DEEPCRAFT™ Ready Model for Snore Detection

Introduction

In this document, we describe the DEEPCRAFT™ Ready Model for Snore Detection, an audio-based AI model developed by Imagimob, an Infineon Technologies company. We provide details about the technical specifications of this machine learning model, its performance in common scenarios, and various test results for the model including the real-time testing on an Infineon PSOC™ 6 board.

imagimob

Model Specification

Model Overview

The DEEPCRAFT™ Ready Model for Snore Detection is designed to detect snoring when the user is sleeping. Such a model can be run for example inside of a wearable device that has a built-in microphone. The model is designed to detect the snores in different kinds of typical sleeping environments.
The model is designed and tested up to 2 meters away. At further distances the performance will begin to deteriorate, and some snores may not be detected.

Model Tech Specs

The DEEPCRAFT™ Ready Model for Snore Detection is able to detect a snore from sound data with the following characteristics:

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

Model Deployment on a PSOC™ 6 Board

The has the following memory footprint:

- RAM: 24.3KB
- FLASH: 53KB

And its inference time is about 23 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 62 ms.

Data Properties

The Ready Model has been built using various positive and negative sounds. The positive sounds are different kinds of snoring from different individuals occurring in different indoor environments. The negative data represents different kinds of sounds that could happen indoors. The sounds are listed in the next sections.

Positive Data

The positive data consist of different types of snoring. We also augmented them with different background noise such as white noise as well as different distances. The positive data has 1 or many snoring events per file, with sound file length ranging from 5 seconds up to 3 minutes.

Sounds used include snoring from individuals around the world at different distances, ages and with different background sounds for example:

- TV sounds
- Fan/Air Conditioning
- Ambient Sounds
- etc.

Negative Data

Many different non-snore, or negative sounds, were used in building the model to ensure that it performs well in different environments.

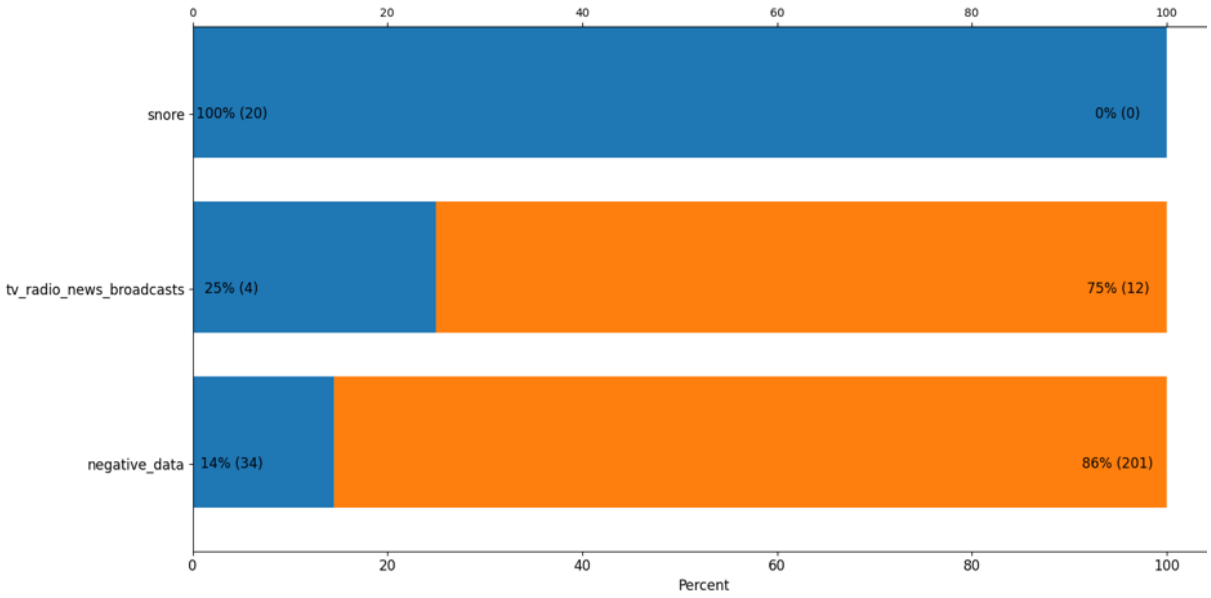
The model has been built using sound recordings belonging to the following non-snore or negative categories:

Indoor sounds

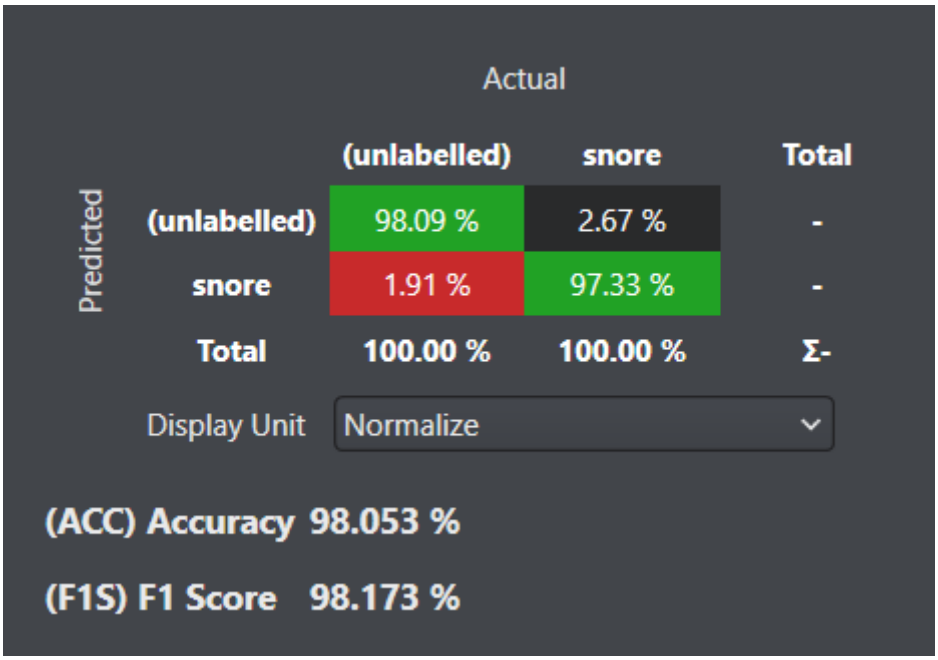
- Cat sounds
- Door sounds
- Generic kitchen sounds
- Blender
- Dishes, cups, cutlery sounds
- Alarms
- Dish washing
- Dog sounds
- Pig sounds
- Sneezing
- Adult laughing
- Baby sounds
- Adult talking
- Dishwasher
- Electrical shaver
- Vacuum cleaner
- Frying food
- Glass break
- Clapping
- Running water
- TV/radio
- Music
- Traffic
- Fan sounds
- Crowd sounds

Testing

Validation Set Results



The plot above shows the predictions of validation set on the file level. As we can see from the plot, the model predicts all of the snore files correctly. But there are some false triggers on the negative data including tv sounds. The 34 false positives among the negative data mainly are alarm, angry cat noises, electrical shaver and loud people talking sounds.

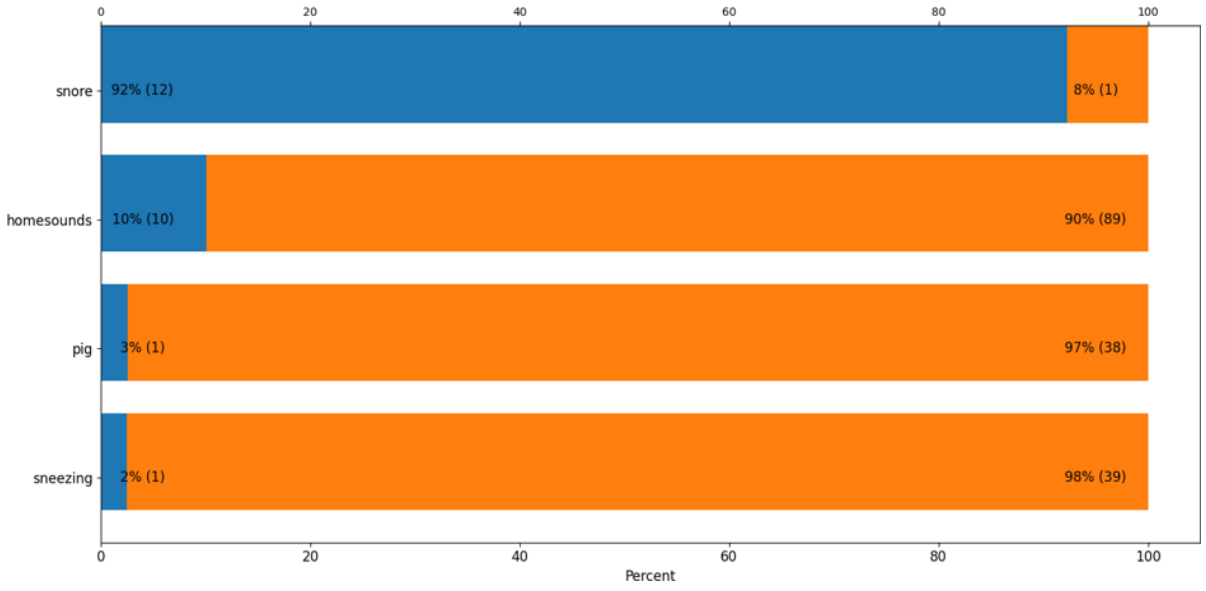


The picture above is the confusion matrix of validation set from our studio. The meaning of the percentages/values is as follows:

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

Here we can see that the model correctly caught 98% of these timesteps as true negatives among the non-snoring data. For the snoring sound, it was able to catch around 97%.

Test set



As we can see from the prediction on the test set, the model can detect 89% of snoring. The 10 false positives among home sounds mainly are sounds of a crowded room of people having conversations.

Test Results on 20 People's Recordings

We tested the model on recordings from 20 participants. For each participant we randomly look at five two-minute segments from the entire sleep duration. We verified that the segments contain snoring and tested if the snores are detected. The testing result is shown at Appendix II.

As shown in the result, the model detected 85% of snoring within 2 minutes. And the model behaved well even with a noisy background. The only participants that the model doesn't react well with are the ones with insects in the background.

On-device testing

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

1. Obtain the Ready Model library from imagimob.com
2. Obtain a PSOC™ 6 board with a microphone. E.g. PSOC™ 62S2 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

We performed on-device testing with 5 people in total. Three of them contain snore sounds and the rest two are normal sleeping sounds without snoring. The testing result is shown below. We can see that the model detects snoring well and does not react to normal sleeping data. However, there are false positives when it comes to loud music and tv sounds.

People	Testing Sounds	Testing Result

Person 1	Around 8 hours of live testing, including snoring sounds, people talking and partially TV background.	Can detect the snore well but there are high false positives on people talking with tv background.
Person 2	Around 6 hours of live testing, mainly quiet sleeping sound without any snoring sound.	0FP on these non-snore sleeping sounds.
Person 3	Around 10 hours of live testing including home sounds, online meeting, music and keyboard typing.	NoticedFPs on loud people talking sound from speaker and loud music sound from speaker. Around 10- 15 FPs in total.
Person 4	Test on 7 hours sleeping sounds that were recorded by the board. Including door opening and closing a few times and maybe some very silent breathing,	3 FPs on these non-snore sleeping sounds.
Person 5	Test on 8 hours sleeping sounds that were recorded by the board. Including partially loud TV sound, snore sound.	The model performs well on snore sounds. But the model doesn't detect every snore, especially at the beginning of the snore period. And the model does not react to the TV sounds.

Appendix I - Additional Details

The positive data has one or many snore events per file, with sound file length ranging from two seconds and up. These files have been downloaded from the following sources:

- Freesound - https://freesound.org/

Freesound is a well established sound database website.

The negative data has been downloaded from the following sources:

- Freesound - https://freesound.org/
- DESED - https://project.inria.fr/desed/

Appendix II - Detailed Test Results on 20 participants' recordings

Participant	Result
30	Doesn't detect short snore, but detects all strong snore
17	Detects all of snore in file level
34	Can see 2 FPs within 10 minutes of background noise without any snoring. Can detect the snore in this noisy background in another 10 minutes file.
35	Detects most of snore and detects all of snore in file level

19	Doesn’t detect any snore in this participant, has cricket sound in the background
6	Detects most of snore and detects all of snore in file level
7	Detects most of snore and detects all of snore in file level
12	Detects most of snore and detects all of snore in file level
13	Doesn’t detect all of the snore but it performs well in terms of file level
11	Doesn’t detect any snore in this participant, which is the same case as participant 19, has cricket sounds in the background
1	Detects most of snore and detects all of snore in file level
14	Doesn’t detect any snore in this participant, which is the same case as participant 19, has cricket sounds in the background
16	Really noisy background, doesn’t detect all of the snore but it performs well in terms of file level

15	Detects most of snore and detects all of snore in file level
48	Detects all of snore in file level
49	Detects all of snore in file level
40	Detects most of snore and detects all of snore in file level
41	Detects most of snore and detects all of snore in file level
42	Detects most of snore and detects all of snore in file level