Sanidhya Mangal

Email: sanidhya.v.mangal@vanderbilt.edu
Portfolio: sanidhyamangal.github.io
Mobile: +1-6159558605
Github: github.com/sanidhyamangal

Summary: A graduate student (masters) at Vanderbilt University working on intersection of machine learning and computer vision processing, holding special interest in generative modeling, natural language processing and reinforcement learning is looking forward to work as machine learning engineer/software developer.

Education

- **Vanderbilt University** — Nashville, TN
  Master of Science - Computer Science; GPA: 3.9/4 — *Fall 2021 - Present*
  Courses: *Advance Machine Learning, Reinforcement Learning, Medical Image Processing, Adv. Algorithms, NLP, Geometric Deep Learning*
- **Medi-Caps University** — Indore, India
  Bachelor of Technology - Computer Science; GPA: 8.41/10 — *August 2016 - June 2020*
  Courses: *Artificial Intelligence, Calculus and Discrete Mathematics, Natural Language Processing, Pattern Recognition, Networking, DBMS, Data Structures and Algorithms, Computer Architecture, Teaching Assistantship: Machine Learning, Computer Graphics and Multimedia*

Skills Summary

- **Languages**: Python, C++, SQL, Bash
- **Frameworks**: Scikit, TensorFlow, Keras, Django, Flask, OpenAI Gym, PyTorch
- **Tools**: Kubernetes, Docker, GIT, MySQL, SQLite, Slurm, Singularity
- **Platforms**: Linux, Web,Accre, AWS, GCP, Microsoft Azure
- **Soft Skills**: Leadership, Event Management, Writing, Public Speaking, Time Management

Experience

- **Maizie Zhou Lab** — Vanderbilt University
  Research Associate — *May 2021 - Present*
  - Working on a research project DeepAquliaFilter to detect deletion structural variants(SVs) in human genome using computer vision.
  - In this project we are running a CNN based classifier to differentiate low confidence and positive SVs.
- **Engineerbabu** — Indore,India
  Machine Learning Engineer — *June 2020 - June 2021*
  - Carried out experimentation and research on diagnosing lung and colon cancer using histopathological images using Deep Convolution Neural Network(CNNs)
  - Lead a team to develop a machine learning models and pipeline to perform facial recognition using selfie and 3D video- By far reducing the inference time by 30% and bumping the overall accuracy to 87.83%. Later serving the models using Django based REST APIs.
  - Developed and deployed a web framework for performing Edge AI ops for object tracking and generation of analytical reports for the same using Tensorflow and DjangoRestFramework.
- **Greater Kailash Hospitals** — Indore, India
  Machine Learning Engineer(Intern) — *Jan 2020 - April 2020*
  - Developed CNN models to diagnose pneumonia and malaria using X-Ray and blood slides respectively with an area under curve (AUC) of 0.94.
  - Deployed the models as a web application using Flask which acted as a second opinion for the doctors reducing false positive by 23%.

Projects

- **PAC(Computer Vision, Semi-Supervised Learning, Domain Adaptation)**: Implementation of the paper Surprisingly Simple Semi-Supervised learning to test out effects of different representational learning (unsupervised rotation, supervised contrastive) on variations of semi-supervised and unsupervised domain adaptation over a OfficeHome Dataset.(December '21)
- **BrainEmbeds(Computer Vision, Generative Modelling, Medical Image Processing)**: The aim of the project was to perform inverse image retinotopy to generate images of stimuli based on the ROI masks in the brain using variational auto encoders.(December '21)
- **MountainCarSolver(Reinforcement Learning,Python)**: It is a collection of multiple deep reinforcement learning methods such as reinforce,DDPG and PPO to solve to OpenAI's mountain car environment. The entire implementation is environment agnostic hence could be used in any settings.(November '21)
- **GPPy(Statistical Machine Learning, Python, JAX)**: The project is a part of advance ml coursework where, I worked progressively to develop a library implementing gaussian process regressor and classificaion. To scale the computation variational sparse technique method is used.(November '21)

- **GAN(Computer Vision, Generative Modelling)**: (Work in Progress) A lightweight framework for tooling Generative Adversarial Networks, built on top of Tensorflow. Currently, in alpha mode continuous work is carried out to add more models and architectures
- **Pacman-Gridworld(Reinforcement Learning, Python)**: A simple value iteration and Q-learning based solution for solving Gridworld problem and playing Pacman game.(September '21)
- **DeepPathoLab(Computer Vision, Medical Image Processing)**: A collection several deep learning models using convolution architecture as a backbone, some notable models are skin cancer lesion detection, pneumonia and CovSars-2 virus screening using lung's computer tomography (CT) scan. (November '20)
- **GAN Farm (Generative Modelling, Computer Vision)**: This project consists of multiple DCGAN architectures and various training strategies such as WGAN, Vanilla GAN, LSGAN, etc on fashion mnist dataset. In addition to this same architecture was used for generation of anime faces and pokemons. (September '20)
- **Image Colorization using GANs(Computer Vision, Deep Learning, Unet)**: Images clicked using drones, provided by ISRO were stitched together using distributed public compute nodes, effectively bringing down processing time exponentially. Tech: PHP, C++, Java, Python (March '18)
- **Deep Vision Lab(Computer Vision, Deep Learning)**: This project combines various computer vision based algorithms such as image classification, neural style transfer, image noise reduction leveraging transfer learning paradigm on top of convolution neural networks (March '18)
- **Neural Machine Translation(Natural Language Processing, Deep Learning)**: A shallow copy of GNMT for developing a real time machine translator for translating English language to Hindi. We achieved a BLEU score of 14.2 after successful training of a model.
- **FacTrack(Natural Language Processing, Deep Learning, Image Processing Flask)**: A project built on top of TF-IDF algorithm to perform passive text summarization of Wikipedia articles, It is a web based interface ingesting the input keyword and then summarizing the parsed content from a Wikipedia webpage and later morphing the facts in template based images.
- **Text Generation(Natural Language Processing, Recurrent Neural Networks)**: A LSTM model trained over the Wikipedia articles to synthesis text data.
- **Attendance Analyzers(Machine Learning, Decision Trees)**: A co-companion for the professors to detect anomaly in attendance computation, it can reduce the human mediation by 20%.
- **Color Compression(Machine Learning, Unsupervised Learning)**: A K-Means based algorithm to compress the topography of images by transforming them into a 8 color model, reducing the image size by approximately 60 %

Publications

- **Paper:An ensemble deep learning framework to refine large deletions in linked-reads**: Got accepted in BIBM 2021. The objective of this paper is to detect deletion linked reads in human chromosome using computer vision. bioRxiv
- **Paper: CNN for diagnosing colon and lung cancer histopathological images**: The objective is to diagnose lung and colon cancer using histopathological images, all the trained models are accurate enough to perform early screening and diagnosis. arXiv Link (September 2020)
- **Paper: LSTM vs. GRU vs. Bidirectional RNN for script generation**: Used distinct sequence-to-sequence models to generate scripts for TV series, The model can generate dialogues for any number of episodes, available on arXiv Link (August 2019)
- **Paper: LSTM based Music Generation System**: The objective is to generate a suite of musical notes with the help of a single layer of LSTM. The model is capable of generating music in MIDI file format, arXiv Link (August 2019)

Volunteer Experience

- **Amazon**
  Alexa Student Influencer *June 2019 – June 2020*
  - First batch of India advocating voice-first technologies.
  - Leading the chapter in central India, impacting and training students and professionals on developing voice-first products using Alexa Skill Kit.
- **Microsoft**
  Student Partner *March 2017– June 2020*
  - Kickstarted a Microsoft community in my university, organizing workshops, hackathons and events.
  - Being ML enthusiast, trained hundreds of student developers on cloud technologies and Azure ML studio.
- **Google**
  Explore ML Facilitator *June 2019 – March 2020*
  - Group leader of machine learning ecosystem community in central India.
  - Trained hundreds of students on cutting-edge technology i.e., Machine Learning, Deep Learning and TensorFlow, harboring their interest in academia research.