MUHAMMAD SAAD

AI Researcher

+971 50 6506593 | muhamamdsaadicup@gmail.com | LinkedIn | Google Scholar | Github | Portfolio

SUMMARY

My research focuses on advancing AI and deep learning methodologies for applications in medical imaging, activity recognition, and immersive technologies. I have developed models for medical image analysis, facial emotion recognition, and violence detection, leveraging this expertise to integrate AI-driven perception into immersive systems. Recently, I have been designing interactive AI avatars in the metaverse as digital twins of educators, capable of responding to questions with realistic animations and lip-syncing, transforming virtual learning and human-computer interaction

EDUCATION

Bachelor of Science, Software Engineering Aug 2017 – Sep 2021

Islamia College Peshawar (ICP), Pakistan

- Undergraduate research student supervised by Dr. Muhammad Sajjad and Dr. Jamil Ahmad
- **Thesis**: "Visual explanation of deep learning-based breast cancer classification via gradient localization"
- **Major Courses**: Object-Oriented Programming (OOP), Data Structure and Algorithms, Software Architecture, Artificial Intelligence

RESEARCH INTERESTS

- **Medical Image Analysis:** Developing advanced methods to extract meaningful insights from medical images, enabling improved patient care and accurate diagnosis
- **Action Recognition:** Exploring advanced techniques for detecting and classifying human actions in videos, focusing on applications in surveillance, sports analytics, and human-computer interaction
- **Computer Vision and the Metaverse:** Developing innovative applications to enhance human-computer interaction and contribute to the growth of immersive technologies

EXPERIENCE

Graduate Research Assistant **Abu Dhabi, UAE**

Metaverse Center, Mohamed Bin Zayed University of Artificial Intelligence *Jan 2023 – Present*

Research topics: Digital twin, Metaverse, Violence Detection, LLMs for Interactive Avatars

- Worked on real-time violence detection on Jetson Nano at the Technology Innovation Institute (TII).
- Interactive avatar animation with Mixamo and real-time lip-syncing using JavaScript and TypeScript (GitHub).
- Developed a complete React-based dashboard for the Malaria No More (MnM) project, designed for seamless data visualization and improved decision-making.
- Created a visual avatar assistant powered by a fine-tuned LLaMA 3 model, customized with haptics and multimedia data to enhance educational experiences and interactive multimedia books (GitHub).
- Built a custom virtual learning platform named ZapAura, built on Mozilla Hubs, featuring full-body avatars, real-time lip-syncing, and an AI teaching assistant powered by ChatGPT for multilingual interactions.

Undergraduate Research Assistant **Peshawar, Pakistan**

Digital Image Processing (DIP) Lab ICP *Dec 2020 – 2022*

Research topics: Medical Imaging, Activity recognition, Facial emotion recognition (FER)

- Contributed to NTNU's implementation of the facial emotional recognition module assigned by the ALAMEDA AI Toolkit to analyze facial expressions for pain assessment and emotional state monitoring in neurological healthcare.
- Attention-based CNN-LSTM, CNN-GRU, and Video Vision Transformer (ViViT) Models for Complex Activity Recognition in Cricket (GitHub)
- Mentored new students and interns in their final-year projects, providing guidance and technical expertise for project completion.
- Teaching assistant for Python programming course.

PUBLICATIONS

- **M. Saad**, M. Ullah, H. Afridi, F. A. Cheikh, and M. Sajjad. BreastUS: Vision Transformer for Breast Cancer Classification Using Breast Ultrasound Images, 2022 16th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Dijon, France, 2022.
- **M. Saad**, M. Khan, M. Saeed, Abdulmotaleb El Saddik, and Wail Gueaieb. Combating Counterfeit Products in Smart Cities with Digital Twin Technology, 2023 IEEE International Smart Cities Conference (ISC2), Bucharest, Romania, 2023.
- M. Saeed, A. Khan, M. Khan, **M. Saad**, Abdulmotaleb El Saddik, and Wail Gueaieb. Gaming-Based Education System for Children on Road Safety in Metaverse Towards Smart Cities, 2023 IEEE International Smart Cities Conference (ISC2), Bucharest, Romania, 2023.
- M. Khan, **M. Saad**, Abbas Khan, Wail Gueaieb, Abdulmotaleb El Saddik, Giulia De Masi, and Fakhri Karray. Action Knowledge Graph for Violence Detection Using Audiovisual Features, 2024 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 2024.
- **Co-Authored Submissions (Under Review):**
  - *All Languages Matter: Evaluating LMMs on Culturally Diverse 100 Languages*, submitted to **CVPR 2025**.
  - *CP-Diffusion: Conditional Prompt-Based Diffusion Models for Video Generation*, submitted to **CVPR 2025**.

PROJECTS

COVID-19 Progression Visualization (GitHub)
- Applied pre-trained CNNs models and proposed lightweight CNN for COVID-19 X-ray classification.
- Used Grad-CAM to visualize disease progression on X-rays over time, enabling model interpretability.
- Gained insights into CNN performance and critical regions in medical imaging.

Maize Leaf Disease Detection Using Conventional Machine Learning Algorithms (GitHub)
- Implemented a machine learning pipeline for maize leaf disease classification using conventional algorithms, including SVM, KNN, and Random Forest.
- Extracted key features from maize leaf images using texture, color, and shape-based feature extraction techniques.
- Evaluated and compared the performance of algorithms, optimizing for precision, recall, and overall accuracy.

Sequential Models for Video Analysis (GitHub)
- Extracted video features using CNNs, applied LSTM, GRU, and attention-based models for temporal analysis.
- Enhanced temporal analysis with LSTM-Attention and GRU-Attention.
- Explored Vision Transformers for video analysis, assessing their effectiveness compared to traditional CNN-sequential model pipelines.

Violence Detection Using Stacked Ensemble Learning (GitHub)
- Developed a stacked ensemble framework combining advanced architectures (CNN-LSTM, CNN-GRU, ViViT, X3D, and GNN) for real-time violence detection in video surveillance systems.
- Leveraged temporal attention mechanisms and ViViT to enhance interpretability, enabling precise detection of violent actions and patterns in video sequences.
- Designed a robust and scalable solution for diverse surveillance applications, demonstrating adaptability to varying video qualities and environmental conditions.

TECHNICAL SKILLS

- **Programming:** Python, MATLAB, C++, SQL, JavaScript, HTML/CSS
- **Libraries:** Pandas, NumPy, Matplotlib, OpenCV, Hugging Face
- **Frameworks:** Keras, TensorFlow, PyTorch, Scikit-learn, NodeJS, React, A-Frame, ThreeJs
- **Tools:** PyCharm, VS Code, Git, Blender, Docker, LaTeX, AWS (S3, EC2, SageMaker), Digital Ocean, Weights & Biases

HONORS AND AWARDS

- Awarded the University of Ottawa Graduate Studies Scholarship for September 2025
- Award of appreciation for securing 1st position in Youth Talent Expo
- Awarded with a data science certificate by the government of Pakistan (NAVTTC)

ACADEMIC SERVICES

- Attended a virtual Omniverse meeting on its academic and research applications.
- Attended a virtual talk on the metaverse, gaining valuable insights and understanding of its applications.

- Mentor and project evaluator for new batch of DIP Lab, guiding and assessing final-year students' projects.

Extracurricular Activities

- **Academic:** Engaging in reading and writing scholarly articles to stay updated with the latest advancements.
- **Physical Fitness and Sports:** Actively participating in cricket and maintaining physical fitness through regular gym workouts.

References

- Will be provided upon request.