ZHOU BAICHUAN

Beihang University, Beijing
baichuanzhou@buaa.edu.cn ◇ Website ◇ GitHub

EDUCATION

B.Eng, Beihang University Expected 2024
School of Automation Science and Electrical Engineering
Major: Automation
Grades: 88/100 (3.64/4)
Relevant Coursework: Deep Learning(96), Probability and Statistics(100), Linear Algebra(87), Adversarial Machine Learning(90), Control Theory I&II(92 & 100).

RESEARCH EXPERIENCE

Review of Foundation Models Jul. 2023 - Aug. 2023
Project Lead Advisor: *Lei Huang*

- Reviewed and summarized papers in Vision Language Foundation Models/Agents in the last 5 years.
- Created and maintained a highly comprehensive glossary in vision language pretraining, including a comprehensive list of vision-language models and agents, a summary of related tasks and datasets, and a collection of learning resources, such as related talks and surveys. (GitHub Repo)
- A formal review paper of Foundation Models (in progress)

TinyLLaVA Dec. 2023 - Now

- Pretrained and finetuned a family of small-scale large multimodal models(LMMs): TinyLLaVAs.
- Comparable performance on VQA, GQA, SQA, TextVQA with standard 7B and 13B models.
- Open-sourced model weights, and evaluation results on HuggingFace.
- Over 3000 downloads(rapidly growing) in total on HuggingFace community.

PROJECTS

Intuitive Classification Mar. 2023 - May. 2023

- Independently built an interactive visualization tool to help users build intuitions about training a simple classification neural network, which is capable of visualizing the output representation space of neural networks.
- Provided interface to interactively tune neural networks hyper-parameters, including initialization methods, optimizer choice, normalization layer, activation function, learning rate and visualization layer. Support defining arbitrary number of layers and neurons.
- Code, related blog post and interactive demo have been open-sourced.

Give eyes to GPT-2 Jun. 2023 - Jul. 2023

- Independently pretrained and evaluated a vision-language foundation model.
- Connected GPT-2 language abilities with CLIP's vision abilities through a linear projection layer.
- Pretrained the model on CC3M, COCO2017, and evaluated the model on COCO2017 and zero-shot VQAv2.
- Code, related blog post and interactive huggingface space have been open-sourced.

SKILLS

Languages	English(IELTS 8.0, CET-6 663, CET-4 667), Chinese(native)
Programming Languages	Python(proficient), C/C++/MATLAB/Java(experience)
Deep Learning Frameworks	PyTorch, Transformers, DeepSpeed
MISC	Javascript, React, TensorFlow.js