Bipin Saha

bipinsaha.bd@gmail.com | bipin-saha.github.io | +8801828-647005
bipin-saha | bipinsaha | Google Scholar

Education

University of Rajshahi Graduated: Sept. 2022
BSc. in Electrical and Electronic Engineering ***GPA: 3.38/4.00 (3.68 in last 60 credits)***
Notable Coursework: Control System, Digital Signal Processing, Microprocessor and Embedded System, Power Electronics, Biomedical Engineering, Computational Methods, and MATLAB Programming.

Research

- Rahman U, Ferdousi R, **Saha B**, Hossain MS, Islam MJ, Ahamed S, Reaz MIB, Islam MR. **FORS-EMG: A Novel sEMG Dataset for Hand Gesture Recognition Across Multiple Forearm Orientations**. Submitted in **IEEE Transactions on Instrumentation and Measurement**. November 2024.
- Taton TK, **Saha B**, Islam MJ, Mostaque SK. **Prediction of Waiting Time in Queues: An Ensemble Learning Approach**. Submitted in International Journal of Intelligent Computing and Cybernetics. October 2024.
- Taton TK, **Saha B**, Akter A, Islam MJ, Mostaque SK. **Waiting Time Prediction in Queue Management: Leveraging Machine Learning Approach**. Accepted in IEEE ICRPSET 2024 Conference. December 2024.
- Bikasuzzaman M, Polok AP, **Saha B**. **A Transformer-Based Approach for Summarizing Employee Logs**. Submitted in Progress in Artificial Intelligence. August 2024.
- **Saha B**, Islam MJ, Mostaque SK, Bhowmik A, Karmakar T, Chowdhury NH, Reaz, MIB. **Bangladeshi Native Vehicle Detection in Wild**. Under Review in **IEEE Transaction on Intelligent Transportation Systems**. June 2024.
- **Saha B**, Mondal BK, Mostaque SK, Hossain M, Hossain J. **Numerical Modeling of CuSbSe2- based Dual-Heterojunction Thin Film Solar Cell with CGS Back Surface Layer**. AIP Advances. 2023 Feb 1;13(2):025255.
- **Saha B**, Islam MJ, Dipto AS, Mostaque SK. **An Efficient Approach for Appearance-Based Eye Gaze Estimation with 13 Directional Points**. In 2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2) 2021 Dec 26 (pp. 1-5). IEEE.

Experience

Machine Learning Engineer Dec. 2023 - Present
Business Automation Ltd. *Generative Transformers, LLM, VLM, Computer Vision*

- Digitized Handwritten Prescriptions by introducing efficient layout segmentation and OCR (CER 0.0951), leveraged generative LLM models for output formation (IDs, Diagnosis, Medication, Frequencies, and medical histories), optimized both training via implementation of custom loss function and inference time to around 2.28 seconds and minimized human workload by 52.68%, impacting more than 50k patients.
- Implemented Bangla Law Consultancy chatbot using Llama-3 with a contextual RAG system and TF-IDF indexing, optimized training process through QLoRA, QDoRA, and PEFT techniques to achieve faster and more efficient interactions, configured model interaction API and contenraized using Docker.
- Extracted Land Records information using OCR, Bangla Document OSD functioning, and Name Entity Recognition employing the Mistral Model with 4-bit quantization for effective memory allocation.
- Executed BIDA and EBS Employee Behavioral Log Text Summarization enhancing efficacy in task-based accomplishment utilizing Mistral, Pegasus, and ML Models.

Assistant Robotics Engineer Jan. 2023 - Nov. 2023
Get-Aid Ltd. *Computer Vision - Detection, Segmentation, Mobile Robots, Manipulator, PID Control*

- Developed an Object Detection pipeline using YOLOv5 and achieved grocery product recognition with @mAP.50-0.995 and @mAP.50:0.95-0.872, impacting the super shop automation with a mobile robot.
- Implemented 5DOF inverse kinematic manipulator to navigate 3D shelf space using visual perception.
- Implemented PID control algorithm, resulting in precise and accurate slot navigation for the robotic system.

Computer Vision Research Intern Oct. 2021 - Jan. 2022
Brainekt AI Lab

- Developed and annotated image data to prepare an Emotion Recognition Dataset, and trained a deep learning algorithm for 9 facial expression recognition, impacting the field of emotion analysis.

Skills

Languages: Python, C/C++, SQL, MATLAB
AI Stack: Multimodal LLM, VLM, VAE, GPT, Computer Vision, DL, ML
Frameworks: PyTorch, Tensorflow 2.0, Transformers, Scikit-Learn
Modules: Spacy, NLTK, Pandas, OpenCV, Pillow, Numpy, Matplotlib, Seaborn, Flask
DevOps: Docker, Git, GitHub Actions CI/CD
Tools: Coppeliasim, Simulink, Proteus, Easy-EDA
Platforms: Arduino, BluePill, ESP32, 8266, Raspberry Pi, Linux
Documentation: Origin Pro, LaTeX, Adobe Photoshop, Premier Pro

Notable Projects || Complete Project List

Handwritten Prescription Digitalization using Layout Analysis and VLM OCR

Utilizing vision-based layout segmentation and an OCR model, the project digitizes handwritten prescriptions, reducing the CER score for drug names to 0.0951. It accurately extracts IDs, diagnosis frequencies, quantities, and patient histories, converting raw data into structured outputs in 2.28 seconds, minimizing human workload by 52.68%, and improving healthcare record-keeping.

End to End Behavioral Cloning of Self Driving Car

Imitation learning-based self-driving car system utilizing cameras and a specially designed neural network. The network analyzes camera images to steer the car, and a separate PID controller manages the car's speed. This system combines deep learning with control engineering to achieve real-time self-driving navigation. [GitHub] [YouTube]

KUKA LRB IIWA 14 R820 Robot's Object Sorting on Conveyor

Object sorting is implemented using a KUKA robot on a conveyor system. Utilizing inverse kinematics, the robot efficiently handles objects. A color sensor in CoppeliaSim identifies green objects on the conveyor. The robot's gripper then picks up these objects and transfers them to another conveyor belt for seamless sorting. [YouTube]

Slab Bend Detection and Tracking (Collaboration : Tata Steel India)

This real-time industrial video monitoring system, built with Python, uses multiple cameras and a YOLOv8 model to detect defects like metal slab bends. Anomalies are displayed on a user-friendly interface with camera details and confidence scores. The system also logs data, manages users, and connects to IP cameras. [YouTube]

Teleoperated Quadcopter for Aerial Mapping (Gravity Destroyer)

The Gravity Destroyer is a quadcopter drone designed for stable navigation and functioning aerial mapping photography with brushless motors and autopilot. The radio and LiPo battery have a flight time of 20-25 minutes within a 2km radius. Lifting up to 2kg and works with a landmine detector bot. [YouTube]

Vision Based Lane Detection and RRT Path Planning

Develop a lane detection system using the Hough Transformation, preprocessing images with grayscale, Gaussian blur, and Canny edge detection to isolate and detect lane markings. Implement the RRT algorithm in a simulation to plan optimal paths while avoiding obstacles. [GitHub]

Landmine Detection and Disposal Robot

A differential drive robot mounted with a metal detector sensor covering a 180-degree search space performed a grid search and avoided obstacles using an ultrasonic sensor in challenging terrains. [YouTube]

ChatPDF with RAG System - Gemini Pro (Specialized for Research Paper QnA)

ChatPDF with RAG is designed to answer questions about given research papers in a chat-like interface. It uses semantic Retrieval-Augmented Generation (RAG) to answer specific information. Uniquely, it can also translate the research paper's concepts into code, potentially helping implement its ideas. [GitHub] [Deployment]

Text2Text Generation and Summarization using Mistal-7B and Pegasus Model

Using Pegasus and Mistral, employee work logs are automatically being summarized. Managers receive concise summaries highlighting key points, progress, and challenges by feeding verbose logs into the system. The text has been augmented using GPT models for better robustness, streamlining processes for improved decision-making, and performance evaluation. ROUGE-1 score of 0.613, ROUGE-2 score of 0.373, ROUGE-L score of 0.557, and ROUGE-Lsum score of 0.556 [GitHub]

Bengali Social Media Sentiment Analysis using LLama-3, BERT LSTM, BiLSTM Models

This research builds a highly accurate (90.31% test accuracy) sentiment analysis model for Bangla text. It uses word and sentence augmentation techniques to make robust the sentiment (positive, negative, neutral) of Bangla writing, even when the way people express themselves is complex. This is useful for any application that needs to understand how people feel in Bangla text. [GitHub]

Leadership and Volunteering

- **Chairperson**, IEEE Robotics and Automation Society, University of Rajshahi SBC. (December 2020 - October 2022)
- **Secretary**, IEEE University of Rajshahi Student Branch. (December 2019 - December 2020)
- **Volunteer**, Bangladesh Innovative Education Society (April 2018 - March 2019)
- **Instructor**, Design You PID Controller Webinar (August 2021). [YouTube]
- **Instructor**, Arduino Workshop Series 2, 3, 4. Electronics Club, Dept. of EEE, RU (October 2019 - January 2020)

References

Shaikh Khaled Mostaque (misha@ru.ac.bd), Associate Professor, Dept. of Electrical and Electronic Engineering, University of Rajshahi
Md. Johirul Islam (johirul@phy.ruet.ac.bd), Associate Professor, Dept. of Physics, Rajshahi University of Engineering and Technology

Last Update: November 26, 2024