Shourov Joarder

joardershourov60@gmail.com | shourovj | shourovj | shourovj | shourovj

Education

Bangladesh University of Engineering & Technology (BUET) Bangladesh
B.Sc. in Electrical and Electronic Engineering *Feb 2020 – March 2025*
Major: Communication and Signal Processing (*CSP*)
CGPA: **3.88/4.00** (**2nd** in *CSP major*)

Relevant Courses: Artificial Intelligence and Machine Learning | Digital Image Processing I | Computer Programming | Linear Algebra | Probability and Statistics | Random Signals and Processes | Microprocessors and Embedded System | Digital Signal Processing | Control Systems | Digital Electronics

Research Interests

Computer Vision | Multimodal LLM (VLMs) | Medical Image Processing | Autonomous Vehicle

Publications

1. S. Hasan, **Shourov Joader**, A. Nayem, H. Hasan, and S. A. Fattah, *"Multilingual Voice-Controlled Smart Wheelchair with Advanced Features," Published at IEEE ECCE*. (DOI: *10.1109/ECCE64574.2025.11013785*)
   TLDR: Developed a full scale wheelchair integrated with multilingual voice-control, collision detection and avoidance.

2. A. Dhar, D. Sikder, A. Shovon, and **Shourov Joarder**, *"Skin Cancer Semantic Segmentation," Published at IEEE ECCE*. (DOI: *10.1109/ECCE64574.2025.11013785*).
   TLDR: Unet-based Stacked Hourglass model converts cartesian image to polar image which is then fed to a **TransUnet** model for estimating the semantic segmentation on skin cancer **ISIC** dataset.

Research Experiences

Undergraduate Thesis Student, EEE, BUET *Mar 2024 - Mar 2025*

Supervisor : Dr. Kamrul Hasan

*Unsupervised End-to-End Sequential Deep Learning Method for Ultrasound Strain Elastography (***Manuscript in preparation***)*
A medical imaging technique that determines the strain field by tracking displacements between pre and post ultrasound RF frames to detect tumor/lesion in the examined tissue by their strain.

- Developed a two-stage unsupervised model consisting of a new **Contextual Feature Encoder**, a novel **TriCrossAttention**, a **Sequential PWC Decoder**, and proposed two novel losses for Displacement Field and Strain Estimation in Strain Elastography and implemented this in Pytorch.
- My proposed model beats SOTA unsupervised ReUSENet (implemented) in terms of SNR, CNR, NRMSE metrics, and also improves the strain image quality by enhancing the lesion SNR.

Ongoing Researches

Video Difference Understanding *Mar 2025 - Present*
This is an ongoing research on how well VLM's can understand the difference between two videos and how their performance can be improved.

Biasness of VLM on Medical Tasks *Mar 2025 - Present*
Do the state-of-the-art VLM models has any biasness towards medical data? This raises questions on the fairness on these multimodal models.

Work Experience

Adjunct Lecturer, Dept. of CSE, BRAC University. *June 2025 - Present*
Conducting multiple theory and lab courses.

Machine Learning Engineer, ACI Ltd. *Apr 2025 - Present*
Ongoing Projects

- Medical-VLM with explainability and personalized assistance for patients.
- Built a robust OCR and annotation tool using state-of-the art VLM to extract medicine details and bank cheque details from images of a handwritten prescription and cheques.

Competitions

- **1st** Runner-Up of Undergraduate Project Idea Contest at 25th ICCIT 2022, Bangladesh. [Certificate]
- **57th** in the public leaderboard of DL Sprint - BUET CSE Fest 2024, Bengali AI Math Olympiad a LLM based competition.
- **Best Notebook Award** at the DL Sprint - BUET CSE Fest 2022, Bengali ASR Competition.

Honors and Awards

- University Merit Scholarship (3 times) *2020, 2021, 2023*
- University Dean's List Scholarship (2 times) *2021, 2022*
- University Stipend (2 times) *2021, 2022*

Selected Projects

Coding Google's PaliGemma VLM from Scratch Github
Implemented Google's opensource **PaliGemma** Vision Language Model (VLM) from scratch. Implemented the **SigLip** vision encoder, **KV-cache** and the **GemmaLM** with the multimodal projector. This project was inspired by Umair Jamil's paligemma open-source project.

Autonomous Inventory Robot Github
Developed an autonomous robotic system capable of executing real-time voice commands to identify, retrieve, and transport specified objects. The system integrates **Google Speech API** based speech recognition, **YOLOv5-s** for real-time object detection with robotic arm manipulation for precise object grasping with a **Rasphberry-pi** as the processor. Following successful acquisition, the object is returned to a predefined base location using a Line Following Robot (**LFR**) navigation system.

Deep-Learning-based-Breast-Cancer-Classification-Using-VGGIN Github
Trained the **VGGIN** model—a custom deep learning architecture that integrates **VGG-19** with the **Inception** module, on the BreakHis histopathology dataset. Achieved a test accuracy of 99.628%, demonstrating the model's effectiveness in classifying breast cancer subtypes from histopathological images.

Voice Controlled Wheelchair for Disabled Patients Github
Developed a full scale voice-controlled wheelchair for physically impaired people. **GMM**-based trained **VoiceRecognitionV3** module takes voice commands from the patients in any language and in any accent and moves accordingly. In addition, the wheelchair collision avoidance and emergency help feature.[Video].

Machine Learning Based Electrical-Fault-Classification-with-GAF-image Github
ML algorithms like **Decision Tree Classifier**, **Random Forest** and **CNN** based deep learning method with **GAF** transformed images were used to classify 5 different types of electrical fault *(eg. LL, LLL, LG, LLG and No-fault)* from the BUS voltage and current data.

Extracting Audio from Muted Video Github
The main goal of this project was to extract the audio signal from a muted video using signal processing methods in MATLAB. The local and global pixel motions were captured using **Complex Steerable Pyramid** decomposition. This was originally a project by Abe Davis, MIT [Visual Microphone].

Car Theft Detection and Prevention with Automatic GPS Tracking Github
This is an IoT based project. **GSM** and **GPS** technology was used to track and send the car location, and **ESP32-CAM** module was used to capture the photo of the thief and shut the car down.[Video Presentation].

Skills

Programming	Python, MATLAB, C/C++, Verilog, Assembly, Processing3, Linux, LaTeX, Git
Framework	PyTorch, TensorFlow, Jupyter, OpenCV
Simulations and Tools	Raspberry Pi, Arduino, Proteus, Pspice, Quartus, Keil
Frontend Backend	FastAPI, Flusk, Gradio, Android Studio, Kotlin (Basic).

Extra Curriculum

President, BUET Photographic Society *Aug 2024 - Mar 2025*
President, BADHAN-(Ahsanullah Hall Unit, BUET) *Aug 2024 - Mar 2025*