

# SHANG-YI CHUANG

kagaminccino.github.io ♦ +886-920-335-066 ♦ 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**, **Computer Vision**, and **Natural Language Processing**.
- Strong expertise in deep learning frameworks including **PyTorch**, **Keras**, **TensorFlow**, and **scikit-learn**.

## EDUCATION

---

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

## WORK EXPERIENCE

---

- |  |             |
|--|-------------|
| <b>Research Assistant at Academia Sinica</b> in Taiwan   | 2018 – 2021 |
| <ul style="list-style-type: none"><li>• <b>Audio-Visual Multimodal Learning Projects</b> (IEEE/ACM TASLP, INTERSPEECH 2020)<ul style="list-style-type: none"><li>· Improved the system robustness against insufficient hardware or inferior sensors in a car-driving scenario.</li><li>· Minimized additional multimodal processing costs while addressing privacy problems of facial data.</li><li>· Significantly reduced the size of data to 0.33% without sacrificing the speech enhancement performance.</li></ul></li><li>• <b>EMA (Electromagnetic Midsagittal Articulography) Projects</b> (ISCAS 2021, EUSIPCO 2021)<ul style="list-style-type: none"><li>· Addressed silent speech for patients with vocal cord disorders or high-noise environments.</li><li>· Improved the character correct rate of automatic speech recognition by 30% in speech enhancement tasks.</li><li>· Incorporated EMA into speech synthesis systems and achieved 83% preference in a subjective listening test.</li></ul></li><li>• <b>Cross-Lingual Movie QA System</b><ul style="list-style-type: none"><li>· Focused on reducing unfavorable inequalities in technology caused by limited data in minority languages.</li><li>· Implemented transfer learning with additional English corpus to enhance a Mandarin QA System.</li><li>· Achieved zero-shot learning on Mandarin Movie QA tests.</li></ul></li><li>• <b>Self-Supervised Learning on Speech Enhancement</b><ul style="list-style-type: none"><li>· Aimed at realizing speech enhancement without limited intrusive paired data.</li><li>· Improved 43% of speech quality by applying a denoising autoencoder with a linear regression decoder.</li><li>· Greatly encouraged the realization of unsupervised deep learning systems.</li></ul></li><li>• <b>Construction of Multimodal Datasets</b><ul style="list-style-type: none"><li>· Highly addressed multimodal common problems of asynchronous devices.</li><li>· Supervised crucial environment setups for collaborative labs, schools, and hospitals.</li><li>· Published Taiwan Mandarin Speech with Video, an open source dataset including speech, video, and text.</li></ul></li></ul> |             |

## SKILLS

---

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