* Model Details

MODEL CARD

**Audio Sound Event** 一張含有 文字, 美工圖案 的圖片

自動產生的描述

This model has approximately 6.2 thousand parameters. There 4 predetermined sound events including "Background noise", "baby crying ", “glass breaking” and “gun shot”. It can be applied for real-time inference, with an inference time of approximately 225ms on the VA8801 platform.

* Model Specifications

Model Type : Convolutional Neural Network

Model Architecture : A modification network based on ResNet for VA8801

Input : 1\*192\*24

Output : 4 classes

* Application

Smart speakers, mobile phones.

Application example: Devices is able to recognize the sound event and then executes the follow-up instructions.

* Limitation
  1. Hard to distinguish the sound of baby talking or baby laughing from the sound of baby crying.
* Training Data

Dataset : Background noise, baby crying, glass breaking, gun shot.

Total : 26330 clips

Preprocess: sample rate is 8000, number of mel bins is 24.

* Reference

[基于Pytorch实现的声音分类系统](https://github.com/yeyupiaoling/AudioClassification-Pytorch)