Dang Le Dang Khoa

(+65) 8398 3951 dangledangkhoa@gmail.com github.com/dangkhoadl linkedin.com/in/dangledangkhoa

TECHNICAL SKILLS

- Languages: Python (Pandas, Matplotlib, Numpy, Scikit-learn, Pytorch, Keras, PySpark), C/C++ (STL, Libtorch, OpenMP), Bash/Linux (shell scripting, awk, Dockerfile)
- Tools: kaldi-asr, ESPnet-e2e, jupyter notebook, git, huggingface

EXPERIENCE (SELECTED)

Senior Research Engineer

Institute for Infocomm Research (I2R) - A*STAR

February 2020 - Present

Speech-to-text (Automatic Speech Recognition)

- Trained and maintained Bahasa Malay core Speech-to-Text engine and language model, improved Word-Error-Rate by 10% from the baseline model
- Integrated and finetuned Singapore English Speech-to-Text engine into domain-specific services, including Legal, Healthcare, and Education. Intelligence Court Transcription System featured on Straitimes.
- · Transformed and augmented audio and text data for model training.

Natural Language Processing - Audio Analytics and Classifications

- Trained and integrated ECAPA-TDNN deep learning model for speaker diarization service, meeting minute note-taking application.
- Trained *U-net* audio enhancement model for Satellite-based VHF communication; performance improved by 3% compared to RNN baseline model.
- Modeled sentence embedding inference and logic backend for Vietnamese-voiced FAQ chatbot.
- Contributed to other team-scaled *Natural Language Processing* and *Classification* tasks: Sentiment Analysis, Embedded Sound Classification, and Text Processing.

Researcher Satellite Research Centre August 2016 - August 2018

- Implemented a novel Star Tracking Algorithm onto a Programmable System-on-a-chip for mini satellites.
- Optimized the pattern recognition algorithm runtime, improved the accuracy by 5% on average.
- Designed a tree-based pattern searching algorithm, improved runtime processing by 25%.

Honors & Awards (selected)

Al Singapore - Trusted Media Challenge 2021

- Challenged for the detection of audio-visual fake media, a competition hosted by AI Singapore, sponsored by Singapore Press Holdings Ltd. with prize monies of up to SGD 700,000.
- · Lead and contributed chief engineering work for aasrali team to compete with other 474 teams.
- Ranked 6th on the final stage leaderboard Top 1% teams Final Round Leaderboard.

ConferencingSpeech 2021 Challenge

- Challenged for Far-field Multi-Channel Speech Enhancement for Video Conferencing, hosted by Interspeech and Tencent Ethereal Audio Lab.
- · Ranked in top 10 with I2R-ALI team.

Google Code Jam 2018

- Competed individually in the problem solving contest organized annually by Google with over 27,000 competitors.
- Advanced to round 2 Top 10% candidates.

EDUCATION

Master of Engineering

Nanyang Technological University Singapore

August 2015 - August 2018

- School of Electrical and Electronics Engineering Research.
- Thesis title: Implementation of An Autonomous Star Recognition Algorithm using Hardware-Software Co-Processing Approach.
- Source code: git.io/vpucY
- Publication: https://doi.org/10.32657/10220/48371

Bachelor of Engineering

Vietnam National University

August 2010 - April 2015

- Ho Chi Minh City University of Technology Second Upper Honour.
- · Major: Electrical and Electronics Engineering, Minor: Automation System Design.
- THESIS TITLE: Applying of Fuzzy Logic Algorithm on Legged Locomotion Robot.

OPEN-SOURCE PROJECTS (SELECTED)

WER-in-CPP

- Developed an open-source API to calculate Word-Error-Rate for ASR project based on Minimum-Edit-Distance Dynamic Programming problem
- · Added the new features of text-based visualization and WER-per-utterance for statistical analysis.
- Source code: https://github.com/dangkhoadl/WER-in-cpp

Machine Learning and Algorithms Repositories

- Maintain GitHub repositories for personal note-taking, introducing and mentoring students to Data Structures Algorithms, and Machine Learning.
- Solved multiple problems in Computer science (Traveling Salesman, Stable Marriage Matching, etc.) and Machine Learning (Random Walks, AI bot playing the trending mobile game Lyto Different Color).
- Source code: https://github.com/dangkhoadl

CERTIFICATIONS (SELECTED)

- Coursera Discrete Optimization Certificate
- Coursera Machine Learning Specialization Certificate
- Coursera Data Structures and Algorithms Specialization Certificate
 Udacity Machine Learning Engineer Nanodegree Certificate