Shreekantha Nadig

shreekantha.nadig@iiitb.ac.in | +919738410350
https://github.com/sknadig | https://sknadig.dev | https://www.linkedin.com/in/sknadig/

EDUCATION

MS BY RESEARCH - DATA SCIENCE | IIIT - BANGALORE Bengaluru, IN | Jan 2017 - Dec 2019
GPA: 3.74/4

B.E. - TELECOMMUNICATION ENGINEERING | JNNCE (VTU) Shivamogga, IN | Aug 2011 - Jul 2015
GPA: 75/100

MS THESIS

MULTI-TASK LEARNING IN END-TO-END ATTENTION-BASED AUTOMATIC SPEECH RECOGNITION | ESPnet, kaldi, PyTorch, TensorFlow Bengaluru, IN | Jan 2017 – Dec 2019

- Developing state-of-the-art systems for end-to-end ASR using Joint CTC and Attention-based models
- Study how pure data-driven models can be blended with knowledge-based models for reducing model complexity, faster training/inference and extracting deeper insights into speech recognition
- Use of ASR toolkits like Kaldi, ESPnet with PyTorch and TensorFlow to build end-to-end ASR models
- Study of various Attention mechanisms and how they can be modelled efficiently for interpretability, explainability of end-to-end models.
- Multi-target hybric CTC/Attention network for joint phoneme-grapheme recognition
- Jointly learning to align and transcribe using attention-based alignment and uncertainty-to-weigh losses
- Keyword-spotting using attention-based end-to-end ASR models

OPEN SOURCE CONTRIBUTIONS

ESPNET | Recipes and bug fixes (ESPnet1 and ESPnet2)

EXPERIENCE

SPEECH RECOGNITION ENGINEER 3 | Dialpad Bengaluru, IN (Remote) | Dec 2019 – Present

- Lead the R&D on end-to-end ASR for conversational telephony and videoconferencing speech
- Building and benchmarking various end-to-end ASR architectures like RNN-CTC, RNN Attention-based Encoder-Decoder (AED), RNN-Transducer, Transformer-Transformer and Conformer-Transformer with hybrid ASR models and external ASR services.
- Automating the data preparation pipeline for ESPnet and Kaldi
- Developing pronunciation-assisted sub-word models using fast-align, GIZA++ and Pynini
- Integrating sub-word and word-level RNNLMs with sub-word level end-to-end ASR models
- Developing interfaces for shallow fusion of multi-level (sub-word and word-level) RNNLMs
- Demonstrated that the end-to-end ASR models can outperform the best hybrid models on various English accents
- Post training quantization of end-to-end ASR models and comparing RTF, WER with FP32 models
- Changes to ESPnet to make end-to-end models pure PyTorch with the use of torchaudio features
- Developing performance monitoring techniques for end-to-end ASR models based on RNN-AED and CTC confidence scores, and their efficacy in semi-supervised and self-supervised learning techniques.
- Developing and integrating Kaldi endpoint detection with in-house hybrid models to achieve 4% relative WERR
- Developing noise robust hybrid ASR models by using a front-end VAD
- Improving the performance of hybrid ASR models on different accents using hypotheses from end-to-end ASR models
- Developing a web-app to visualize decoding hypothesis CTM files using wavesurfer-js

MACHINE LEARNING INTERN - ASR | Observe.AI Bengaluru, IN | May 2019 – Aug 2019

- Developing a feature extraction pipeline using tf.signal and tf.data that is compatible with Kaldi extracted features
- Building End-to-end ASR systems with Joint CTC and Attention on large conversational speech datasets in ESPnet with features extracted from TensorFlow
- Implementing different keyword-spotting papers - Deep-KWS, CTC KWS
- Converting an ESPnet PyTorch model to TensorFlow
- Deploying the KWS model using TensorFlow serving with an RTF of 0.05 on GPU

RESEARCH SCHOLAR | IIIT-Bangalore Bengaluru, IN | Jan 2017 – Dec 2019

- Machine Intelligence and Robotics Center | ESPnet, PyTorch
  - Intelligent multi-lingual speech-technology based e-gov applications
  - Developed multilingual end-to-end ASR models for English-Tamil and English-French
  - Development of an unrestricted vocabulary keyword-spotting system using a phone-level, attention-based Encoder-Decoder model
  - Semi-supervised and self-supervised learning pipelines for end-to-end ASR models
- Intel AI Academy Student Ambassador | Intel NCS, Intel AI DevCloud, ASR, TensorFlow
  - Small-footpring keyword spotting on-the-edge using Intel Neural Compute Stick 2
- Virtual lab for NPTEL | Django, Embedded C, HTML, CSS, Javascript
  - A remote hardware lab to study the signal response of stepper motors.
  - Developed web-server using Python Django and embedded system to control the motors using Embedded C.
- Graduate Teaching Assistant Bengaluru, IN | Jan 2018 – Dec 2019
  - Deep Learning for Automatic Speech Recognition | Kaldi, PyTorch, TensorFlow
    - State-of-the-art in end-to-end ASR technologies
    - 1-pass and 2-pass decoding using joint CTC+Attention and LM
    - Endpointing for end-to-end ASR
  - Automatic Speech Recognition | Kaldi, scikit-learn, cloud computing
    - Phonetics and phonology, signal processing for ASR
    - GMM-HMM, DNN-HMM tutorials in Kaldi
  - Introduction to Robotics | Python, Robot operating system, Gazebosim
  - Digital and the Everyday: from codes to cloud | MOOC (NPTEL)

SVT ENGINEER | Sonus Networks Bengaluru, IN | Aug 2015 – Jan 2017

- Worked as a part of Sustaining SVT on Real-Time communication products Sonus Insight (EMS) and SBC
- Developed automated test frameworks in Python, Perl, Linux and Java
- Worked with CentOS, Red Hat Enterprise Linux and Solaris to develop and test the products
- Developed tools which reduced team effort from many hours to a couple of minutes
- Collaborated with overseas teams in testing/fixing the product for potential security breaches

AWARDS

THIRD PRIZE MUCS 2021: MUltilingual and Code-Switching ASR Challenges for Low Resource Indian Languages Bengaluru, IN | Aug 2021

`https://github.com/dialpad/mucs_2021_dialpad`

Team contributions to multilingual and low-resource ASR for Indian Languages. Benchmarking and open-sourcing various end-to-end methods and studying effects of channel distortions on language identification.

BEST STUDENT PAPER AWARD – HONORABLE MENTION SPCOM 2020 Bengaluru, IN | Jul 2020

Jointly learning to align and transcribe using attention-based alignment and uncertainty-to-weigh losses

THIRD PRIZE JUincubator Hackathon powered by GMASA Bengaluru, IN | Jul 2017

Developed a web app "iCarto" - a serious game for urban planning

SKILLS

Programming	Python, Bash, JavaScript, C++
Libraries	Keras, PyTorch, TensorFlow, scikit-learn, Django, OpenCV
Toolkits	Kaldi, ESPnet, K2, NeMo, WeNet

INVITED TALKS

IIIT-B SAMVAAD TALK
Bengaluru, IN | Dec 2020

- Multi-task learning in end-to-end attention-based automatic speech recognition (MS Thesis)
- Open challenges in multi-task learning for ASR

IIIT-B GUEST LECTURE SERIES Deep learning for ASR
Bengaluru, IN | Sep 2020 - Dec 2020

- Discussions on RNN-CTC, RNN-AED, RNN-T and Transformer models for ASR
- Discussions on model quantization and weight sparsity in RNN-T models for low computational resource and latency constraints
- Unpacking and analyzing the Pixel Recorder app to showcase how tflite models are packed with custom TensorFlow ops

ARTIFICIAL INTELLIGENCE : A WAY FORWARD
Bengaluru, IN | Sep 2019

- Faculty development programme at Dayananda Sagar College of Arts, Science and Commerce, Bangalore
- Discussions on the use of AI in speech and language technology

TCS THINK LABS
Bengaluru, IN | Feb 2019

- Motivation and introduction to end-to-end ASR
- Discussion on the topics of RNN, CTC, Attention and LM fusion

IIIT-B AI READING GROUP
Bengaluru, IN | Nov 2018

- Discussions on various attention models in end-to-end ASR
- Semi-supervised learning with end-to-end ASR models

BMSCE AI WORKSHOP
Bengaluru, IN | Sep 2018

- Artifical Intelligence and Deep Neural Networks Workshop for undergraduate students at BMS College of Engineering, Bangalore
- Code examples and tutorials in TensorFlow Keras

PUBLICATIONS

- Vasundhara Gautam, Wang Yau Li, Zafarullah Mahmood, Frederic Mailhot, Shreekantha Nadig , Riqiang Wang and Nathan Zhang, "Avengers, Ensemble! Benefits of ensembling in grapheme-to-phoneme prediction" In Proceedings of the 18th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology, 2021.
- Shreekantha Nadig , V. Ramasubramanian, Sachit Rao, "Multi-target hybrid CTC-Attentional Decoder for joint phoneme-grapheme recognition," 2020 International Conference on Signal Processing and Communications (SPCOM), Bangalore, India, 2020
- Shreekantha Nadig , Sumit Chakraborty, Anuj Shah, Chaitanay Sharma, V. Ramasubramanian, Sachit Rao, "Jointly learning to align and transcribe using attention-based alignment and uncertainty-to-weigh losses," 2020 International Conferenceon Signal Processing and Communications (SPCOM), Bangalore, India, 2020 (Best Student Paper Award – Honorable Mention)
- Abhijith Madan, Ayush Khopkar, Shreekantha Nadig , K. M. Srinivasa Raghavan, V. Ramasubramanian, "Semi-supervised learning for acoustic model retraining: Handling speech data with noisy transcript," 2020 International Conference on Signal Processing and Communications (SPCOM), Bangalore, India, 2020