

SHANG-YI CHUANG

kagaminccino.github.io ◇ +1-917-478-6519 ◇ sc2357@cornell.edu ◇ www.linkedin.com/in/sychuang/

SUMMARY OF QUALIFICATIONS

Extremely self-motivated engineer with excellent understanding of machine learning algorithms

- 5+ years experience in developing software programs for scientific research
- 3+ years experience in **Speech Processing and Language Processing**
- Strong expertise in deep learning frameworks including **PyTorch**, **TensorFlow**, **Keras**, and **scikit-learn**

EDUCATION

- | | |
|---|----------------|
| Cornell Tech in New York, United States | 2021 – Present |
| <ul style="list-style-type: none">• M.Eng. in Computer Science• Cornell Tech Merit-Based Scholarship• Courses: Machine Learning Engineering, Algorithms and Data Structures for Applications | |
| National Taiwan University in Taipei, Taiwan; GPA: 3.86/4.30 | 2012 – 2017 |
| <ul style="list-style-type: none">• B.S., Major in Mechanical Engineering, Minor in Electrical Engineering• Dean's List Award (Top 5% of the class in GPA) | |
| Osaka University in Osaka, Japan; Grade: Highest grade | 2016 – 2017 |
| <ul style="list-style-type: none">• Frontier Lab Special Auditor in Adaptive Machine Systems• Japan Student Services Organization Scholarship | |

WORK EXPERIENCE

- | | |
|---|-------------|
| Research Assistant at Academia Sinica in Taipei, Taiwan | 2018 – 2021 |
| <ul style="list-style-type: none">• Audio-Visual Multimodal Learning for On-device Systems (IEEE/ACM TASLP, INTERSPEECH 2020)<ul style="list-style-type: none">• Improved system robustness against insufficient hardware or inferior sensors by a data augmentation scheme• Minimized additional multimodal processing costs by applying an autoencoder and data quantization techniques• Significantly reduced the size of data to 0.33% without sacrificing speech enhancement performance• Cross-Lingual Movie QA (Question Answering) System<ul style="list-style-type: none">• Reduced unfavorable inequalities in technology caused by limited data in minority languages• Applied transfer learning to a Mandarin system by incorporating translated corpus in dominant languages• Achieved zero-shot learning on Mandarin movie QA tests by using pre-trained multilingual models• EMA (Electromagnetic Midsagittal Articulography) Systems (ISCAS 2021, EUSIPCO 2021)<ul style="list-style-type: none">• Designed silent speech for patients with vocal cord disorders by joint training mel-spectrogram and deep feature loss• Improved the character correct rate of automatic speech recognition by 30% in speech enhancement tasks• Incorporated EMA into end-to-end speech synthesis systems and achieved 83% preference in subjective listening tests• Self-Supervised Learning on Speech Enhancement<ul style="list-style-type: none">• Realized speech enhancement by applying a denoising autoencoder with a linear regression decoder• Enhanced 43% of speech quality without limited intrusive paired data• Potentially empowered the realization of unsupervised dereverberation | |

SELECTED PUBLICATIONS

- [1] **S.-Y. Chuang**, Y. Tsao, C.-C. Lo, and H.-M. Wang, "Lite Audio-Visual Speech Enhancement," in *Proc. INTERSPEECH 2020*.
- [2] Y.-W. Chen, K.-H. Hung, **S.-Y. Chuang**, J. Sherman, X. Lu, and Y. Tsao, "A Study of Incorporating Articulatory Movement Information in Speech Enhancement," in *Proc. EUSIPCO 2021*.

SKILLS

- | | |
|-----------------------------|---|
| Programming Language | Python, C, MATLAB, Bash, Visual Basic, SQL, LabVIEW, Verilog |
| Toolbox | Dlib, OpenCV, FFmpeg, Hugging Face, SoX, Praat, librosa, pandas |
| Visualization | visdom, Matplotlib, plotly, gnuplot, Inkscape, Visio |