与测试客户端之间的异质性。该工作投稿 ICML 2025。

获得奖项	2018 年美国数学建模大赛 Honorable Mention Award; 多次获得大连理工大学学习优秀奖学金; 大连理工大学优秀本科毕业生; 哈尔滨工业大学研究生学业奖学金 (一等); 硕士研究生国家奖学金。
关注方向	<div>多模态大语言模型Dec. 2024 – Present</div> <div>Qwen、DeepSeek、Intern 等 mllm/llm 模型的各阶段创新成果及应用探索。</div>