

Project description

Given an Audio clip the model should extract the unique sounds present in the audio clip and develop a context of the environment. For e.g. if the children playing at park clip is the give we should be able to extract the sounds of the children, birds (if any), cars to understand the context about the environment.

Data set:

The plan is to use the Audio set outsourced by google

<https://research.google.com/audioset/dataset/index.html>

for the initial prototypes and see how it turns out. This contains good amount of labelled data for the use case.

Use cases:

Improve the decision-making process of self-driving cars by perceiving the environment through sound along with the images.