

## EDUCATION

**Princeton University – B.S.E., Computer Science** **August 2020 – May 2024 (Expected)**  
*Certificates in Statistics and Machine Learning, Optimization and Quantitative Decision Science*

## RESEARCH EXPERIENCE

**Undergraduate Researcher, Princeton Computational Imaging Lab** **June 2023 – Present**  
*Department of Computer Science, Princeton University* *Advisor: Prof. Felix Heide*  
**Research Assistant, Princeton Natural Language Processing Group** **May 2022 – Present**  
*Department of Computer Science, Princeton University* *Advisor: Prof. Karthik Narasimhan*  
**Summer Lab Assistant, Princeton Visual AI Lab** **June 2021 – August 2021**  
*Department of Computer Science, Princeton University* *Advisor: Prof. Olga Russakovsky*

## INDEPENDENT RESEARCH PROJECTS

**Interpretable Object Tracking via Inverse Neural Rendering** **June 2023 – Present**  
Julian Ost\*, Tanushree Banerjee\*, Mario Bijelic, Yuval Bahat, Felix Heide  
*Princeton Computational Imaging Lab, Princeton University* *Advisor: Prof. Felix Heide*

- Recast visual inference tasks as inverse neural rendering (INR) problems. Evaluated our approach on multi-object tracking of vehicles in the wild using the Waymo and nuScenes datasets.
- To realize this approach efficiently, we first generatively pre-trained a neural representation model on cars in the ShapeNet v1 dataset and finetuned this model on images of vehicles from the Waymo and nuScenes datasets. We then performed INR by optimizing over the latent space of this pre-trained model to find a latent code that most closely represents the observed image of the detected car from the object detection stage in the canonical 3D object tracking pipeline. Finally, the latent code thus obtained served as an interpretable feature in the object association stage of the canonical 3D object tracking pipeline.
- Our method achieved performance comparable with other monocular MOT baselines. This work is in submission at CVPR 2024. I aim to improve our object representation model to better generalize to realistic cars as observed in the Waymo and nuScenes datasets.

**Lie Detection in the Diplomacy Game using Large Language Models** **May 2022 – Present**  
Tanushree Banerjee, Richard Zhu, Runzhe Yang, Denis Peskov, Brandon Stewart, Karthik Narasimhan  
*Princeton Natural Language Processing Group, Princeton University* *Advisor: Prof. Karthik Narasimhan*

- Proposed a bootstrapping framework that leverages self-generated feedback to enhance LLM reasoning capabilities for nuanced natural language inference tasks. This framework consists of three stages: (i) the *suggestion* stage, where a cost-effective LLM generates an initial prediction; (ii) the *feedback collection* stage, where an LLM provides feedback on initial predictions; (iii) the *modification* stage, where an LLM refines the initial predictions based on the auto-generated feedback.
- Investigated our proposed framework for detecting deception in Diplomacy games by collecting a novel dataset of human feedback on initial predictions and comparing the modification stage performance when using human feedback rather than LLM-generated feedback. Our LLM-generated feedback-based approach achieved superior performance, with a 39% improvement over the zero-shot baseline in lying-F1 without any training required.
- Submitted work to ARR August 2023, and plan to revise and resubmit the manuscript to ACL 2024 with experiments using LLMs outside of the GPT series.

**Reducing Object Hallucination in Visual Question Answering** **January 2023 – May 2023**  
*Independent Work Project, Spring 2023, Princeton University* *Advisor: Prof. Olga Russakovsky*

- Studied the tendency of Visual Question Answering (VQA) models to predict non-existent objects when queried with an image and an unrelated question. Devised a procedure to quantify the extent to which VQA models can identify unrelated questions based on the tradeoff between the risk of failing to identify an irrelevant question and the number of related questions incorrectly identified as unrelated.
- Proposed several approaches to identify questions unrelated to an image. The best approach involves quantifying the extent to which an off-the-shelf VQA model relies on the text input relative to the image input to determine whether the input image-question pair is unrelated. This approach achieves a 40% improvement over the random baseline.

**Bias in Skin Lesion Classification** **January 2022 – May 2022**  
*Independent Work Project, Spring 2022, Princeton University* *Advisor: Prof. Olga Russakovsky*

- Trained a linear classifier head on top of the pre-trained AlexNet model on the Fitzpatrick17k dataset consisting of images of 114 different categories of skin lesions on a diverse set of skin tones. Evaluated the degree to which this trained model is biased on darker versus lighter skin tones in this dataset using a bootstrapping method to find the discrepancy in the classification accuracy on lighter vs. darker skin tones.
- Found that the confidence interval of the classification accuracy on darker skin tones was much larger than the confidence interval

of the classification accuracy on lighter skin tones for the same number of examples, suggesting that although the discrepancy in accuracy was within the error, the performance of the model on darker skin tones was much more inconsistent, making the model more unreliable for skin lesion classification on darker skin tones.

### **Police Body-Worn Camera Project**

*Princeton Visual AI Lab, Princeton University*

**June 2021 – August 2021**

*Advisor: Prof. Olga Russakovsky*

- Worked towards curating a novel human action recognition dataset consisting of videos “collected in the wild” from noisy police body-worn camera footage, aimed to be a new, more challenging application-oriented video action recognition benchmark rather than existing “curated” benchmarks for human action recognition.
- Built a system to clean up ~20TB of raw body-worn camera footage to capture relevant, interesting clips while discarding clips with no action occurring.

### **SELECTED COURSE PROJECTS**

#### **Counterfactual Analysis for Spoken Dialogue Summarization**

**November 2023 – Present**

Tanushree Banerjee\*, Kiyosu Maeda\*, Sanjeev Arora

*Fundamentals of Deep Learning Course Project, Fall 2023*

*Advisor: Prof. Sanjeev Arora*

- Exploring how upstream task errors, such as speaker identification and Automatic Speech Recognition accuracy, impact the quality of spoken dialogue summarization. Generate simulated “counterfactual” errors to control the level of error introduced. Exploring how well LLMs can summarize spoken dialogues with varying error levels.

#### **Towards Efficient Frame Sampling Strategies for Video Action Recognition**

**January 2023 – May 2023**

Tanushree Banerjee\*, Brian Lou\*, Ameya Vaidya\*, Olga Russakovsky

*Computer Vision Course Project, Spring 2023*

*Advisor: Prof. Olga Russakovsky*

- Proposed optical flow-based frame sampling strategies for video action recognition models to help select the most salient frames for computationally efficient action recognition.
- The proposed strategy achieved an 11.7% improvement over a uniform random sampling, indicating that optical flow contains some signal for the saliency of a frame for action recognition.

#### **[Re] Double-Hard Debias: Tailoring Word Embeddings for Gender Bias Mitigation**

**January 2022 – May 2022**

Tanushree Banerjee\*, Jessica-Ann Ereyi\*, Kevin Castro\*, Danqi Chen, Karthik Narasimhan

*NLP Course Project, Spring 2022*

*Advisors: Prof. Danqi Chen, Prof. Karthik Narasimhan*

- Reproduced the results from “Double-Hard Debias: Tailoring Word Embeddings for Gender Bias Mitigation” (Wang et al., 2020) to reduce gender bias present in pre-trained word embeddings. In addition, we compared the original paper’s results to five additional baselines on two benchmarks on a more robust bias metric. We also evaluated the Double-Hard Debias method on Spanish word embeddings to see if this method generalized well to languages other than English.
- Although our results were similar to the original paper on English word embeddings, the Double-Hard Debias method performed significantly worse on Spanish word embeddings, indicating that this method may not generalize well to other languages.

### **ACADEMIC SERVICE**

#### **Statistics and ML Certificate Ambassador, Princeton University**

**June 2023 – Present**

- Organize events, such as lunches with faculty, and guide students for course selection.

#### **Princeton AI4ALL Research Instructor, Princeton University**

**July 2022**

- Taught AI technology and policy to rising 11th graders from underrepresented groups. Led the NLP workshops and developed coding tutorials and lectures in preparation for an NLP-based capstone project. Organized guest lectures by Princeton faculty.

### **OTHER WORK EXPERIENCE**

#### **Circulation Desk Assistant, Princeton University Engineering Library**

**August 2021 – November 2021**

- Manage the main circulation desk, guide patrons, and answer visitors’ questions about the library space.
- Assist visitors with checking books in and out of the library.

#### **Software Engineering Intern, Nautilus Software Technologies, Hong Kong**

**December 2020 – January 2021**

- Recreated the code-breaking game ‘Mastermind’ in chatbot format on Facebook Messenger using JavaScript.

### **HONORS**

#### **Nominee, Computing Research Association (CRA) Outstanding Undergraduate Research Award**

**October 2023**

Nominated by the Princeton Department of Computer Science for contributions to research as an undergraduate

### **EXTRA-CURRICULAR ACTIVITIES**

#### **Strawberry Milk (Student Rock Band)**

**September 2021 – August 2022**

- Founding member and guitarist. Performed at student social events, including Lawnparties Spring 2022.