Khawar Islam

3 Gasan-ro Geumcheon-gu, 08577, Seoul, South Korea
+82-010-4835-0804
khawar-islam.github.io

Education

2nd Mar-20 – 18th Feb-22 **Sejong University**, Seoul, South Korea.
- M.S in Computer Science and Engineering
- Thesis: Face Pyramid Vision Transformer
- GPA: 4.25/4.5, Presentation Link

1st Feb-14 – 23rd Apr-18 **Federal Urdu University of Arts, Sciences & Technology**, Karachi, Pakistan.
- B.S. in Computer Science
- Title: Framework for Passenger Seat Availability Using Face Detection in Passenger Bus
- GPA: 3.53/4.0

Publications

CVPR 2024 **DiffuseMix: Label-Preserving Data Augmentation with Diffusion Models**.
Khawar Islam, Muhammad Zaigham Zaheer, Arif Mahmood, Karthik Nandakumar
In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2024
[Paper] [Supplementary] [Project Page] [Code] [Poster] [Dataset] [Video]

BMVC 2022 **Face Pyramid Vision Transformer**.
Khawar Islam, Muhammad Zaigham Zaheer, Arif Mahmood
In British Machine Vision Conference (BMVC), 2022
[Paper] [Supplementary] [Project Page] [Code] [Poster] [Video]

CVPRw 2021 **Image Compression with Recurrent Neural Network and Generalized Divisive Normalization**.
Khawar Islam, Dang Lien Minh, Sujin Lee, Hyeonjoon Moon
In IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRw), 2021
[Project Page] [Code] [CVF Open Access] [IEEE Xplore]

IVCNZ 2021 **Face Recognition Using Shallow Age-Invariant Data**.
Khawar Islam, Sujin Lee, Dongil Han, Hyeonjoon Moon
In Image and Vision Computing New Zealand (IVCNZ), 2021
[PDF] [IEEE Xplore]

IMAVIS 2020 **Person Search: New paradigm of Person Re-Identification: Recent Works**.
SCI Q1 Khawar Islam
IF 4.7 In journal of Image and Vision Computing (IMAVIS), Elsevier, 2020
[Sciencedirect, PDF]

JBAS 2017 **Huge and Real-Time Database Systems: A Comparative Study and Review for SQL Server 2016, Oracle 12c & MySQL 5.7 for Personal Computer**.
Khawar Islam, Kamran Ahsan, SAK Bari, Muhammad Saeed, Syed Asim Ali
In Journal of Basic & Applied Sciences, 2017

Research and Work Experience

4th-Ap-22 – Present **Research Scientist - Generative AI**, COREMAX, Seoul, South Korea.
Tools: YOLO v8, Stable Diffusion, Training, Fine-Tuning, Dataset Development, Hugging Face, Prompt Learning, Pix2Intruct, RNNs, GANS, Diffusion Models, OCR and Docker
- Working on image classification, image segmentation and object detections methods.
- Worked on Data Parallel, Distributed Data-Parallel, and Multi-Node Training..
- Experience building deep neural network models using CNNs, GANs, Transformers.
- Optimize GPU acceleration, and parallel processing to optimize the system performance.
- Work on image classifiers, object detectors, multi-modal classifiers, etc.
- Training, fine-tuning CV models, transfer learning, and evaluating their performance.
- Experiment with different DL architectures, loss functions, and regularization methods.
- Proposing lightweight self-supervised learning method for abuse content recognition.
- Write research papers, and technical reports, and present the work at conferences and meetings.

15th Mar-20 **Computer Vision Engineer (PTE)**, Sejong University, Seoul, South Korea.
28th Feb-22 **Tools:** Vision Transformers, Various Pre-Trained Models, Face Recognition Models.

- Proposed a new SOTA method that were published in CVPRw hosted by Google Research.
- I worked on neural network-based image compression with Prof. Hyeonjoon Moon.
- Worked on face recognition with limited data and resources with Prof. Hyeonjoon Moon.
- Hands-on experience on recent object detection models (YOLO, FasterRCNN, etc.).

15th Ap–19 **Research Associate (FTE)**, NUCES-FAST, Karachi, Pakistan.
16th Mar-20 **Tools:** VGG-Face, Google FaceNet, OpenFace, Facebook DeepFace, DeepID, ArcFace, Dlib.

- Performed research and development to solve missing children problems across the country.
- Develop and implement computer vision algorithms and systems to solve complex problems.
- Optimize CV models for real-time applications, object detection, tracking, and classification.
- Perform code reviews and provide constructive feedback to ensure high-quality software.

1st Aug-18 **Research Assistant (PTE)**, Information System Research Group, FUUAST Karachi.
15th Feb-19 **Tools:** C#, Visual Studio, DevExpress for UI Design

- Planning and execution of research projects, e.g literature reviews, data collection, and analysis.
- Assisting in the preparation of research proposals, grant applications, and progress reports.
- Maintaining complete research records, including databases and other documentation.
- Analyze research data using statistical tools and present findings in a clear and concise manner.

1st Mar-17 **Software Engineer (FTE)**, Axact Softwares, Karachi, Pakistan.
30th Jul-18 **Tools:** Swift, Xcode Editor, Alamofire, RealmSwift, SDWebImage, MessageKIT

- Designing, developing, and maintaining iOS applications using Swift programming languages.
- Writing clean, efficient, and maintainable code that follows best practices and coding standards.
- Troubleshooting and debugging issues, and implementing fixes to ensure smooth functioning.
- Providing technical guidance and mentorship to junior team members, as needed.
- Worked on MVC and singleton pattern, real-time push notifications, Firebase, realm
- Optimizing computer vision algorithms for performance and memory usage on mobile devices.

Honors and Achievements

2021 **NeurIPS Support**, Sponsored by Naver Labs Europe, DeepMind and Google AI.
2020 **Full Tuition Fee Waiver**, School of Computer Engineering, Sejong University.
2020 **Prof. Stipend**, CVPR Lab, Sejong University.
2020 **Conference Travel Grant**, Ministry of Planning Commission, Pakistan.
2018 **Merit Scholarship for Bachelor Degree**, Quaid-e-Azam Aligarh Society, Karachi.
2016 & 2017 **Most Valuable Professional Award**, C-SharpCorner, India.

English Proficiency and Relevant Courseworks

IELTS-AC 6.5 Listening 7.5, Speaking 6.5, Writing 6.5, Reading 6.0
A+ Introduction to Deep Learning, Advanced Deep Learning, Advanced AI
A0 Machine Learning, Topics in Machine Learning, Data Science
A Data Structure and Algorithms, Linear Algebra, Numerical Computing

Technical Skills

IP Algorithms Segmentation, Filter Design, Noise Removal, Compression, Super Resolution
Object Det SSD, YOLO-NAS, YOLOX, Faster R-CNN, DETR, GroundingDINO, YOLOv8
DL techniques Tracking, Optical Flow, Particle Filtering, Pose Estimation
CV Algorithms Image Classification and Detection, Tracking, Siamese Networks
Soft. Control JIRA, Agile, Scrum and GIT or Subversion
Lang Python, Keras, JAVA, Unix Shell Scripting
Deployment Docker and Kubernetes
Editor Visual Studio Code, PyCharm, Jupyter Notebook
OS WINDOWS 10, Ubuntu 18.04, Ubuntu 20.04, MAC OSX
APIs PyTorch, TensorFlow, Caffe, OpenCV, Flask, Django

Recent Research/Industrial Projects

2024 Present **High-Quality Video Generation from Diffusion Models (Text-2-Video).**
- Examined latest research on Magic-Me, ConsistI2V, Video-LaVIT, Lumiere and Boximator.
- Designed identity specific, consistency, subject-aware rich visual-motions prompts.
- Fine-tune above models and quantitative comparison with baseline models (CLIP, DINO).
- This structure facilitates easy extensions, such as integrating additional analysis tools.

2024 Present **Image Restoration and Enhancement with GANS and Diffusion Models.**
- Worked on Image-Super Resolution, Image DeBlurring, Image In-painting, and Image Debazing.
- Utilized recent models to remove motion out-of-focus & blur, restored image clarity.
- Creating algorithms capable of filling missing or damaged parts of images with realistic content.
- Developing cutting-edge methods to remove haze, fog, atmospheric distortions,and visibility.

2023 **Image Editing Instructions with Diffusion Models (Image-to-Image).**
- Integrating Imagic, SINE, and LEDITS into existing platforms to edit high-resolution image.
- Improving the models' capabilities in understanding and interpreting complex text inputs
- Designing user interfaces that allow non-experts to leverage text-based image editing.
- Exploring applications beyond traditional image editing, such as in gaming, and virtual reality.

2023 **Foundation Multimodal Vision Language Models (CLIP, Visual ChatGPT).**
- Worked on fashion projects and datasets like KAGL, F-MNIST, and DEEP datasets
- Generate captions from Kosmos-2, FuseCap, Blip by Salesforce, PromptCap & Pix2Struct.
- Evaluation with Relevance, Coherence and Fluency, Completeness, Creativity and Diversity.
- Image Generation with Stable Diffusion v2, RealCompo, GLIGENn and SD v1.5.

2023 **Vision Language Models To Extend Unseen Domains & DomainShift.**
- Utilized Latent Augmentation using Domain descriptions leverages a multimodal knowledge.
- Fine-Tuning under Distribution Shift, Semantic Augmentation, and Removing Dataset Bias.
- Verbalizing the training domain "photos of birds" to"paintings of birds" can improve robustness.
- Performed experiments on ROBIN, CUB-Paintings, DomainNet dataset and achieved 81%.

2023 **Reliable Lightweight Real-Time Open-Vocabulary Object Detection.**
- Investigated YOLO-World, MDETR, GLIP, GLIPv2, Grounding DINO, and DetCLIP.
- Used dataset for experiment such as Objects365V1, GQA, Flickr and CC3M†
- Fine-tuned w/ RepVL-PAN and w/o RepVL-PAN to achieve state-of-the-art performance.
- Further tune YOLO-World on COCO dataset (including 80 categories) with mask annotations.

2022 **Online Continual Learning with Blurry Data and Incorrect Labels.**
- Considered a framework for diversity and purity in memory updates and usage strategy.
- Used CIFAR-100, Food-101N, and a subset of real-world noisy data ImageNet, and WebVision.
- Utilized first online blurry CL set-up with contaminated data stream, and realistic.
- Because of semi-supervised learning with label cleaned up samples and unreliable samples.

2021 **Vision Transformer for General and Age-Invariant Face Recognition.**
- Introduced Shallow Age-Invariant Face Learning (SAIFL) approach
- Generated positive and negative age-invariant face pairs with large age gap.
- Proposed new ViT architecture for FR under limited training data and resources.
- Deep face models were implemented from the scratch using PyTorch.
- Investigated FaceNet, OpenFace, VGG-Face, ArcFace, and DeepID.
- Face aging benchmark datasets (e.g., FG-NET, MORPH, and Age-DB)