# 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)

**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*
***Deep Learning:*** *Transformers/ViTs, Diffusion, Trojan Networks, Adversarial Examples*

## WORK EXPERIENCE

PERCEPTION AND INTELLIGENCE LAB (UMD) - **PhD Student Research Assistant** Apr 2020 – Present
- Computer Vision PhD Student working with Professor Abhinav Shrivastava.
- Performing research on transformer network interpretability and representation learning for various tasks.

META - **Research Internship** May 2023 – Nov 2023
- Conducted research in video representation learning for fine-grained tasks for a conference publication.
- 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 and pipelines for hearing assistance.

SRI INTERNATIONAL - **Summer Internship** Jun 2021 – Aug 2021
- Studied backdoor attacks in multimodal neural networks as part of the TrojAI program.
- 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 published at ECCV 2020.
- Programmed on a team to develop, refine, and improve a modular framework for multimedia analysis.

## SELECT PUBLICATIONS

**Multi-entity Video Transformers for Fine-Grained Video Representation Learning** FGVC 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.
- This work was presented at the CVPR 2025 Workshop on Fine-Grained Visual Categorization.

**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.

**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.

**Dual-Key Multimodal Backdoors for Visual Question Answering** CVPR 2022
- Performed the first exploration of backdoor attacks in multimodal VQA neural networks.
- Developed a novel attack strategy for multimodal models and released a large dataset for defense research.

**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.
- Investigated defensive strategies for detector models to combat real-world patch attacks.