

神经网络与深度学习课程设计

2024-秋季

一、概述

本课程设计中，你将有机会将所学知识应用到你感兴趣的问题中，课程设计选题范围如下：

□ 视觉注意力模型的分析与比较

- 在图像分类任务上针对深层卷积神经网络实现 SE, Transformer 等视觉注意力模型
- 对不同视觉注意力模型进行参数调优、分析比较各种视觉注意力模型的性能，给出合理、可靠结论
- 尝试对现有视觉注意力模型进行改进

□ 基于深度生成模型的图像生成

- 基于预训练 Diffusion 模型的图像生成改进方法
- 对 Diffusion 模型进行参数调优、分析影响图像生成质量的关键因素，给出合理、可靠结论
- 至少实现一种深度生成模型（GANs（不建议做卡通头像生成和手写数字生成任务）, Pix2Pix, CycleGAN, StyleGAN, DDPM, Stable Diffusion 等），完成图像生成

□ 多模态信息融合（尤其是大模型，可参考本课题组相关工作）

- 视觉与文本融合
- 语音与视频融合
- 多模态图像融合

□ 自选课题

- 尝试对未开源代码的顶会论文进行复现，分析其可改进的点，并实验验证
- 参与实验室的项目转化（分数适当降低，核心点贡献点必须与神经网络/深度学习相关）

选题限制：这是一门深度学习课程，因此选题应该采用深度学习方法解决问题。例如：一个传统机器学习项目或者物联网项目就不是一个好的选择。

二、提交内容要求及重要时间节点

2.1 研究建议

- ✓ 所研究的问题是什么？为什么这个问题是有趣的？（motivation）
- ✓ 你打算阅读哪些文章以提供研究背景？（related works）
- ✓ 你打算使用哪些数据？或者你打算如何收集新数据？（data）
- ✓ 你打算提出什么方法或算法或者进行什么样的改进？即使目前还没有一个确切的答案，但你应该有一个大致的了解。（introduction）
- ✓ 你打算如何评估实验结果？定性的结果还是定量的结果（例如，哪些性能指标或使用什么样的统计检验？）（evaluation）

2.2 课程报告（Final Report）

4-6 页 word 或者 pdf，采取类似深度学习会议（CVPR、NeurIPS、ICML）论文的组织结构。下面是报告的建议结构以及评分标准。不一定必须使用这样的结构，但对于大多数项目来说，这可能是一个很好的起点。

- ✓ 标题， 作者
- ✓ 摘要：简要描述研究问题、方法和关键结果，应该小于 300 字。
- ✓ 简介（10%）：描述试图解决的问题、为什么这个问题重要以及结果概述

- ✓ 相关工作（10%）：讨论相关的已发表论文。你的方法有何相似或不同之处？
- ✓ 数据（10%）：描述所使用的数据。使用的是什麼类型的数据？来自哪里？是否进行了一些预处理步骤？
- ✓ 方法（30%）：讨论解决问题使用的具体方法。为什么方法是正确的做法？你应该证明已将本学期积累的想法和技能应用于解决所选择的问题。
- ✓ 实验（30%）：介绍做的实验来证明你的方法确实解决了问题。采取的实验可能会因具体地选题而异，可以和先前的方法进行比较，进行消融实验以确定系统各种组件的作用，尝试不同的超参数来了解参数选择对结果的影响，使用可视化技术深入了解模型的工作原理，讨论算法的局限性等。这部分应该包含图片、表格来展示实验结果。
- ✓ 结论（5%）：总结主要实验结果。你学到了什么？为将来的扩展提出建议。
- ✓ 写作/格式清晰度（5%）：报告是否用词严谨且格式符合要求？
- ✓ 补充材料：不计入 4-6 页限制，作为一个单独的文件进行提交。补充材料可以包括：额外的一个 pdf 来展示更多实验结果或理论分析，源代码，短小且炫酷的演示视频等。
- ✓ 这些东西不要放在补充材料中：任何大于 200 MB 的数据集、文件，模型权重文件（特殊情况提前联系助教）。

2.3 展示 PPT（Slides）

能够快速帮助读者建立起对你课程设计的直观理解的幻灯片（ ≤ 3 页）

提交方式：提交到智慧树

提交时间：2025 年 1 月 14 日

三、参考资料

你可以从最新的高水平学术论文中寻找灵感：

- [CVPR](#): IEEE Conference on Computer Vision and Pattern Recognition
- [ICCV](#): International Conference on Computer Vision

- [ECCV](#): European Conference on Computer Vision
- [NIPS](#): Neural Information Processing Systems
- [ICLR](#): International Conference on Learning Representations
- [ICML](#): International Conference on Machine Learning
- Publications of [Bing Cao](#)
- Homepage of Bing Cao (<https://bcaosudo.github.io/>)
- [Awesome Deep Vision](#)
- [Kaggle challenges](#): An online machine learning competition website. For example, a [Yelp classification challenge](#).

还可以查看往年的课程设计。

四、学术道德

只要在文章中清楚地引用来源，可以查阅任何论文、书籍、在线参考资料或公开可用的代码。如果你使用了其他课程中做过/正在做的项目，必须清楚的标注出工作的增量部分。本课程不会容忍任何学术不端行为（每年我们都会很遗憾地看到一些同学在这门课程中越过了底线，希望各位同学将维护自己的学术信誉视为极其重要的事情）。