Vedant JOSHI,
Vision & Imaging Engineer-I

🏛 Tonbo Imaging Pvt. Ltd, Bangalore
in linkedin.com/in/vedant-joshi-b822bb169
🔗 https://vedrocks15.github.io/
📱 +91-8879891562   @ vedrocks15@gmail.com

"**How to make machines think ?**", is the question that drives my research every morning. Based on this question, my fundamental goal is to understand the inner workings of deep models & make a sincere contribution towards building dynamic as well as modular mechanisms that match the level of human intelligence by efficiently representing a multitude of input modalities into a single structured latent space & are contextually modifiable based on the situation at hand.

🎓 RESEARCH INTERESTS

> Representational learning
> Self-supervised learning
> Building multi-modal latent spaces
> Ensemble learning
> Large scale data retrieval
> Model quantization & pruning
> Model explainability & interpretability

➕ HIGHLIGHTS

> Gold medalist for highest GPA in graduating class of 2017.
> Received Best Paper award at 8th ICSCC 2021 by IEEE.
> SIH-2020 Finalist under Airport Authority of India.
> Oracle Certified Java Associate Programmer
> Kaggle open source contributor for implementing latest research papers & frameworks.

💼 WORK EXPERIENCE

Vision & Imaging Engineer-I   FEBRUARY 2023 - CURRENT
🔗 Tonbo Imaging Pvt. Ltd.

> **IR object detection** : Worked on improving the performance of YOLO FIRI on IR images, from **0.5 mAP to 0.7 mAP** on the FLIR-v2 dataset. The focus was to improve the propagation of weak features present in long wavelength IR images.
  > Added higher resolution feature maps in the **detection head** to improve the model's detection rate for tiny objects.
  > Along with wider **CSP modules**, additional range of kernels were used in the **SK attention** module to model dynamic receptive fields for multi-scale detection.
> **RGB to IR mapping** : Developed an **attention based UNET model**, that learns the mapping of images between RGB and IR domain to generate psuedo-thermal data for regular driving videos.
  > Implemented **subpixel-convolution** along with **ICNR initialization** in the decoder phase to generate checkered free, high resolution IR images. An additional combination of **perceptual loss** as well as L1 loss was used to improve the details generated in the final image.
> **Self-supervised Depth Estimation** : Working on an experiment to develop a learning framework, to utilize **PoseNet** to predict the ego motion of a camera in a monocular IR video & use adjacent frames to retrieve the depth information.

Data Scientist - I   JULY 2022 - DECEMBER 2022
🔗 Vedantu Innovations Pvt. Ltd.

> **Voice Nudges** : Developed a concatenative synthesis engine using **spectral smoothing methods** that stitch the name of an individual student to create natural sounding audio nudges in the voice of the class teacher.
  > The nudges ran the gamut of scenarios from motivational to concerning based on the user behaviour in an AI live session.
  > Experimental results indicated a **10% improvement** in student engagement & concentration after receiving personalised nudges during the sessions.
> **User Cohort Generation** : Implemented the tabular self-supervised framework **SCARF** & TabTransformers from scratch in order to build latent representations based on the click-stream & user interaction features.
  > The latent representations were repurposed to build user behaviour clusters (cohorts) using UMAP & K-Means.
  > The exercise of cohort generation allowed us to isolate users & identify key parts of our product that a particular student was not utilising as well as the actions that could be taken to improve his/her engagement.
> **Profanity Filtering** : Developed a novel profanity detection module via **contrastive learning** to train a LSTM module that generated embedding vectors which were invariant to the spectrum of ways in which an abuse can be posted on the chat platform.
  > The pipeline allowed us to control the size of the profane vocabulary to a limited number of key abuses & also supported easy addition of new abuses without retraining the model.
  > The efficiency of the entire system was further boosted by maintaining a modular cleaning pipeline that handled removal of domain specific tokens & close matches to reduce the number of false positives predicted.
> **Year On Year Retention** : Developed user engagement metrics for computing platform stickiness & month on month **unsupervised clusters** to improve year end prediction scores of shallow learning models.

ML Researcher

 Vedantu Innovations Pvt. Ltd.

> **Cluster Cleaner** : Developed a NLP based cleaning pipeline using a parametric, ensemble mechanism of **n-gram Sim-Hashing & Levenstein distances**, whose goal was to robustly remove irrelevant strings in doubt clusters due to OCR misreads. The improved homogeneity of doubt clusters improved the quality of matches returned by the search engine.
> **SSL Search Engine** : Developed the diagram matching module of the doubt search pipeline using image based self supervised frameworks to create a compressed & structured embedding space that places similar looking diagrams close to each other. The module was made scalable by using approximate K-NN search on the embedding space for the large scale doubts repository.
>> Developed a domain specific augmentation pipeline that modelled noises generated from the camera of the student & these noises forced the model to learn invariant representations that improved the quality of returned matches.
>> Implemented SOTA frameworks such as **BYOL & SimCLR** with an innovative cross temperature scaling mechanism for InfoNCE loss. The new parameter improved the model's ability to focus on hard negative samples during pretraining.
> **Doubts Clustering Module** : Repurposed the search engine to develop an unsupervised clustering via **UMAP & HDBSCAN** to reduce redundancy from **9.2 lakh** data points to **2.5 lakh** clusters & help academic experts answer more unique doubts.

Deep Learning Intern

 Vedantu Innovations Pvt. Ltd.

> **Image Denoisers** : Developed a class of augmentation functions that mimicked the noise & skews present in our doubts image repository. These functions were utilised for random corruptions on clean images & used to train **UNET segmentation & variational autoencoder models** to generate cleaner binarised images. The cleaner images occupied less space in the repository & improved the text extraction accuracy of Tesseract OCR.
> **Image 2 Latex Markup** : Porting legacy Tensorflow code to 2.0 for the paper **Image to Markup Generation with Coarse to Fine Attention** to generate latex markup for images containing mathematical texts. This allowed us to create a smaller, latex only doubts search space.
> **Subjects Classifier** : For doubts categorisation, developed a NLP pipeline to classify OCR read text into corresponding subject using **BERT vectors & bi-directional attention based LSTMs**. Further a text pre-processing module was added to normalise OCR reads & map the high frequency tokens to a common keyword based on the output generated from LDA for topic modelling.
> **Siamese Search Pipeline** : Developed an image search module in a lower dimension space using supervised training.
>> Trained a Scaled YOLOv4 to detect & extract diagrams from doubt images to reduce the false positive matches casued due to background noises. The model was able to achieve a **mAP score of 90%**.
>> MobileNetV1 was trained on the diagram dataset generated by YOLO via a hardness aware supervised triplet loss mechanism. The generated vectors were rapidly searched through HNSWlib framework.

Research Intern

 TCS Rapid Labs

> **Vision based lip reading** : Implemented a word only version of **Deep Mind's lipNet model** for the large scale Oxford BBC LRW dataset. The encoder-decoder architecture learnt to make an optimal mapping from frame to character via greedy CTC decoding.
>> Developed functions to achieve speaking speed normalisation by frame count streching & pre-processed each video frame to lip region only via the Dlib package.
>> Added the **Bahdanau soft attention mechanism** in between the encoder-decoder architecture to allow the model to capture complex interactions between various frames & visualise its focus while making a prediction at each time step.

Quantum Computing Intern

 UST Global

> **Artificial face mask synthesis** : Implemented **classical GANs** to generate synthetic images in order to increase the size of the dataset used to train face mask detectors in public during Covid-19 pandemic.
>> Curation of multiple face-mask datasets along with normal face capturing datasets in order to train generator networks to capture the underlying face structure distribution.
> **Quantum Discriminators** : Implemented a POS on using **random quantum circuits** with no error correction as initial filters in order to build better discriminator networks. This made the task of achieving NASH equilibrium for the GAN system even more challenging thus improving the quality of images generated by the generator.

📖 EDUCATION

Bachelor of Technology Honours Computer Science   AUGUST 2017 TO APRIL 2021
🔗 IIIT Kottayam

> Gold medal for receiving the highest grade in the batch of 2017. Acquired CGPA : $3.94/4.0(9.82/10.0)$
> 3 out of 8 semesters with perfect grade of $4.0(10.0)$
> Head of the IoT group in the Technical Club.
> Received grant from Orzico, USA for a project on Autonomous Fruit picker.

☰ SKILLS

Programming languages	Python, Java, C, C++
Deep Learning Frameworks	TensorFlow, Keras, PyTorch, NLTK, Scikit-learn, Nvidia-TensorRT
Database	Microsoft SQL Server
Development Tools	Visual Studio Code, Google Colab
Operating Systems	Windows, Linux, Mac OS X
Embedded Boards	RaspberryPi, Arduino (Nano, UNO, Mega), Nvidia Jetson Nano, NodeMCU
Domain Knowledge	AI, Micro Ariel Vehicles, IoT

💻 ACADEMIC PROJETS

Coco Layers : Edge Based Inference

Orzico Funded Project   December 2019 - April 2021

Developed an edge based real time object detection system which served as a pluggable module for any robotic implementation to pluck coconuts directly from trees.

> Curated a novel, small scale, manually annotated coconut images dataset using drones at multiple locations in Kerala. The dataset successfully captured a spectrum of **natural lighting conditions, occlusions & coconuts of various shapes & sizes** that occur in the wild.
> Completed a comparative study on model quantization & weights pruning for **SSD MoileNetV2, YOLOv3 & tensorRT optimised tiny YOLOv4** for real time inference on RaspberryPi 3B+ & Nvidia Jetson Nano.
> Developed a model deployment pipeline for quantized models with efficient frame buffer handling mechanisms & thread level multiprocessing to achieve a **frame rate of 22 FPS** in real scenarios.

`Python` `TensorFlow Lite` `Nvidia TensorRT` `OpenCV` `Linux`

AAI IoT Dashboard

🔗 AK13_aviato   SIH 2020 Finals   May 2020 - July 2020

Developed a real time, React based dashboard to display **live IoT services** data for the terminal manager & handled artificially generated data using MongoDB.

> Implemented **real time pipelines** to communicate with IoT devices using multiple protocols & updated the database as well as the dashboard in a synchronized manner.
> Implemented **Facebook's prophet model** to adapt to the trend, periodicity & seasonality of resource consumption data at the terminal of an airport & predict the service demand by different cohorts of flyers.

`Python` `SciKit` `React.js` `Flask` `MongoDB` `TimeSeries Forecasting`

Lemon Picking Drone

🔗 Project Report   IIIT-Kottayam   August 2019 - December 2019

Created an **autonomous drone with a 3D-printed peripheral arm** that assisted in picking lemons suspended in mid-air along with a live video feed streaming service.

> Developed a quad-copter from scratch & an inter-process communication system between pixHawk flight controller system & Raspberry Pi to complete **GPS guided autonomous flights**.
> Implemented a light weight **color segmentation** script to perform lemon tracking by edge devices. A gstreaming pipeline was written to broadcast the live detection to a user interface.
> The 3D arm control was handled by an embedded C subroutine that controlled the **micro controller Arduino UNO** along with a combination of stepper motors.

`Python` `Embedded C` `PixHawk` `MissionPlanner` `Parallel Computing`

Internet of Ledgers

[⧉] Project Report   IIIT-Kottayam   August 2019 to November 2019

Developed a secure IoT home automation communication system where the data generated in the network is secured & tracked by a Blockchain platform, **Hyperledger** to achieve transparency between user & service providers. Also use of smart contracts was implemented to perform M2M communications between the IoT nodes via MQTT protocol.

Embedded C   Hyperledger   Shell Script

[⌖] PUBLICATIONS

1. **Vedant Sandeep Joshi**, Sivanagaraja Tatinati, Yubo Wang, *Looking For A Match : Self-supervised Clustering For Automatic Doubt Matching In e-learning Platforms* preprint : arXiv:2208.09600, https://doi.org/10.48550/arXiv.2208.09600

2. **Vedant Sandeep Joshi**, Sivanagaraja Tatinati, Yubo Wang, *YZR-net : Self-supervised Hidden representations Invariant to Transformations for profanity detection* preprint : arXiv:2211.15532, https://doi.org/10.48550/arXiv.2211.15532

3. **Vedant Sandeep Joshi**, Ebin Deni Raj, *FYEO : A Character Level Model For Lip Reading*, IEEE/8th International conference on Smart Computing and Communications, Accepted Sep 6 2021, https://doi.org/10.1109/ICSCC51209.2021.9528104

4. **Vedant Sandeep Joshi**, Ebin Deni Raj, Jeena Thomas, *Quantized Coconut Detection Models with Edge Devices*, Journal of Interconnection Networks [Scopus Indexed Journal], Accepted November 2 2021, https://doi.org/10.1142/S0219265921440102

[❝] REFERENCES

Dr. Sivanagaraja Tatinati
Manager - Data Science, VEDANTU INNOVATIONS PVT. LTD
[@]   tatinati.sivanagaraja@vedantu.com
[☎]   On Request

Dr. Ebin Deni Raj
Associate Dean (Academic Affairs), IIIT KOTTAYAM
[@]   ebindeniraj@iiitkottayam.ac.in
[☎]   +91 (0) 482-2202154

Dr. J. V. Bibal Benifa
Assistant Professor, IIIT KOTTAYAM
[@]   benifa@iiitkottayam.ac.in
[☎]   +91 (0) 482-2202163