

Abdelrahman Eldesokey, PhD

Email: abdo.eldesokey@gmail.com

Mobile: +966 54 844 99 60

Pages: [LinkedIn](#) | [Google Scholar](#)

I am a Postdoctoral Fellow at KAUST, Saudi Arabia, specializing in Generative AI. I work with a team of PhD students conducting research on advanced topics in generative models and Multi-Modal Large Language Models (MLLMs). I bring over 11 years of combined academic and industrial experience in computer vision and machine learning across Sweden, Egypt, and Saudi Arabia. I hold a Ph.D. in Computer Vision and Deep Learning from Linköping University in Sweden.

AREAS OF EXPERTISE

- **Computer Vision & Machine Learning** Deep Learning, Diffusion Models, Multi-Modal Large Language Models (MLLM), AI for Sports, Depth Estimation, Object Detection/Tracking, Uncertainty Estimation for Deep Learning, Image/Video Object Segmentation, Optical Flow
- **Software Management** Agile, Version Control, Cloud (Azure, AWS), Distributed Systems (Redis, IROS)

PERSONAL INFORMATION

- Date of Birth: 06/11/1989
- Citizenship: Swedish, Egyptian
- Residence: KAUST, Saudi Arabia

PROFESSIONAL EXPERIENCE

- Mar 2023 - Present: **Postdoctoral Fellow** at KAUST, Saudi Arabia
 - **3D Layout Control for Text-to-Image Diffusion Models:** Introduced a novel approach for controlling image generation through a user-provided 3D layout [\[Project Page\]](#) [\[Paper\]](#)
 - **Consistent Characters Animation Synthesis using Diffusion Model:** Developed a zero-shot approach for animating characters based on image and motion diffusion model [\[Project Page\]](#) [\[Paper\]](#)
 - **Diffusion-Based Video Semantic Segmentation:** Supervised a project for segmenting videos using pre-trained diffusion models under a zero-shot setting [\[Project Page\]](#) [\[Paper\]](#)
 - **3D Sematic Correspondence using Vision and Large Language Models:** Supervised a project for matching non-isometric 3D shapes based on vision and language models [\[Project Page\]](#) [\[Paper\]](#)
 - **Head Avatars Generation with Multi-Modal Conditioning:** Supervised a project for generating and editing head avatars by conditioning a GAN using a diffusion model [\[Paper\]](#)
 - **Fine-Grained Image Editing using Diffusion Models:** Proposed an approach for editing fine parts of images to produce complex image concepts [\[Coming Out Soon\]](#)
 - **Improving Reasoning of Multi-Modal LLMs through Enhanced Vision Encoders:** Developing a visual chain-of-thought paradigm for improving the visual reasoning of MLLM [\[In progress\]](#)
 - **Multi-View Consistent Vision Features:** Co-developing a vision encoder based on image foundation models that is consistent across different views [\[In progress\]](#)
- Aug 2021 – Feb 2023: **Senior Machine Learning Engineer** at Signality, Sweden
 - **AI-based Algorithms for Sports Analytics:** Led a team of 3-5 Machine Learning engineers and data scientists to develop algorithms for player detection/tracking and automated events detection (Auto VAR).
 - **Scrum Master:** Managed and tracked the progress of the R&D team to meet the deadlines and deliverables in a timely manner.

- Apr 2022 – Feb 2023: **Adjunct Lecturer** at Linköping University, Sweden
 - **Multi-Modal Uncertainty Estimation in Regression Tasks**: Supervised a project for estimating uncertainty in regression tasks with multi-modal predictions [Paper]
- May 2016 – Jul 2021: **PhD Student** at Linköping University, Sweden.
 - **Uncertainty Estimation for Sparse Data**: Developed a novel Convolutional Neural Network (CNN) architecture for uncertainty estimation in sparse data denoted as *Normalized CNNs* with several applications in autonomous driving, structure-from-motion, and robotics [Video]
 - **Vehicles Detection in Thermal Imagery**: A collaboration project with the Swedish Defense Research Agency (FOI) where we developed a GAN-based approach for generating synthetic datasets to train object detection models in the thermal domain.
 - **Cyclist Detection in Driving Environments**: An industrial/academic collaboration where I developed vision-based cyclist detection algorithms in challenging driving environments.
- Oct 2013 – Apr 2016: **Research Assistant** at Nile University, Egypt.
 - **TraffiSense**: Developed a smart vision-based traffic analytics system for the Ministry of Interior in Egypt. This included vehicle count, velocity estimation, density analysis, violation detection, and License Plate Recognition (LPR) [Video]

EDUCATION

- Nov 2016 – July 2021: **PhD in Computer Vision and Deep Learning** at Linköping University, Sweden
 - **Thesis**: “Uncertainty-Aware Convolutional Neural Networks for Vision Tasks on Sparse Data”
 - **Supervisors**: [Prof. Michael Felsberg](#) & [Prof. Fahad Shahbaz Khan](#)
- Oct 2013 – Apr 2016: **MSc in Communication and Information Technology** at Nile University, Egypt
 - **Thesis**: “The AERIAL tracker: A Robust Visual Tracker for Micro-Drones” [Videos]
 - **cGPA**: 3.97/4.0
 - **Supervisor**: [Prof. Mohamed Elhelw](#)
- Sep 2007– Aug 2011: **BSc in Computers and Systems Engineering** at Mansoura University, Egypt
 - **Graduation Project**: “Augmented Reality Platform on Android for Educational Applications”
 - **Grade**: Excellent with honors (Rank 4/180)

SELECTED PUBLICATIONS

- [2024] **Eldesokey, A.**, & Wonka, P. “Build-A-Scene: Interactive 3D Layout Control for Diffusion-Based Image Generation”. *arXiv preprint arXiv:2408.14819*.
- [2024] Wang, Q., **Eldesokey, A.**, Mendiratta, M., Zhan, F., Kortylewski, A., Theobalt, C., & Wonka, P. “Zero-Shot Video Semantic Segmentation based on Pre-Trained Diffusion Models”. *arXiv preprint arXiv:2405.16947*.
- [2024] Para, W. R., **Eldesokey, A.**, Li, Z., Reddy, P., Deng, J., & Wonka, P. “AvatarMMC: 3D Head Avatar Generation and Editing with Multi-Modal Conditioning”. *arXiv preprint arXiv:2402.05803*.
- [2024] **Eldesokey, A.**, & Wonka, P. “LATENTMAN: Generating Consistent Animated Characters using Image Diffusion Models”. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7510-7519).
- [2024] Xiong, Z., Jonnarth, A., **Eldesokey, A.**, Johnander, J., Wandt, B., & Forssén, P. E. “Hinge-Wasserstein: Estimating Multimodal Aleatoric Uncertainty in Regression Tasks”. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3471-3480).
- [2023] Abdelreheem, A., **Eldesokey, A.**, Ovsjanikov, M., & Wonka, P. “Zero-shot 3d shape correspondence”. In *SIGGRAPH Asia 2023 Conference Papers* (pp. 1-11).
- [2021] **Eldesokey, A.**, & Felsberg, M. “Normalized Convolution Upsampling for Refined Optical Flow Estimation”. In 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021), (Vol. 5, pp. 742-752).

- [2021] **Eldesokey, A.** “*Uncertainty-Aware Convolutional Neural Networks for Vision Tasks on Sparse Data*” (Doctoral dissertation, Linköping University Electronic Press).
- [2020] **Eldesokey, A.**, Felsberg, M., Holmquist, K., & Persson, M. “Uncertainty-aware CNNs for depth completion: Uncertainty from beginning to end”. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12014-12023).
- [2019] **Eldesokey, A.**, Felsberg, M., & Khan, F. S. “Confidence propagation through CNNs for guided sparse depth regression”. *IEEE transactions on pattern analysis and machine intelligence*, 42(10), 2423-2436.
- [2018] Nyberg, A., **Eldesokey, A.**, Bergstrom, D., & Gustafsson, D. “Unpaired thermal to visible spectrum transfer using adversarial training”. In *Proceedings of the European conference on computer vision (ECCV) Workshops*.

AWARDS AND HONORS

- Best Paper Award at VISAPP conference (2021)
- Received an Affiliation to the Wallenberg AI, Autonomous Systems and Software Program (2017)
- Graduate Scholarship from Nile University for master’s degree (2013)
- Honor award for outstanding performance in BSc from Mansoura University (2011)
- One of the Top 10 BSc graduation projects in Egypt by IEEE Gold (2011)

TECHNICAL SKILLS

- **Programming**

| Language | Years of Experience | Last Used |
|----------|---------------------|-----------|
| Python | 10 Years | Present |
| Matlab | 10 Years | 2019 |
| C/C++ | 7 Years | 2016 |

- **Deep Learning Frameworks** PyTorch (Expert), TensorFlow & Keras (Intermediate)
- **Data Science** Pandas, Scikit-learn, Seaborn
- **Cloud** AWS, Azure
- **Versioning Control** Git
- **Distributed System** Redis, IROS
- **Agile Management**
 - Jira/Confluence
 - Trello

SOFT SKILLS

- **Languages**
 - Arabic (Mother Tongue)
 - English (Fluent)
 - Swedish (Advanced)
- **Self-motivated:** Can steer projects from the planning phase to delivery autonomously
- **Team Leading:** Can lead small teams of engineers/students and exploit their individual potential
- **Presentation Skills:** Poses excellent presentation skills and can adapt concepts to fit audiences of different backgrounds
- **Social:** Can effortlessly interact with people from different backgrounds and get them engaged