Md Mahedi Hasan

PhD Researcher in ML and CV | Morgantown, WV | No Sponsorship Required

GitHub | LinkedIn | Scholar | mahedi-61.github.io | mh00062@mix.wvu.edu | +1 (681) 212-9391

Summary

- Doctoral Researcher with 7+ years of experience in computer vision and machine learning, specializing in pre-training and fine-tuning vision–language models (VLMs) and multimodal large language models (MLLMs) for fine-grained visual understanding and text-based image retrieval;
- Experienced in industry- and government-funded research (DoD, FBI, Qualcomm, and NSF).
- Published in top-tier conferences (WACV, BMVC) and journals (TAES, TBIOM);
- Research Interests: Vision–Language Models | Multimodal Representation Learning | Foundation Models | Continual Learning | Weakly-Supervised & Self-Supervised Fine-tuning.

Technical Skills

Programming Languages: Python (expert), C++ (proficient), Java (advanced)

Frameworks & Tools: PyTorch, TensorFlow, Hugging Face, Transformers, Scikit-learn, Numpy

Systems & Deployment: AWS, GCP, Kubernetes, Docker, MLflow, TorchServe, TensorRT

Algorithms: Multi-Modal Learning (VLMs, MLLMs) Fine-tuning & Adaptation (LoRA, Q-LoRA)

Funded Research Projects

1. **Optimizing Foundation Models for Edge Computing Platforms** (Qualcomm, 2025-2028)
   - Designed a Conv-LoRA–based domain adaptation technique to fine-tune the Segment Anything Model (SAM), for defect segmentation, reducing training cost by 40% and enabling edge-device deployment.
   - Designing self-supervised learning approaches to optimize foundation models for tasks with limited manually labeled data.
2. **Super-Resolution Object Characterization in Low Earth Orbit** (DoD, 2023-2025)
   - Designed an SRGAN-based super-resolution framework for diverse target chips detected and cropped by the Innovative Target Chipping (ITC) module, enhancing real-time detection of space objects.
   - Developed an adapted **Multi-Scale SRGAN** module to perform super-resolution with up-scaling factors of 4 and 8, improving the fidelity of satellite image patches for downstream analysis.
3. **A Perpetual Deep Face Recognition System** (NSF, 2022-2023)
   - Designed and implemented the complete research pipeline, including data preprocessing, model development, training, and evaluation.
   - Built a class-incremental continual learning framework, **CLFace**, for face recognition that enables consistent performance improvement across sequential tasks, and mitigates catastrophic forgetting.
4. **One-to-One Face Recognition with Human Examiner in the Loop** (NSF, 2022-2023)
   - Developed **CaptionFace**, a text-guided face recognition framework that enhances state-of-the-art face recognition models by integrating facial attribute information through natural language descriptions.
   - Proposed **GPTFace**, a vision–language captioning model that combines ViT-B/16 and GPT-2 to generate semantic descriptions from low-resolution facial images, achieving a 7% improvement in zero-shot retrieval performance.
5. **2023 Incomplete Fingerprint Records and Matching** (FBI, 2023-2024)
   - Conducted multi-finger fusion across all ten fingers using commercially off-the-shelf fingerprint matchers and performed comprehensive performance analysis.
   - Developed and trained a CycleGAN-based generative model to generate synthetic fingerprint images.
6. **Evaluation of the Performance of Multi-Finger Contactless Fingerprint Matching** (NSF 2021-2022)
   - Developed a **Coupled-GAN–based** framework to improve multi-finger contactless fingerprint matching and enhance contact-to-contactless fingerprint interoperability, achieving an 11.68% improvement in TAR at FAR = 0.01% over prior state-of-the-art methods.

Selected Research Publications

Peer-reviewed Journals

1. S. M. Sami, **M. M. Hasan**, N. M. Nasrabadi, and R. Rao, "FDCT: Frequency-Aware Decomposition and Cross-Modal Token-Alignment for Multisensor Target Classification," in IEEE Transactions on Aerospace and Electronic Systems (TAES), vol. 61, no. 4, pp. 9036-9057, August 2025.
2. **M. M. Hasan**, S. M. Sami, N. M. Nasrabadi, and J. Dawson, "Learning Multi-Scale Knowledge-Guided Features for Text-Guided Face Recognition," in IEEE Transactions on Biometrics, Behavior, and Identity Science (TBIOM), vol. 7, no. 2, pp. 195-209, April 2025.
3. S. M. Sami, **M. M. Hasan**, N. M. Nasrabadi, and R. Rao, "Contrastive Learning and Cycle Consistency-Based Transductive Transfer Learning for Target Annotation," in IEEE Transactions on Aerospace and Electronic Systems (TAES), vol. 60, no. 2, pp. 1628-1646, April 2024.
4. **M. M. Hasan**, N. M. Nasrabadi, and J. Dawson, "On improving interoperability for cross-domain multi-finger fingerprint matching using coupled adversarial learning." IET Biometrics, vol.12, pp.194-210, 2023
5. **M. M. Hasan**, and Hossen Asiful Mustafa. "Learning view-invariant features using stacked autoencoder for skeleton-based gait recognition." IET Computer Vision, vol. 15, no. 7, pp. 527-545, 2021.

Peer-reviewed Conference Papers

1. **M. M. Hasan**, S. M. Sami, N. M. Nasrabadi, and J. Dawson, "FaceCPT: Toward Cross-Modal Facial Representation Learning with Face-Caption Pre-Training," In 36th British Machine Vision Conference (BMVC), 2025.
2. **M. M. Hasan**, S. M. Sami, and N. M. Nasrabadi, "CLFace: A Scalable and Resource-Efficient Continual Learning Framework for Lifelong Face Recognition." In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 5082-5091. IEEE, 2025.
3. **M. M. Hasan**, S. M. Sami, and N. M. Nasrabadi, "Text-guided face recognition using multi-granularity cross-modal contrastive learning." In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 5784-5793. 2024.
4. **M. M. Hasan**, and N. M. Nasrabadi, "Improving face recognition from caption supervision with multi-granular contextual feature aggregation." In 2023 IEEE International Joint Conference on Biometrics (IJCB), pp. 1-10. IEEE, 2023.
5. **M. M. Hasan**, N. M. Nasrabadi, and J. Dawson, "Deep coupled GAN-based score-level fusion for multi-finger contact to contactless fingerprint matching." In 2022 International Conference of the Biometrics Special Interest Group (BIOSIG), pp. 1-7. IEEE, 2022.
6. M. S. Islam, **M. M. Hasan**, S. Abdullah, et al., "A deep Spatio-temporal network for vision-based sexual harassment detection." In 2021 Emerging Technology in Computing, Communication and Electronics (ETCCE), pp. 1-6. IEEE, 2021.
7. **M. M. Hasan**, M. S. Islam, and S. Abdullah, "Robust pose-based human fall detection using recurrent neural network." In 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON), pp. 48-51. IEEE, 2019.
8. **M. M. Hasan**, M. M. Abir, M. Ibrahim, et al. "AIBangla: A benchmark dataset for isolated Bangla handwritten basic and compound character recognition." In 2019 International Conference on Bangla Speech and Language Processing (ICBSLP), pp. 1-6. IEEE, 2019.
9. S. Abdullah **M. M. Hasan**, S. M. S. Islam, "YOLO-based three-stage network for Bangla license plate recognition in Dhaka metropolitan city." In 2018 International Conference on Bangla Speech and Language Processing (ICBSLP), pp. 1-6. IEEE, 2018.
10. S. M. S. Islam, and **M. M. Hasan**. "DEEPGONET: Multi-label prediction of GO annotation for protein from sequence using cascaded convolutional and recurrent network." In 2018 21st International Conference of Computer and Information Technology (ICCIT), pp. 1-6. IEEE, 2018.

Research Experience

Graduate Researcher at West Virginia University, Morgantown, USA June 2021 - Present

- Built end-to-end training and evaluation pipelines for multiple funded research projects, including visual representation learning using weakly supervised pretraining and vision–language modeling.
- Developed **FaceCPT**, a weakly supervised, vision–language pretraining framework for learning general facial representations with strong zero-shot transferability.
- Proposed and developed **ADFace**, which generates discriminative, attribute-driven captions from low-resolution facial inputs via a novel reward mechanism, improving text-to-face retrieval performance.

Deep Learning Engineer at AIBangla, Dhaka, Bangladesh January 2018 - February 2019

- Developed a general recognition system using OpenPose and LSTM-based networks for real-time sign gesture and pose understanding.

Education

PhD in Computer Engineering (June 2021 - Present)

- West Virginia University (WVU), Morgantown, USA (GPA: 4.0/4.0)
- Courses: Applications of Neural Networks, Pattern Recognition, Deep Learning, Computer Vision

M.Sc. in Information and Communication Technology (ICT), (October 2014 - September 2020)

- Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh (GPA: 3.58/4.0)

B.Sc. in Electrical and Electronic Engineering (EEE) (March 2009 - September 2013)

- Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh (GPA: 3.48/4.0)

Fellowships and Awards

- **Best Poster Award:** Center for Identification Technology Research (CITeR) workshop, MI, USA, 2022
- **University Merit Scholarship:** Govt. of the People's Republic of Bangladesh, 2009-2012

Selected Presentations & Talks

1. **Talk**: "A Deep Dive into Vision Transformer and Its Multimodal AI Applications," School of Mathematical and Data Sciences, West Virginia University (WVU), Morgantown, WV, November 2025
2. **Presentation**: A Perpetual Deep Face Recognition System, Fall 2023 Program Review, organized by Center for Identification Technology Research (CITeR), Potsdam, NY, October 25, 2023.
3. **Presentation**: Deep Multi-Finger Contact-to-Contactless Fingerprint Matching for Increased Interoperability, organized by Computer Science and Electrical Engineering, WVU, March 2023.
4. **Talk**: "The Engine Behind ChatGPT," Cognitive Mind, Dhaka, Bangladesh, December 2023
5. **Presentation**: Text-Guided Face Recognition using Contrastive Learning, EAB & CITeR Biometrics Workshop, organized by European Association for Biometrics (EAB), Martigny, Switzerland, April 2023.
6. **Presentation**: One-to-One Face Recognition with Human Examiner in the Loop, Fall 2022 Program Review, organized by CITeR, East Lansing, MI, November 2022. [**Best Poster Award**]
7. **Presentation**: Evaluation of the Performance of Multi-Finger Contactless Fingerprint Matching, Spring 2022 Program Review, organized by CITeR, Buffalo, NY, May 2022.

Activities

- Contributor: Research work on SROC LEO at **O Analytics,** August 2023– December 2025
- Participant: Attended EAB and CITeR Biometrics Workshop in Martigny, Switzerland, April 2023
- Reviewer at CVPR, AAAI, WACV, BMVC, TPAMI, TIP, TIFS, TASE, TPD