

Yerlan Sharipov: Speech Processing

This paper discusses the major architectures that are utilized in many state-of-the-art speech processing technologies. The three main architectures that the paper focused on are:

- Convolutional Neural Network (CNN)
- Recurrent Neural Network (RNN)
- XLNET

Convolutional Neural Network: This is one of the most widely used neural network architecture. It is widely used in image classification.

Recurrent Neural Network: This is used extensively in speech processing and ASR systems. It's main advantage is that it can be used on long sequences of data as opposed to single data points.

XLNET: This is the new state of the art algorithm in ASR systems that can recognize speech patterns based on a "context" word.

Shuwei Li: CNN based on Digital Signal Processing

This paper discusses how convolutional neural network (CNN) is integrated into digital signal processing. The paper discusses how CNNs can be used to overcome the issue of overfitting data that comes with Multilayer Perceptions making it one of the prime architectures used in digital signal processing. Also, due to the approach employed by CNNs, long short-term memory can be developed.

Erin O'Neill: Multi-speaker Identification State of the Art

This paper discusses the difficulties involved in multi-speaker identification and a technique used to overcome some of them. Multi-speaker identification is a relatively nascent field that shows a lot of promise. This is as a result of the many potential application of this technology. With efficient multi-speaker identification systems, the issue of noise in many audio processing systems including automatic speech recognition (ASR) systems becomes a non-issue.

Zuxiong Tan: Speech Processing

This paper discusses the full architecture of a multi-speaker identification system. The author discusses the various methods employed at each stage. The multi-speaker identification system involves utilizing a multichannel input that is then tested against an acoustic model in which judgment is made to determine which speaker is being “talked to.” For the feature extraction state, there are two widely used technologies which are I-Vector and MFCC. MFCC is modeled using the human hear and as shown to be very effective as discussed in the paper. The author also discusses the various types of technologies that are used for training and testing the dataset. These include:

Radial Neural Networks: Radial Neural Networks are a special class of Neural networks that utilize a radial basis to form inputs and neuron parameters to form outputs as a linear combination of these basis. It is used widely in function approximation, system control and time series classification.

Probabilistic Neural Networks: Probabilistic Neural Networks are a class of Neural networks used mainly for classification. It uses Bayes rule to assign weights to an input and classifies the input based off the result.

Long Short memory Networks: This is one of the most utilized algorithms for automatic speech recognition. It is a special class of deep neural networks algorithms that is based on an artificial recurrent neural network architecture. It is especially useful in ASR as it can be used on not just a single data point at a time but on long sequence of data making it a very important algorithm in speech processing.

Zhou Shen: Speech Processing: Neural Network in DSP

This paper discusses a particular kind of Neural Network algorithm called STOCHASM. This algorithm consists of three layers. An input layer, middle layer and output layer.