Name - Mayank Kumar Singh

Homepage - https://mayank-git-hub.github.com
GitHub - https://github.com/mayank-git-hub

One Step at a time is the mantra with which I have been developing my Research profile. I have worked on Computer Vision as well as Speech and Music domain. Currently, I am interested in Singing Voice Conversion and its applications to reviving the singers who are today not with us.

Publications

- **Deep Audio Visual Source Separation** - *Guide Prof. Rajbabu Velmurugan and Naoya Takahashi, SONY Corp.*
  - Published the paper "Improving Voice Separation by Incorporating End-to-End Speech Recognition" in the ICASSP 2020 conference.
  - Modified ConvTasNet to incorporate ASR features for enhanced **Audio Source Separation** on the AVSpeech dataset
  - Achieved State of the Art Results on it, beating Google and Oxford's implementation, increasing the SI-SNR by 3.7dB.
- **Hierarchical Disentangled Representation Learning for Singing Voice Conversion** - *Guide Naoya Takahashi, SONY Global Corp.*
  - Published a paper in IJCNN 2021 Conference.
- **NENET: An Edge Learnable Network for Link Prediction in Scene Text** - *Guide Prof. Shubhasis Chaudhari, IIT Bombay*
  - Posted the paper "NENET: An Edge Learnable Network for Link Prediction in Scene Text" on arxiv.
  - Proposed a novel method of linking characters by creating a graph of characters and applying GNN.
  - Proposed a novel modification of GNN which outperforms other methods on link prediction task.
- **ISBI 2018: Diabetic Retinopathy, Segmentation of lesions**- *Guide Prof. Amit Sethi*
  - Aim - Segmentation and classification of the lesions in patients of Diabetic Retinopathy
  - Applied state-of-the-art algorithm **fusion-net** for segmentation and **Zoom-In Net** for classification.
  - Competition Paper- https://mayank.autonise.com/pdf/ISBI2018.pdf

Professional Experience

- **SONY Research Japan, Research & Development Engineer** : *Jan - Ongoing 2021*
  - Working on Improving Quality and Speed of Singing Voice Conversion.
- **SONY Research India, Research & Development Engineer** : *Oct - Dec 2020*
  - Worked in collaboration with Audio Technology Research Department in SONY Japan, Osaki to develop Singing Voice Conversion Models.
  - Published a paper in the **IJCNN 2021** named **Hierarchical disentangled representation learning for singing voice conversion**.
- **Director at Autonise AI** : *Sep 2018 - Dec 2020*
  - Founded a team of 8 with the vision to act as Technical Consultant in the field of Machine Learning.
  - Targeted the domains -
    Text Detection and Recognition, **Quant Algorithms**, **Facial Segmentation**
- **SONY Japan, Research Internship** : *May - July 2019*
  - Worked with Audio Technology Research Department in SONY Japan, Osaki to improve Deep Audio Visual Source Separation
  - Used **WaveNet** like architecture, Temporal Convolution for audio speech separation and used visual features for improving separation **SISNR** and surpassed current SOTA implementations.
- **HDFC Life, Research Internship** : *May - July 2018*
  - Automated customer interaction by automating questions asked using **Reinforcement Learning**.
  - **Feature Engineering** and **Clustered Customer data** for extracting useful statistics and analysis of the algorithm

Projects

Research Projects

- **Segmentation of Medical Image** - *Guide Prof. Amit Sethi*

- Applied **NN**, **SSNMF**, **NMF**, **SVM** algorithms to do pixel-level segmentation on Hyperspectral Images
- Implemented the initial steps for detecting cancer by segmenting epithelium, stromal and goblet cells.

- **Whole Slide Image Stitching using DC motor video** - *Guide Prof. Amit Sethi*

Other Projects

- **Kaggle Competition: *iMaterialist Challenge (Furniture) at FGVC5***
  - An orthodox classification competition with **highly skewed class size** and high intra class and low inter class variation.
  - Trained **ResNet-152**, **NASNet** model using extensive class specific data augmentation.
  - Got a **rank of 30** under the team name 'Artificial incoherence'
- **Text Detection and Recognition on Documents**
  - Implemented Pixel-Link for Text Detection on https://github.com/mayank-git-hub/Text-Recognition
  - Achieved an F1-score of 74% which is **6% more than Google's on our custom dataset** consisting of passports, aadhar cards, driving license cards and other docs which we annotated using our annotation tool built using javascript.
- **Web Development**
  - Designed and deployed website https://www.primeacademypune.com

Education

Institution	Specialisation	Year	GPA/Percentage
Indian Institute of Technology, Bombay	Electrical Engineering, B.Tech	2020	8.84
Air Force School, VN, Pune (HSC)	Computer Science	2016	93.8%
Air Force School, VN, Pune (SSC)	None	2014	10

Courses Undertaken

Computer Vision	Probability and Random Process	Data Analysis and Interpretation
Network Theory	Data Structures & Algorithms	Linear Algebra
Computer Networks	Signals and Systems	Micro-Processors

Technical skills

- **Programming Languages:**
  Proficient in: C, C++, Python, JAVA, Javascript, Angular
  Specific libraries for Machine Learning - Tensorflow, Pytorch
  Also basic ability with: MATLAB, Shell Script, Arduino, NgSpice, VHDL, AutoCad, Solidworks.
- **Deep learning models:**
  Classification - Res-Nets, Inception-Net, Alex-Net, Capsule-Net, Zoom-In-Net, NASNet
  Segmentation - U-net(Variants - ResNet-UNet, Fusion-Net)
  Feature Extraction - Siamese doublet/triplet networks, AutoEncoders, Variational AutoEncoders.
  Audio Separation - WaveNet, ConvTasNet, TasNet
- **Web & Android Development:**
  Server Side - Django, NGINX, Flask, AWS, Kotlin, Django-Channels
  Client Side - Ionic, Android-Studio(JAVA & Kotlin), HTML, JS, D3JS, Three JS, ES6, React, Angular

Scholastic Achievements

- Secured All India Rank of **108 in JEE Advance 2016 in General Category** among 198,228 candidates
- Secured All India Rank of **1484 in JEE Mains 2016 in General Category** among 1,207,058 candidates
- Recipient of prestigious **KVPY fellowship with All India Rank of 363 (/60,000)**