

KHONDKER SALMAN SAYEED

salkhon050@gmail.com | salkhon.github.io

🎓 Khondker Salman Sayeed |  salkhon |  salkhon |  SalmanKhondker

RESEARCH INTEREST




- Natural Language Processing
- Vision Language Modeling
- Multimodal Learning
- Low Resource Deep Learning

My research goal is to advance research and development in Natural Language Processing and Multimodal Learning, with a focus on Vision-Language Models. Seeking opportunities to apply AI for real-world impact, leveraging strong technical skills and a proven track record of success in deep learning competitions.



EDUCATION


- **Bangladesh University of Engineering and Technology (BUET)** April 2019 - June 2024
Bachelor of Science in Computer Science and Engineering Dhaka, Bangladesh
CGPA: 3.9/4.0, placed on the dean's list in all terms. (academic session delay due to COVID-19 pandemic)

PUBLICATIONS

- | | |
|--------------|--|
| Accepted | Haz Sameen Shahgir [†] , Khondker Salman Sayeed[†] , Abhik Bhattacharjee, Wasi Uddin Ahmad, Yue Dong, Rifat Shahriyar (Joint first-author)
Title: "IllusionVQA: A Challenging Optical Illusion Dataset for Vision Language Models".
[arXiv]
 COLM 2024: The 1st Conference on Language Modeling |
| Accepted | Haz Sameen Shahgir, Rownok Zahan Ratul, Md Toki Tahmid, Khondker Salman Sayeed , Atif Hasan Rahman
Title: "RNA-DCGen: Dual Constrained RNA Sequence Generation with LLM-Attack".
[bioRxiv]
 NeurIPS 2024 MLSB: Workshop on Machine Learning in Structural Biology |
| Under Review | Tamzeed Mahfuz, Satak Kumar Dey, Ruwad Naswan, Hasnaen Adil, Khondker Salman Sayeed , Haz Sameen Shahgir
Title: "Too Late to Train, Too Early To Use? A Study on Necessity and Viability of Low-Resource Bengali LLMs".
[arXiv]
 COLING 2025: The 31st International Conference on Computational Linguistics |
| Competition | H.A.Z. Sameen Shahgir [†] , Khondker Salman Sayeed[†] , Tanjeem Azwad Zaman [†] , Md. Asif Haider [†] , Sheikh Saifur Rahman Jony, M. Sohel Rahman (Joint first-author)
Title: "Ophthalmic Biomarker Detection Using Ensembled Vision Transformers – Winning Solution to IEEE SPS VIP Cup 2023".
[arXiv]
 ICIP 2023: IEEE SPS Video and Image Processing Cup |

WORK EXPERIENCE

- **IQVIA**  June 2024 - Present
Machine Learning Engineer Remote
 - Joined the Next Best Action ML team at the IQVIA Orchestrated Analytics project.
 - Working on training, testing, and deploying IQVIA's in-house Large Language Models for generating expert domain-specific insights on the pharmaceutical industry.
- **BUET CSE NLP Lab**  March 2023 - June 2024
Research Assistant Dhaka, Bangladesh
 - Prepared a benchmark illusion dataset that tests the visual understanding capabilities of Vision Language Models to identify visual illusions in images and correctly describe them. Published this first-author work at COLM 2024.
 - Worked on extending the XLSum dataset from BUET CSE NLP group to include images and BBC article-summary pairs.
 - Used multimodal models using images and articles to generate summaries from the extended dataset to research potential improvements in the quality of generated summaries as a result of the inclusion of images

 Supervisors: [Prof. Rifat Shahriyar](#), CSE BUET. [Wasi Uddin Ahmad](#), Senior Research Scientist, NVIDIA. [Prof. Tahmid Hasan](#), CSE, BUET.

HONORS AND AWARDS

- **IEEE Upsilon Pi Epsilon Honor Society Award 2023 (Awardee)**

February 2024

IEEE Computer Society



- Received this prestigious award in recognition of my academic results and success in national and international deep learning competitions.
- Awarded to only 4 applicants internationally.
- Received a \$1000 scholarship.

- **IEEE Video and Image Processing Cup 2023 (Champion)**

October 2023

IEEE Signal Processing Society, ICIP 2023



- Champion in this annual international computer vision competition. 2023's edition was based on biomarker prediction from retina OCT scans.
- Received travel grant to present our solution at ICIP 2023, Kuala Lumpur, and a \$5000 scholarship prize.
- **Summary:** In this work, we employed two vision transformer-based models: MaxViT and EVA-02. Our ensembled solution achieved a patient-wise F1 score of 0.81 in the first phase and 0.85 in the second and final phase, setting the state-of-the-art ophthalmic biomarker detection using OCT images.

- **DL Sprint 2022 (Champion)**

September 2022

Optimizely, Bengali.AI, and CSE BUET



- Champion in this Kaggle community competition based on Automatic Speech Recognition on the Bengali Common Voices Dataset.
- Received a \$1300 scholarship.
- **Summary:** In this work, we fine-tuned wav2vec 2.0 for Bengali speech recognition using the Bengali Common Voice Speech Dataset. Our model achieved impressive results, with a word error rate (WER) of 0.25 on the validation set and a Levenshtein Distance of 2.61 after further training. This work bridges the gap for a widely spoken language.

- **EEE Day Datathon 2023 (First Runner-up)**

March 2023

EEE BUET, Apurba Technologies Ltd.



- Runner-up in this Kaggle community competition based on Automatic Grammatical Error Tagging on Bengali Text.
- Received a \$500 scholarship.
- **Summary:** In this work, we propose a method for detecting grammatical errors in Bangla using a Text-to-Text Transfer Transformer (T5) Language Model. We fine-tune the small variant of BanglaT5 on a corpus where errors are bracketed by the dedicated symbol \$. Despite the T5 model's primary design for translation, we achieve low Levenshtein Distance in tagging grammatical errors in Bangla.

- **AI For Bangla 2023 (Honorable Mention)**

March 2023

EBLICT, Bangladesh Computer Council (BCC)



- Prize winner in this national competition calling for advancements in AI for Bengali for creating the first Bengali Sign Language video dataset and preparing a baseline on that dataset.
- Received a \$500 scholarship.
- **Summary:** In this research, we introduce a new word-level Bangla Sign Language dataset, BdSL40, comprising 611 videos across 40 words. We prepare two baselines for classification: one using a 3D Convolutional Neural Network (CNN) model and another employing a novel Graph Neural Network (GNN) approach.

- **Robi Datathon 2024 (First Runner-up)**

May 2024

Robi Axiata Ltd, Huawei



- Runner-up in this national data-science competition based on customer purchase behavior prediction in a data-scarce setting.
- Received a \$3000 scholarship.

LEADERSHIP EXPERIENCE

- **Lead Organizer & Instructor of DL Sprint 2.0 2023**

August 2023

BJIT Group, CSE BUET



- Served as an organizer of this open-for-all Computer Vision Competition on Automatic Document Layout Analysis for Bengali.
- Conducted two workshops as an instructor on common competition practices, computer vision, and image segmentation.
- Took ownership, being involved in the competition's conception to execution.

PROJECTS

-  **Ticketing** – A microservices web app that lets users buy and sell tickets
 - Backend: NodeJS, Typescript, MongoDB, NATS Streaming. Consists of 6 independent services, coordinated by an eventbased architecture.
 - Frontend: NextJS, TailwindCSS
 - Orchestration: Kubernetes
-  **Sub-C-Compiler** – A compiler for a subset of the C programming language
 - Tools: Flex (Lexer), GNU Bison (Parser)
 - Languages: C, C++
-  **Mobile-Doc** – A telemedicine web app
 - Backend: FastAPI, MongoDB, Redis, BigQuery
 - Frontend: React, MUI
 - Deployment: Netlify
 - Deployment: [Pokedoc](#)
-  **Suntech** – An e-commerce web app
 - Backend: Flask, OracleDB
 - Frontend: Jinja2, HTML, CSS, VanillaJS
-  **Ray-Tracing** – A rendering pipeline implementing ray tracing in OpenGL scenes
 - Language: C++
 - Algorithms: Phong Lighting Model, Recursive Reflection
-  **Enhanced xv6 OS** – Copy-On-Write and Memory Page-Replacement in the xv6 OS
 - Language: C
 - System Calls, Process Schedulers, Page Replacement
-  **twitt3r** – A Twitter clone using the t3 stack
 - Framework: NextJS
 - Stack: T3 (Typescript, trpc, TailwindCSS)
 - Persistence: PlanetScale, Prisma ORM
 - Deployment: Vercel
 - Deployment: [twitt3r](#)
-  **DocumentQA** – A langchain LLM document chatbot
 - Language: Python
 - Libraries: langchain, OpenAI
 - UI: Chainlit
 - Database: ChromaDB Vector Database
-  **IllusionVQA** – A Challenging Optical Illusion Dataset for Vision Language Models
 - This work is a part of my undergraduate thesis
 - Published at **COLM'24**
 - Paper: [IllusionVQA: A Challenging Optical Illusion Dataset for Vision Language Models](#)
 - Website: [IllusionVQA](#)

REFERENCES

1. **Rifat Shahriyar (PhD)**
Professor, Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology (BUET)
Email: rifat@cse.buet.ac.bd
Relationship: Undergraduate Thesis Supervisor
2. **AKM Ashikur Rahman (PhD)**
Professor, Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology (BUET)
Email: ashikur@cse.buet.ac.bd
Relationship: Academic Advisor