DEEPCRAFT™ Ready Model for Factory Alarm and Siren Detection

Introduction

This document describes the DEEPCRAFT™ Ready Model for Factory Alarm and Siren Detection, an audio-based AI developed by Imagimob, an Infineon Technologies company, that detects alarms and sirens happening inside a factory. This report also provides the technical specifications of the machine learning model, its performance in common scenarios, and various test results for the model including the real-time testing performance on an Infineon PSOC™ 6 board.

Contents

Introduction....1
Model Specification....2
Model Overview....2
Expected Performance....2
Model Tech Specs and Deployment....2
Data Properties....2
Positive Data....3
Negative Data....3
Testing....4
Test Set Results....4
On-device testing....5
Test Results on 20 Alarm/Siren Audio Files....6
Test Results on One Minute Alarm Sound in a Factory....6
Test Results on 46 Minutes of Factory Sound....6
Appendix I - Data Sources....7
Appendix II - Factory Sound Examples....7
Appendix III - Results of Model Testing on Factory Sound Files....7
Appendix IV - Results of Model Testing on 20 Alarm/Siren Audio Files....8

imagimob
An Infineon Technologies Company

Model Specification

Model Overview

Upon activation, the DEEPCRAFT™ Ready Model for Factory Alarm and Siren Detection can be used, for instance, to disable the noise cancellation function in an environment with diverse factory sounds. It's important to note that the model does not differentiate between different types of alarm and siren signals.

Expected Performance

This model focuses on alarm and siren sounds, which are defined as those with prolonged acoustic components longer than 1.2 s. Multiple subsequent detections count as one alarm and/or siren event.

The model performs 89% on the test set and the model on-device testing result is 80%. More detailed testing results are described below. An alarm/siren event, compared to its background, needs to have a Signal To Noise Ratio (SNR) which is at least 10 decibels (dB) in order to be detected. The model is expected to be robust against:

- different distances and angles from the sound source.
- diverse factory background sounds.

Model Tech Specs and Deployment

The alarm/siren detection DEEPCRAFT™ Ready Model is able to detect an alarm/siren from sound data with the following characteristics:

- Sample rate: 16000 Hz
- Channels: 1 (Mono)
- Bit Depth: 16bit

The Flash size of the model is 32.216 KB and RAM size of the model is 27.025KB.

Its inference time is about 144 ms when running on a PSOC™ 6 (model CY8CKIT-062S2-43012) mounting a Sense shield with a microphone (model CY8CKIT-028-SENSE). The model outputs a prediction every 800 ms.

Data Properties

This Ready Model has been constructed using a variety of positive and negative sounds, with the positive sounds consisting of different types of alarm and siren

signals. The negative data encompasses a range of sounds that could potentially occur within factory environments.

Positive Data

The positive data used to build the model consists of sound recordings with one or more alarm/siren events per file. Many different types of alarms/sirens are used, such as ambulances, police car sounds from different countries, fire alarms and smoke alarms.

Negative Data

The model has been built using sound recordings belonging to the following non-alarm/non-siren, or negative data, categories:

- Factory sounds
- Human talking with and without factory noise
- Hammering
- Keyboard typing
- Music
- Office
- Rain
- TV/radio
- Vacuum cleaner
- Vehicle engine
- Wind
- Chainsaw
- Circular saw
- Drilling
- String trimmer
- Wood chopping
- Human breathing with and without factory noise
- Human clapping with and without factory noise
- Human clearing throat with and without factory noise
- Human laughing with and without factory noise
- Human screaming with and without factory noise
- Human shouting with and without factory noise
- Human sneezing with and without factory noise
- Human speech with and without factory noise
- Human whistling with and without factory noise
- Hydroelectric power plant sounds
- Coal mine sounds
- Blast furnace sound

Testing

Test Set Results

The test set includes 176 non-alarm/non-siren files and 128 alarm/siren files. The average length of a non-alarm/non-siren is 184 seconds, and the average length of an alarm/siren is 58 seconds. The performance of the model on the test set indicates that a good balance between true positives and false positives is achieved when using a confidence threshold equal to 90%. The model's performance is summarised in the table below, in so-called confusion matrices. The meaning of the percentages/values reported in both of them is as follows:

- Top Left Value (True Negatives): actual negative/non-alarm/non-siren data predicted as negative/non-alarm/non-siren data
- Bottom Left Value (False Positives): actual negative/non-alarm/non-siren data predicted as positive/alarm/siren data
- Top Right Value (False Negatives): actual positive/alarm/siren data predicted as negative/non-alarm/non-siren data
- Bottom Right Value (True Positives): actual positive/alarm/siren data predicted as positive/alarm/siren data

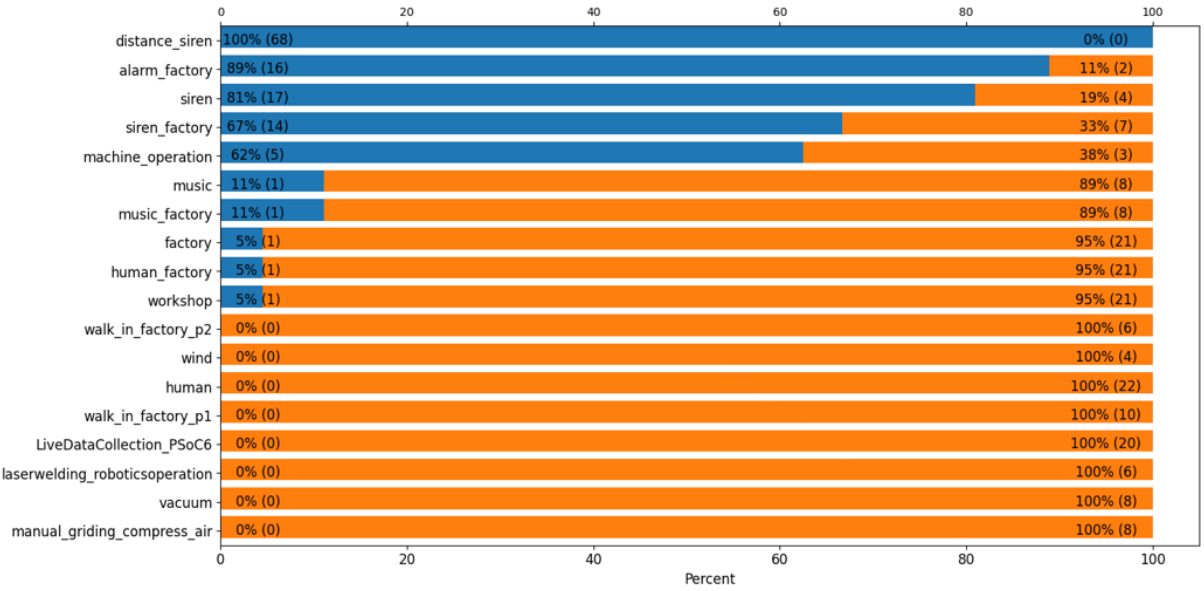
File-Based Confusion Matrix	**Actual non-alarm/non-siren**	**Actual alarm/siren**
Predicted non-alarm/non-siren	94.32% (166 files)	10.16% (13 files)
Predicted alarm/siren	5.68% (10 files)	89.84% (115 files)

According to the file-based confusion matrix, the model achieves 92% accuracy on this data. Specifically, the model has a True Positive rate of approximately 89.84%. A False Positive rate of 5.68% for non-alarm/non-siren files was caused by the following sound categories:

- Human in factory
- Music in factory
- Factory
- Music
- Machine Operation

The histogram plot below provides a detailed breakdown of false positives per category. The blue bar and the values on the left side refer to the number of files in each category where at least one alarm/siren event is detected. The orange bar and the values on the right side indicate the number of files where non-alarm/non-siren was detected.

It is important to note that the test set includes 20 audio files recorded by two different PSOC™ 6 boards. Each file contains five minutes of car factory sounds. The model successfully produces zero false positives when performing on this data. This functional test result is shown in the Appendix III.



From this histogram, it is clear that the machine operation sound and music trigger the most false positives. The music triggering false positives contains siren-like sounds. The machine operation sounds closely resemble alarms or sirens at a distance. For specific examples of these sound profiles, refer to the provided Appendix II.

On-device testing

The on-device testing was done on:

- One device with 10 different kinds of siren sounds, and 10 different kinds of alarm sounds, augmented with factory background sound played by a speaker.
- Three devices with one minute of real time alarm factory sound.
- Three devices with approximately 40 minutes of real time factory sound.

For these tests, the model is deployed on the PSOC™ 6 (model CY8CKIT-062S2-43012) mounting a Sense shield with a microphone (model CY8CKIT-028-SENSE) and the confidence threshold is set to 90%. The result summary is listed below.

To test this model, the following steps need to be done:

1. Obtain the ready model library from Imagimob
2. Obtain a PSOC™ 6 board with a microphone, such as the PSOC™ 062S2 Wi-Fi BT Pioneer Kit
3. Create an example project that samples the microphone.
4. Use the provided API calls and example code in the library header
5. Create a UI for displaying the library outputs. E.g. a printf statement to a terminal

Test Results on 20 Alarm/Siren Audio Files

The PSOC™ 6 with the alarm/ siren detection model deployed was tested on 20 completely different alarm/siren audio files using the speaker. The result table is provided in Appendix IV.

As noted, four false negatives out of 20 audio files are seen in the test result. 80% of alarms/sirens have been detected in this test.

Test Results on One Minute Alarm Sound in a Factory

Three PSOC™ 6 boards with the DEEPCRAFT™ Ready Model for Factory Alarm and Siren Detection deployed were tested on one minute long alarm sounds inside of a factory in real time. The result table is provided in Appendix IV.

As noted, zero false negatives were seen in the test result. 100% of alarms/sirens were detected in this real time testing. The volume of the alarm sound was relatively low in a loud factory environment, however, the model still managed to trigger during this alarm testing.

Test Results on 46 Minutes of Factory Sound

Three PSOC™ 6 boards with the DEEPCRAFT™ Ready Model for Factory Alarm and Siren Detection deployed were tested inside of a factory for 46 minutes in real time. The result table is provided in Appendix IV.

As noted, there are two types of machine sounds inside of the factory that triggered the model. The sounds that triggered false positives lasted around one minute, which is around 2% out of 46 minutes testing time. Moreover, these machine operations sound similar to alarm/siren sounds. One of the sound examples is provided in Appendix II - Factory Sound Examples.

Appendix I - Data Sources

Below are the data sources that were used in addition to our own data collection.

The positive data has been downloaded from the following sources:

- Freesound - https://freesound.org/
- Shapingwaves - https://www.shapingwaves.com/
- Surroundsoundlab - https://www.surroundsoundlab.net/

The negative data has been downloaded from the following sources:

- Freesound - https://freesound.org/
- Sonicsalute - https://sonicsalute.com/
- Surroundsoundlab - https://www.surroundsoundlab.net/

Appendix II - Factory Sound Examples

Machine operation sound examples that trigger false positives are provided in the link below:
Machine operation in factory

Appendix III - Results of Model Testing on Factory Sound Files

Test Results

Factory Files	Result
Session1	True Negative
Session2	True Negative
Session3	True Negative
Session4	True Negative
Session5	True Negative
Session6	True Negative
Session7	True Negative
Session8	True Negative

Session9	True Negative
Session10	True Negative
Session11	True Negative
Session12	True Negative
Session13	True Negative
Session14	True Negative
Session15	True Negative
Session16	True Negative
Session17	True Negative
Session18	True Negative
Session19	True Negative
Session20	True Negative

Appendix IV - Results of Model Testing on 20 Alarm/Siren Audio Files

Test Results

Siren	Result
670115_police-siren-echoing-through-c_FMA2_135.3	True Positive
BrazilAmbulance_3a8a_factory_797.8	True Positive
ChinaPolice_95af_factory_2153.6	True Positive
ItalyAmbulance_5c31_factory_829.8	False Negative

GermanyPolice_9dbf_factory_2217.3	True Positive
FranceAmbulance_7a38_factory_1382.8	True Positive
USAPolice_4c6f_factory_1844.4	True Positive
UKFireEngine_7d13_factory_4.8	True Positive
USAPolicewithHowler_1b6e_factory_1374.6	False Negative
UKAmbulance_2d8a_factory_905.2	True Positive

Alarm	**Result**
SW_alarm_703d+ factory sounds	True Positive
SW_alarm_218b + factory sounds	True Positive
SW_alarm_94f2 + factory sounds	True Positive
SW_alarm_9fa9_v+ factory sounds	True Positive
SW_alarm_81c1_v+ factory sounds	True Positive
SW_alarm_668e + factory sounds	False Negative

Alarm_bell_ringing_3p3u-Utoj8Y_0_10	True Positive
SW_alarm_372b_v+4.0_aug2_3570.1	False Negative
Fire Alarm Test in Abandoned Factory	True Positive
Fire Alarm Siren - Factory	True Positive

Device	Sound	Test Result
Board 1	One minute alarm	True Positives
Board 1	46 minutes factory	False Positives on two types of machine operation sounds
Board 2	One minute alarm	True Positives
Board 2	46 minutes factory	False Positives on one type of machine operation sound
Board 3	One minute alarm	True Positives
Board 3	46 minutes factory	False Positives on one type of machine operation sound