中国科学院

随机复杂结构与数据科学重点实验室

学术报告

报告题目： Inductive Biases of Deep Convolutional Networks: A Theoretical Perspective

报 告 人： 吴磊(北京大学)

时　　间： 2023 年 10 月 12 日 (周四)　15:00 - 16:00

地　　点： 数学院南楼 620

报告摘要：

In this talk, we'll discuss the inductive biases of deep convolutional neural networks (CNNs), which are believed to be vital drivers behind CNNs' exceptional performance on vision-like tasks. Specifically, we'll analyze the universality of CNNs and show that achieving it requires only a depth of $O(\log d)$, where d is the input dimension. Additionally, we'll demonstrate that CNNs can efficiently capture long-range sparse correlations with only $O(\log^2d)$ samples. These are achieved through a novel combination of increased network depth and the utilization of multi-channeling and down-sampling.

We'll also explore the inductive biases of weight sharing and locality through the lens of symmetry group by introducing locally-connected networks (LCNs), which can be viewed as CNNs without weight sharing. We'll compare the performance of CNNs, LCNs, and fully-connected networks (FCNs) on a simple regression task and highlight the cruciality of weight sharing and the importance of locality. Our findings demonstrate that weight sharing and locality break different symmetries in the learning process, leading to provable separations between the two biases.