imagimob
An Infineon Technologies Company

Benchmark Report: DEEPCRAFT™ Ready Model for Cough Detection

Introduction

The DEEPCRAFT™ Ready Model for Coughing Detection is a commercial solution for consumer electronics manufacturers. We compare its performance against cutting-edge models and solutions to ensure real-world effectiveness. We evaluate our model against the most recent research paper with the highest accuracy and an industrial solution integrated into common mobile phones.

Evaluation

We evaluate our model against the following models:

1. CRN[1]: a recently published research paper that adopts a CNN on the Mel spectrogram of the audio signals
2. a software feature integrated into common mobile phones

The comparisons were performed in the following ways:

- Model hardware performance
  - We compared the memory footprint and inference time of the Ready Model against the CRN model
- Model accuracy performance
  - We compared the recall and the outlier accuracy against the model found in common mobile phones

We deployed our model, generated using DEEPCRAFT™ Studio code generation, onto Infineon's PSOC™ 6 with an IoT sense expansion kit (CY8CKIT-028-SENSE).

The CRN model does not provide a pre-trained model, so we cannot evaluate the model accuracy. The model found in common mobile phones is not publicly available, so it's not included in the hardware evaluation.

Model Hardware Performance

In Figure 1, we present the performance of our model in terms of inference time and memory footprint. The CRN model requires 3.6 times more memory than the

[1] https://www.frontiersin.org/articles/10.3389/frobt.2021.580080/full

DEEPCRAFT™ Ready Model. Moreover, the DEEPCRAFT™ model runs 120 times faster. Note that the actual inference time of our model may differ when deployed in the microcontroller. This is a relative comparison in a CPU, given that the CRN model cannot be deployed in a PSOC™ 6 microcontroller.
.

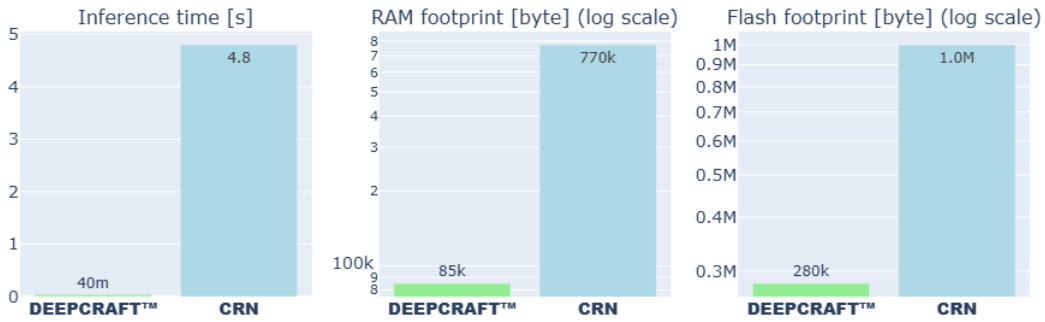

Figure 1 Evaluation of our model against state-of-the-art research model regarding inference time and memory footprint.

Model Accuracy Performance

Figure 2 compares our model's performance with a model found in common mobile phones. The other model shows better recall than our model, because the other model can detect short coughs, such as throat clearing, etc., and coughs from longer distances. However, the DEEPCRAFT™ model is focused mainly on detecting coughs from sick people, which are not considered short. However, the other model is more sensitive to outlier noises than the DEEPCRAFT™ model.

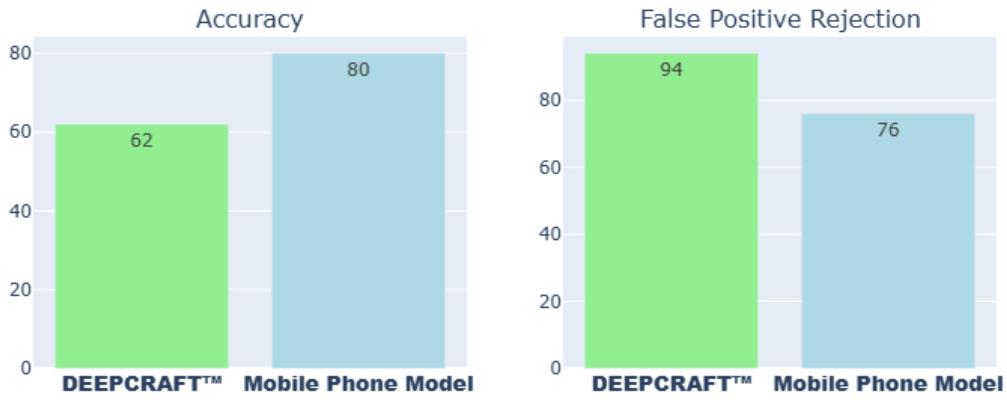

Figure 2 Evaluation of our model against the industrial model from common mobile phones.

Accuracy is the number of correct detections (true positives) divided by the total number of events occurring (true positives). The higher the number, the better. In this case, the event is a coughing sound.

$$Accuracy = \frac{True\ Positives}{Total\ Positives}$$

False Positive Rejection is the number of negative sounds correctly rejected or not triggered divided by the total number of negative sounds. The higher the number, the better. In this case, the negative sounds are any sound that is not a cough, i.e., talking, laughing, etc.

$$False\ Positive\ Rejection = \frac{True\ Negatives}{Total\ Negatives}$$