MADHAV AGARWAL

Edinburgh, UK
Email: madhav14130@gmail.com, madhav.agarwal@ed.ac.uk,
Google Scholar: Link ◇ Presonal Webpage: Link ◇ LinkedIn: madhav1ag

EDUCATION

PhD, Computer Science — Jan 2025 - Present
University of Edinburgh, UK

MS by Research, Computer Science and Engineering — Jan 2021 - June 2023
International Institute of Information Technology Hyderabad, India — CGPA: 9.17

Bachelor of Technology (Hons.), Information Technology — Aug 2014 - June 2018
Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India — Percentage: 76.00

SKILLS

Technical Skills	Computer Vision, Deep Learning, Machine Learning, Pattern Recognition, Natural Language Processing
Programming Languages	Python, C, C++
Deep Learning Framework	PyTorch, TensorFlow
Generative AI	Large Language Models (LLM), Vision Language Models (VLM), Gaussian Splatting, Diffusion Models, Attention & Transformers, GAN, NeRF, CNN

WORK EXPERIENCE

Applied Machine Learning Engineer | VidLab7 GmbH (now: Moonscale) — April 2024 - Nov 2024

- Synthetic Avatar Generation, LipSync, Voice Cloning, Deep Fake Detection

Artificial Intelligence Researcher | Technical University of Munich — Sept 2023 - Dec 2023

- 3D Motion Modelling and Generation, Face and Full Body Reenactment.

Computer Vision Consultant | Gradient Ascent AI, Artflow AI, Stealth Startups — Sept 2022 - Aug 2023

- Computer Vision, Deep Learning, Image and Video Segmentation.

Computer Vision Consultant | GuiseAI — Jan 2022 - Jun 2022

- Optimisation of Deep Learning Models for Edge Devices, Security and Surveillance at Edge Devices.

Senior Data Scientist | CAW Studios — Dec 2019 - Aug 2020

- Document information extraction and digitalization of the tabular data using deep learning and OCR.

Data Scientist | Innefu Labs Pvt. Ltd. — Jul 2018 - Sept 2019

- Smart Security and Surveillance Solutions using deep learning, deployed at multiple large-scale locations.

Machine Learning Engineer | Tidyquant Pvt. Ltd. — Jan 2018 - April 2018

- Predictive analytics and forecasting using Machine Learning.

PROJECTS

Spatio-Temporal Understanding in VLMs: Exploring how VLMs process spatio-temporal information to understand their current behavior and design enhanced models capable of downstream tasks such as emotion recognition in videos and 3D scene understanding.

3D Talking Heads: Generate wobble-free 3D talking head videos with accurate lip-sync from audio using Gaussian Splatting, modeling fine-grained head and facial motion for realistic performance.

Image Manipulation using Diffusion Models Image synthesis and out-of-distribution image inversion using spherical interpolation in latent space of diffusion models.

Face Reenactment and Talking Heads: Transfer motion and expression information on a static image from a video. Proposed various novel techniques to use audio and visual modalities along with facial cues. Build a pipeline to use these methods with face super-resolution and frame interpolation for efficient transmission.

Page Object Detection and Localization: Detect Tables, Figures, Charts, etc., in unstructured document images using a cascade architecture and deformable convolutions with high accuracy. Extended this work to incorporate NLP and OCR for end-to-end document digitalization.

Smart Security and Surveillance Solutions: Developed AI enabled modules like Facial Recognition, Trespassing Detection, Person Count, Abandon Objects, Group Formation, Heat- Map Generation, etc. These modules have custom object detection and tracking capabilities to handle objects like Weapons, Persons, Faces, Helmets, Shoes, Safety Harnesses, etc.

Person Attribute Recognition and Classification: Classifying the attributes of person like age, gender, clothes etc. for identifying customer behavior in retail markets and provide them with personalized experience.

Satellite Images: Detection and classification of ships, submarines using real-time satellite images.

ACHIEVEMENTS & EXTRA-CURRICULAR ACTIVITIES

- Graduate Scholarship from IIIT-Hyderabad to cover the full tuition fee for the MS degree.
- Coordinator in 6th Summer School of AI (2022) organized by IIIT-Hyderabad.
- Recipient of IIIT-Hyderabad's Non-Academic Award for spreading Mental Health Awareness on campus.
- Member and Student Head Coordinator (2017-18) of Big Data Centre of Excellence, R&D Lab of AKGEC, for three years (2015-2018).
- Winner of 'BigWar' event of BigDataThon'17 (24-hour hackathon on Big Data technologies) at ABESIT, Ghaziabad (April 2017).
- Organized Workshops on "Machine Learning and Neural Networks" (Feb 2018) and "Data Science" (March 2017) as a Mentor.

PUBLICATIONS

- GaussianHeadTalk: Wobble-Free 3D Talking Heads with Audio Driven Gaussian Splatting. *Winter Conference on Applications of Computer Vision (WACV), 2026.* **Madhav Agarwal**, Mingtian Zhang, Laura Sevilla-Lara, Steven McDonagh. (Link)
- Understanding the Generalization of Pretrained Diffusion Models on Out-of-Distribution Data. *AAAI Conference on Artificial Intelligence (AAAI), 2024.* **(Oral)**. Sai Niranjan Ramachandran*, Rudrabha Mukhopadhyay*, **Madhav Agarwal***, C.V. Jawahar & Vinay Namboodiri. (Link)
- Audio-Visual Face Reenactment. *Winter Conference on Applications of Computer Vision (WACV), 2023.* **Madhav Agarwal**, Rudrabha Mukhopadhyay, Vinay Namboodiri & C. V. Jawahar. (Link)
- Compressing Video Calls using Synthetic Talking Heads. *33rd British Machine Vision Conference (BMVC), 2022.* **Madhav Agarwal**, Anchit Gupta, Rudrabha Mukhopadhyay, Vinay Namboodiri & C. V. Jawahar.(Link)
- Dataset Agnostic Document Object Detection. *Pattern Recognition, 2023.* Ajoy Mondal, **Madhav Agarwal** & C. V. Jawahar. (Link)
- CDeC-Net: Composite Deformable Cascade Network for Table Detection in Document Images. *25th International Conference on Pattern Recognition (ICPR), 2020.* **(Oral)**. **Madhav Agarwal**, Ajoy Mondal & C. V. Jawahar. (Link)