张潇

电话: (+86) 189-9551-1421 **邮箱**: xiao_zhang@hust.edu.cn 个人主页 **GitHub** 谷歌学术

教育经历

2018.09-2021.06	硕士	华中科技大学 人工智能与自动化学院	加权：90.3/100	排名：12/189
2014.09-2018.06	学士	华中科技大学 光学与电子信息学院(卓越班)	加权：3.91/4.0	排名： 5/318

技术能力

- 熟悉 python 编程以及算法相关知识。
- 熟悉深度学习(侧重泛化能力与对抗鲁棒性)、机器学习领域的相关理论和算法。
- 熟悉主流的深度学习框架，如 Keras、Tensorflow、PyTorch。
- 具有良好的英语阅读写作交流能力（托福102）。

项目&论文

神经网络泛化能力

1) 从神经网络线性区域角度分析了**Batch Normalization (BN)、Dropout**等优化技巧对模型泛化的影响。发现相同的神经网络使用BN、Dropout等优化技巧后，其**切分特征空间的线性区域变得更加的小而密**，从而使其能够更好的拟合复杂的预测平面。

2) 通过特征空间模型预测曲面的地形分析来解释神经网络的泛化与记忆现象。传统观点认为训练过程中**神经网络逐渐学习目标函数在特征空间地形上从低频到高频的频率分量**。我们在**误差双下降实验设置**中重新思考了这个现象，发现**训练后期神经网络高频分量会再次消失**，说明神经网络架构本身就具有隐式的正则化作用。

3) 通过**偏差方差分解**分析训练过程中的泛化误差曲线，发现对于0-1误差(错误率)而言，**分解出的方差与误差具有一样的变化趋势**。我们以此为基础设计了新的指标来度量采样噪声引入的方差，使其能够在**不使用校验集的情况下预测训练过程中泛化性能曲线的变化**。该研究申请到了2020年“CCF-百度松果基金”项目。

[1] X. Zhang and D. Wu, "Empirical Studies on the Properties of Linear Regions in Deep Neural Networks," ICLR 2020. (Poster)

[2] X. Zhang, D. Wu, and H. Xiong, "Rethink the Connections among Generalization, Memorization and the Spectral Bias of DNNs," IJCAI 2021 .

[3] X. Zhang, D. Wu, and *et al*, "Optimization Variance: Exploring Generalization Properties of DNNs," work in progress.

脑机接口对抗样本

1) 针对常见脑机接口脑电信号打字机中的特征提取方法和机器学习模型构建了相应的对抗噪声。用Tensorflow重建了脑机接口中传统的特征模块以及分类器，方便计算梯度构建对抗扰动。结果表明**攻击者能够以极高的概率任意篡改脑机接口打字机的字符输出**。该研究被中国科学杂志社、EurekAlert等国内外媒体报道。

2) 针对EEGNet等脑机接口深度学习模型构建了相应的对抗噪声，分析了对抗噪声干扰的脑电信号时频区域。

3) 考虑针对时序信号构建对抗噪声时的时序因果性，提出了新的**构建通用对抗扰动噪声**的方法。比起常见的基于DeepFool的通用对抗扰动噪声，具有**更小的扰动幅度以及更高的攻击成功率**。

[1] X. Zhang, D. Wu, and *et al*, "Tiny Noise, Big Mistakes: Adversarial Perturbations Induce Errors in BCI Spellers," NSR 2020. (IF=16.69)

[2] X. Zhang and D. Wu, "On the Vulnerability of CNN Classifiers in EEG-Based BCIs," TNSRE 2019. (IF=3.34)

[3] Z. Liu*, **X. Zhang***, and D. Wu, "Universal Adversarial Perturbations for CNN Classifiers in EEG-Based BCIs ," TNSRE 2020, *submitted*. (IF=3.34)

比赛经历

2019世界机器人大赛第三届中国脑机接口比赛技术赛|一等奖 2019.06-2019.08

简介：队长。负责P300打字机的字符识别。从在线传输的EEG脑电信号中实时分析被试注视的字符（共36个）。采用了欧拉数据对齐、黎曼流形特征提取、通道选择、LR分类、bagging集成等方法，获得了第一的成绩。该成果也被长江日报、华中科技大学新闻网等媒体报道。

IEEE WCCI Open Source Intelligence Discovery for Cybersecurity Threat Awareness|第一名 2018.06-2018.07

简介：队长。对官方提供的Twitter的短文本数据进行二分类，为安全分析师提供其负责的IT基础架构安全威胁的及时信息。使用带词性标注的词袋模型构建了带阈值的在线LR模型，比起常用的如TextCNN等模型获得了更好的效果。

所获荣誉

硕士国家奖学金（2次） 三好研究生（2次） 汇顶科技奖学金 华中科技大学本科优秀毕业生

华中科技大学启明学院荣誉学生 本科国家励志奖学金 第七届全国大学生数学竞赛二等奖