

SAKIB CHOWDHURY

<https://sakibchowdhury131.github.io/>

+1 (201) 448-6568

sakibchowdhury131@gmail.com

SUMMARY

Robotics and machine learning researcher specializing in generative policies and learning from human demonstrations for robotic manipulation. Experienced in multimodal policies, world models and force-conditioned policy learning, with a focus on learning generalizable robot behaviors. Developing and deploying learning-based policies across simulation and real robotic platforms.

EDUCATION

Ph.D. in ECE, *CGPA: 3.97/4.00* (September 2023 - May 2028 (Expected)) Stevens Institute of Technology

Research Focus: Learning multimodal behaviors from human demonstrations using generative policies.

M.Sc. in ECE, *CGPA: 3.97/4.00* (September 2023 - December 2025) Stevens Institute of Technology

Relevant Coursework: 3D Computer Vision, Applied Modeling & Optimization; Autonomous Navigation for Mobile Robots; Mobile Manipulators; Wearable Robotics;

B.Sc. in EEE, *CGPA: 3.65/4.00* (February 2017 - March 2022) Bangladesh University of Engineering and Technology (BUET)

Relevant Coursework: Linear Control Systems; Digital Logic Design; Digital Signal Processing, Microprocessors, C Programming;

RESEARCH EXPERIENCE

Graduate Researcher (Robotics & Automation Lab)

Stevens Institute of Technology, NJ, USA

Sept 2023 – Present

- Developed a **force-conditioned generative visuomotor policy for robotic tool use from human demonstrations**, introducing robot-replay-based external force capture from visual demonstrations and onboard joint torque sensing, eliminating the need for external F/T sensors. Used gaussian splatting to render novel views of the scene. Demonstrated generalizability across **diffusion, ACT, and flow-matching policies**, with force-based admittance correction improving manipulation performance. [\[LINK\]](#)
- Developed a **neural motion planner for time critical tasks**. showed that joint space trajectories generated by tree-based classical planners are piecewise linear in nature, and learning the breakpoints of these piecewise linear trajectories can make the planner faster, **achieving 6.5 ms average planning time** for rapid-response tasks such as robotic ball catching. [\[LINK\]](#)
- Developed a **residual learning method for projectile trajectory prediction** to capture unmodeled dynamics such as friction, air resistance, and restitution; demonstrated improved interception performance in a simulated robotic table-tennis environment using PyBullet. [\[LINK\]](#)

ML Research Engineer

Celloscope, Dhaka, Bangladesh

Dec 2021 – Aug 2023

- Developed **Bangla automatic speech recognition (ASR) and text-to-speech (TTS) systems**, fine-tuning Wav2Vec and Whisper models and training a **Conformer-based ASR** using ESPnet/Kaldi-style recipes; achieved **21% WER on raw production audio**, outperforming Wav2Vec and Whisper on Bangla; Trained a **Tacotron-based TTS** system for Bangla; [\[LINK\]](#)
- Built an **audio-to-dialogue extraction pipeline** combining speaker diarization, ASR, and downstream processing to automatically transform multi-speaker audio into structured dialogue.
- Designed a **data collection and annotation pipeline using Label Studio** for image-to-text extraction, supporting dataset creation and model development for document and license-plate recognition.
- Trained and deployed **DONUT, a document understanding Vision Transformer (ViT)**, for Key Information Extraction (KIE) from **Bangladeshi national ID cards and vehicle license plates**; [\[LINK\]](#)

SELECTED PROJECTS

- **Entropic Flow Policy:** Developed an entropy-guided policy learning framework that replaces reward/action-matching objectives with goal-reachability entropy, guiding a diffusion-style action sampler toward lower-entropy states. Evaluated the approach on robotic reaching, obstacle navigation, and Push-T manipulation tasks. [\[LINK\]](#)
- **Heat Dissipation Policy:** Developed a diffusion policy that replaces Gaussian corruption with a heat-equation-based forward process, progressively removing high-frequency trajectory structure before low-frequency structure. [\[LINK\]](#)
- **Gaussian Splatting of a Robot Arm:** Reproduced a Gaussian-splatting-based simulation pipeline to render novel views of a real robotic workspace while moving the robot in simulation. Integrated real-world workspace reconstruction with simulated robot interaction to improve visual realism and reduce the sim-to-real gap for learning-based robotic manipulation. [\[LINK\]](#)
- **Synthetic Episode Generation using COSMOS3:** Used the NVIDIA Cosmos3 video world model to generate synthetic videos of tool-use interactions for training robotic manipulation policies. [\[LINK\]](#)
- **OpenVLA vs. OpenVLA-OFT:** Evaluated OpenVLA and OpenVLA-OFT vision-language-action (VLA) policies locally on NVIDIA RTX A6000 GPUs using Franka Panda simulation, comparing their performance on robotic manipulation tasks. Achieved superior success with OpenVLA-OFT and analyzed how fine-tuning and differences between simulation and real-world data affect VLA performance. [\[LINK\]](#)
- **Heat Dissipation Image Generation:** Developed an image generation model that replaces Gaussian diffusion noise with a heat-equation-based corruption process. Demonstrated image reconstruction and iterative generation while preserving the mass-conservation property of the heat equation. [\[LINK\]](#)

TECHNICAL SKILLS

ML Frameworks	PyTorch, TensorFlow, Scikit-Learn, Numpy
Robotics Softwares and Simulators	ROS, Bullet, MuJoCo, IsaacSim, Gazebo, OMPL, MoveIt
Programming Languages	Python, C/C++, Bash, MATLAB, Simulink
Embedded Microcontrollers	ATmega328, STM32, ESP32, MSP430
Others	Linux, Docker, Git, Flask, OpenCV, FastAPI, PostgreSQL

SELECTED PUBLICATIONS

- **S. Chowdhury**, R. Huang, V. Chen, Y. Guo, “Force-Conditioned Imitation Learning from Vision-Only Human Demonstrations via Robot Replay”, (Submitted at ICRA 2027). [\[LINK\]](#)
- **S. Chowdhury**, Y. Guo, “FlashPlan: A Fast Neural Motion Planner for Robotic Ball Catching”, (Submitted at ICRA 2027). [\[LINK\]](#)
- **S. Chowdhury**, Y. Guo, “Learning to Strike for Robotic Table Tennis”. [\[LINK\]](#)
- **S. Chowdhury**, M. Morshed, S. A. Fattah, “SpectroCardioNet: Attention-Based Triple-Spectrogram PCG Network for Cardiac Disease Detection”, IEEE Sensors Journal, 2022. [\[LINK\]](#)
- **S. Chowdhury**, D. K. Sikder, A. Roy, “A Simulated Intelligent Pixelated Electrode Array for Surface Electromyography Sensors”, IEEE Sensors Journal, 2023. [\[LINK\]](#)
- S. M. Monsur, Shariar Kabir, **S. Chowdhury**, “SynthNID: Synthetic Data for Bangla Key Information Extraction”, EMNLP Bangla Workshop 2023. [\[LINK\]](#)
- S. M. Monsur, **S. Chowdhury**, M. S. Fatemi, et al., “SHONGLAP: A Large Bengali Open-Domain Dialogue Corpus”, LREC 2022. [\[LINK\]](#)
- T. Mahmud, M. J. Alam, **S. Chowdhury**, et al., “CovTANet: Hybrid Tri-Level Attention Network for COVID-19 Lesion Segmentation and Severity Prediction”, IEEE Transactions on Industrial Informatics, 2021. [\[LINK\]](#)