Mithilesh Vaidya | CV

mithilesh.vaidya@iitb.ac.in • methi1999.github.io
Final year student at IIT Bombay

Education

Indian Institute of Technology Bombay, Mumbai, India 2017-Present
- Bachelor and Master of Technology in Electrical Engineering, **CPI 9.47/10**
- **Minor** in Computer Science and Engineering

Academic Achievements

- Ranked **2nd** in Electrical Engineering Dual Degree Programme among **61 students** [2021]
- Awarded **AP grade** in **Control Systems** course for **exceptional performance** [2020]
- Secured an **All India Rank of 388** in JEE Advanced 2017 among 2,00,000 candidates [2017]
- Awarded the **Kishore Vaigyanik Protsahan Yojana (KVPY)** Fellowship by Govt. of India [2015]
- Recipient of the prestigious **National Talent Search Examination (NTSE)** scholarship by National Council of Educational Research and Training (NCERT), Government of India [2015]
- **Silver medal** in Homi Bhabha Young Scientist Examination [2011]

Publications

- **M. Vaidya**, K. Sabu, P. Rao, "Deep Learning for Prominence Detection in Children's Read Speech", *arXiv (2021)*, 2110.14273. *To be published at ICASSP 2022.*

Research Experience

- **Prominence Detection in Children's Read Speech** Jan'21 - Oct'21
  Guide: Prof. Preeti Rao, IIT Bombay Master's Thesis I
  - Replaced a Random Forest Classifier baseline with a **CRNN** framework for predicting the degree of prominence for each word in children's read speech
  - Explored inputs at various hierarchies: raw waveforms, acoustic contours and word-level aggregates
  - Demonstrated an improvement in the acoustic features extracted from word segments using Sinc convolution
  - Exploited phrase boundary labels in various **multi-task learning** paradigms
  - Used part-of-speech tags and various **NLP** embeddings such as **GloVe** and **BERT** for incorporating complementary lexical information
- **Comprehensibility Rating for Children's Read Speech** May'21 - Present
  Guide: Prof. Preeti Rao, IIT Bombay Master's Thesis II
  - Extracted acoustic features from **raw waveforms** using various deep convolutional architectures
  - Experimenting with **transfer learning** by using pre-trained models trained on Emotion Recognition datasets
  - Exploring Attention mechanisms for **fusing** lexical and acoustic features
  - Exploring self-supervised learning for extracting **prosody embeddings** from a vast unlabelled dataset
- **Character Animation from Video for Blender** July'21 - Present
  Guide: Prof. Parag Chaudhuri, IIT Bombay Research Project
  - Working on a Blender plugin consisting of an **integrated pipeline** for extracting **3D human pose** from a video using various deep learning backends and **retargeting** it to a rigged character in Blender
  - Added VIBE and MediaPipe to the pose extraction backend of the plugin
  - Exploring a **self-supervised** graph neural network framework for dynamic mapping of animations from source to dissimilar target skeletons

- **SIRD Dynamics** Aug'20 - Dec'20
  Guide: Prof. Sharayu Moharir, IIT Bombay Research Project
  - Studied the **SIRD** model which is widely used for studying the outbreak of epidemics
  - Simulated the model with various underlying **network topologies** in place of random mixing
  - Formulated multiple mathematical models for calculation of **precise dynamics**

Professional Experience

- **Verification of FPGA-based High Frequency Trading Platform** Apr'20 - June'20
  APT Portfolio Pvt. Ltd. | Guide: Mr. Vivek Pannikar, Senior Verification Engineer Internship
  - Implemented **DPI**, a protocol for exchanging data between SystemVerilog and other languages, for **speeding up verification** of testbenches using Cocotb, Quartus and Riviera by **3x**
  - Used Python **metaclasses** for automatically generating Python, SystemVerilog and C DPI header and implementation files from high-level JSON inputs
- **Autonise AI** Sep'18 - May'19
  Machine Learning Startup Co-Founder
  - Implemented **PixelLink** and a GRU for word-level text detection, **invariant** to font size, colour, background, orientation, etc. and demonstrated an accuracy of **74%** on a proprietary dataset of documents like Aadhar Card, Passport, Driving Licence, etc.
  - Implemented a robust model with a **UNet** backbone for **segmenting** out spots, patches and wrinkles in selfies and exposed it through **AWS for demonstration**

Key Projects

- **Legendre Memory Unit** | Course Project: Advanced Machine Learning Jan'21 - May'21
  - Implemented and analysed the performance of Legendre Memory Units (LMU), **an improved sequential model**, on various tasks and datasets such as JSB Chorales, Mackey-Glass dynamics, etc.
  - Suggested modifications to the **core equations** by studying various **basis** functions
- **Audio Steganography** | Course Project: Automatic Speech Recognition Jan'21 - May'21
  - Exploited **adversarial attacks** on ASR systems for hiding any given sequence of tokens in any audio file
  - Analysed performance as a function of various token sequence properties such as **length and perplexity**
  - Demonstrated **high PESQ scores** which indicate low perceptibility of deviation from original audio
- **Video Toonification** | Course Project: Digital Image Processing Aug'20 - Dec'20
  - Used **Mean Shift Segmentation** across both time and spatial dimensions for toonification of videos
  - Benchmarked results with standard techniques such as Bilateral Filtering
- **Auction Theory** | Course Project: Game Theory Aug'20 - Dec'20
  - Studied various models in Auction Theory such as **Vickrey Auction** and **First Price Sealed Bid** Auctions
  - Discussed equilibrium and optimal auction design analysis
- **FMX Rendering and Animation** | Course Project: Computer Graphics Nov'20 - Dec'20
  - Designed and rendered an FMX track with obstacles of varying shapes such as **cylinders and ramps**
  - Designed, rendered and animated a rider and a motorbike on the track using **keyframing**
  - Employed **Phong Shading, Texture Mapping** and used a **Skybox** for a realistic look
- **Pipelined RISC Processor** | Course Project: Microprocessors Oct'19 - Nov'19
  - Designed a 16-bit, 8-register, 6-stage **Pipelined** RISC processor in VHDL
  - Employed **Branch Prediction** and **Hazard Mitigation** techniques for optimizing the performance

- **FindIt** | Self Project: Audio Fingerprinting May'19 - June'19
  - FindIt is a Python program for identifying a song given a **short noisy segment**, similar to Shazam
  - An **audio fingerprint** consisting of constellations of major time-frequency peaks is stored in a hash table
- **Handwriting Recognition Pen** | Summer Project May'18 - June'18
  - Built a pen which can instantly convert **handwriting strokes** on ordinary paper into text
  - Designed the pen from scratch in AutoCAD and **3D printed** it
  - Generated own training data for each letter using a custom OpenCV script

Technical Skills

- **Programming Languages:** Python, C++, C, Bash, Verilog, VHDL, OpenGL
- **Softwares:** Matlab, Arduino, LaTeX, Blender, VHDL, AutoCad, Solidworks, Android Studio
- **Data Science:** PyTorch, Pandas, Numpy, OpenCV, TensorFlow, MATLAB

Key Coursework

- **Electrical Engineering:** Digital Communication, Digital Systems, Digital Signal Processing, Data Analysis and Interpretation, Control Systems, Information Theory and Coding, Markov Chains, Microprocessors
- **Computer Science:** Automatic Speech Recognition, Advanced Machine Learning, Computer Graphics, Foundations of Intelligent and Learning Agents, Data Structures and Algorithms, Digital Image Processing, Network Security, Computer Networks, Operating Systems
- **Miscellaneous:** Calculus, Complex Analysis, Linear Algebra, Differential Equations, Biology, Chemistry, Economics, Psychology, Engineering Drawing, Environmental Studies

Positions of Responsibility

- **Editorial Board Member, Insight** Apr'21 - Present
  Insight is IIT Bombay's student media body with over 10,000+ readers
  - Surveyed the **effectiveness** of the Faculty Advisor program by taking inputs from both students and faculty and **suggested various reforms**
  - Initiated a series on **startups from research labs** at IIT Bombay as part of the **LinkedIn** team
  - Interviewed authorities and current international students for understanding the causes behind **poor international representation** at IITB and suggested **remedies** for the same
- **Department Academic Mentor** Apr'21 - Present
  - Selected as part of a 35-member team on the basis of **ethics, peer-reviews** and an **interview**
  - **Mentoring** 6 sophomores in academic and co-curricular activities
- **Teaching Assistant** July'21 - Present
  - TA for EE679, a graduate-level course on Speech Processing which covers speech production, analysis techniques and applications such as ASR, speech synthesis, etc.
  - In charge of assisting the instructor in conducting the evaluation of the course

Extra-Curricular Activities

- National-level quarter-finalist at Bournvita Quiz Contest; appeared on **National TV** for the same
- **Won 2nd prize** in **Android app development** competition organised by Web and Coding Club
- Successfully completed a 12-month **Lawn Tennis** course under National Sports Organisation and represented Hostel 4 in inter-hostel tournaments
- Awarded **Best Outgoing Student** of the year 2014-15 by Nirmala Convent High School