



Projects

CSC3160/MDS6002 teaching team

Project Topics

▶ **Three Recommended projects (Baseline provided)**

- Singing voice conversion (POC: Xueyao Zhang)
- Detecting generated abstract (POC: Feng Jiang)
- Voice Spoofing detection (POC: Li Wang)

▶ **Other projects including: (POC: Xi Chen)**

- Conversation summarization
- Accent conversion
- Audio quality prediction
- Commonsense QA
- Fake news detection
-

If you do NOT choose one of the three recommended projects, please contact Xi Chen (xichen7@link.cuhk.edu.cn) before Feb 24

Singing Voice Conversion Project

▶ POC: Xueyao Zhang (xueyaozhang@link.cuhk.edu.cn)

▶ Background

- Singing Voice Conversion (SVC) is to convert singing voice to our desired targets. It has been paid more and more attention due to its potential applications in imitation and entertainment, singing voice beautification, vocal education, and creative art.
- You can read this tutorial to know more about it (including the related papers and datasets): <https://www.zhangxueyao.com/data/SVC/tutorial.html>

▶ Task

- You need to conduct an “**any-to-one**” conversion task. Specifically, given any singing voice source, you are required to convert it to *Opencpop* singer’s singing voice.

▶ Baseline

- WORLD-based SVC: <https://github.com/SLPcourse/Singing-Voice-Conversion>

Singing Voice Conversion Project

- ▶ Potential ideas to work on
 - Reproduce 1-3 published papers
 - Inject singing voice **domain knowledge** (eg: pitch, duration, score, timbre features) in to the model
 - Explore the impact of different **vocoders**
 - Design better **evaluation metrics**. For example, more robust objective evaluation, or more efficient and friendly subjective evaluation
 - Improve the **explainability** of SVC. For example, **visualize** the different module of the deep learning framework.
 - Bring your own idea

XU

provide some ideas for singing voice conversion research



Here are a few ideas for singing voice conversion research:



1. Developing new machine learning models for singing voice conversion that can handle a wider range of singing styles and languages.
2. Improving the quality of the converted singing voice, by researching ways to better preserve the expressive elements of the singing, such as vibrato and phrasing.
3. Investigating the use of unsupervised or semi-supervised learning techniques for singing voice conversion, which would require less labeled data.
4. Developing a singing voice conversion system that can work in real-time, so that it can be used in live performance settings.
5. Exploring the use of transfer learning techniques to improve the performance of singing voice conversion models on low-resource languages and genres.
6. Investigating the use of singing voice conversion for creating more expressive synthetic singing voices for digital music composition, or to enhance the singing of people with speech and singing disorders.
7. Investigating the use of singing voice conversion for creating expressive synthetic singing voices for virtual assistants and voice assistants.
8. Investigating the use of singing voice conversion to enhance the singing in games, movies, and other interactive media.
9. Investigating the use of singing voice conversion to enhance the singing in karaoke and singing games.

These are just a few examples, and there are many other possibilities for singing voice conversion research.

Suggestions from ChatGPT

Detecting Generated Abstract Project

- ▶ POC: Feng Jiang (jeffreyjiang@cuhk.edu.cn)

Voice Spoofing Detection Project

- ▶ **POC: Li Wang (wangli@cuhk.edu.cn)**
- ▶ **Background**
 - Intelligent voice systems, such as speaker recognition and automatic speech recognition, are vulnerable to spoofing attacks, limiting their further development. Therefore, voice spoofing detection techniques are needed to ensure the safety the system.



Voice Spoofing Detection Project

▸ Task

- The voice spoofing detection task is essentially a binary classification task. Specifically, given an audio, label it as either bonafide or spoof.

▸ Baseline

- High Frequency Cross-Subband: <https://github.com/SLPcourse/Voice-Spoofing-Detection>

Voice Spoofing Detection Project

► Potential ideas to work on

- Reproduce 1-3 published papers
- Try generative methods: In the ASVSpooof2021 competition, the one class classification method was used, which is a generative method that shows the highest spoof detection accuracy, outperforming the baseline system.
- Data augmentation: text-to-speech, voice conversion, adversarial generation, and audio clipping are the ways to generate attack samples, and the dataset can be augmented using the above methods.
- Exploring network structures: Different network structures, such as convolutional neural networks and attention mechanisms, are good at identifying different types of spoofed samples due to their different modeling properties.
- Bring your own idea.