

Muhammad Ferjad Naeem

Curriculum Vitae

✉ ferjad.naeem@vision.ee.ethz.ch
📄 ferjad.github.io

Education

- 2021–2023 **Ph.D. Computer Vision and Machine Learning**,
ETH Zurich, Switzerland
Advisor: Prof. Luc Van Gool, PD. Dr. Federico Tombari (Google)
Google Ph.D. Fellowship with appointment as Research Consultant at Google.
- 2018–2021 **M.Sc. Computer Science (Biomedical Computing)**,
Technical University of Munich, Germany
Advisor: Prof. Nassir Navab, Prof. Zeynep Akata

Selected Publications

- Journal • “*Learning graph embeddings for open world compositional zero-shot learning*”.
Massimiliano Mancini*, **Muhammad Ferjad Naeem***, Yongqin Xian, Zeynep Akata.
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 2022.
- Conference • *I2MVFormer: Large Language Model Generated Multi-View Document Supervision for Zero-Shot Image Classification (Highlight top 2.5%)*.
Muhammad Ferjad Naeem, ..., Luc Van Gool, Federico Tombari.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2023 .
- *I2DFormer: Learning Image to Document Attention for Zero-Shot Image Classification*.
Muhammad Ferjad Naeem, Yongqin Xian, Luc Van Gool, Federico Tombari.
Neural Information Processing Systems (NeurIPS) 2022.
- “*Learning Graph Embeddings for Compositional Zero-shot Learning*”.
Muhammad Ferjad Naeem, Yongqin Xian, Federico Tombari, Zeynep Akata.
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2021.
- “*Reliable Fidelity and Diversity Metrics for Generative Models*”.
Muhammad Ferjad Naeem, Seong Joon Oh, Yunjey Choi, Youngjung Uh, Jaejun Yoo.
International Conference on Machine Learning (ICML) 2020.

Experience

- March 2021 – Present **Google Zürich + ETH Zürich**: Doctoral Researcher
Research Consultant at Google and Doctoral Researcher at ETH Zurich
I am pursuing my PhD Research jointly between Google and Computer Vision Lab at ETH. My research focuses on developing interpretable and robust Vision-Language Models (VLM). My work has utilized web image-text and Large Language models to improve the zero-shot generalization of VLM. My most recent work, SILC, improves the CLIP training objective. Our model SILC is the most powerful contrastive image-language model at Google and sets SOTA on multiple classification, retrieval, detection and segmentation benchmarks.

Key achievements.

- Developed the most powerful contrastive image-text foundational model at Google. Currently under review, preprint available here.
- Fluent in training models parallelized over large compute clusters. My largest training has utilized 2048 TPU chips in a session.
- Proposed one of the first works that utilizes a Large Language Model to improve the generalization of Vision-Language Models. This work received a highlight at CVPR 2023 (top 2.5% of submitted papers).

Dec 2020 – **Nvidia**: Computer Vision Intern

Feb 2021 *Remote from Munich, Germany (COVID) with Nvidia, Santa Clara, USA*

I worked in the DriveIX team on various deep learning based solutions for the AI enabled cockpit for driver assistance.

April 2020 – **Eberhard Karls University of Tübingen**: Visiting Researcher

Nov 2020 *Tübingen AI Research Center, Germany*

I worked on Compositional zero shot learning with Prof. Zeynep Akata. I proposed new SOTA methods and benchmarks that were published as two papers at CVPR2021.

September **Naver Corp. Clova AI Research**: Intern

2019 – *Naver Green Factory, Seongnam, South Korea*

December 2019 I worked on evaluation metrics for generative models under the supervision of Dr. Seong Joon Oh and Dr. Jaejun Yoo. The work was aimed at improving the precision and recall baseline from NeurIPS19 against outliers and bias towards ImageNet domain. The work was presented at ICML2020.

August 2018 – **Technical University of Munich**: Research Assistant

April 2020 *Computer Aided Medical Procedures (CAMP@TUM), Munich, Germany*

I worked on problems relating to robustness in deep neural networks and generative models under the supervision of Prof. Nassir Navab. In this direction I explored Adversarial Attacks, Uncertainty in Deep Learning, Interpretable Networks and Data Augmentation. I also worked on distribution trimming for robustness to outliers in generative models.

May 2015 – **National University of Sciences and Technology**: Research Assistant

July 2018 *TUKN-NUST R&D Center, Islamabad, Pakistan*

I worked on sequence to sequence modeling for problems relating to Optical Character recognition of printed and handwritten text under the supervision of Prof. Faisal Shafait. In this direction, I proposed several architectural improvements and collected new dataset that allowed for further research in the field. My work here led to two conference and one journal publications.

June 2017 – **Hochschule RheinMain**: Research Assistant (DAAD Scholar)

September 2017 *Computer Vision and Mixed Reality Lab, Wiesbaden, Germany*

I worked on the challenging problem of Underwater Fish detection and tracking under DAAD grant FIBEVID supervised by Prof. Adrian Ulges. In this direction, I proposed an end to end pipeline using FasterRCNN for detection and a Kalman Filtering based algorithm for tracking.

Achievements and Grants

May 2020 Google AI Residency 2020 finalist

- August 2018 "International Undergraduate Excellence Award" from Computer Aided Medical Procedures(CAMP) group at Technical University of Munich, Germany to join Prof. Nassir Navab's group.
- November 2018 "Korea Advanced Institute of Science and Technology - KAIST" EE-Camp Travel award for exceptional performance during undergraduate studies.
- June 2017 DAAD grant for "Fish Biodiversity Estimation by Low-Cost Non-Destructive Video Based Sampling", German Academic Exchange Service (DAAD), 2017, to move to Wiesbaden Germany for the Summer of 2017
- November 2017 ICDAR Travel Grant sponsored by International Association of Pattern Recognition (IAPR) to attend ICDAR 2017 in Kyoto Japan

Other Publications

- Conference
- *"Introducing Language Guidance in Prompt-based Continual Learning"*. Muhammad Gul Zain Ali Khan, **Muhammad Ferjad Naeem**, Luc Van Gool, Didier Stricker, Federico Tombari, Muhammad Zeshan Afzal. IEEE International Conference on Computer Vision (ICCV) 2023.
 - *"Open World Compositional Zero-Shot Learning"*. Massimiliano Mancini*, **Muhammad Ferjad Naeem***, Yongqin Xian, Zeynep Akata. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2021.
 - *3D Compositional Zero-shot Learning using DeCompositional Consensus*. **Muhammad Ferjad Naeem**, E.P. Ornek, Yongqin Xian, Luc Van Gool, Federico Tombari. European Conference on Computer Vision (ECCV) 2022.
- Journal
- *"I2DFormer+: Learning Image to Document Summary Attention for Zero-Shot Image Classification"*. **Muhammad Ferjad Naeem**, Yongqin Xian, Luc Van Gool, Federico Tombari. International Journal of Computer Vision (IJCV) 2024.
- Preprint
- *"SILC: Improving Vision Language Pretraining with Self-Distillation"*. **Muhammad Ferjad Naeem**, Yongqin Xian, Xiaohua Zhai, Lukas Hoyer, Luc Van Gool, Federico Tombari.
 - *"Learning to Prompt with Text Only Supervision for Vision-Language Models"*. Muhammad Uzair Khattak, **Muhammad Ferjad Naeem**, Muzammal Naseer, Luc Van Gool, Federico Tombari.
 - *"SemiVL: Semi-Supervised Semantic Segmentation with Vision-Language Guidance"*. Lukas Hoyer, David Joseph Tan, **Muhammad Ferjad Naeem**, Luc Van Gool, Federico Tombari.