

我叫张皓然，是一名山东大学计算机学院的大三学生。我用大学前两年时间系统性地学习了计算机学科的基础知识。大三时前往青岛国创智能家电研究院（海尔）实习，以探究自己学到的计算机知识在产业中是如何被使用的。

而这次实习经历让我第一次直面“智能化”的现实困境——家电的所谓“智能”往往停留在预设规则的机械响应。我认为真正的智能应该理解人的意图，而非仅仅识别动作。这一发现促使我转向机器学习，尤其是吴恩达机器学习课程中“让机器理解人类”的愿景深深吸引了我。

从 Kaggle 寻找数据集进行分析到 MNIST 手写数字识别实验，再到用 BERT 做电影评论情感分析，我逐渐明白：单模态的感知如同盲人摸象。如何让 AI 像人类一样综合理解文字、语音、表情、姿态甚至传感器数据，成了我新的执念。

今年春季我加入了学院教授的智能体工程实践小组，以 AI 优化办公流水线自动化为目的，学习调研了智能体的相关研究与论文，并且独立完成了一个简历分析智能体的项目。这一经历使我科研实践能力得到了提升，激发了我对 ai 科研的热情。

我对这一方向的研究十分着迷，但自己探索的局限性与为数不多的本科时光都使我举步维艰。我希望在前辈的指导下，探究多模态具身推理的相关课题，走上科研道路。我认为多模态大模型的未来可能在于几何理解与语言模型的推理能力，比如让虚拟助手不仅能回答用户问题，还能根据用户实时的表情和姿势调整交互策略。

未来我希望能够打造能真正“站在用户身边”的 AI 助手——它既能理解人的物理存在，又能像 ChatGPT 般理解意图，最终模糊数字世界与现实的界限，助力世界模型。