

M Sakib Rahman

☎ +880-1981410182 | ✉ contactwithsakib.rahman.arfin@gmail.com

msakibr.portfolio | github.com/MSakibR | huggingface.co/sakib04 | Google Scholar: M Sakib Rahman

Research Interests

Natural Language Processing · Large Language Models · Multimodal Large Language Models · Prompt Engineering · Low-Resource Bangla NLP · Deepfake Detection · Computer Vision · Code Intelligence

Education

University of Asia Pacific (UAP)

2023 – Present

B.Sc. in Computer Science and Engineering, CGPA: 3.84 / 4.00 (up to 4th Year, 1st Semester)

Dhaka, Bangladesh

Publications

[1] M. Sakib Rahman, Shahrier Alam Shimul, Salman Nur Ramim, Nasima Begum, Tasnim Jahin Mowla. **“BanglaAccento: A Benchmark Dataset for Bangla Regional Dialect Speech Recognition.”** *Data in Brief, Elsevier*. Submitted, July 2026 (Under Review).

[2] M. Sakib Rahman, Md. Noushad Jahan Ramim, Sajjad Ullah Fahim. **“Visual Grounding Repair for Multimodal Large Language Models (MLLMs).”** Manuscript in Preparation. Target: IEEE Conference.

Research Experience

BanglaAccento — Bangla Regional Dialect Speech Recognition System

[2026] – Present

NLP Research, University of Asia Pacific

Ongoing

- Building an end-to-end speech-to-speech communication system for Bangla regional dialects, using 6 hours of curated dialectal speech data (Barishal, Noakhali, Rangpur, Sylhet, Chittagong)
- Co-authoring **“BanglaAccento: A Benchmark Dataset for Bangla Regional Dialect Speech Recognition”** (submitted to *Data in Brief, Elsevier*, under review), extending the released Bangla Regional Speech Dataset

Visual Grounding Repair for Multimodal Large Language Models (MLLMs)

[2026] – Present

CV Research, University of Asia Pacific

Ongoing

- Proposing **Visual Grounding Repair**: detecting, diagnosing, and correcting grounding hallucinations (wrong object, partial overlap, wrong attribute/relation) in MLLMs such as LLaVA, Qwen2.5-VL, and InternVL
- Constructing a VGR dataset by mining prediction errors via IoU thresholds on RefCOCO/RefCOCO+/RefCOCOg/Visual Genome; designing a lightweight repair module combining CLIP re-ranking, SAM/Grounding DINO region proposals, and a verifier

ShutkiVision — A Visual Dataset of Traditional Bengali Dried Fish

[2026] – Present

CV Research, University of Asia Pacific

Ongoing

- Curated and released a 1,537-image dataset across 4 dried fish types under 8 backgrounds and 4 lighting conditions; trained a classification model achieving **98.94% accuracy**

Projects

Bangla OCR using TrOCR | PyTorch, Hugging Face, TrOCR

2026 | GitHub

- Built an end-to-end Optical Character Recognition (OCR) system for printed Bangla documents using Microsoft's TrOCR transformer architecture.
- Implemented image preprocessing, text extraction, and performance evaluation using Character Error Rate (CER) and Word Error Rate (WER).

RAG Chatbot for PDF Question Answering | LangChain, FAISS, LlamaIndex, Llama 3

2025 | GitHub

- Developed a Retrieval-Augmented Generation (RAG) chatbot capable of answering natural language questions from PDF documents.
- Integrated vector embeddings, semantic search, and LLM-based response generation for accurate document retrieval.

Visual Question Answering (VQA) | Qwen2.5-VL, Hugging Face

2026 | [GitHub](#)

- Built a multimodal question answering system capable of reasoning over images and natural language queries.
- Evaluated performance across object recognition, scene understanding, visual reasoning, and text-image understanding tasks.

YOLOv11 Road Damage Detection | YOLOv11, OpenCV, PyTorch

2025 | [GitHub](#)

- Developed a real-time road damage detection system for identifying potholes, cracks, and pavement defects using YOLOv11.
- Performed dataset annotation, augmentation, model training, and evaluation using Precision, Recall, and mAP metrics.

Bangla Named Entity Recognition (NER) | BanglaBERT, Transformers, Hugging Face

2026 | [GitHub](#)

- Fine-tuned transformer-based language models for Bangla Named Entity Recognition to identify persons, organizations, and locations.
- Evaluated sequence labeling performance using Precision, Recall, and F1-score while comparing multiple transformer architectures.

Software & Embedded Systems Projects

Hire Your Wheels — Car Hiring Web Application (Django) | [GitHub](#)

- Feature-rich car rental platform with OTP verification, custom user model, reservation system, and admin dashboard

Easy Hisab — Invoice & GST Automation Web App (Next.js + Django API) | [GitHub](#)

- Automated invoicing platform with tax calculations, customer management, product catalog, and reporting dashboard

NoteMan — Educational Note Sharing System (Next.js + React Native) | [Web](#) | Mobile

- Platform for schools/colleges to upload and distribute notes, study materials, and documents

UCAM System — University Automation (Java + JavaFX) | [GitHub](#)

- GUI-based university class automation system with record management and admin features

Microcontroller Projects — Embedded Systems & IoT Development | [GitHub](#)

- Hands-on microcontroller projects including sensor integration, actuator control, and automation using Arduino and ESP32; implemented IoT-based solutions with real-time data monitoring and system optimization

Open Source Contributions

Bangla Regional Speech Dataset | [Mendeley](#)

[2026]

- 5 hours of mono, 16kHz WAV audio spanning 5 Bangla regional dialects — Barishal, Noakhali, Rangpur, Sylhet, and Chittagong — for dialect identification and low-resource Bangla speech research

A Visual Dataset of Traditional Bengali Dried Fish | [Kaggle](#)

[2026]

- 1,537 images across 4 dried fish types, captured under 8 background and 4 lighting conditions, for image classification and computer vision research

Leadership

Team Lead — ShutkiVision (ML/DL-Based Dried Fish Classification)

Jan 2026 – Jun 2026

University of Asia Pacific

Dhaka, Bangladesh

- Led a 6-member team to develop an image classification system for identifying different types of dried fish (Shutki) using deep learning techniques.
- Managed project planning, delegated responsibilities across data collection, preprocessing, model training, evaluation, and documentation, while ensuring timely project completion.
- Collaborated on dataset preparation, optimized model performance through experimentation, and presented the final solution as part of the university capstone project.

Team Lead — Easy Hisab (Business Management & Invoice System)

Jan 2025 – Jun 2025

University of Asia Pacific

Dhaka, Bangladesh

- Led a 4-member team to design and develop a web-based business management platform featuring invoice generation, customer management, expense tracking, and financial reporting.
- Designed the overall system architecture, assigned frontend, backend, database, and testing tasks, and coordinated development throughout the project lifecycle.
- Oversaw feature integration, code reviews, and deployment preparation while ensuring the application met functional and usability requirements.

Technical Skills

Languages: Python, JavaScript, Java, C++

NLP / LLM: PyTorch, Prompt Engineering, LLM Fine-tuning, BERT, T5, HuggingFace Transformers

CV / Deep Learning: EfficientNet, ResNet, GradCAM++, CNN, MobileNet

Full-Stack: Next.js, React, Django, Tailwind CSS, Vercel

Tools: Git, Selenium, Gemini API, HuggingFace Hub, Google Colab, Conda