Venkata Sai Ritwik Kotra

+1 470-430-4362 | vkotra3@gatech.edu | linkedin.com/in/ritwik | github.com/sairitwik27 | scholar.google.com/ritwik

EDUCATION

Georgia Institute Of Technology GPA: 4.0/4.0
MS in Electrical and Computer Engineering *Aug. 2022 – May 2024*
Relevant Courses: *Statistical Machine Learning, Online Machine Learning, Advanced Programming Techniques, Random Processes, Medical Image Processing*

National Institute Of Technology,Karnataka,Surathkal Cummulative GPA: 9.21/10.00
BTech in Electonics and Communication Engineering *Aug. 2018 – May 2022*
Relevant Courses: *Speech and Audio Processing, Image and Video Processing, Intro To Artificial Intelligence, Graph Theory, Statistical Analysis,Advanced DSP*

TECHNICAL SKILLS

Languages: Python, C/C++, Bash, SQL
Frameworks: PyTorch, Tensorflow, Kaldi, ESPNet, sklearn
Developer Tools: Git, VS Code, Visual Studio
Libraries: pandas, NumPy, scipy, Matplotlib, HuggingFace, NLTK

RESEARCH EXPERIENCE

Robot Code Generation using Large Language Models Jan 2023-Present
AVA Laboratory, Georgia Tech *Supervisor: Prof. Larry Heck*

- Using code to generate robot policies from Large Language Models. Also working on code as a symbolic reasoning unit to aid Language Models. Improving Microsoft's Airsim capable ChatGPT to be more robust to changes in prompts.(Work in Progress)

Accent Robust approaches for End-to-End ASR June 2021-July 2022
LEAP Laboratory,Indian Institute Of Science,Bangalore *Supervisor: Dr.Sriram Ganapathy*

- Trained Transformer,Conformer and RNN-Transducer on the datasets from scratch. Achieved WER of 12.5% on Commonvoice dataset and 18% on NISP dataset with the Conformer.
- Used LSTM with K-Means clustering with baseline ASR as Domain Adversarial Training(DAT) Network. WER fell to 15% on NISP and 10% on Commonvoice using the DAT+ technique. [Code]
- Explored wav2vec2.0,CPC, APC and PASE for learning accent aware representations. Designed APC-PASE, a modified APC framework for accent adaptation. Found significant improvements on adding embeddings from an accent classifier trained on wav2vec2.0 as cross attention input to conformer encoder

Image and Audio Processing for Cancer and COVID-19 detection December 2019 -May 2022
SMILE Laboratory, NITK,Surathkal *Supervisor: Dr.Deepu Vijayasenan*

- Designed an unsupervised algorithm with morphology,thresholding and clustering methods for automatic generation of labels for image segmentation. UNet used to segment with manual and algorithm annotations as ground truths. Automatic Labels outperformed manual labels by 1.25%. **Published in EMBC 2020 conference**. (Computer Vision, TensorFlow,sklearn)[EMBC Paper]
- COVID-19 Detection using Speech from YouTube Videos(May 2020) Built an SVM classifier with utterance and phoneme level supervectors as features on audio data scraped from YouTube. Documented in arXiv.(Speech For Healthcare, Kaldi,sklearn)[Paper]
- Used spectral features as input to SVM classifier. Achieved AUC of 0.752. **Published in INTERSPEECH 2021**. **(UG Thesis)** Compared CPC with popular FBANK based techniques for detecting Covid-19 and Cold. Got SOTA results(AUC of 0.890 and 0.857 for Covid-19 and Cold) while jointly predicting 2 kinds of diseases with a single representation learning framework. (Speech for Healthcare,Representation Learning,PyTorch,fairseq)[Interspeech Paper][[Interspeech Code]][DiCOVA 2 Report][UG Thesis]

PROJECTS

Analysis of Speech Transformer SSL Models for Speaker ID | *PyTorch,Huggingface* December 2022

- Did a layer wise analysis of performance on finetuning the hubert-base and wav2vec2-base models on VoxCeleb1 dataset.
- Found not passing attention mask and padding with silence for 5s audio results in 10% boost in accuracy over released Huggingface pre-trained model. [Report][Code]

COVID-19 Infected Region Segmentation from Lung-CT Images | *PyTorch,Computer Vision* December 2021

- Built novel Residual U-Net and SegNet with Attention Modules for segmentation of Ground Glass Opacities from Healthy Lung region. Residual U-Net achieved the best dice score of 0.779.[Report]

Emotion Detection from EEG Signals | *SKLearn,Time Series* December 2021

- Used SEED-IV dataset to do a 4 class emotion classification using Granger causality, VAR as well as 1d Convolutional Neural Networks. Achieved an F1-Score pf 0.70. [Code]

Publications

Ritwik, K. V. S., Kalluri, S. B., Vijayasenan, D. (2021). COVID-19 Detection from Spectral Features on the DiCOVA Dataset. In Interspeech (pp. 936-940).

Lakshmi, S., Ritwik, K. V. S., Vijayasenan, D., Sreeram, S., Suresh, P. K. (2020, July). Deep learning model based Ki-67 index estimation with automatically labelled data. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC) (pp. 1412-1415). IEEE.

Ritwik, K. V. S., Kalluri, S. B., Vijayasenan, D. (2020). COVID-19 patient detection from telephone quality speech data. arXiv preprint arXiv:2011.04299.