

Sadab Shiper

 researchgate.net/profile/Sadab-Shiper
 scholar.google.com/pAK8RoYAAAAJ
 linkedin.com/in/sadabshiper04
 github.com/SadabShiper

 sadabshiper.github.io
 sadabshiper@iut-dhaka.edu

RESEARCH INTERESTS

Computer Vision, Natural Language Processing, Vision-Language Models, Multimodal Learning, and Low-Resource Language AI.

EDUCATION

Islamic University of Technology (IUT), Gazipur, Bangladesh
B.Sc. in Computer Science & Engineering

July 2021 - October 2025
CGPA: 3.24/4.00

PUBLICATIONS

Conference Publications

- **Enhancing Vision Language Corruption Robustness using Cross Distribution & Prompted Denoisers**
SS Latif*, **Sadab Shiper***, KMR Kiran*, MF Ishmam*, MA Hossain, ARM Kamal, MH Ashmafee.
IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2026.
[Paper] [Supplementary Material] [Code]

Under Review

- **BANGLAWILD: An In-the-Wild Bengali Scene Text Recognition Benchmark for OCR and Vision-Language Models**
Sadab Shiper*, TT Dipto*, MM Inzamam, E Tanzeem.
Submitted to ACL Rolling Review (August 2026)
[Paper] [Code] [Data]
(Code & data will be available upon acceptance.)

Work in Progress

- **Re-Think: Interactive Two-Stage Self-Regulated Reasoning for LLMs.**
TT Dipto, Abdullah, **Sadab Shiper**, RI Salvi, Amina, H Mahmud.
(Work in progress)

INDUSTRY & RESEARCH EXPERIENCE

ARI - Arbitrary Research Institute

July 2026 - Present

Research Fellow

- Conducting research and technical training in Advanced Natural Language Processing, Machine Translation, Systems, and AI Agent Evaluation.
- Completed structured training in scientific writing, manuscript preparation, systematic reviews, data visualization, and LaTeX through the Elsevier Researcher Academy.

Apurba Technologies Ltd.

October 2025 - May 2026

Machine Learning Engineer

- Engineered a **4-stage document deskewing and orientation correction pipeline** for Bengali OCR: skew detection via Hough probabilistic line transform with statistical mode estimation; affine rotation with adaptive canvas expansion; upright/upside-down classification using a 5×5 spatial grid voting scheme based on ink-density asymmetry across text-line halves; achieving robust correction on low-quality, rotated, and mixed-orientation scanned documents.
- Designed a **single/multi-column layout detection algorithm** using vertical ink projection analysis with adaptive Gaussian smoothing and 15th-percentile gap thresholding; achieved **100% accuracy on 22 newspaper layouts** and **93%+ on 69 single-column documents**.
- Profiled fp8-quantized OCR model across 26 documents, discovering progressive RAM accumulation (73.9% → 98.3%) as the primary production bottleneck versus consistently low GPU utilization (~9.7%), directly informing the team's memory management strategy.
- Benchmarked **10 state-of-the-art OCR models** on Bengali document data (olmOCR-2-7B, HunyuanOCR, PaddleOCR, DeepSeek-OCR, GLM-OCR, FireRed-OCR, and others) for production pipeline selection.
- Drove the team's VLM fine-tuning strategy via literature review across 5 pipelines (LoRA, QLoRA, PEFT, SFTT), identifying optimal frameworks (TRL, Hugging Face) for the OCR fine-tuning roadmap.

- Built the **Apurba RAG Chatbot** pipeline from scratch - sliding window chunking, 6 embedding models, TinyLLaMA validation; R&D showed Qwen outperforms Gemma for Bengali-context generation.
- Developed the **Apurba HRM Suite Portal** backend using FastAPI with 40+ REST API endpoints with JWT authentication, timezone-aware ISO-8601 timestamps, and role-based access control.
- Architecting an **AI-driven Employee Evaluation microservice** with GitHub API client (REST/GraphQL), cron-based data ingestion scheduler, LLM-based productivity scoring engine, and Dockerized deployment.

SELECTED PROJECTS

Football Game Analysis | Player & Ball Tracking with Team Detection [Code](#)

Computer vision pipeline detecting and tracking players, referees, and ball using **YOLOv8** + **ByteTrack**; custom team classification via KMeans jersey-color clustering; visual annotations with ellipses, triangles, and track IDs.
Technologies: Python, OpenCV, YOLOv8, ByteTrack, scikit-learn, Roboflow, supervision

SahityaBot | Multilingual Bangla RAG Chatbot [Code](#)

RAG chatbot for Bengali/English queries over HSC'26 Bangla literature using **Gemini 1.5 Flash**; OCR extraction via pdf2image + pytesseract, MiniLM dense embeddings, FAISS retrieval, CrossEncoder reranking; groundedness/relevance evaluation pipeline with auto-generated HTML reports.
Technologies: Python, FastAPI, Gemini API, Tesseract OCR, FAISS, sentence-transformers, CrossEncoder, Jinja2, Tailwind

Bangla News Article Classification [Code](#)

GRU/LSTM/CNN models for Bangla news categorization; **GRU: 91.72%**, LSTM: 90.99%, CNN: 90.60% accuracy; weighted loss for class imbalance; tokenization and sequence padding pipeline.
Technologies: Python, TensorFlow, Keras, NumPy, Pandas, scikit-learn

AWARDS & ACHIEVEMENTS

- **3rd place** among ~30 teams at **Datathon, 4th DIU National Data Science Summit 2024** *December 2024*
- **8th place** among 108 teams at **Game of Datathon, 3rd KUET CSE National Fest** *December 2024*
- Selected for 4-month hackathon: **Geeky Solutions Learnathon 3.0**, sponsored by Brain Station 23 *January 2025*
- **Top 20** among ~83 teams at **FOREWARN Disaster Hackathon 2.0** *May 2024*

TECHNICAL SKILLS

AI & LLM: Hugging Face Transformers, LangChain, FAISS, LoRA, QLoRA, PEFT, RAG Pipelines
Computer Vision & OCR: PyTorch, OpenCV, TensorFlow, YOLOv8, PyMuPDF, Tesseract OCR, pdf2image
Data Science & ML: Pandas, NumPy, scikit-learn, NLTK, spaCy, Weights & Biases
Languages & Tools: Python, C++, FastAPI, Docker, AWS, Git, Linux

ALGORITHMIC PROBLEM SOLVING

Codeforces: [shielder04](#) - 375 problems solved **LeetCode:** [sadabsamishiper](#) - 120 problems solved