

## Education

- 2<sup>nd</sup> Sep-24 **The University of Melbourne**, Victoria, Melbourne, Australia.  
Present
  - PhD in Computer Science and Engineering
  - Advisor: Prof. Naveed Akhter (Computer Vision)
  - Thesis: Context-Guided Diffusion Models for Data Augmentation
- 2<sup>nd</sup> Mar-20 **Sejong University**, Seoul, South Korea.  
18<sup>th</sup> Feb-22
  - M.S in Computer Science and Engineering
  - Thesis: Face Pyramid Vision Transformer
  - GPA: 4.25/4.5, [Presentation Link](#)

## Key Publications

- ICLR 2026  **$S^2$ -FracMix: Self-Saliency Fractal Mixup.**  
[Khawar Islam](#), Arif Mehmood, Xin Jin, Naveed Akhter  
International Conference on Learning Representations (ICLR 2026) ([Under Review](#))
- ICLRw 2025 **Context-Guided Responsible Data Augmentation with Diffusion Models.**  
[Khawar Islam](#), Naveed Akhter  
In International Conference on Learning Representations (ICLRw), 2025
- 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
- BMVC 2022 **Face Pyramid Vision Transformer.**  
[Khawar Islam](#), Muhammad Zaigham Zaheer, Arif Mahmood  
In British Machine Vision Conference (BMVC), 2022
- CVPR 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

## Research and Work Experience

- 24<sup>th</sup>-June-25 **Senior Scientist - Autonomy AI**, Neubility, Seoul, South Korea.  
30<sup>th</sup>-Aug-25 **Tools:** YOLO-E, Video Depth Anything, Zero-Shot, VLMs, OWL-ViT, Panoptic DETR, Promptable Segmentation, SegFormer, Vision-Language-Action Models (VLAMs), Swin-UniNet,
  - Addressed **class-imbalance**, boosting minority class  $6 \times 200$  to **1,200** ( $+6 \times$  data boost)
  - Improved **O.Det** performance by introducing **FLUX.1 Aug**, raising **mAP** from **0.72** to **0.80**
  - Designing single quantized model **Object Detection + Segmentation** for delivery robot.
  - Introducing adversarial perturbation concept in current Object Detection model **FGSM+Fractal**.
  - Working on depth estimation via **Depth Anything** to estimate depth from footpath to road.
  - Data collection processes, ensuring **high-quality datasets** for training and validating models.
  - Plan and execute research projects, ensuring timely delivery of milestones and deliverables.
  - Stay up-to-date with latest R&D in CV and DL, and apply latest techniques to solve problems.
  - Implemented **DINO v3** for selecting diverse scene understanding images from public dataset.
- 1<sup>th</sup>-Ap-24 **Chief Technology Officer (CTO) - Gen AI**, Upendi, Seoul, South Korea.  
30<sup>th</sup>-June-25 **Tools:** Stable Diffusion XL Refiner, AnimeGANv2, FLUX, IP-FaceID-Plus, PhotoMaker, Realistic Vision, IP-Adapter, T2IAdapter, ControlNet, ADetailer, EmoStyle, PromptCap
  - I was employed in a part-time capacity from April 1, 2024 to March 31, 2025.
  - Personalized/Consistent character generation for **visual storytelling** across different scenes.
  - Lead team of **three developers** to build full-fledge product with agile methodology.
  - Introduced **character captioning dataset** to fine-tune Stable Diffusion in Generative AI domain.
  - Introduced a new Automated diary with specific character generation like **BTS diary**.
  - Meeting with different **CEOs and Investor** to illustrate the demo and current competition.
  - Data collection processes, ensuring **high-quality datasets** for training and validating models.

- 4<sup>th</sup>-Ap-22 **Research Scientist - AI**, Coremax Technology, Seoul, South Korea.
- 30<sup>th</sup>-Ap-25 **Tools:** YOLO v8, ResNet, LLMs, Chat GPT, GPT-Neo, Claude, Anime Style, CLIP, Log Analysis, IP-Adapter, Stable Diffusion, SmartEdit, FLUX, InstructPix2Pix, PhotoMaker, StoryDiffusion
- Worked on [Large Language Models \(GPT, NEO\)](#) for parsing Unsupervised log data.
  - Developed Memory, CPU and Hardware forecasting model via [Vision Transformer](#).
  - Construct [Hangul dataset \(10M\)](#) and trained different OCRs to recognize handwritten samples.
  - Learning Visual Features without [Label/Image-Text](#) pairs data via Self-Supervised methods.
  - Described project ideas to [Team Lead/CTO](#) related to project with diagrammatic form.
  - Introduced [EleutherAI's/Pythia LLM](#) to analysis DDOS attacks in unstructured data.
  - Trained different ViTs, CNNs for [Image Classification and Object Detection](#) problems.
  - Built UI via [Gardio/Streamlit](#) for projects to show full-fledged demo to the client.
  - Managed code on internal [GitHub](#) with proper commit messages and branches for easiness.
- 15<sup>th</sup> Ap-19 **Research Associate**, Video Surveillance Lab, NUCES-FAST, Karachi, Pakistan.
- 10<sup>th</sup> Mar-20 **Tools:** VGG-Face, Google FaceNet, OpenFace, Facebook DeepFace, DeepID, ArcFace, Dlib.
- Developed [Age-Invariant](#) Deep Neural Network for Asian faces via OpenFace (CMU).
  - Worked on [Person Re-Identification/Person Search](#) via videos captured from CCTV
  - Implemented ECCV2018 for [License Plate Recognition](#) for videos in unconstrained scenarios
  - Meeting with government/private institution for [Data Collection](#) and guidance.
  - Data collection and preprocessing of missing children from [Sarim Burney Trust International](#).
  - Teaching and mentoring [BS students](#) to conduct research related to specific topic.
  - Write clear and concise [technical documentation](#) and present solutions to stakeholders.
  - Evaluate the performance of CV models and develop methodologies for [model comparison](#).
  - Lead [two developers](#) to design full-fledge demo with Web Portal and Mobile App.
- 1<sup>st</sup> Mar-17 **Senior Executive - Application Development**, Axact, Karachi, Pakistan.
- 30<sup>th</sup> 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.
  - Participate in [Agile development process](#), attend sprint planning, stand-ups, and retrospectives.
  - 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.
  - Maintaining and improving existing codebases, including refactoring and bug fixing.
  - Developing and maintaining [image processing pipelines](#) and data processing workflows.

## 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.
- 2016 **Laptop Awarded, Prime Minister Pakistan Program**, Karachi, Pakistan.

## Summer and Winter Schools

- 2022 **OxML School**, Oxford Machine Learning Summer School.
- 2022 **NYU AI School**, New York Artificial Intelligence School, USA.
- 2021 **MLSS**, Machine Learning Summer School, Taipei, Taiwan.
- 2021 **EEML**, Eastern European Machine Learning Summer School.

## 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

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

---

## References

### **Dr. Zaigham Zaheer**

Research Fellow, Mohamed bin Zayed  
University of Artificial Intelligence, UAE  
Research Mentor  
✉ [zaigham.zaheer@mbzuai.ac.ae](mailto:zaigham.zaheer@mbzuai.ac.ae)

### **Prof. Arif Mahmood**

Information Technology University,  
Lahore, Punjab  
Research Mentor  
✉ [arif.mahmood@itu.edu.pk](mailto:arif.mahmood@itu.edu.pk)

### **Prof. Muhammad Atif Tahir**

National University of Computer &  
Emerging Sciences, Karachi, Pakistan  
Research Mentor  
✉ [atif.tahir@nu.edu.pk](mailto:atif.tahir@nu.edu.pk)