






# SYED TALAL WASIM

 [wasimtalal@gmail.com](mailto:wasimtalal@gmail.com)  [wasimsyedtalal](https://www.linkedin.com/in/wasimsyedtalal)  [TalalWasim](https://github.com/TalalWasim)  
 [scholar.google.com/citations?user=uHySarAAAAAJ](https://scholar.google.com/citations?user=uHySarAAAAAJ)  [talalwasim.github.io](https://github.com/talalwasim)

## EDUCATION

---

### University of Bonn

*Ph.D. in Computer Vision*

**Jan 2024 – Ongoing**

*Bonn, Germany*

- Working on Long-Term Multimodal Video Understanding
- Supervised by Professor Dr. Jürgen Gall

### Universidad Autónoma de Madrid

*MS Image Processing and Computer Vision - CGPA - 8.78/10*

**Sep 2019 – Jun 2021**

*Madrid, Spain*

- Funded by the Erasmus Mundus Joint Masters Degree (EMJMD) Scholarship Program
- MS Thesis: Automatic Typography Analysis on Figurative Content
- Supervised by Dr. Mathieu Salzmann (CVLAB) at EPFL

### ETH Zurich

*Robotics Summer School and Symposium*

**Jul 2021**

*Zurich, Switzerland*

- Highly selective program that admits 40 participants (M.S. or Ph.D. students)
- Attended a week-long intensive theoretical and practical education classes in autonomous robotics

### Habib University

*BS Electrical Engineering - Minor in Computer Science - CGPA - 3.86/4.00*

**Sep 2015 – Jun 2019**

*Karachi, Pakistan*

- BS Thesis: SquadBot: A Multi-Agent Robotics Teaching and Research Platform
- Graduated with 1<sup>st</sup> position in program

### Stanford University

*Summer International Honors Program - CGPA - 3.87/4.00*

**Jun 2017 – Aug 2017**

*Stanford, USA*

- Coursework: Technology Entrepreneurship, Leading Trends in IT, Smart Cities & Communities

## WORK EXPERIENCE

---

### Intelligent Visual Analytics Lab, Mohamed Bin Zayed University of AI

**Apr 2022 – Dec 2023**

*Associate Researcher*

*Abu Dhabi, UAE*

- Supervisor: Dr. Salman Khan
- Focused on multimodal video representation learning and out-of-distribution generalization

### Computer Vision LAB, EPFL

**Feb 2021 – July 2021**

*MS Thesis*

*Lausanne, Switzerland*

- Supervisor: Dr. Mathieu Salzmann
- Thesis Title: Automatic Typography Analysis on Figurative Content

### Empathic Computing Laboratory, University of South Australia

**Jul 2020 – Mar 2021**

*Research Intern*

*Remote*

- Supervisor: Dr. Mark Billinghurst
- Focused on Multimodal Emotion Recognition using Facial Micro-Expressions
- Including modalities such as Video, EEG and GSR

### Habib University

**Jan 2018 – Jun 2018**

*Undergraduate Research Assistant*

*Karachi, Pakistan*

- Supervisor: Dr. Muhammad Farhan
- Detection of Mitosis in Breast Cancer Histopathology Images with Deep Learning

## PUBLICATIONS

---

1. A. Shaker, **S. T. Wasim**, M. Danelljan, S. Khan, M.-H. Yang, and F. Khan, “Efficient video object segmentation via modulated cross-attention memory,” in *WACV*, 2025
2. J. Yi\*, **S. T. Wasim\***, Y. Luo\*, M. Naseer, and J. Gall, “Video-Panda: Parameter-efficient alignment for encoder-free video-language models,” in *Under Review*, 2024
3. D. Velayudhan, A. Ahmed, M. Alansari, N. Gour, A. Behouch, T. Hassan, **S. T. Wasim**, N. Maalej, M. Naseer, J. Gall, M. Bennamoun, E. Damiani, and N. Werghi, “STING-BEE: Towards vision-language model for real-world x-ray baggage security inspection,” in *Under Review*, 2024
4. A. Shaker, **S. T. Wasim**, S. Khan, J. Gall, and F. Khan, “Groupmamba: Parameter-efficient and accurate group visual state space model,” *arxiv preprint, arxiv:2407.13772*, 2024
5. H. Suleman\*, **S. T. Wasim\***, M. Naseer, and J. Gall, “Distillation-free scaling of large ssms for images and videos,” *arxiv preprint, arxiv:2409.11867*, 2024
6. **S. T. Wasim**, M. Naseer, S. Khan, M.-H. Yang, and F. Khan, “VideoGrounding-DINO: Towards open-vocabulary spatio-temporal video grounding,” in *CVPR*, 2024
7. M. Z. Yousuf, **S. T. Wasim**, S. N. Hasany, and M. Farhan, “AR-VPT: Simple auto-regressive prompts for adapting frozen vits to videos,” in *VISAPP*, 2024
8. **S. T. Wasim**, K. H. Soboka, A. Mahmoud, S. Khan, D. Brooks, and G.-Y. Wei, “Hardware resilience properties of text-guided image classifiers,” in *NeurIPS*, 2023
9. **S. T. Wasim\***, M. U. Khattak\*, M. Naseer, S. Khan, M. Shah, and F. Khan, “Video-FocalNets: Spatio-temporal focal modulation for video action recognition,” in *ICCV*, 2023
10. M. U. Khattak\*, **S. T. Wasim\***, M. Naseer, S. Khan, M.-H. Yang, and F. S. Khan, “Learning self-regulating prompts for vision-language models,” in *ICCV*, 2023
11. **S. T. Wasim**, M. Naseer, S. Khan, F. Khan, and M. Shah, “Vita-CLIP: Video and text adaptive clip via multimodal prompting,” in *CVPR*, 2023
12. **S. T. Wasim**, R. Collaud, L. Défayes, N. Henchoz, M. Salzmann, and D. Ribes, “Toward automatic typography analysis: serif classification and font similarities,” *Journal of Data Mining in Digital Humanities (JDMDH)*, 2023
13. N. Saffaryazdi, **S. T. Wasim**, K. Dileep, A. F. Nia, S. Nanayakkara, E. Broadbent, and M. Billingham, “Using facial micro-expressions in combination with eeg and physiological signals for emotion recognition,” *Frontiers in Psychology*, 2022
14. **S. T. Wasim**, S. N. Hasany, K. Abbasi, H. Feroz, A. A. Ahmed, M. H. Shaikh, and M. Farhan, “Sim-to-real transfer for object detection and localization on animals,” in *CV4Animals CVPR Workshop*, 2021

## ACADEMIC SERVICES

---

### • Journal Reviewers

- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- IEEE Transactions on Image Processing (TIP)

### • Conference Reviewers

- Computer Vision (CVPR, ICCV, ECCV, WACV, ACCV)
- Artificial Intelligence and Machine Learning (NeurIPS, ICLR, ICML, AAAI)

### • Project Supervision

- Co-supervise undergraduate projects in computer vision at Habib University
- Co-supervise high-school students in Pakistan for the International Science and Engineering Fair (ISEF)

## HONORS AND AWARDS

---

### Ph.D.

- **Compute Award for Large-Scale Open-Vocabulary Video Understanding and Anticipation:** Co-authored EuroHPC Benchmark and Regular access grants with Professor Dr. Juergen Gall totaling nearly 3 million GPU compute hours.
- **Compute Award for Large-Scale Robust Vision Language Models:** Co-authored EuroHPC Benchmark access grant with Professor Dr. Ernesto Damiani totaling 0.2 million GPU compute hours.

### Masters

- **Erasmus Mundus Scholarship:** Two year fully funded scholarship for MS studies

### Bachelors

- **Dean's Medal:** For graduating with the highest CGPA in Electrical Engineering program
- **Best Capstone Award:** Awarded the best capstone project award in the Electrical Engineering program
- **Summer Program Scholarship:** Among 8 students selected for funded International Honors Program at Stanford University
- **President's Honor List:** For maintaining position on Dean's Honor List in consecutive semesters
- **Dean's Honor List:** The top 10% students in the program each semester
- **High Academic Achievement Scholarship:** Additional 10% Scholarship for the Top 3 students in the school each semester
- **Merit Scholarship:** Awarded 65% scholarship for 4 years

### High School

- **Intel ISEF:** Fully funded opportunity to represent Pakistan at the Intel International Science and Engineering Fair (ISEF), 2014 in LA, California

## SKILLS

---

**Research Areas:** Video Understanding, Large Language Models, Vision-Language Models, Efficient Neural Architectures, Multimodal Learning, Representation Learning

**Programming:** Python (Advanced), C/C++ (basic), C# (basic), Java (Basic)

**Common ML Tools:** Pandas, Numpy, Scikit-Learn, Tensorflow/Keras, OpenCV, Pytorch

**AR/VR and Game Engines:** Unity 3D, HoloLens 1

**Languages:** English: C2 (Expert), Urdu: Native