ETH zürich

Cough Counting

Habib Ben Abda

Machine Learning on Microcontroller Project

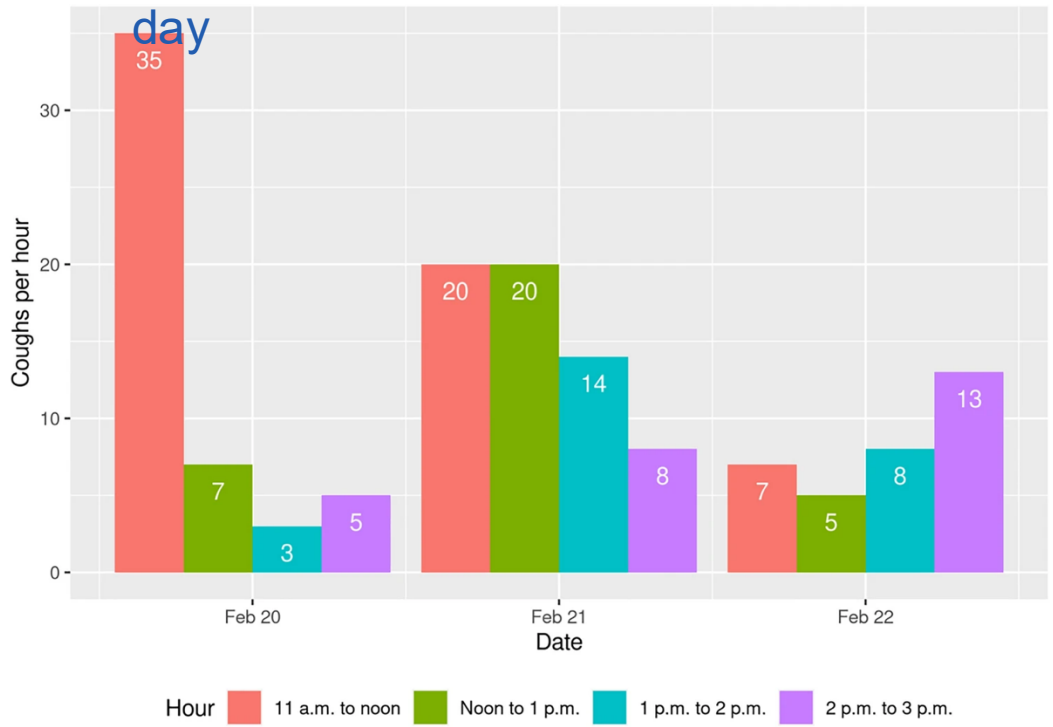
13/06/2025

Motivation

Continuous measurement of number coughs:

- Personalized patient care
- Monitor severity of respiratory diseases
- Assess treatment efficacy of novel antitussive therapies

Cough study measuring only 4 hours per day



Rudd et al., Lung (2022), “The Statistics of Counting Coughs: Easy as 1, 2, 3?”, Fig. 1

Motivation

Continuous measurement of number coughs:

- Personalized patient care
- Monitor severity of respiratory diseases
- Assess treatment efficacy of novel antitussive therapies

24/7 measurement

https://councils.forbes.com/blog/wearable-tech-in-healthcare

Dataset: COUGHVID

EPFL Embedded Systems Laboratory

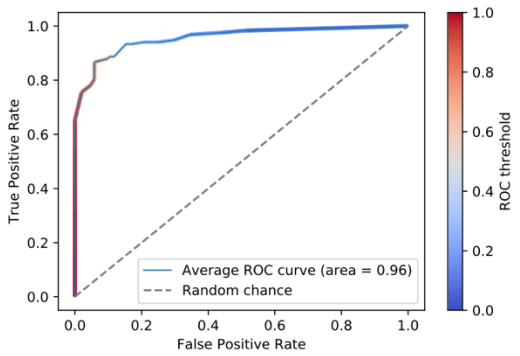
- **COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms**
- **~30’000 sample**
- **eXtreme Gradient Boosting (XGB) cough classifier**

Data Preprocessing

Crowdsourced data

Contains bad samples !

EPFL XGB model
Threashold = 0.8
=> **95.4% precision**



Cough dataset

- Segmentation
- Downsampling to 16kHz
- Cropping to 1s

Preprocessed Cough dataset

Final dataset: ~14'000 sample

(10% testing, 5% validation)

3x Augmentation with:

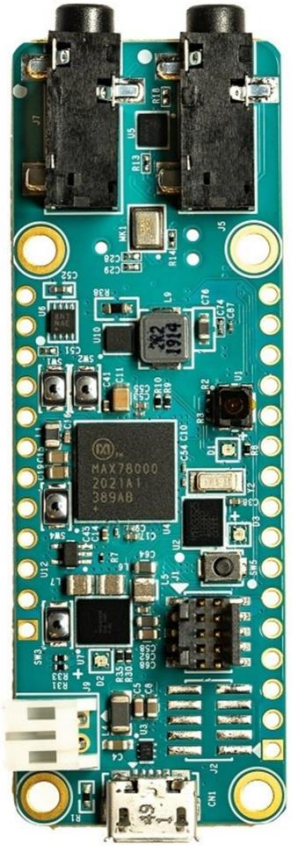
- Stretching
- Shifting
- Noise addition

Augmented Cough dataset

Orlandic, et al., "The COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms". Sci Data 8, 156 (2021).

Target Device

- MAX78000 Microcontroller
  - Dual Core: Arm Cortex-M4 Processor with FPU, 100MHz, RISC-V Coprocessor, 60MHz
  - 512KB Flash Memory
  - 128KB SRAM
  - 16KB Cache
  - Convolutional Neural Network Accelerator
  - 12-Bit Parallel Camera Interface
  - MAX20303 Wearable PMIC with Fuel Gauge
  - Charge from USB
  - On-Board DAPLink Debug and Programming Interface for Arm Cortex-M4 processor with FPU
  - Breadboard Compatible Headers
  - Micro USB Connector
  - Micro SD Card Connector
- Integrated Peripherals
  - RGB Indicator LED
  - User Pushbutton
  - CMOS VGA Image Sensor
  - Low-Power Stereo Audio CODEC
  - Digital Microphone
  - SWD Debugger
  - Virtual UART Console
  - 10-Pin Cortex Debug Header for RISC-V Coprocessor

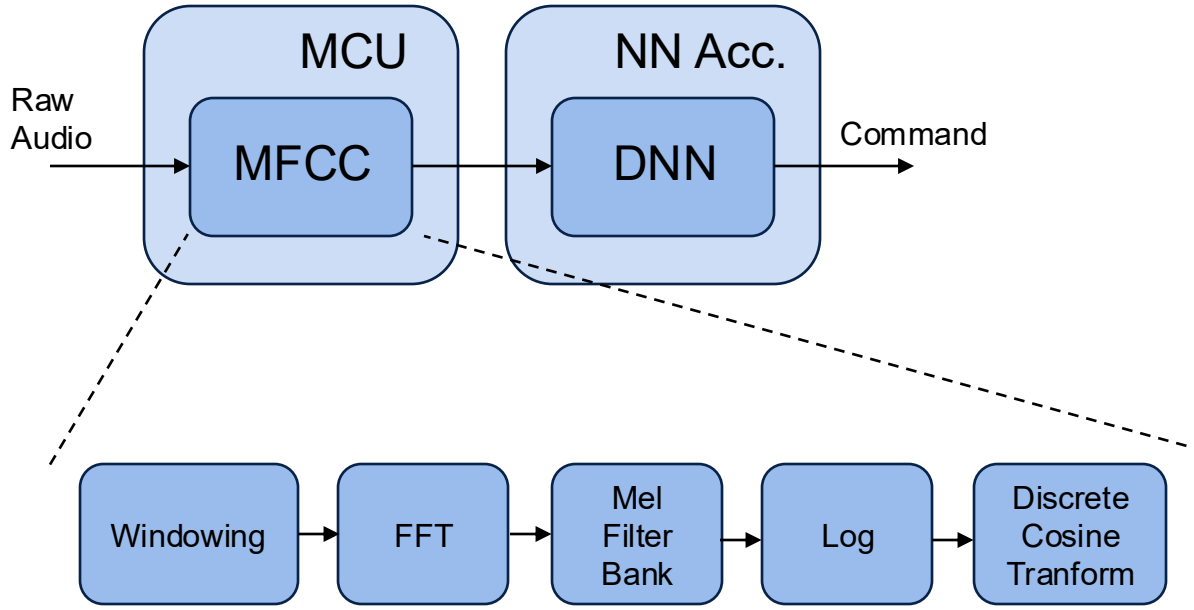


Model Design

- Hardware accelerator ➡ CNN architecture

- Feature extraction?

  First approach:



>2x more power consumption than NN inference !

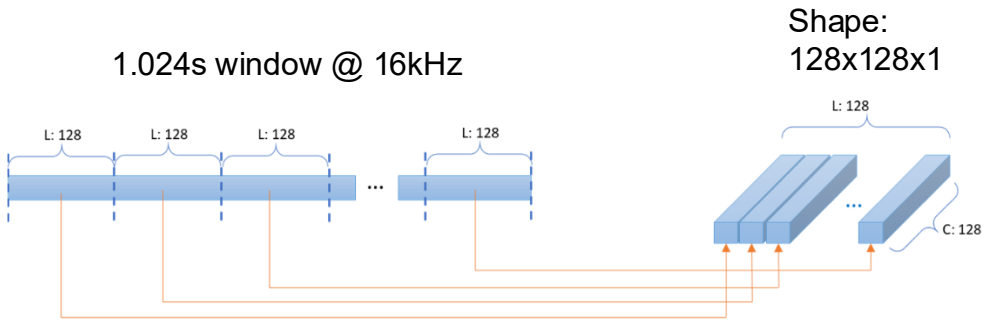
Ulkar M. G., Okman O. E., (2021) “Ultra-Low Power Keyword Spotting at the Edge”, arXiv:2111.04988

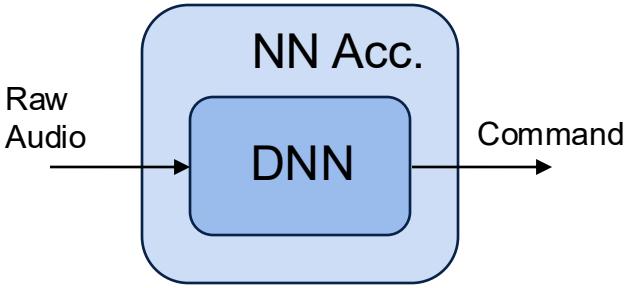
Model Design

- Hardware accelerator ➡ CNN architecture

- Feature extraction?

  Better approach:





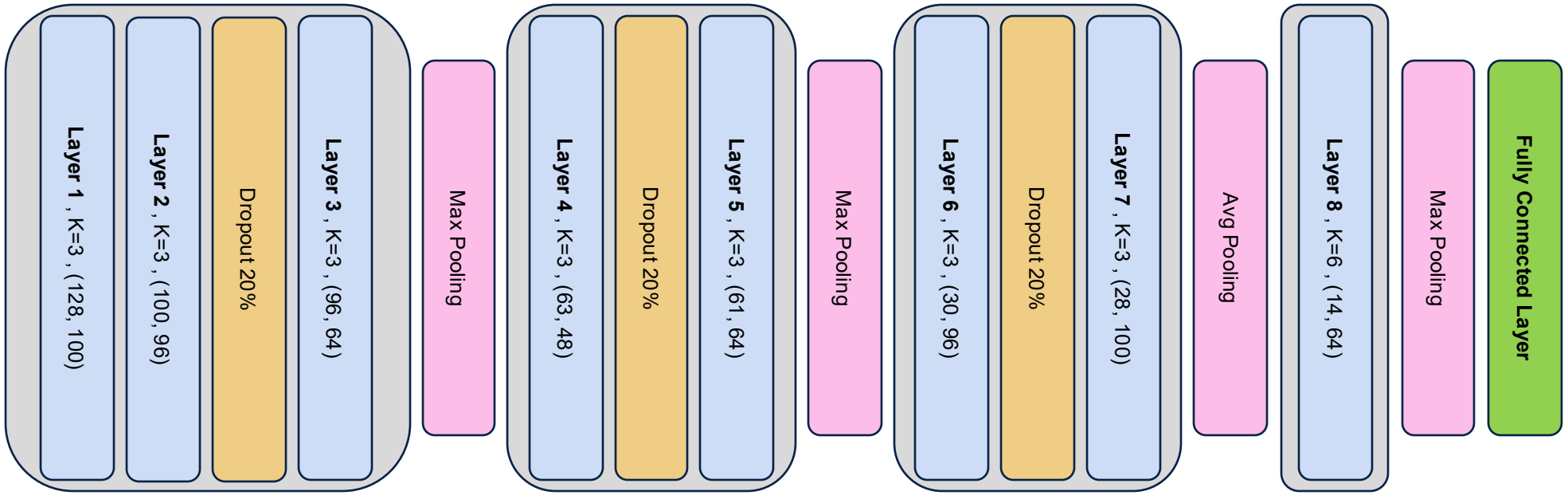
(+) Zero cost (change order of memory locations)

(+) Using 1D conv can extract both short time and long time relations

Ulkar M. G., Okman O. E., (2021) “Ultra-Low Power Keyword Spotting at the Edge”, arXiv:2111.04988

CoughNet v1

- 8 1D-Convolution layers + ReLu activation function
- 4 Max pooling layers
- 1 Fully connected layer (linear)



Inspired from Maxim KWS examples

CoughNet v1 (8 classes)

Nb classes	8 (6 Keywors + Cough + Unknown)
Nb epochs	100
Quatization	Quantization Aware Training
Accuracy	91.5 %
Precision (cough)	0.99
Sensitivity (cough)	0.98
F-1 score (cough)	0.98
Size	428.7 KB (390.6 FLASH + 38.1 SRAM)
Latency	1.8 ms
Nb of operations	8,400,224 ops

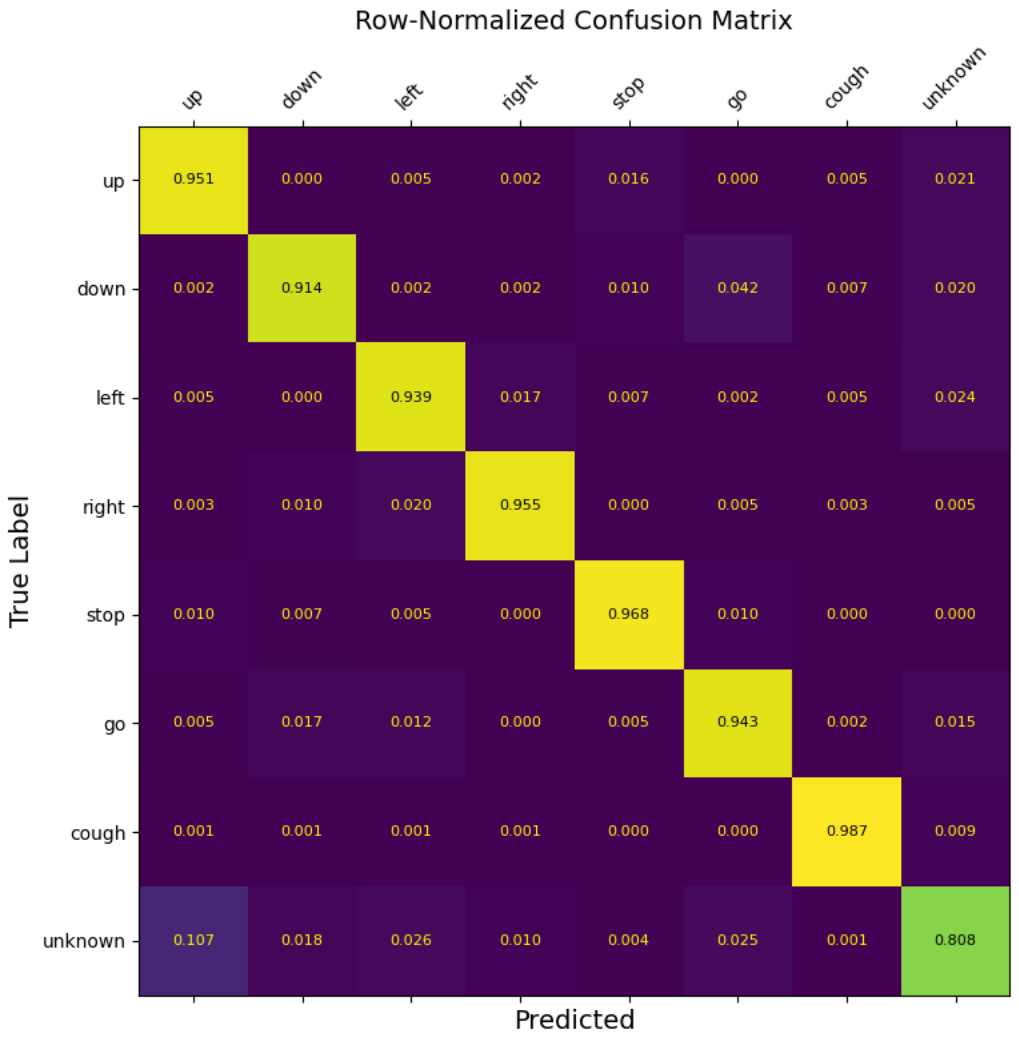


Figure 1: Quatized network, 8 Classes (6 + cough + unknown), 91.5% Accuracy

CoughNet v1 (20 classes)

Nb classes	20 (18 Keywors + Cough + Unknown)
Nb epochs	100
Quatization	Quantization Aware Training
Accuracy	87.9 % (-3.6%)
Precision (cough)	0.98 (-0.01%)
Sensitivity (cough)	0.98
F-1 score (cough)	0.98
Size	432.6 KB (394.4 FLASH + 38.1 SRAM) (+3.9KB)
Latency	1.8 ms
Nb of operations	8,402,272 ops

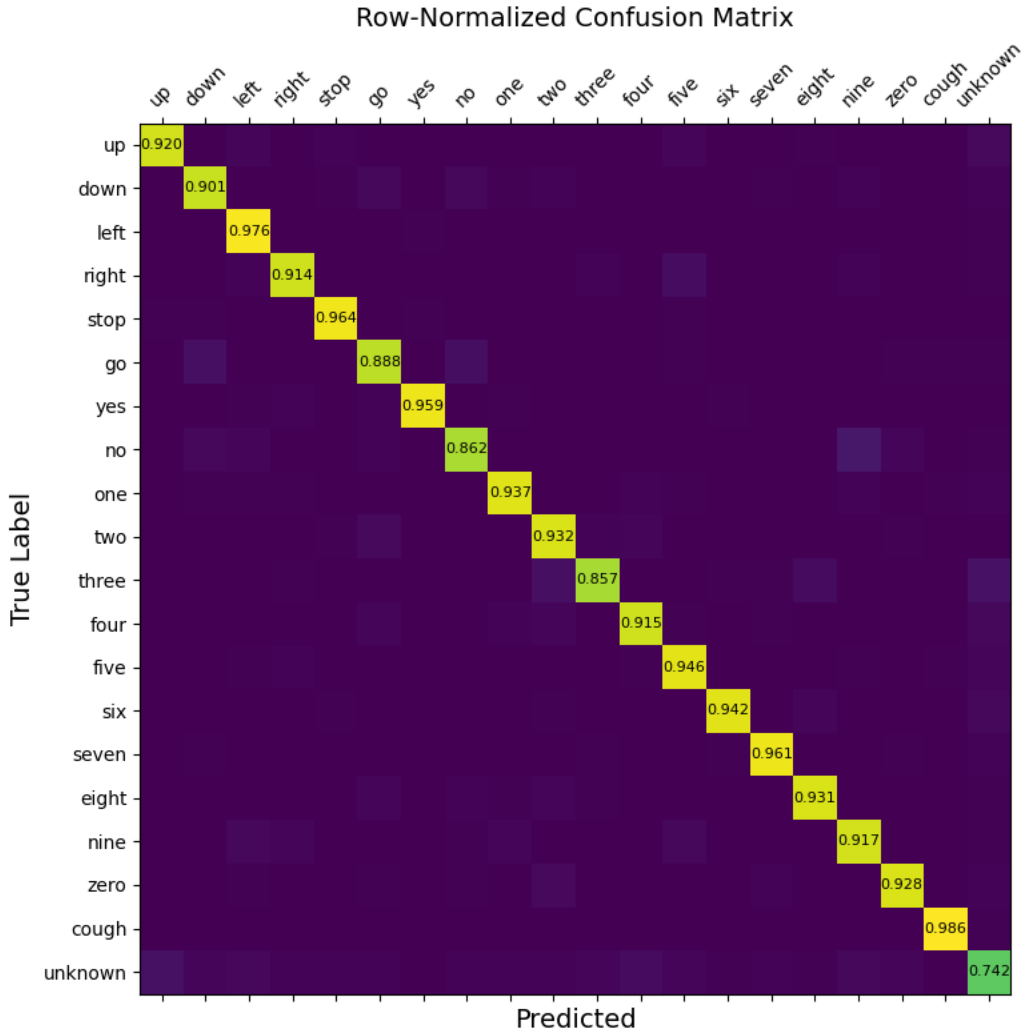
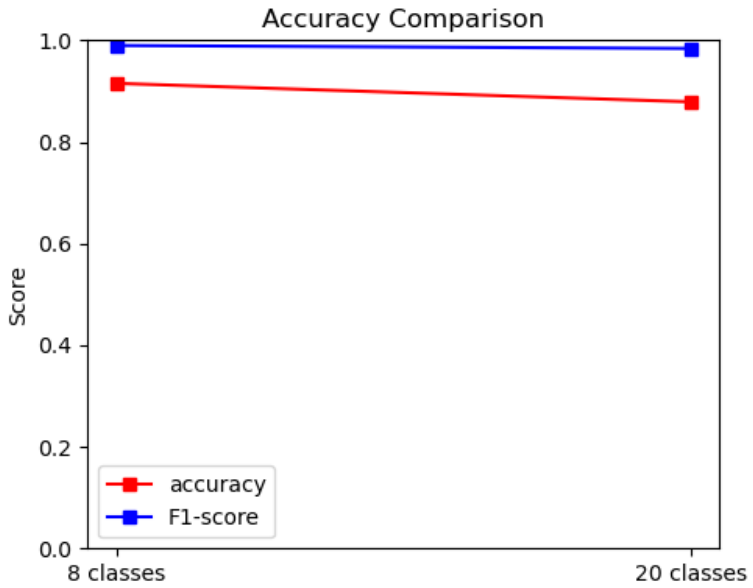
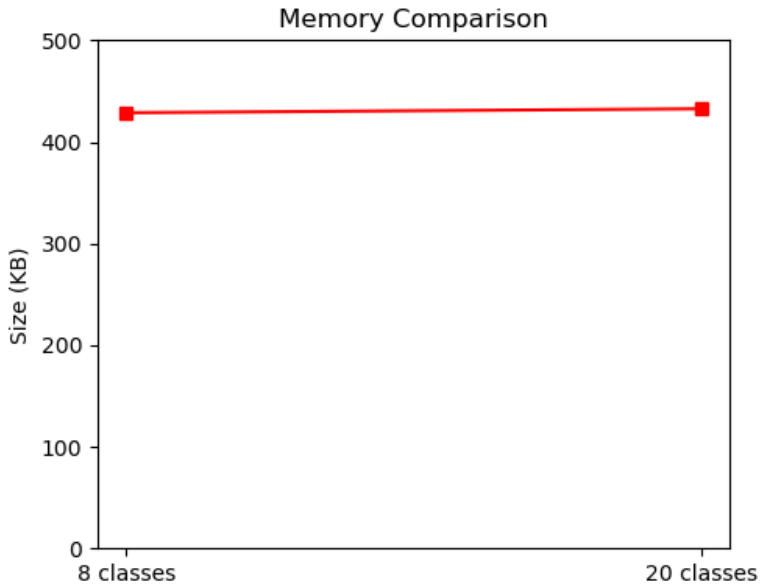


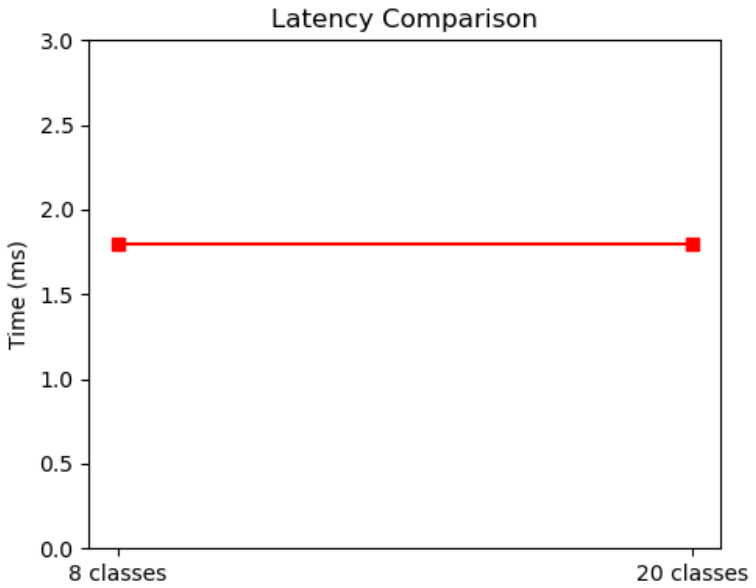
Figure 2: Quatized network, 20 Classes (18 + cough + unknown), 87.9% Accuracy

8 vs 20 classes

Same model, more classes:



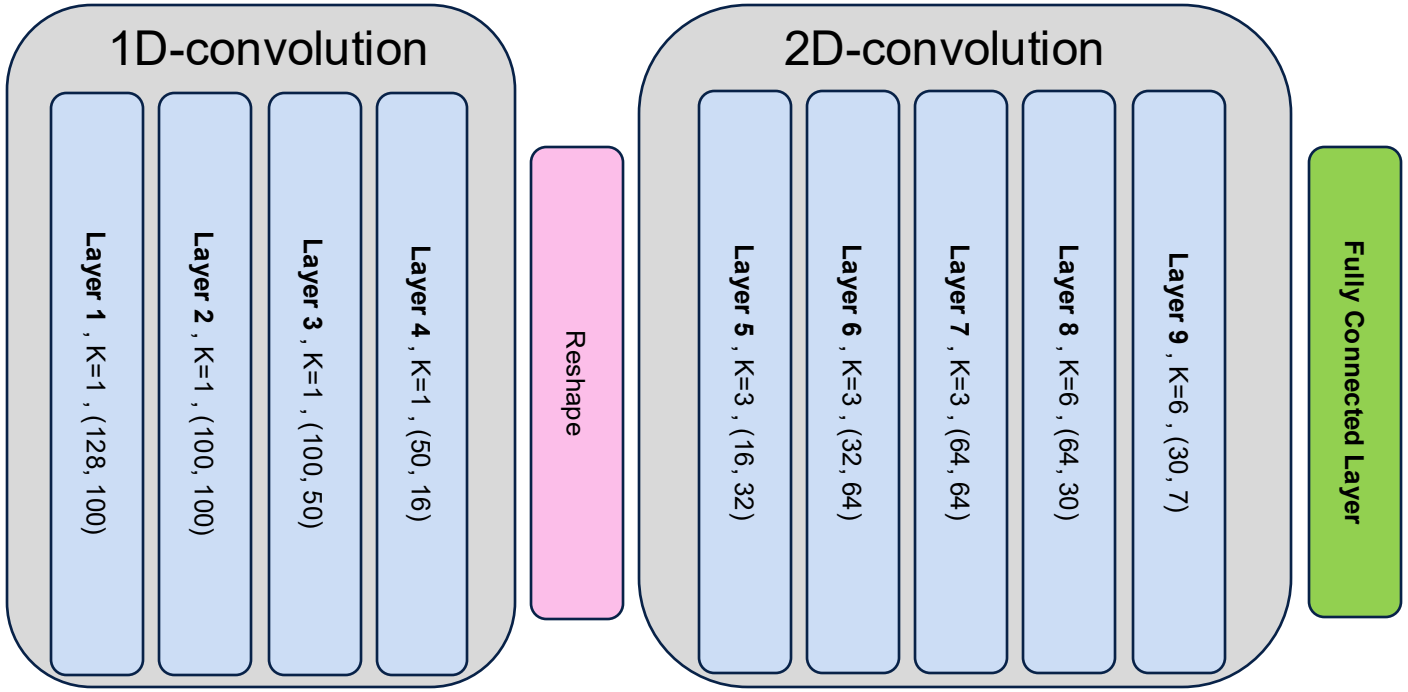




➡ The unknown classes cause most of the accuracy drop, but good for robustness and generalization.

CoughNet v2 (1D-conv then 2D-conv)

- 9 1D-Convolution layers + ReLu activation function
- 1 Fully connected layer (linear)



CoughNet v2 (20 classes)

Nb classes	20 (18 Keywors + Cough + Unknown)
Nb epochs	100
Quatization	Quantization Aware Training
Accuracy	88.1 % (+0.2%)
Precision (cough)	0.96 (-0.02%)
Sensitivity (cough)	0.98
F-1 score (cough)	0.97 (-0.01%)
Size	384.4 KB (346.3 FLASH + 38.1 SRAM) (-48.2 KB)
Latency	2.6 ms (+0.8ms)
Nb of operations	13,859,456 ops

★ Expected to also consume more energy

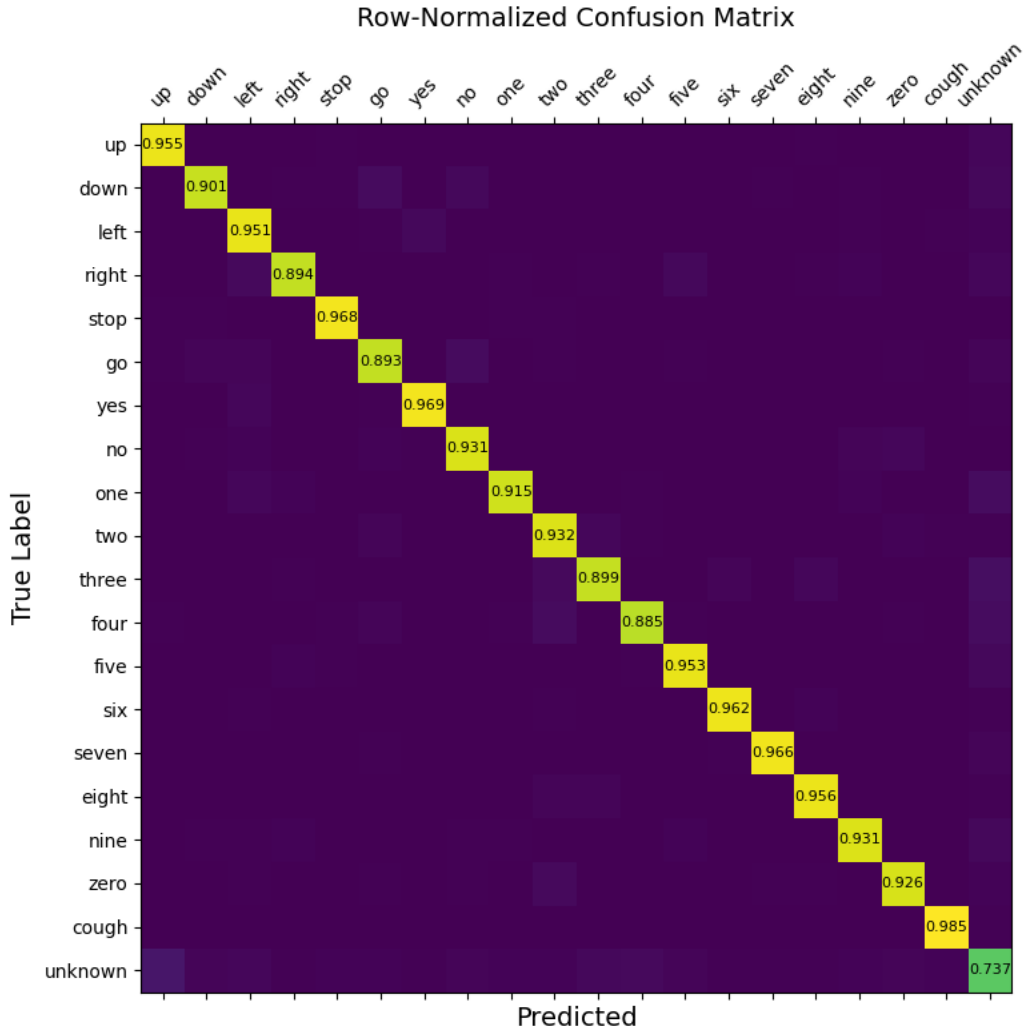
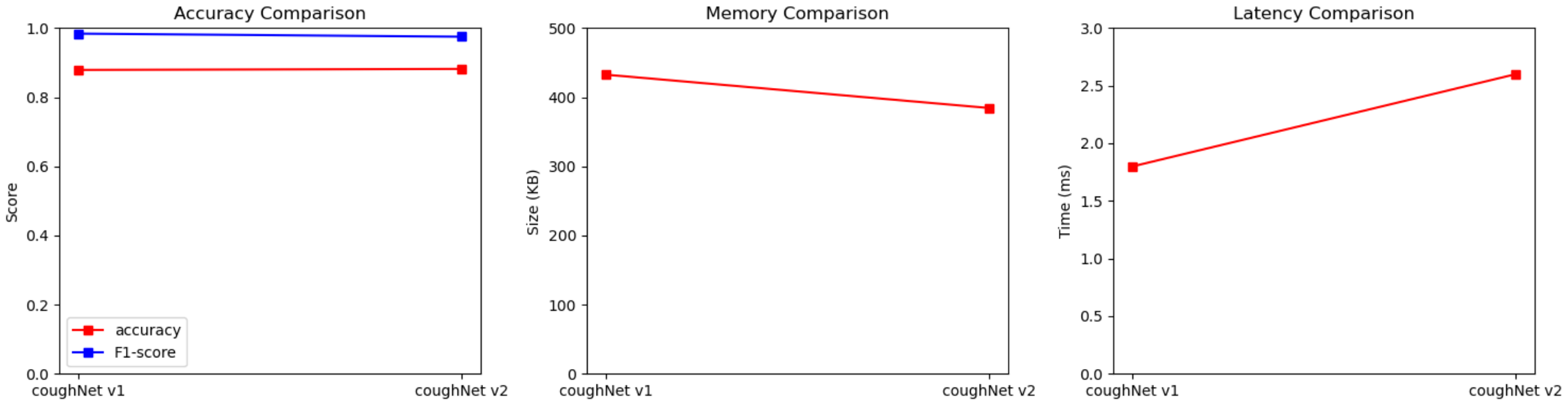


Figure 3: Quatized network, 20 Classes (18 + cough + unknown), 88.1% Accuracy

CoughNet v1 vs. v2

1D-conv model vs 2D-conv model:



➡ Can adapt the model to optimize for memory or latency depending on the need

Comparison with other work

CoughNet vs EPFL XGB model vs Other :

Ref.	Edge-AI execution	HW-aware design	ML performance	Sensors	ML model	Environmental noise
Drugman, 2020, [11]	No	No	*Sens. = 94.7% *Spec. = 95.0%	audio microphone contact microphone	2 Neural Networks	Yes
Otoshi, 2021, [14]	No	No	*Sens. = 92.0% *Spec. = 96.0%	3-axis accelerometer stretchable strain sensor	Variational Autoencoder + K-means clustering	Yes
Nemati, 2021, [28]	No	No	*Sens. = 83.0% *Spec. = 91.7%	microphone kinematic	Random forest + DWT template matching	No
Wang, 2022, [26]	Yes	No	*Acc. = 78.5% *F1 = 77.0%*	4 microphones	CNN	Yes
Jokic, 2022, [21]	No	No	*Acc. = 80.0%	4 microphones	CNN	No
Zhang, 2022, [27]	Yes	No	*AUROC = 77.0%	microphone kinematic	Multi-Nearest Center Classifier	Yes
Kuhn, 2023, [20]	No	No	*Sens. = 88.5% † / 84.15% ǂ *Prec. = 99.97% † / 99.7% ǂ	microphone vibrations sensor	DNN	Yes
EPFL XGB model → **Cough-E (ours)**	Yes	Yes	Sens. = 71.0% Prec. = 78.0% F1 = 77.7%	microphone kinematic	2 XGB classifiers	Yes

(*)Segment-based performance metrics (cf. Section II-F)

(†) daytime testing
(ǂ) nighttime testing

CoughNet v1	Yes	Yes	Sens. = 98.6% Prec. = 98.1% F1 = 98.3%	microphone	CNN Classifier	Yes

Orlandic, et al., “The COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms”. Sci Data 8, 156 (2021).

Comparison with other work

CoughNet vs EPFL XGB model vs Other :

Cough-E (ours)	Yes	Yes	Sens. = 71.0% Prec. = 78.0% F1 = 77.7%	microphone kinematic	2 XGB classifiers	Yes

(*)Segment-based performance metrics (cf. Section II-F)

(†) daytime testing
(ǂ) nightime testing

CoughNet	Yes	Yes	Sens. = 98.6% Prec. = 98.1% F1 = 98.3%	microphone	CNN Classifier	Yes

!! Comparison metrics slightly different

- **CoughNet**: fixed window (1s), metrics deduced from confusion matrix
- **Cough-E**: variable window
  - True Positive: start and end of a predicted cough event overlap with a ground-truth
  - False Positive: predicted cough with no overlap or overlap outside tolerance windows
  - False Negative: ground truth that does not overlap with a prediction

Comparison with Analog Devices

Analog devices paper demonstrating performance of keyword spotting:

TABLE I
COMPARISON OF ACCURACY RESULTS ON THE SPEECH COMMAND DATASET

Model	MFCC	Precision	#Param	Model Size	Acc	
[1]	yes	fp-32	244.2K	976.8 KB	90.2%	
[2]	yes	fp-32	238K	952 KB	95.8%	
[3]	yes	fp-32	61K	244 KB	96.4%	
[4]	yes	int-8	497.6K	497.6 KB	95.4%	
[6]	no	fp-32	122K	488 KB	96.6%	
Analog Devices model → ours	no	**int-8**	419.8K	**419.8 KB**	**96.3%**	200 epochs
CoughNet v1	no	Int-8	432.6 K	432.6 KB	87.9%	100 epochs

Ulkar M. G., Okman O. E., (2021) “Ultra-Low Power Keyword Spotting at the Edge”, arXiv:2111.04988

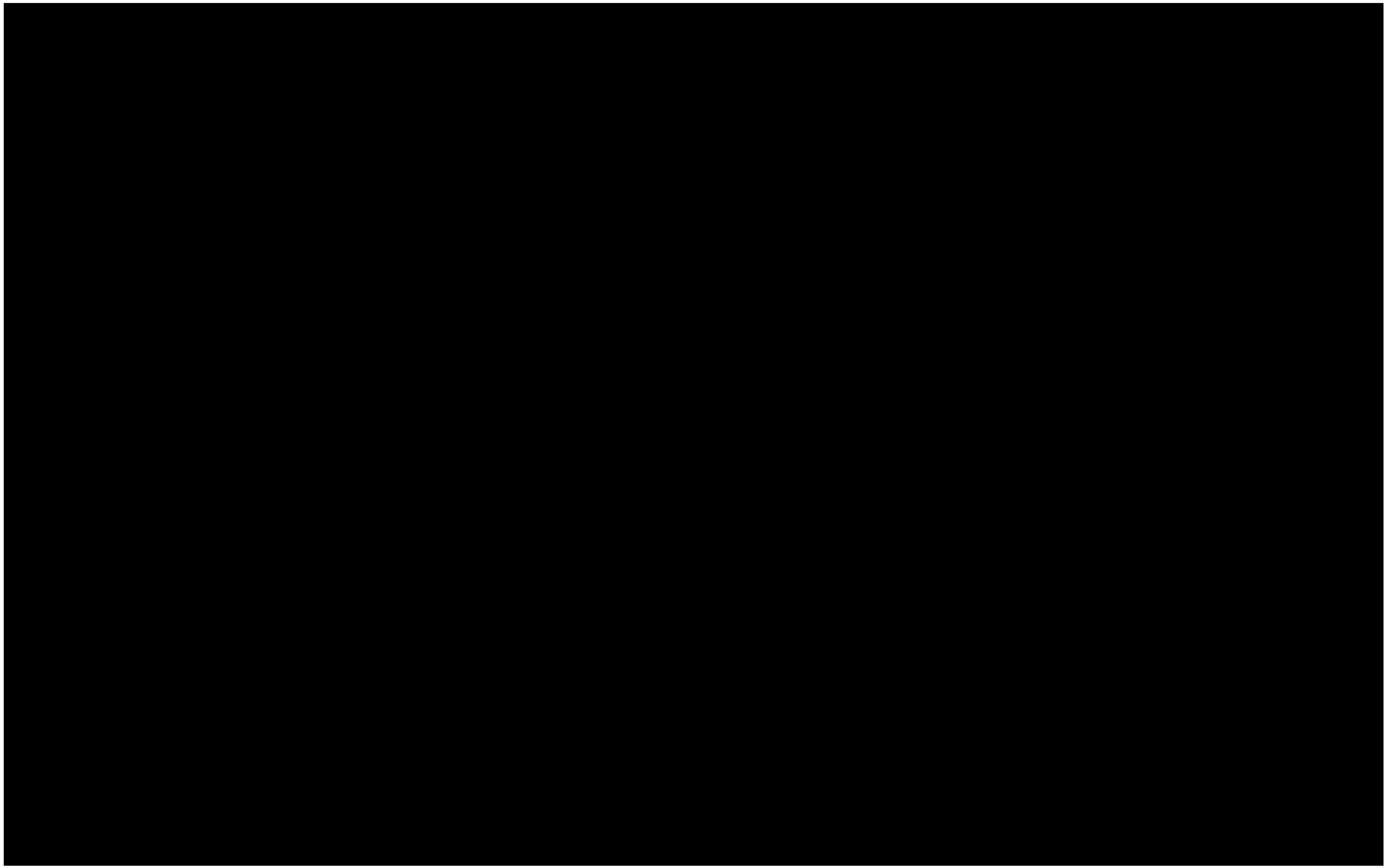
Limitations & Potential Improvements

1. Did not train for long enough due to limited resources

2. Cough detection very good, but cough counting not so accurate
   - Combined coughs
   - Varying duration coughs

3. Could add energy measurements and comparison

ETH zürich



Demo

ETH zürich

```
** New Inference **
1595904: Starts CNN: 40
1595904: Completes CNN: 40
CNN Time: 1850 us

Classification results:
[-0072400] -> Class 00       Up: 0.4%
[-0128208] -> Class 01     Down: 0.0%
[-0127296] -> Class 02     Left: 0.0%
[-0199856] -> Class 03    Right: 0.0%
[-0091440] -> Class 04     Stop: 0.2%
[-0115504] -> Class 05       Go: 0.1%
[-0198704] -> Class 06      Yes: 0.0%
[-0126752] -> Class 07       No: 0.0%
[-0167056] -> Class 08      One: 0.0%
[-0049280] -> Class 09      Two: 1.5%
[-0106832] -> Class 10    Three: 0.1%
[-0115744] -> Class 11     Four: 0.1%
[-0134048] -> Class 12     Five: 0.0%
[-0113936] -> Class 13      Six: 0.1%
[-0085728] -> Class 14    Seven: 0.4%
[-0146256] -> Class 15    Eight: 0.0%
[-0157616] -> Class 16     Nine: 0.0%
[-0141712] -> Class 17     Zero: 0.0%
[+0052320] -> Class 18    Cough: 96.2%
[-0056576] -> Class 19  Unknown: 0.8%
----------------------------------------
Detected word: Cough (96.2%)
Number of coughs today: 6
```

Demo

Thank you for Listening!

Q&A