Identifying sound labels from spectrograms using deep convolutional neural networks

Creative and Innovative Project
CS6611
06/03/2021

Lab No:1 DATE: 06/03/2021

TEAM NUMBER: 12

S.NO	ROLL NUMBER	ROLL NO
1	2018103569	NIRANJAN K
2	2018103063	SHANKAR KUMAR S
3	2018103620	VEDANTH S

Problem Domain

Real World Scenario, Machine Learning, Deep Learning, Computer Vision, Auto encoders, Acoustics.

Abstract

AI plays an important role in acoustics recognition. Deep learning techniques can assist in recognition of sounds which we come across in our day-to-day life. Deep convolutional neural networks have been used to classify environmental sound recently. The classification system with high performance often requires a large well-labeled dataset. The cost of tagging audio segments correctly and completely is quite high thus the deep learning models need to have high generalization ability if a weakly-tagged dataset is used.

In our work, we use a dataset which has sounds from 50 different classes. The dataset has 2000 samples of audio recordings of real-life sounds such as dogs, cats, rain, coughing, sea waves, wind and so on. Within this context, we present an automatic system to detect and classify sounds, especially those generated by birds and insects among other sounds that can be heard in a natural environment. Traditional approach involved the usage of normal Deep Neural Networks. However, analyzing the collected data by humans is a big challenge and for that reason, it is necessary to develop automated technologies in order to help experts on that task. In order to tackle this invasive involvement of humans we propose a new method of using spectrograms.

We process the audio signals in such a way that we create waveforms present in the audio signals along with the labelling of the corresponding signals. We plan on using a dense convolutional neural network for the processing of audio signals.

Deliverables

Input

Audio samples of Animals, Natural soundscapes & water sounds, Human and non-speech sounds, Interior and domestic sounds, Exterior and urban noises.

Output

Predicting the sound label for the given audio sample.

Language and packages

PYTHON, TENSORFLOW, KERAS, PANDAS, NUMPY, PYTORCH, LIBROSA

Base Papers

1) Deep Neural Networks for Identifying Cough Sounds

https://ieeexplore.ieee.org/document/7570164

EVALUTATION

OBSERVATION (5)	
SPOT (5)	
EXECUTION (15)	
TOTAL (25)	
SIGNATURE	