Shreekantha Nadig

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

EXPERIENCE

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

- Architect and built a next-gen Speech Recognition product end-to-end from R&D to production that is toolkit agnostic and performing better than HMM hybrid ASR models
- Lead the R&D on streaming end-to-end ASR for conversational, telephony, and videoconferencing speech under low latency and multi accent scenarios
- Built and benchmarked various end-to-end ASR architectures with CTC, Attention-based Encoder-Decoder (AED), Transducer, Transformer and Conformer models with hybrid ASR models and external ASR services
- Developed interfaces for shallow fusion of multi-level (sub-word and word) RNNLMs and n-gram LMs
- Developed methods to bias the models towards a list of keywords, resulting in an absolute WERR of 7%
- Automated the data preparation pipeline for training ASR models, reducing the turnaround time for experiments and increasing productivity of the team
- Developed pronunciation-assisted sub-word models using fast-align, GIZA++ and Pynini, resulting in an absolute WERR of 3% compared to BPE sub-words
- Post training quantization of ASR models to achieve 50% faster RTF and 75% smaller models on disk
- Implemented the ASR inference in ONNX runtime, reducing the latency by 3x
- Developed 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
- Developed better endpoint detection for hybrid models and achieved 4% relative WERR
- Developed a web-app for internal users to query production calls and visualize hypotheses using wavesurfer-js

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

- Developed a feature extraction pipeline using tf.signal and tf.data
- Implemented different keyword-spotting (KWS) papers - Deep-KWS, CTC KWS
- Developed methods to convert a custom PyTorch model to TensorFlow
- Deployed 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
  - Developed multilingual end-to-end ASR models for English-Tamil and English-French
  - Developed an unrestricted vocabulary keyword-spotting system
  - Semi-supervised and self-supervised learning pipelines for end-to-end ASR models
- Graduate Teaching Assistant — Bengaluru, IN | Jan 2018 – Dec 2019

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

SKILLS

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

OPEN SOURCE CONTRIBUTIONS

ESPNET | Recipes and bug fixes (ESPnet1 and ESPnet2)

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

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. **(Best overall submission)**
- **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

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