Agneet Chatterjee
Ph.D. Student (1^{st} Year)
Computer Science
Arizona State University

Email: agneet@asu.edu
Website
Google Scholar
LinkedIn

Education

Arizona State University — Tempe, AZ
Ph.D. in Computer Science — *Spring 2023 –*

- **Advisors**: Dr. Chitta Baral, Dr. Yezhou Yang
- **CGPA**: 4/4
- **Awards**: SCAI Doctoral Fellowship

Jadavpur University — Kolkata, India
B.E (Hons) Computer Science — *2015-2019*

Research Interests

I work on problems in the Vision and Language domain. My current research interests are :

- Exploring the role of language-guided knowledge in aiding low-level computer vision tasks
- Benchmarking and improving instruction-following abilities of generative models

Technical Skills

Languages : Python, C++, JAVA, R, MATLAB

Libraries & Tools : PyTorch, CUDA, FSDP, NCCL, TensorBoard, NumPy, Matplotlib, Scikit-Learn, NLTK, OpenCV, Kornia, Huggingface, Diffusers, GIT, Linux, Amazon Mechanical Turk

Selected Projects

1. Training-Free Improvement of Spatial Fidelity in Text-to-Image Generation — Under Review, 2023

- Introduced a training-free T2I method that leverages spatially accurate image as additional guidance for diffusion models. Demonstrated state-of-the art performance on Stable Diffusion and ControlNet in generating spatially faithful images, outperforming existing open-source methods on the VISOR benchmark.
- Proposed a dataset of 190k text-image pairs, where each image is guaranteed to follow the spatial instruction mentioned in the corresponding text. Developed an extendable framework which currently covers 80 MS-COCO objects, 3 diverse backgrounds, and includes GPT-4 as an additional coordinate generator.
- Examined the trade-off between diversity and controllability introduced by our method on T2I generation. Through human studies, showed that the method is generalizable to out-of-distribution scenarios and complex prompts.

2. Quantifying and Benchmarking the Efficacy of Language-Guided Depth Estimators — Under Review, 2023

- Quantified the impact of low-level language as additional conditioning for depth estimation, in both zero-shot and supervised settings.
- Empirically showed that current use of language for depth estimation works only in a constrained setting. With adversarial inputs and domain shift, language-guided estimators exhibit lower robustness compared to vision-only depth estimators like IDisc and AdaBins.
- Systematic generation of additional knowledge for depth estimators through ground-truth segmentation and depth maps along with object-level image masking.

3. Reliability Checklist - An unified framework for reliability of NLP Systems — Under Review, 2023

- Developed a versatile tool that comprehensively assesses the reliability-related aspects of NLP systems. This tool is not tied to any specific model, can be customized for different datasets, and comes with strong visualization features.
- This tool currently supports 5 datasets from the GLUE benchmark (e.g. MNLI, MRPC), 6 discriminative models (e.g. ALBERT-base, BERT-base-uncased) and 7 task-independent augmentations.

- Through evaluations on calibration (ECE, OCE) and sensitivity based metrics, we find previously unreported failures of language models and document the trade-off between performance and robustness.

4. Autoregressive generation of interleaved image and text pairs

- To generate long range and consistent generation of image-text pairs, our method introduces learnable image tokens during training of a decoder-only autoregressive LLM's.
- Our architecture involves generating image tokens learned from an independently trained VQGAN, a MLP based adapter and example packing.

Pre-Doctoral Publications

1. A two-phase gradient based feature embedding approach J. of Information Security and Applications, 2021
Agneet Chatterjee, Soulib Ghosh, Anuran Chakraborty, Sudipta Kr Ghosal, Ram Sarkar

2. LSB based steganography with OCR: an intelligent amalgamation Multimedia Tools and Applications, 2020
Agneet Chatterjee, Sudipta Kumar Ghosal & Ram Sarkar

Experience

Salesforce June 2019 - July 2022
Software Engineer 2 *Hyderabad, India*

- Lead developer on a 15-person global team in the QTC domain, aiming to unlock a revenue of $1B in India by FY27.
- Spearheaded the development and deployment of a full-stack application to automate 70k monthly emails, creating a modular and configurable microservice-based subscription model that was adopted organization-wide, with potential earnings of $7M+.
- Developed a scalable, OAuth 2.0-secure, data extraction solution utilising Mulesoft and Heroku, facilitating 35M+ transactions and achieving a 200x performance improvement over the preceding reporting-based solution.

Samsung R&D Institute May 2018 – July 2018
Research Intern *Bangalore, India*

- As part of the the Graphics and VR team, worked on Physically Based Rendering (PBR) of 3D models creating a C# and Java-based application using Unity's Rendering API's.

Open-Source Contributions

SunPy 2017
Open-source Contributor

- Python-based implementation of a Multi-Scale Gaussian Normalisation Algorithm, as well as refactoring of existing codebase, released as part of v0.8

Honors and Awards

- SCAI Doctoral Fellowship, 2023 (ASU)
- 1st position - 2020, Summer Hackathon (Salesforce)
- INSPIRE Scholarship - 2020, Govt. of India

Service

- Reviewer for CVPR workshop (Open-Domain Retrieval Under a Multi-Modal Setting)
- Project Mentor, CSE 576 - Natural Language Processing, ASU

Courses

Natural Language Processing | Robot Learning | Artificial Intelligence | Mathematical Foundations of ML