

# 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**, **TensorFlow**, **Keras**, and **scikit-learn**.

## EDUCATION

---

### Cornell Tech

2021 – Present

- M.Eng. in **Computer Science**
- Merit-Based Scholarship
- 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

### National Taiwan University, GPA: 3.86/4.30

2012 – 2017

- B.S., Major in Mechanical Engineering, Minor in Electrical Engineering
- Dean's List Award (Top 5% of the class in GPA)

### Osaka University, Grade: Highest grade

2016 – 2017

- Frontier Lab Special Auditor in Adaptive Machine Systems
- Japan Student Services Organization Scholarship

## WORK EXPERIENCE

---

### Research Assistant at Academia Sinica in Taiwan

2018 – 2021

- **Audio-Visual Multimodal Learning Projects** (IEEE/ACM TASLP, INTERSPEECH 2020)
  - Improved the system robustness against insufficient hardware or inferior sensors in a car-driving scenario.
  - Minimized additional multimodal processing costs while addressing privacy problems of facial data.
  - Significantly reduced the size of data to 0.33% without sacrificing the speech enhancement performance.
- **EMA (Electromagnetic Midsagittal Articulography) Projects** (ISCAS 2021, EUSIPCO 2021)
  - Addressed silent speech for patients with vocal cord disorders or high-noise environments.
  - Improved the character correct rate of automatic speech recognition by 30% in speech enhancement tasks.
  - Incorporated EMA into speech synthesis systems and achieved 83% preference in a subjective listening test.
- **Cross-Lingual Movie QA System**
  - Focused on reducing unfavorable inequalities in technology caused by limited data in minority languages.
  - Implemented transfer learning with additional English corpus to enhance a Mandarin QA System.
  - Achieved zero-shot learning on Mandarin Movie QA tests.
- **Self-Supervised Learning on Speech Enhancement**
  - Aimed at realizing speech enhancement without limited intrusive paired data.
  - Improved 43% of speech quality by applying a denoising autoencoder with a linear regression decoder.
  - Greatly encouraged the realization of unsupervised deep learning systems.
- **Construction of Multimodal Datasets**
  - Highly addressed multimodal common problems of asynchronous devices.
  - Supervised crucial environment setups for collaborative labs, schools, and hospitals.
  - Published Taiwan Mandarin Speech with Video, an open source dataset including speech, video, and text.

## 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