Yasufumi Moriya

Address: 528 Collins Avenue West, Whitehall, Dublin9, Dublin, Ireland

Mobile: +353 089 706 5388

Email: yasufumi1006@gmail.com

EDUCATION & TRAINING

Dublin City University, Dublin, Ireland **Apr. 2017 - present**

PhD Candidate at the School of Computing

Supervisor: Prof. Gareth. J. F. Jones

Thesis: Augmenting Automatic Speech Recognition (ASR) and Spoken Content Retrieval (SCR) for User-generated Video Collection

Summary:

- Developed hybrid hidden Markov model (HMM) deep neural network (DNN) ASR systems using Kaldi for spoken content of user-generated videos.
- Improved transcription accuracy by the use of visual context features and human face features from videos for neural language modelling for ASR.
- Investigated semi-supervised training and acoustic model adaptation for ASR user-generated videos. Resulted in improved speech transcripts and better search accuracy on a spoken video collection.
- Implemented neural ranking systems using ASR transcripts for SCR research and improved search accuracy on a video collection.
- Participated in the Text REtrieval Conference (TREC) 2020 Podcasts Search task. Achieved 2nd best results for official submissions.
- Gained hands-on experience of working on video data involving data collection and cleaning and data annotation using Amazon Mechanical Turk.
- Experience using high performance computing (HPC) for processing large video dataset.

University of Edinburgh, Edinburgh, United Kingdom **Sept. 2014 - Aug. 2015**

Master of Science in Speech and Language Processing with distinction

Supervisor: Dr. Zhizheng Wu

Thesis: Stacked Denoising Autoencoder for the Front-end of DNN-based Speech Synthesis

Summary:

- Gained a wide variety of knowledge about speech processing and natural language processing, as well as linguistics and machine learning methods.
- Hands-on experience of building machine translation, natural language understanding, text-to-speech and automatic speech recognition systems through assignments.
- Developed a denoising auto-encoder using Theano for text analysis of text-to-speech for master's dissertation project.
- Developed a webpage using PHP to collect user assessment of the proposed method for dissertation.

Sophia University, Tokyo, Japan **Apr. 2009 - Mar. 2014**

Bachelor of Foreign Studies (Russian) with concentration in linguistics

Dissertation: Occurrence of English Technical Words in Japanese

CAREER HISTORY

Japanese Language Engineer, Amazon Ireland (part-time) **Sep. 2018 - present**

- Working on improvement of pronunciation for the Japanese text-to-speech (TTS) system.

Japanese Linguist, Ivona Software Amazon Development Center **Nov. 2015 - Mar. 2017**

Summary:

- Gained experience of agile development of commercial neural TTS software. Contributed to the release of three Japanese voices.
- Created guidelines of phonetic annotation, and supervised the quality of annotation performed by contractors.
- Prepared test cases and identified issues for quality assessment of a Japanese TTS text processing engine.
- Improved text expansion modules using C language for the text-analysis process of the Japanese TTS system.
- Worked in a team with software engineers and speech scientists to improve the voice quality of Japanese TTS system.
- Created recording prompts for voice recording sessions and supervised recordings of voice actors.

SELECTED PUBLICATIONS

- Y. Moriya, and G. J. F. Jones, "LSTM Language Model Adaptation with Images and Titles for Multimedia Automatic Speech Recognition", In Spoken Language Technology,, 2018
- Y. Moriya, and G. J. F. Jones, "Multimodal Speaker Adaptation of Acoustic Model and Language Model for ASR using Speaker Face Embedding" In International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019
- Y. Moriya and G. J. F. Jones “DCU-ADAPT at the TREC 2020 Podcasts Track”, In Text REtrieval Conference (TREC), 2020
- Y. Moriya and G. J. F. Jones “An ASR N-best Transcript Neural Ranking Model for Spoken Content Retrieval, Submitted to Automatic Speech Recognition and Understanding (ASRU) 2021
- Y. Moriya and G. J. F. Jones “Augmenting ASR for user-generated videos with semi-supervised training and content genre adaptation” Submitted to International Conference on Statistical Language and Speech Processing (SLSP) 2021

TECHNICAL SKILLS

I use Python for simple tasks such as data preprocessing and for more complex tasks of building machine learning models. I can comprehend C/C++ language and have experience of feature implementation for the production system at Amazon. Day-to-day, I use Linux and Windows, and I’m familiar with Mac OS from my experience at Amazon.

Python (fluent)

- Familiar with various libraries such as PyTorch, NumPy, Scipy and PyLucene.

C/C++ (intermediate)

- Modified Kaldi code to incorporate extralingual features.
- Implemented text expansion for the Japanese TTS.

Other Programming languages I have experience with

- Bash script, Perl, R and PHP

OS

- Windows, Mac, Linux

LANGUAGES SKILLS

Japanese (native), English (fluent), Russian and Polish (conversational)