

# Siddhant Bikram Shah

 SiddhantBikram |  siddhantbikram |  siddhantshah3000@gmail.com |  Google Scholar

## EDUCATION

---

- **Delhi Technological University** New Delhi, India  
*Bachelor of Technology* 2019 - 2023
  - Major: Computer Engineering
  - Minor: Machine Learning
  - CGPA: 9.09/10

## PUBLICATIONS

---

- Bhandari A.\*, **Shah S.B.\***, Thapa S.\*, Naseem U., Nasim M.  
CRISISHATEMM: MULTIMODAL ANALYSIS OF DIRECTED AND UNDIRECTED HATE SPEECH IN TEXT-EMBEDDED IMAGES FROM RUSSIA-UKRAINE CONFLICT  
*In 2023 MMCM Workshop, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1993-2002).*
- **Shah S.B.**, Garg S., Bourazeri A.  
EMOTION RECOGNITION IN SPEECH BY MULTIMODAL ANALYSIS OF AUDIO AND TEXT  
*In 2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 257-263).*
- **Shah S.B.**, Bhandari A., Shambharkar P.G.  
DEEP LEARNING METHODS FOR COVID-19 MITIGATION: APPLICATIONS, CHALLENGES AND FUTURE IMPLICATIONS  
*In 2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) (pp. 775-781).*
- Bhandari A., **Shah S.B.**, Shambharkar P.G.  
THE CONFLUENCE OF AI AND BLOCKCHAIN IN TACKLING THE COVID-19 PANDEMIC  
*In 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-6).*

\* denotes equal contribution.

## EXPERIENCE

---

- **Research Intern** Remote  
*University of Illinois Urbana-Champaign; Advisor: Dr. Haohan Wang* Jul'23 - Present
  - Creating a counterfactual image generation method to solve the problem of long-tailed features simultaneously in inter-class and intra-class conditions.
  - Working on a contrastive learning approach to learn variant and invariant features within classes.
- **Research Intern** Remote  
*University of Essex; Advisor: Dr. Aikaterini Bourazeri* Jul'22 - Present
  - Proposed a multimodal framework to diagnose depression by analyzing audio, text, and video data.
  - Incorporating the model into a counseling chatbot to provide need-based therapy.
- **Research Intern** Remote  
*BASIRA Lab, Imperial College London; Advisor: Islem Rekik* Jan'23 - May'23
  - Created a Graph-U-Net for simultaneous connective brain template generation and graph reconstruction from a population of multi-view brain networks.
  - Implemented SIMLR-based clustering for more representative sample selection.

- **Undergraduate Researcher** Delhi, India  
*Delhi Technological University; Advisor: Prashant Giridhar Shambharkar* *Jan'22 - May'23*
  - Investigated the applications, challenges, and future implications of AI, DL, and Blockchain in dealing with the COVID-19 pandemic.
  - Developed a multimodal system to diagnose Alzheimer's disease by combining speech and text modalities.
- **Computer Vision and Software Engineering Intern** Punjab, India  
*Trident Group* *May'22 - July'22*
  - Developed an application to supervise and streamline textile production by using edge detection on real-time videos.
  - Built OCR models to solve captcha challenges for business-process automation.
- **Application Development Intern** Remote  
*Molog Media* *Sep'21 - Nov'21*
  - Developed a large-scale website "So'Ham" for the National Gallery of Modern Art, New Delhi.
  - Used Flask to facilitate remote MySQL communication and manage various website functionalities.

## PROJECTS

---

- **CrisisHateMM: Hate detection in text embedded images**
  - Released a comprehensively annotated multimodal dataset for text-embedded images pertaining to the Russia-Ukraine conflict.
  - Conducted experiments with various architectures and measured baseline results.
- **Image Generation using a Text Prompt**
  - Developed a framework to generate images based on a text prompt by using VQGAN-CLIP.
  - Used VQGAN as the generator, CLIP as the discriminator, and ESRGAN for super-resolution.
- **3D Brain Tumour Segmentation using 3D U-Net**
  - Implemented a U-Net model for 3D brain tumor segmentation in multimodal MRI images from the BraTS2020 dataset.
  - Used a 3D autoencoder to extract representations from segmented images and fed this data to an SVR model to predict the age and number of days of survival of patients.
- **Speech Synthesis by using Deep Learning**
  - Used the SV2TTS framework to develop a speech recognition and synthesis system.
  - The system replicated input voice characteristics by using an encoder, vocoder, and synthesizer.

## ACADEMIC AWARDS

---

- Recipient of the Study in India scholarship provided by the Government of India for undergraduate study at Delhi Technological University (2019–2023)
- Awarded cash prize for superb performance in AISSCE 12th grade (April 2018)
- Awarded cash prize for superb performance in AISSCE 10th grade (April 2016)

## TECHNICAL SKILLS

---

- **Programming Languages:** Python, C++, C
- **Familiar with:** HTML, CSS, Javascript,  $\LaTeX$
- **Frameworks:** Pytorch, Pytorch-Geometric, Numpy, Tensorflow, Flask