

VATSAL AGARWAL

📞 978-328-8283

✉ vatsalag99@gmail.com

🌐 [linkedin.com/vatsalag99](https://www.linkedin.com/in/vatsalag99)

🐙 github.com/vatsalag99

Education

University of Maryland, College Park

Ph.D. Computer Science

Advisor: Abhinav Shrivastava

Sep. 2022 – May 2027 (Expected)

College Park, MD

Overall GPA: 3.9

University of Maryland, College Park

Bachelor of Science in Computer Science

Gemstone Honors College, Departmental CS Honors (Presidential Scholarship)

Sept. 2018 – May 2022

College Park, MD

Overall GPA: 3.96

Research Interests

My research focuses on building effective multimodal models for scene understanding. Recently, I have been working on using explicit memory for open-vocabulary segmentation as well as leveraging diffusion models for building effective MLLMs. Previously, I have worked on improving attention mechanisms for classification and mesh recovery.

Graduate Experience

Apple

May 2024 – Present

Graduate Machine Learning Intern — Eshan Verma

Cupertino, CA

- Building efficient multimodal large-language models using text-to-image diffusion models as task-adaptive visual encoders.

PAII Inc

June 2022 – Nov 2022

Graduate Research Intern — Dr. Youbao Tang and Dr. Mei Han

Remote (San Francisco, CA)

- Developed deep learning model for precise human shape and pose estimation from a single image using Transformers
- Devised method to recover human mesh via a coarse-to-fine approach that is able to maintain performance with up to 2x reduction in parameters and 10x reduction in FLOPs.

Undergraduate Experience

MIT CSAIL

June 2021 – Oct 2021

Undergraduate Research Intern — Dr. Antonio Torralba

Remote (Cambridge, MA)

- Developed methods for Sim2Real task, by translating synthetic indoor scenes from [VirtualHome](#) environment to realistic images using text as an intermediary.
- Built automated pipeline using VirtualHome API to create image-text dataset. Used API-generated segmentation maps to filter out objects and obtain object relations. Wrote templates to ensure diverse caption generation.

Perception and Intelligence Lab

March 2021 – May 2022

Undergraduate Research Assistant — Dr. Abhinav Shrivastava

College Park, Maryland

- Investigated adversarial attacks and defenses in context of frequency domain and experimented with attention methods to discover most important frequencies.
- Devised method to enhance multi-scale interactions by using high-level features to select important information from lower-level features. Integrated module to PvT architecture and demonstrated strong improvement in performance.

National Institutes of Health Clinical Center

June 2019 – Nov 2021

Deep Learning Research Intern — Dr. Youbao Tang and Dr. Ronald Summers

Bethesda, Maryland

- Developed a weakly-supervised attention-based co-segmentation model adapted from [arXiv:1810.06859](#) and applied it for the task of lesion segmentation using the [DeepLesion](#) dataset.
- Experimented with different channel/spatial attention mechanisms and post-processing strategies to improve model performance. Incorporated mechanisms to aggregate multi-scale features and obtain high-resolution features.
- Designed novel attention methods to improve co-segmentation models.

Gemstone Honors College

Jan 2019 – May 2022

Student Researcher — Dr. Anil Deane

College Park, Maryland

- Worked as part of team of 13 students to apply deep learning methods to understand psychiatric illnesses using MRI scans. Leading development of models to discriminate between healthy and mentally-ill patients.
- Built supervised/unsupervised models to identify imaging biomarkers corresponding with mental illnesses.

- Created a UI for more streamlined STORM imaging analysis using PyQt5. Cleaned existing analysis codebase for increased modularity and ported several MATLAB codes to Python for more cohesive pipeline.

Technical Skills

Programming Languages: Python, Java, SQL, C, SAS, Ruby, OCaml, Visual Basic

Tools: Pytorch, Keras, Scikit-Learn, Pandas, NumPy, SciPy, Matplotlib, Plotly, OpenCV, PyQt5

Frameworks: Git, AWS Sagemaker, Google Cloud, Anaconda, Jupyter, Docker

Relevant Coursework: Algorithms, Applied Statistics, Advanced Visual Recognition, Computational Linguistics, Data Science, Deep Learning, Machine Learning, Linear Algebra, Multivariate Calculus, Probability Theory, Statistics

Projects

Deformable 1D Convolutions | *Python, Pytorch*

October 2020

- Adapted [Efficient Channel Attention](#) mechanism and modified it by experimenting with shuffling operation and deformable convolution to more effectively capture global interactions.

Animating Pokemon with Donkey-Net | *Python, Pytorch*

April 2020

- Modified [Monkey-Net](#) model for dual reconstruction of moving key-points using cycle-consistency loss and created custom dataset using affine similarity and frame re-sampling. Experimented with attention mechanisms and multi-scale perceptual losses to improve performance.

Publications

- [Do text-free diffusion models learn discriminative visual representations?](#)

Soumik Mukhopadhyay, Matthew Gwilliam, **Vatsal Agarwal**, Yosuke Yamaguchi, Namitha Padmanabhan, Archana Swaminathan, Tianyi Zhou, Abhinav Shrivastava
Under Review

- [Diffusion Models Beat GANs on Image Classification](#)

S. Mukhopadhyay, M. Gwilliam, **V. Agarwal**, N. Padmanabhan, A. Swaminathan, S. Hegde, T. Zhou, A. Shrivastava

- [A Frequency Perspective of Adversarial Robustness](#)

S. R. Maiya, M. Ehrlich, **V. Agarwal**, S. Lin, A. Shrivastava
BMVC 2023

- [Volumetric super-resolution imaging by serial ultrasectioning and STORM in neural tissue](#)

T. Vatan, J. A. Minehart, C. Zhang, **V. Agarwal**, J. Yang, C M. Speer
Star Protocols, 2021

- [Weakly Supervised Lesion Co-Segmentation on CT Scans](#)

V. Agarwal, Y. Tang, J. Xiao, R. Summers
IEEE International Symposium on Biomedical Imaging (ISBI), 2020

- [Weakly-Supervised Lesion Segmentation on CT Scans using Co-Segmentation](#)

V. Agarwal, Y. Tang, J. Xiao, R. Summers
SPIE Medical Imaging, 2020

Awards

CRA Undergraduate Research Award | *Honorable Mention*

December 2021

Northrop Grumman Cyber-Security Challenge | *Python, Raspberry Pi*

March 2019

- Developed Python application for keyboard dynamics classification using self-organizing maps for anomaly detection and classifying users based on keystroke dynamics.
- Won **third place** award as part of three-person team against undergraduate teams for choice of problem and efficacy of model.

Northrop Grumman Image Recognition Challenge | *Python, Raspberry Pi*

November 2018

- Developed Python application using YOLOV2 algorithm to detect emergency vehicles. Utilized Keras and Tensorflow for data augmentation and design of convolutional neural network for object detection.
- Won **first place** award as part of a three-person team against undergraduate teams for model accuracy and speed