Mayug Maniparambil
SFI Centre for Research Training in Machine Learning
mayugmaniparambil@gmail.com
+00353899754851
Linkedin / Scholar / Website

PROFILE

My research interests span the areas of machine learning and computer vision, specifically in making computer vision models learn with limited annotations by making use of semantic information from language models.

I am currently working on efficient multimodal alignment by leveraging concepts from representational similarity analysis and the hypothesis of universal, or "platonic," representations.

Prior to this, I worked as a computer vision researcher at Qure.ai, Mumbai, where I developed cranial bleed segmentation/ classification algorithms, using weak supervision and active learning techniques. I also have had the privilege of collaborating at the Computational Imaging Lab at IIT Madras, guided by Prof. Kaushik Mitra, where my focus lay on deep image priors for addressing ill-posed inverse challenges in computational imaging.

I hold a B. Tech and M. Tech in Electrical Engineering from the Indian Institute of Technology, Madras.

EDUCATION

2021-2025 PhD in Machine Learning

ML Labs - University College Dublin, Dublin City University

2013-2018 B Tech, M Tech, Electrical Engineering with specialization in signal processing,

Indian Institute of Technology Madras; GPA: 8.5/10

PROGRAMMING & TECHNOLOGY SKILLS

- **Fluent in python, pytorch and tensorflow**
- **Fluent in signal processing, image processing, computer vision and NLP techniques.**

RESEARCH EXPERIENCE

2021 - Vision Language Models and Pretraining under Prof Kevin McGuiness (RIP) and Prof Noel O'Connor

- **From Unimodal to Multimodal: Scaling up Projectors to Align Modalities** (CVPR 2025)

  Developed a novel vision-language model alignment framework that trains lightweight projection layers between frozen unimodal vision and language encoders, leveraging semantic similarities in their embedding spaces. Demonstrated a 76% accuracy on ImageNet with 20x less data and 65x less compute than CLIP, showcasing the framework's efficiency across zero-

shot classification and image-text retrieval tasks. Enabled flexible adaptation to multilingual, cross-domain tasks and enhanced applications in robotics and computer vision systems through modular, accessible model development.

- **Do Vision and Language Encoders Represent the World Similarly? (CVPR 2024)**

  Researched alignment between vision and language encoders not trained jointly, using Centered Kernel Alignment (CKA) to analyze latent space structures. Demonstrated semantic similarities between unaligned and aligned encoders, offering effective methods for matching without linear layer training. Proposed Fast Quadratic Assignment Problem (QAP) optimization and localized CKA metric-based matching. Successfully applied these methods in cross-lingual, cross-domain caption matching, and image classification tasks. Concluded that unaligned encoders have valuable semantic alignment potential. Tested on COCO, NoCaps, and ImageNet-100 datasets. Implemented practical cross-lingual image retrieval using sentence transformers in various languages and a CLIP vision encoder in English.

- **Test-Time Adaptation with SaLIP: A Cascade of SAM and CLIP for Zero shot Medical Image Segmentation (CVPR 2024-Workshop Oral)**

  Developed SaLIP Framework for Medical Image Segmentation: Pioneered the integration of SAM (Segment Anything Model) and CLIP vision encoders into the SaLIP framework for zero-shot medical image segmentation. This innovative approach combines part-based segmentation with targeted mask retrieval, allowing precise organ segmentation without training or fine-tuning. Achieved state-of-the-art results in segmentation accuracy across multiple organs. The method minimizes the need for extensive datasets and domain expertise, enhancing accessibility and efficiency in medical imaging applications. Published resources to encourage further adoption and development.

- **Enhancing CLIP with GPT-4: Harnessing Visual Descriptions as Prompts (ICCV-W 2023).**

  We show that visually descriptive textual (VDT) information can be used to improve the 0-shot domain transfer performance of VLM's like CLIP. We then use GPT-4 to produce such VDT in a scalable way and show SOTA 0-shot transfer performance on a suite of 12 datasets. We also design a few-shot adapter that makes use of this VDT information to produce generalizable classifiers that show SOTA few-shot adaptation performance in the Base-to-New setting on the benchmark suite of 12 datasets.

- **BaseTransformers: Attention over base datapoints for one shot learning (accepted BMVC 2022).**

  Developed a novel few shot learning algorithm that applies attention over base datapoints to improve prototype representation in one shot learning. Our method achieves SOTA in few shot benchmarks Mini-ImageNet and Tiered-ImageNet in the inductive setting.

2017– 2018 Phase Retrieval for Fourier Ptychography (FP) Microscopy Dual Degree Project under Prof. Mitra

- **Developed an auto-encoder based architecture that can be adapted for performing FP phase retrieval under** varying levels of overlap by using same generator network, but different optimization frameworks depending on the amount of overlap.
- For low overlap case, supervised learning-based technique that learns a conditional prior to map the low-resolution measurements to its high-resolution phase and amplitude is used.
- For higher overlap, a novel non-data driven framework, that optimize over the generator parameters by minimizing the forward measurement error of FP is used. The proposed method exploits the low-level image statistics captured by generator network's inherent structure, thereby making it more robust to phase-amplitude leakage.
- Using simulations for uncorrelated phase and amplitude in both low and high overlaps, the proposed algorithm outperformed previously proposed FP phase retrieval

2017– 2018 Medical Image Denoising Dual Degree Project under Prof. Kaushik Mitra

- **Developed a GANs based Generative model to in-paint serial two-photon tomography images of mouse brain** in collaboration with Dr. Pavel Osten's Lab, Cold Spring Harbor Laboratory. Proposed an encoder-decoder network with adversarial loss to remove high expression regions in mice brains by inpainting that matches the context.
- Gained a solid understanding of various generative models (GANs, variational autoencoders, context encoders)

INDUSTRY EXPERIENCE

2024-2025 (6 months) Applied Scientist Intern, Amazon Robotics, Berlin

- Worked on Vision Language Models, and Domain Adaptation for robot induced defects monitoring under very limited data scenarios.
- Trained models for detecting multi-drop and robot induced damage from multi-view videos of the Vulcan pick extraction robot using synthetic datasets and domain adaptation.

2018 – 2021 Data Scientist Qure.ai, Fractal Analytics, Mumbai

- Responsible for improving bleed segmentation and bleed quantification algorithms in head CT scans. Designed a novel model, that makes use of large dataset of coarse annotations and small dataset of fine annotations to give significant improvements in mAP scores.
- Responsible for improving bleed detection algorithms in head CT scans. Measures like composite scaling of depth, width and resolution of networks, auto augmentation search, implementing squeeze-excitation blocks etc. lead to significant improvements in tracked metrics. (AUC, AP)
- Developed models to localize and report textually, the location of fractures in cranial scans. Used a UNet based model to segment out bones in cranial scans and used the segmentation maps in conjunction with predictions from a fracture segmentation model to determine fracture location.
- Developed an efficient workflow for medical image annotations by interns. Used pretrained models to prepopulate the scans with abnormality masks to speed up annotations process.
- Worked on an FDA accepted study for the product that does bleed localization and quantification in head CTs.

TEACHING EXPERIENCE

- Teaching Assistant for Advanced Machine Learning, DCU
- Teaching Assistant for Image Processing, DCU
- Teaching Assistant for Systems and Control Theory, DCU
- Teaching Assistant for Introduction to Java, DCU
- Teaching Assistant for Advanced Electrical Lab for Undergraduate Students, IIT Madras
- Teaching Assistant for Digital Image Processing for Undergraduate Students, IIT Madras

FUNDING AND AWARDS

- ML Labs/SFI funding for Ph.D. in Machine Learning for four years.
- 3rd place in Huawei Data Science Hackathon 2022.
- Finalist in STOIC Covid 19 AI challenge to predict Covid severity from CT scans in 2021.

PUBLICATIONS AND CONFERENCES

- Paper "Understanding Space Is Rocket Science - Only Top Reasoning Models Can Solve Spatial Understanding Tasks" (under review Neurips 2025)
- Paper "Pinpoint Counterfactuals: Reducing social bias in foundation models via localized counterfactual generation" (under review Neurips 2025)
- Paper 'Harnessing Frozen Unimodal Encoders for Flexible Multimodal Alignment' accepted at Conference on Computer Vision and Pattern Recognition 2025
- Paper '**Do Vision and Language Encoders Represent the World Similarly?**' accepted at Conference on Computer Vision and Pattern Recognition 2024
- Paper 'Test-Time Adaptation with SaLIP: A Cascade of SAM and CLIP for Zero shot Medical Image Segmentation' accepted at Conference on Computer Vision and Pattern Recognition 2024 Workshops- Oral Presentation
- Paper 'Enhancing CLIP with GPT-4: Harnessing Visual Descriptions as Prompts' accepted at International Conference on Computer Vision 2023 - 'What's Next in Multimodal Foundation Models' Workshop
- Paper 'BaseTransformers: Attention over base data-points for One Shot Learning24 ' accepted at British Machine Vision Conference 2022.
- Paper 'The STOIC2021 COVID-19 AI challenge: applying reusable training methodologies to private data' accepted Transactions on Medical Imaging 2024
- Paper 'An Ensemble Deep Learning Approach for COVID-19 Severity Prediction Using Chest CT Scans' accepted Irish Machine Vision and Image Processing Conference 2023.
- Paper 'Phase retrieval for Fourier Ptychography under varying amount of measurements' accepted for spotlight presentations at British Machine Vision Conference 2018.
  - Poster on Phase retrieval for Fourier Ptychography under varying amount of measurements presented at CCD workshop, CVPR 2018
- Presented a poster 'Denoising High Density Mouse Brain Images using Deep Generative Models' at Society for Neuroscience Conference 2017, Washington D.C.
- IITMSAT Communications Systems-' A Lean Sat Design Approach ' presented at the 3rd IAA Conference on University Satellite Missions and CubeSat Workshop at Rome, Italy.

INTERESTS

I enjoy cinema, reading, lifting and long-distance running.

REFERENCES

- Prof. Noel O' Connor, Professor, School of EE, DCU
  - noel.oconnor@dcu.ie
- Dr. Kevin McGuiness, Assistant Professor, School of EE, DCU
  - kevin.mcguiness@dcu.ie
- Prof. Kaushik Mitra, Assistant Professor Electrical Engineering Dept. Indian Institute of Technology, Madras
  - kmitra@ee.iitm.ac.in
- Dr. Sanath Narayan, Senior Researcher, Technology Innovation Institute, Abu Dhabi
  - sanath.narayan@tii.ae
- Dr. Pooja Rao, Head R&D, Co-founder at Qure.ai
  - pooja.rao@qure.ai
- Dr. Prashant Warier CEO, Co-founder at Qure.ai
  - prashant.warier@qure.ai