

HAI PHAN

m: (+1)2177502413 ◊ Email: pthai1204@gmail.com

[Linkedin](#) ◊ [Google scholar](#) ◊ [Github](#)

WORK EXPERIENCE

Blue Marble Geographics

Jan 2023 - Present

Senior Research Development Engineer

- **Led a team consisting of two full-time research engineers** in the design and implementation of a machine learning system in C++ Our primary focus was on developing advanced algorithms for object detection and segmentation specifically tailored for 5K high-resolution sensory images. The successful outcome of our work result in the deployment of this system into production.
- Designed comprehensive test cases, collected, and processed thousands of sensory images captured from drone and satellite devices. This endeavor was undertaken with the goal of rigorously assessing and validating the performance of Machine Learning systems.
- Proposed diffusion models in synthesizing more novel sensory dataset, which are used for improving training process in object detection and segmentation.
- **Developed C++ CUDA layers** to enhance object detection during both training and testing phases, resulting in a notable **3×** acceleration of computational processes.
- Proposed and designed advanced Machine Learning and Computer Vision models for instance segmentation and detection applications on Drone Devices. The proposed models achieved an impressive mAP-50 accuracy of **92%** on 5K high-resolution testing images captured from Drones and Satellite devices. Furthermore, the models have demonstrated exceptional speed, with a rate of **9 FPS** on CPU, making them among the most efficient in the industry.

Languages/Technical usage: Diffusion models, GAN, Generative AI, C++, CUDA, Python, Pytorch, Azure, Object Detection, Segmentation, Computer Vision.

Adobe Research, Media Intelligence Lab

Sep 2022 - Dec 2022

Research Scientist Intern

- Proposed a novel Recommendation System to search millions of projects on the Behance platform by using multimodal techniques including large language models and vision models. With an impressive accuracy rate of up to **90%**, our system can effectively rank the top 10 projects from an extensive collection of millions of items. Additionally, our system can increase the speed by up to **3 ×** by re-ranking millions of items.
- Designed a method for efficiently selecting a groundtruth dataset from a vast collection of millions of items. This dataset serves the purpose of training and testing the system, ensuring its accuracy and effectiveness.
- Designed and implemented a Neo4j server on a Linux platform to facilitate the seamless collection and organization of millions of items within a comprehensive knowledge graph.

Languages/Technical usage: AWS, Python, Pytorch, Multimodal, Large Language Models, Computer Vision, Neo4j Database server.

Meta/Facebook AI, Virtual Reality Lab

Jun 2022 - Sep 2022

Research Scientist Intern

- Proposed a new 3D relightable model to render various lighting effects to the human face in a 3D world. Our model can generate highly realistic relit images with an impressively low pixel value error rate of **10** when compared to the original images.
- Proposed a novel diffusion model to manipulate variations in faces (e.g. age, bear, lighting conditions, face angles, etc.).
- Collected hundreds of human 3D scan data, which will serve as valuable resources for the purposes of training and testing. This comprehensive dataset will enable the development and refinement of advanced algorithms and models, allowing for more accurate and reliable outcomes in various applications such as computer vision, virtual reality, and biometric analysis.

Languages/Technical usage: GAN, Diffusion Models, Generative AI, Python, Pytorch, Pytorch3D, Neural Rendering, Internal Frameworks

CyLab biometrics, Carnegie Mellon University

2017 - 2021

Research Associate

- **Led a team consisting of 4 full-time software engineers and 3 master students** to design and implement an end-to-end efficient C++ SDKs for face algorithms (e.g. face detection, landmarking, and recognition) for mobile devices. These SDKs were successfully delivered to sponsors for their production deployment.
- **Led a team consisting of 3 master students** to design comprehensive test cases and collect thousand of real-life face images. This endeavor was undertaken with the goal of rigorously assessing and validating the performance and capabilities of the face C++ SDKs.
- **Efficient deep learning for mobile devices.** Researched and developed facial recognition software for devices such as iOS, Android, JetsonTX, and Xavier, using the CMU face detection and matching C++ SDK. The model achieves a significant $8 \times$ improvement in inference speed. Our facial recognition model was capable of running at **8-10 FPS** on an iPhone 7, **125 FPS** on the Xavier GPU, and **20 FPS** on the Xavier CPU.
- **Developed C++ CUDA layers** to expedite the training and inference stages for tasks encompassing object detection, 3D facial landmarking, and face matching, achieving an impressive $5 \times$ increase in computational speed.
- **Proposed efficient 3D Face Pose Estimation for mobile devices.** The proposed method outperformed previous methods with a Normalized Mean Error rate of just **3%**.

Languages/Technical usage: C++, CUDA, Python, Pytorch, Caffe, MXNet, Objective-C, OpenCV, Machine Learning, Compressed Networks, Computer Vision.

AXON (Part of TASER/AXON acquisition of FOSSIL Vision in 2016)

2017

Senior Research Engineer

- **Built and led a founding engineering team comprising of 8 individuals** to develop end-to-end C++ SDKs for object tracking and object detection and successfully integrate to production systems which are still used now.
- **Led a team consisting of engineers and contractors** in the task of designing comprehensive test cases and gathering a vast amount of video data. This endeavor aimed to thoroughly assess the quality and effectiveness of the proposed object tracking and detection frameworks.
- **Machine Vision:** Developed a real-time correlation filter object tracking C++ software development kit (SDK) that is capable of tracking the movement of people and objects. The SDK achieved an impressive speed of **30 FPS** on iPhone 5 and 6 while maintaining a very high level of accuracy.
- **Developed C++ CUDA enhancements** for Siamese neural networks, thereby optimizing object detection and tracking during both training and testing phases, resulting in a twofold acceleration of the entire process.

Languages/Technical usage: C++, CUDA, Objective-C, iOS, OpenCV, Machine Learning, Computer Vision, Signal Processing.

FOSSIL GROUP (acquired MISFIT WEARABLES in 2015)

2014 - 2017

Senior Research Engineer

Misfit Wearables was a tech start-up, manufacturing and selling wearable hardware devices in 40+ countries. In Nov 2015, Misfit was acquired by Fossil Group for \$260M after just 3 years in operation.

Post - acquisition: Senior ML Research Engineer.

- **Led a team of six full-time software engineers** in developing C++ Computer Vision SDKs to successfully deliver video algorithms (e.g. heart rate detection, object tracking, and object detection) and integrate them into production systems. I played a significant role in deal diligence process for AXON acquisition.
- **Led a team consisting of engineers and contractors** in the task of designing comprehensive test cases and collecting a vast amount of data. This initiative was undertaken to validate and evaluate the accuracy and efficacy of the heart rate estimation algorithm, as well as the object tracking and object detection functionalities.
- **Developed an end-to-end real-time non-intrusive heart rate estimation algorithm** using face or finger video data. The algorithm obtained low error rate of **1%**. The results are output in just **9-10 seconds**.
- **Designed an end-to-end API that streams data to record human activities.** The streaming can last up to **16**

hours and data are pushed to the server efficiently. The stream data are used to create 3D human body simulation in upcoming activities tracking products.

- Designed a backbone Computer Vision C++ SDK of object tracking and heart rate detection for both Android, iOS, and Mac devices.
- Developed Computer Vision algorithms: object tracking and heart rate detection in C++.

Pre - acquisition: Senior Software Engineer (iOS)

- Designed and developed an iOS Misfit production app in early stage that has been widely used by 3M+ users worldwide.
- Designed and developed a end-to-end backbone iOS Bluetooth Low Energy (BLE) SDK which helps the Misfit app search, connect, and update firmware for Misfit hardware devices and widely used in the team until now.
- Developed iOS Misfit themes for various professional partners (e.g. Coca Cola, Victoria's Secret, Swarovski, etc.)

Languages/Technical usage: C++, Objective-C, iOS, OpenCV, Machine Learning, Computer Vision, Signal Processing.

Computer Science Lab, Auburn University

2021 - 2023

Research Associate

- Proposed novel Explainable AI (XAI) techniques to enhance the interpretability of face matching and improve the top-1 accuracy rate by approximately **11%** for Out-of-Distribution faces, such as masked faces, profile faces, sunglasses, and more.
- Developed a new StyleGAN algorithm to manipulate facial attributes (e.g. age, sex, bear, etc.) to do the testing for face matching.
- Designed rigorous test cases and collected data for testing human-AI face recognition system.
- Proposed a novel user-friendly system to improve human interpretation for face recognition. The user studies are conducted through 30 people with different backgrounds and achieved **85%** when they use our proposed system to recognize faces.

Languages/Technical usage: StyleGAN, StyleFlow, Diffusion models, Generative-AI, Python, Pytorch, C++, Computer Vision, XAI, Machine Learning.

EDUCATION

Auburn University, AL, USA

May 2021 - May 2023

PhD candidate, Computer Science and Software Engineering

Carnegie Mellon University (CMU), Pittsburgh, PA, USA

Jan 2019 - May 2020

Master of Science, Electrical and Computer Engineering (ECE)

University of Science, HCMC

Sep 2009 - 2013

Bachelor of Science, Advanced Program In Computer Science (**GPA: 3.61 / 4.0**)

TECHNICAL STRENGTHS

Machine Learning Techniques	Neural Rendering, GAN, Diffusion models, LLM, Gen-AI
Deep Learning frameworks	Caffe, Tensorflow, Pytorch, MXNet, CUDA programming
Cloud Usage	Google Cloud, AWS, Azure
Programming Languages	C++, Objective-C, Python, Matlab

ACTIVITIES

Reviewer of Neurips 2021 and ICLR 2020, 2021

Oral Presentation at WACV 2020, Colorado, USA

Mar, 2020

AWARDS

Travel award, CVPR 2022.

Jun, 2022

EPSCoR graduate research scholar award [News](#) (\$25,000).

May, 2022

Woltosz PhD Fellowship award (\$24,000)

May, 2021

PUBLICATIONS

H. Phan, Anh Nguyen, *DeepFace-EMD: Re-ranking Using Patch-wise Earth Movers Distance Improves Out-Of-Distribution Face Identification*, CVPR 2022 (acceptance rate: 25%). [Arxiv Code Demo](#)

H. Phan, Z. Liu, D. Huynh, Z. Shen, K. Cheng and M. Savvides, *Binarizing MobileNet via Evolution-based Searching*, CVPR 2020 (acceptance rate: 22%). [cvpr2020](#)

H. Phan, D. Huynh, Y. He, M. Savvides, and Z. Shen, *MoBiNet: A Mobile Binary Network for Image Classification*, in WACV 2020. [wacv20](#)

Zhiqiang Shen, Honghui Shi, Jiahui Yu, **Hai Phan**, Rogerio Feris, Liangliang Cao, Ding Liu, Xinchao Wang, Thomas Huang, Marios Savvides. *Improving Object Detection from Scratch via Gated Feature Reuse* 30th British Machine Vision Conference (BMVC), 2019. [bmvc2019](#)

An T. Duong, **Hai T. Phan**, Nam Do - Hoang Le, Son T. Tran. *Hierarchical Approach for Handwritten Digits Recognition Using Sparse Auto-encoders*. In Springer Conference of Advanced Soft Computing 2014. [springer](#)

Hai T. Phan, An T. Duong, Nam Do - Hoang Le, Thai Son Tran . *Hierarchical Sparse Auto-encoder Using Linear Regression-based Feature in Clustering for Handwritten Digit Recognition*. In 8th International Symposium on Image and Signal Processing and Analysis (ISPA) - 2013. (**Oral Presentation**) [ieee](#)