MATTHEW WALMER

College Park, MD 20740 • Email: mwalmer@umd.edu • Website: matthewwalmer.com

EDUCATION

UNIVERSITY OF MARYLAND, COLLEGE PARK, College Park, MD
Pursuing a PhD in Computer Science, focused on Computer Vision, 2019 – Present (GPA 3.97)
Thesis Title: Improving Deep Networks by Learning from their Failures and Successes
Proposal Committee: Abhinav Shrivastava (Advisor), David Jacobs, David Forsyth

JOHNS HOPKINS UNIVERSITY, Baltimore, MD
Master of Science in Engineering (MSE) in Biomedical Engineering, 2017 (GPA 4.00)
Bachelor of Science in Biomedical Engineering, 2016 (GPA 4.00, Dean's List)
With Minor Degrees in Computer Science and Mathematics
Received the Richard J. Johns Award for Academic Excellence

SKILLS

Tools & Libraries: *Python, PyTorch, TensorFlow, LaTeX, OpenCV, Scikit-learn, Image Editing*
Deep Learning: *Transformers/ViTs, Diffusion, Self-Supervised Learning, Adversarial Attack/Defense*

WORK EXPERIENCE

PERCEPTION AND INTELLIGENCE LAB (UMD) – **PhD Student Research Assistant** Apr 2020 – Present
- Computer Vision PhD Student working with Professor Abhinav Shrivastava.
- Published multiple first-author papers at top-tier computer vision conferences.
- My current research focuses on network understanding, interpretability, and enhancement, primarily for self-supervised models and Vision Transformer (ViT) architectures.

META – **Summer Research Internship** May 2023 – Nov 2023
- Conducted research in video representation learning for fine-grained tasks for a conference publication.
- Designed a specialized transformer architecture leveraging multi-entity video scene processing.
- Collaborated with New York-based Meta team members during a summer internship and fall extension.

OMNISPEECH – **Machine Learning Consultant** Jul 2022 – Aug 2022
- Consulted on the development of machine learning algorithms for hearing assistance devices.
- Optimized training pipelines for faster training and design iteration.

SRI INTERNATIONAL – **Summer Research Internship** Jun 2021 – Aug 2021
- Studied backdoor attacks in multimodal neural networks as part of the TrojAI program.
- Investigated the impact of multimodal learning on the effectiveness of backdoor attacks.
- Wrote and published "Dual-Key Multimodal Backdoors for Visual Question Answering" at CVPR 2022.

THE MITRE CORPORATION – Mar 2018 – Aug 2019 (full-time)
Computer Scientist, Computer Vision Specialty Aug 2019 – Sep 2020 (part-time)
- Co-PI of a MITRE Innovation Program research project that developed a publication accepted to ECCV 2020.
- Studied the in-the-wild effectiveness of adversarial patch attacks on object detection models, and released a large dataset that is used as a standard benchmark for defensive research.
- Programmed on a team to develop, refine, and improve a modular framework for multimedia analysis.

NEUROMEDICAL CONTROL SYSTEMS LAB (JHU) – **Master's Student Researcher** Nov 2016 – Jul 2017
- Studied single unit recordings of human neural circuits by applying various modeling techniques.
- Wrote and published a research paper and presented it at EMBC 2017 in Jeju, South Korea.
- This work was also included in a textbook, written by Zhe Chen and lab PI Sridevi V. Sarma.

MATTHEW WALMER

College Park, MD 20740 • Email: mwalmer@umd.edu • Website: matthewwalmer.com

PUBLICATIONS

UPLiFT: Efficient Pixel-Dense Feature Upsampling with Local Attenders IN REVIEW

- Proposed an efficient method to upsample coarse pre-trained visual features into pixel-dense features.
- Proposed a new Local Attender module based on a neighborhood of fixed spatial offsets.
- Achieved state-of-the-art performance on several dense tasks with better efficiency and scaling.

Trokens: Semantic-Aware Relational Trajectory Tokens for Few-Shot Action Recognition ICCV 2025

- Proposed a pipeline to leverage object tracking and smart token selection for video classification.
- Used these trajectory-aligned tokens (Trokens) to model intra- and inter-trajectory dynamics.
- Achieved state-of-the-art performance on several few-shot action recognition benchmarks.

Multi-entity Video Transformers for Fine-Grained Video Representation Learning CVPRW 2025

- Proposed and trained a specialized video representation learning architecture for temporally dense task.
- Achieved state-of-the-art results on several fine-grained video benchmarks with self-supervised learning.
- Demonstrated superior self-supervised foreground actor isolation over base DINO model.
- Presented this research at the CVPR 2025 Workshop on Fine-Grained Visual Categorization (FGVC).

LiFT: A Surprisingly Simple Lightweight Feature Transform for Dense ViT Descriptors ECCV 2024

- Developed a compact support network to enable efficient generation of high-resolution ViT features.
- Demonstrated significant performance benefits for dense tasks using these enhanced features.
- Presented this research at ECCV 2024 in Milan, Italy.

TIJO: Trigger Inversion with Joint Optimization for Defending Multimodal Backdoored Models ICCV 2023

- Consulted on the development of multimodal defenses for Dual-Key Backdoor Attacks.
- Created the TrojVQA dataset, which was used for this defensive analysis.
- Demonstrated the importance of joint multimodal optimization to create effective defenses.

Teaching Matters: Investigating the Role of Supervision in Vision Transformers CVPR 2023

- Conducted an in-depth comparison of Vision Transformers (ViTs) trained with different supervision methods.
- Identified distinct patterns in multi-headed attention layer behavior depending on supervision group.
- Studied trends in learned representation similarity and usability for downstream tasks.
- Provided extensive analytical results to guide feature and layer selection for various applications.
- Presented this research at CVPR 2023 in Vancouver, Canada.

Dual-Key Multimodal Backdoors for Visual Question Answering CVPR 2022

- Work conducted during a summer internship with SRI International.
- Performed the first exploration of backdoor attacks in multimodal neural networks.
- Examined the impact of backdoor learning and input modality bias in a multimodal trigger regime.
- Developed a novel attack strategy for multimodal models and released a large dataset for defense research.
- Presented this research at CVPR 2022 in New Orleans.

APRICOT: A Dataset of Physical Adversarial Attacks on Object Detection ECCV 2020

- Developed and released a dataset to study the “in-the-wild” effectiveness of physical adversarial patch attacks on object-detecting neural networks, available at apricot.mitre.org.
- Investigated defensive strategies for detector models to combat real-world patch attacks.
- The APRICOT dataset is also included as a benchmark in the DARPA GARD program.
- Presented this research at the ECCV 2020 virtual conference.

Neuronal Activity in Human Anterior Cingulate Cortex Modulates with Internal Cognitive State During Multi-Source Interference Task EMBC 2017

- Work conducted with the Neuromedical Control Systems Lab at Johns Hopkins during master’s studies.
- Formulated Poisson point process models for single-unit recordings taken in human subjects.
- Created a behavioral model based on experiment stimulus and subject reaction times and showed that the activity of some neurons correlated with this model with high statistical significance.
- Presented this research at EMBC 2017 in Jeju, South Korea.