Vocal music cheater

The vocal music cheater algorithm shall be designed to generate singing audio from the speech (in your own voice) that reads lyrics or words. The song can be in a specific genre, either something similar with speech like rap, or something far from like opera. This includes correcting the fundamental frequency of sound, adjusting the changing timbre between speech and singing, and aligning the rhythm and score of speech. You can also consider converting the timbre from vocal / instrumental music to instrumental music.You may use frequency estimation, vocoder, rhythm alignment, auto-tune etc.

acoustic holography

Acoustic holography is a method for estimating the sound field near a source by measuring acoustic parameters away from the source by means of an array of pressure and/or particle velocity transducers. The Measuring techniques included in acoustic holography are becoming increasingly popular in various fields, most notably those of transportation, vehicle and aircraft design, and noise, vibration, and harshness (NVH). Project could be a model the acoustic pressure field on a special volume and evaluate your model.

Automatic Speech Verification

ASV technology is used to distinguish between a piece of audio that is saying right now or that has been pre-recorded. Such technology has important implications for information security. Learning acoustic representations based on knowledge of acoustics, signal processing and machine learning etc., and assess the characteristics of poorly classified audio and the possible causes of model errors. You can refer to https://www.asvspoof.org/ for the acoustic features and classification models that are probably useful.

Audio enhancement

The foreground will have a live speech or concert etc. Sometimes, the signal can be passed through a special filter like the mask you wear during COVID or gets an acoustic echo in a room. Sometimes you should reduce background noise. The background noise can be any distractors like radio, television or white noise etc. The goal is to filter out the background & channel distractors in various simulated environments with varying SNR. You can choose a specific task like speech recognition with & without masks. You may use adaptive filtering, clustering, equalization-cancellation models, Acoustic Echo Cancellation etc.

Blind Estimation of Number of Speakers

One of the important prerequisites for blind source separation is knowing the number of sources to be separated. One approach to this could involve finding the number of significant independent formant contours in a spectrogram.

(Non-linear) ICA Algorithm

The success of ML algorithms generally depends on feature extraction, and it is believed that this is because different representations can entangle and hide more or less the different explanatory factors of variation behind the data. Independent component analysis (ICA) is a method to separate multivariate signals into additive subcomponents. This is done by assuming

that the subcomponents are non-Gaussian signals and that they are statistically independent from each other. Your work is to design a new ICA algorithm that has some advantage (in computation efficiency, better result or some innovation like non-linear etc.) over current algorithms (e.g. FOBI, FastICA, JADE, Infomax, etc.). It is also recommended to evaluate where and why your algorithm is. If you suspect the weakness meets the strength of DL, you may combine them ONLY AFTER class (and publish them perhaps (*^▽^*) )

Accent Identification / conversion
One of the key difficulties with speech recognition is handling accents. Develop a robust method for accent identification or conversion.

Automated DJ
Given a collection of recorded music (or radio programmes), play the music successively and harmoniously, without pauses. You may use the following techniques: rhythm alignment (BPM, tempo...etc), finding similar tunes (i.e. instruments, chords,..etc.) between successive music, match the pitch between vocals, use a recommender system (e.g. Spotify API) to find the next music to play

Denoising EEG signals
Use an EEG signal, in any application of interest, and denoise it to easily identify the signal you care about. You may use ICA, PCA, NMF as many of the medical signals have much noise within it which DL methods always do not work and interpreted methods are needed.

Hacker's Keyboard . Sensor
Protection of information from malicious intent is a topic of increasing concern in the security community as advances in machine learning continue to broaden the avenues of attack. One such attack which has generated interest in the last decade has been the potential for reconstruction of a user’s pattern of keystrokes by analyzing the subtle differences between sounds made by individual keys in audio recordings of typing. The project is to implement or defeat such attacks or similar attacks related with signal processing in an unsupervised manner.

Gesture Phase Segmentation
Seek to aid human computer interaction by identifying nonverbal communication through the use of hand gesture analysis. Analyze the positions of speakers’ hands, neck, and spine to classify frames into the types of gestures.

Predicting & Classifying RF Signal Strength
In order to plan and deploy multi transmitter wireless networks, such as cellular networks, engineers must be able to predict the propagation of electromagnetic waves from transmitters. RF signal propagation prediction is usually carried out by human experts using a propagation model, which comes from empirical field measurement to predict RF path loss at a given distance. These models must be selected and adjusted according to the specific environment in which the transmitter will be deployed; Therefore, detailed information about the environment must be collected. After prediction, the resulting signal strength must be measured manually.

This manual modeling and measurement process is very time-consuming; Network deployment is measured in weeks and months. In addition, it is not easy to scale the three-dimensional network deployment supporting UAV. This project aims to automate the signal propagation modeling process.

Non-intrusive load monitoring
Nonintrusive load monitoring (NILM), or nonintrusive appliance load monitoring (NIALM),[1] is a process for analyzing changes in the voltage and current going into a house and deducing what appliances are used in the house as well as their individual energy consumption. Electric meters with NILM technology are used by utility companies to survey the specific uses of electric power in different homes. NILM is considered a low-cost alternative to attaching individual monitors on each appliance.

Few-shot Bioacoustic Event Detection
This challenge focuses on sound event detection in a few-shot learning setting for animal (mammal and bird) vocalisations. Participants will be expected to create a method that can extract information from five exemplar vocalisations (shots) of mammals or birds and detect and classify sounds in field recordings. The main objective is to find reliable algorithms that are capable of dealing with data sparsity, class imbalance, and noisy/busy environments. Refer to data challenge DCASE2021.

Self-awareness in Heterogeneous Multi-Robot Systems
A new Challenge on self-awareness in heterogeneous multi-robot systems has been organized within the first International Conference on Autonomous Systems (ICAS 2021). The research field of this competition is the unsupervised anomaly detection through self-aware autonomous systems, which is an active topic involving IEEE Signal Processing Society, also through the Autonomous Systems Initiative, and Intelligent Transportation Systems Society. You can refer to the signal processing conference ICAS 2021 Challenge

Machine Learning for 3D Audio Signal Processing
3D audio is gaining increasing interest in the machine learning community in recent years. The field of application is incredibly wide and ranges from virtual and real conferencing to game development, music production, autonomous driving, surveillance and many more. In this context, Ambisonics prevails among other 3D audio formats for its simplicity, effectiveness and flexibility. You can refer to the signal processing conference ICAS 2021 Challenge

Time-domain Timbre Transformation
The aim is to transfer timbre from one musical source to another musical source, but entirely in the time domain. Frequency domain methods typically struggle with phase reconstruction issues which time-domain approaches seek to avoid. The most popular dataset for this is the NSynth dataset by Magenta, though a newer dataset such as Amp-Space may be used once it is released.

Audio Texture Synthesis
The aim is to synthesize “audio textures” from a set of parameters and a random seed. Audio textures are often loosely defined and thus benefit from a data-based approach for analysis and synthesis. Such a synthesis model would have applications in sound design for film, music, video games, etc. SynTex/DSynth would be a useful dataset for this application.

Learning Tube Shapes for Vowel Synthesis with FDTD simulation of Webster’s Equation
Physics based modeling provides a good prior for modeling of many situations. It is possible that such priors may greatly improve the quality or plausibility of learned results. The aim of this work is to match tube shapes (that approximate a vocal tract shape) to produce vowel sounds. By parametrizing the tube shape function and the loss function in a differentiable manner, an optimization scheme can be proposed to learn the unknown tube shape in Webster’s equation. See Bilbao, “Numerical Sound Synthesis” for more information on Webster’s equation.

Cymbal Matching using FDTD simulations of Von Karman’s Nonlinear Shell Model
Physics based modeling provides a good prior for modeling of many situations. It is possible that such priors may greatly improve the quality or plausibility of learned results. The aim of this work is to match cymbal sounds to nonlinear shell vibration simulations. By parametrizing the nonlinearities, the cymbal shape parameters and the loss function in a differentiable manner, an optimization scheme can be proposed to learn the unknown parameters in the von Karman equation. See Bilbao, “Numerical Sound Synthesis” for more information on the von Karman equation.