2) Spectral Image based Environmental Sound classification using CNN with meaningful data-augmentation

10.1016/j.apacoust.2020.107581

Sound classification is a tedious task and Neural Networks like RNN, LSTM etc are based on sequence based sound signal classifiers that have been used for some time.

Recents studies have discovered an alternate use of Convolutional Neural Network, originally developed for image based classification, to classify sound signals. In this paper, spectrogram images, which are visual representations of spectrum of frequencies, as it varies with time, are used as an input, the technique of Mel-spectrogram is used to obtain features here.

This work uses 3 data sets

- ESC-10
- ESC-50
- Urban sound 8k (Us8k)

even though there are 3 datasets, the data requirements of a NN is not met. So the technique of data augmentation is utilized here.

Different models with difference in internal architecture employed to compare one-another & optimize the task.

Some of the models used are

- CNN-1 with 7 layers
- CNN-2 with 9 Layers

~~some~~

transfer learning is based on the following models

- Res Net
- DenseNet
- Squeeze Net
- Alex Net
- VGG