

The poster features a blue background with a large circular graphic on the right containing the text "60 ICTP 1964-2024". The main title "Workshop on TinyML for Sustainable Development" is in white text on a dark blue rectangular background. Below it, event details are listed: "22 - 26 July 2024", "São Paulo, Brazil", and "Deadline: 6 May 2024". A "FURTHER INFORMATION" section includes an email address (smr3961@ictp.it), a web link (<https://indico.ictp.it/event/10499/>), and a QR code. A note encourages female scientists to apply. Logos for partner institutions are at the bottom: B, Harvard John A. Paulson School of Engineering and Applied Sciences, IBM, UNIFEI, and TINYML4D.

Temperature Dependence Psychoacoustics

Prof. Marcelo J. Rovai
rovai@unifei.edu.br

UNIFEI - Federal University of Itajuba, Brazil
TinyML4D Academic Network Co-Chair



Temperature Dependence Psychoacoustics

Simple TinyML Proof-of-concept



<https://www.hackster.io/mjrobot/listening-temperature-with-tinyml-7e1325>



Audio Engineering Society

Convention e-Brief 473

Presented at the 145th Convention
2018 October 17 – 20, New York, NY, USA

This Engineering Brief was selected on the basis of a submitted synopsis. The author is solely responsible for its presentation, and the AES takes no responsibility for the contents. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Audio Engineering Society.

Why can you hear a difference between pouring hot and cold water? An investigation of temperature dependence in psychoacoustics.

He Peng¹ and Joshua D. Reiss²

¹Tianjin University

²Queen Mary University of London

Correspondence should be addressed to He Peng, Joshua Reiss (hepeng2018@hotmail.com, joshua.reiss@qmul.ac.uk)

[http://www.eecs.qmul.ac.uk/~josh/
documents/2018/19737.pdf](http://www.eecs.qmul.ac.uk/~josh/documents/2018/19737.pdf)

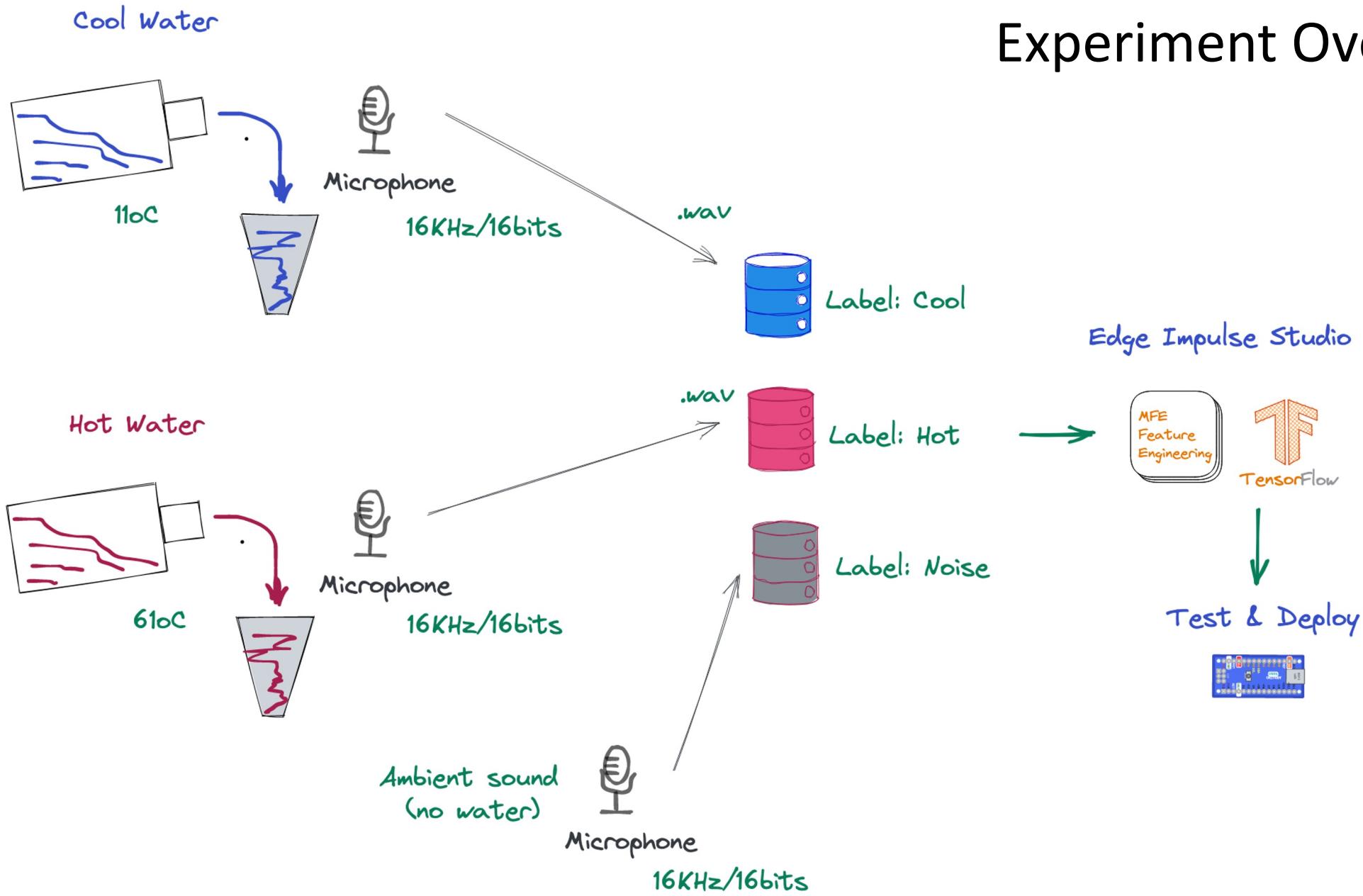


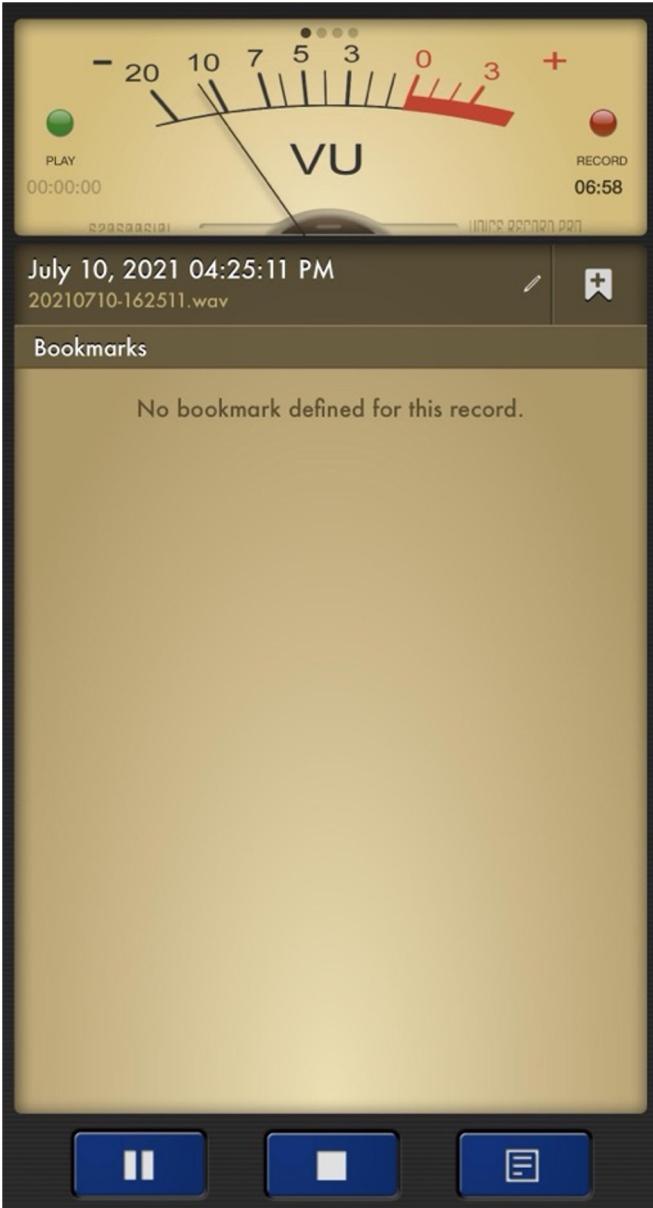
You Can Hear The Difference Between Hot and Cold W...

Tom Scott

[\(min: 0.17 => min 2:37\)](https://www.youtube.com/watch?v=Ri_4dDvcZeM)

Experiment Overview





Voice Recorder

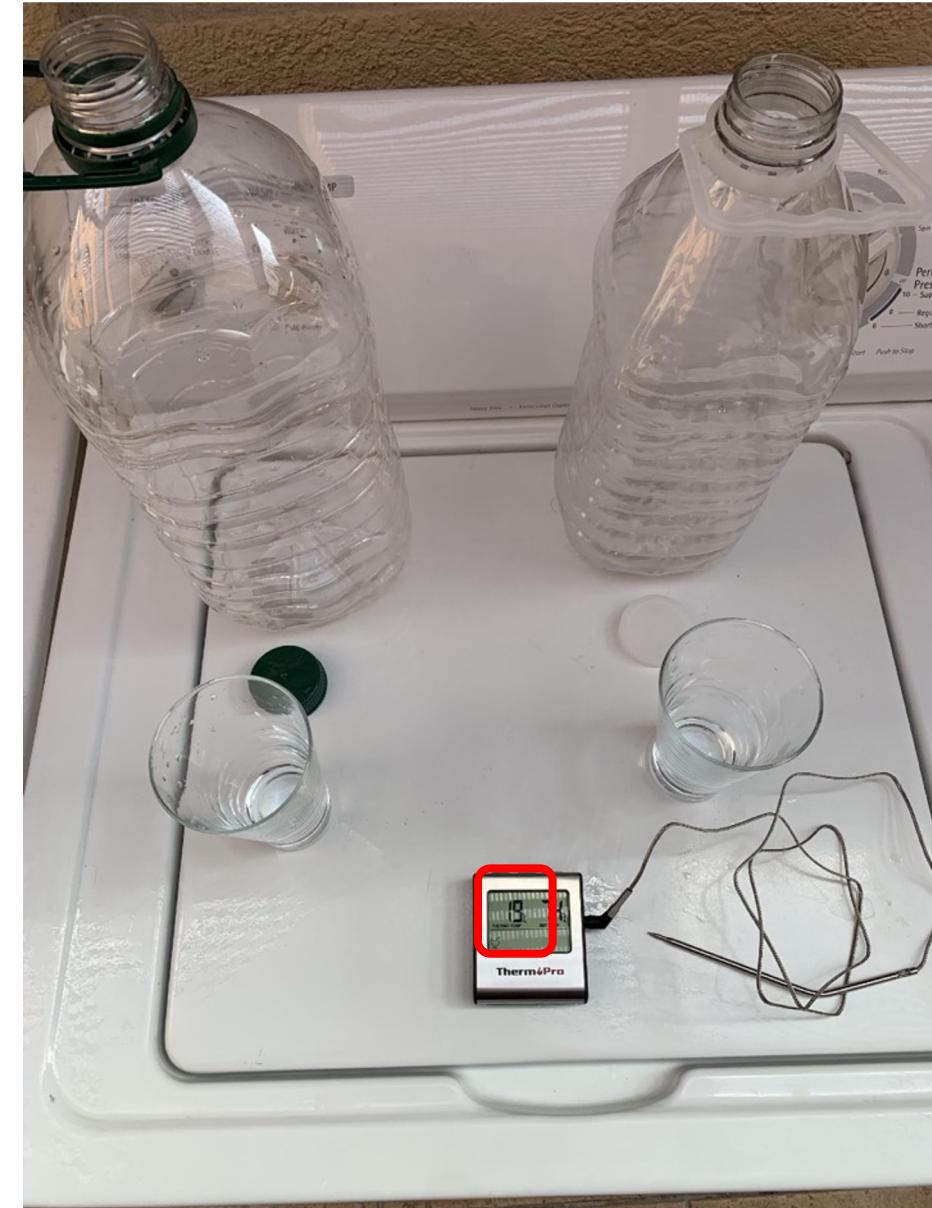


Sample Sound:

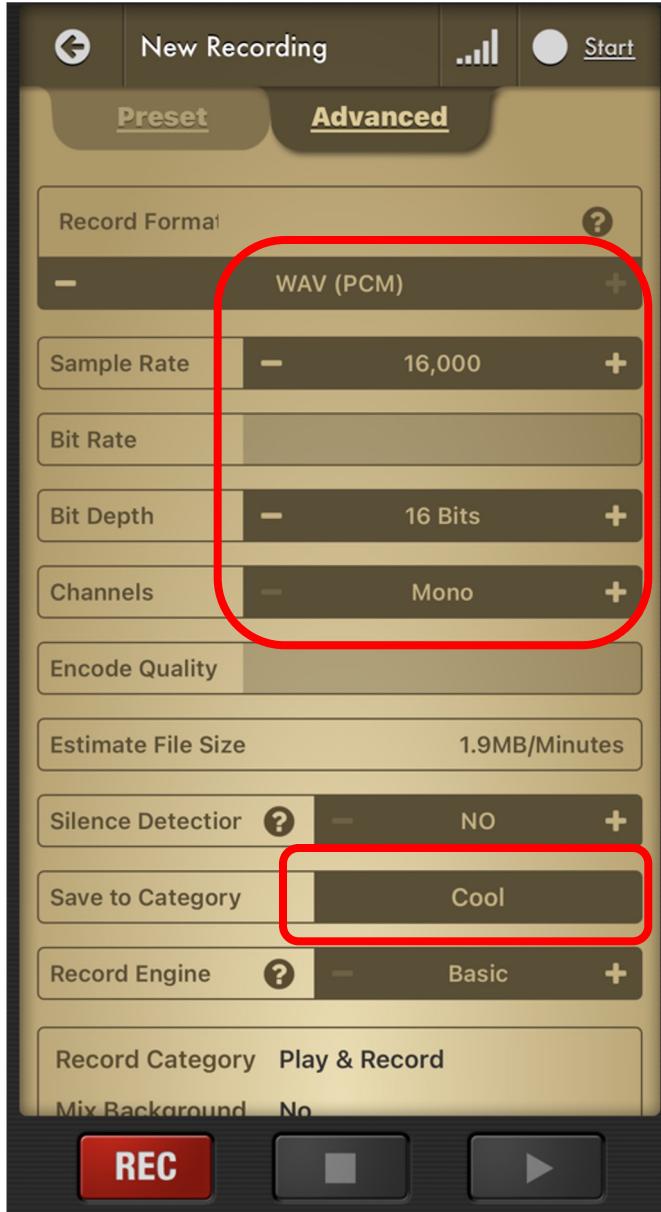
- 16KHz
- PCM – 16bits
- Mono

Classes:

- Hot
- Cool
- Noise



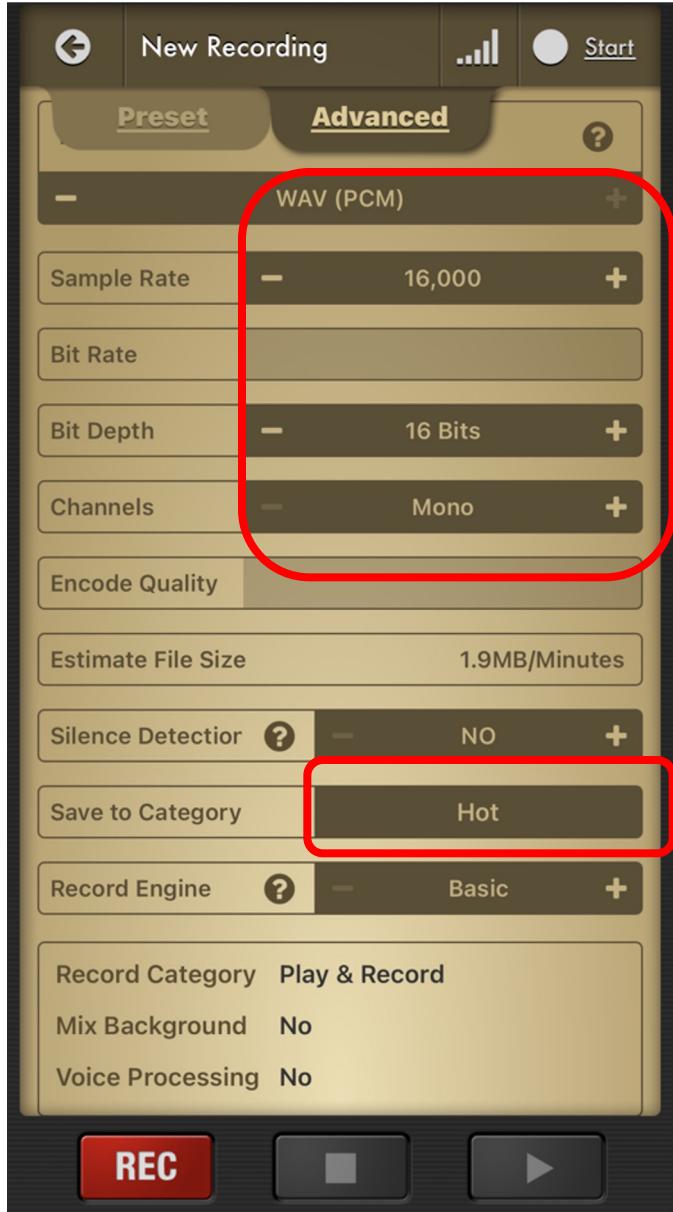
Ambient Temperature: 19°C



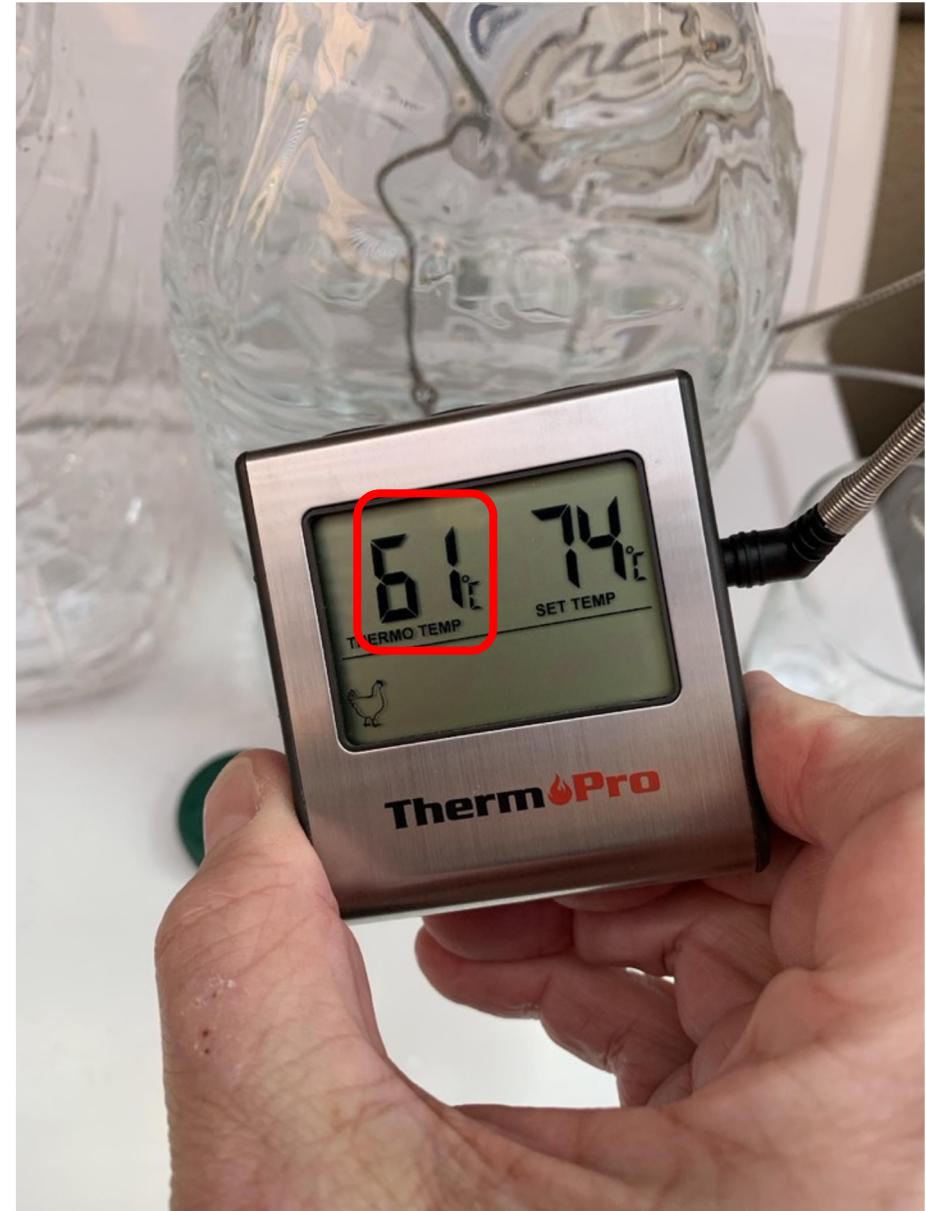
Class: Cool



Cool Water Temperature: 11°C



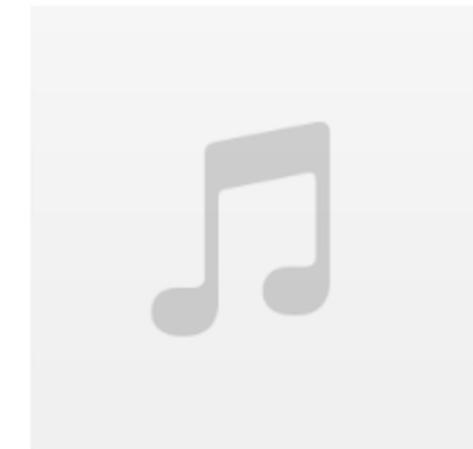
Class: Hot



Hot Water Temperature: 61°C

File tree view showing recorded audio files:

- data
- cool
 - hot
- 20210710-125854.wav
- 20210710-125930.wav
- 20210710-125956.wav
- 20210710-130010.wav
- 20210710-130041.wav
- 20210710-130100.wav
- 20210710-130119.wav
- 20210710-130129.wav
- 20210710-130142.wav
- 20210710-130200.wav
- 20210710-130212.wav
- 20210710-130224.wav
- 20210710-130236.wav
- 20210710-130246.wav
- 20210710-130256.wav
- 20210710-130304.wav
- 20210710-130315.wav
- 20210710-130329.wav
- 20210710-130348.wav
- 20210710-130358.wav
- 20210710-130408.wav
- 20210710-130416.wav



20210710-125854.wav

Waveform audio - 640 KB

Information	Show Less
Created	Today 13:58
Modified	Today 13:58
Duration	00:20
Audio channels	Mono
Sample rate	16 kHz
Bits per sample	16

Data captured using app Voice Recorder and uploaded to Computer



Upload existing data

You can upload existing data to your project in the [Data Acquisition Format](#) (CBOR, JSON, CSV), or as WAV, JPG or PNG files.

Select files

No file chosen

Upload into category

- Automatically split between training and testing ?
- Training
- Testing

Label

- Infer from filename ?
- Enter label:

hot

Upload output

Uploading 14 files...

```
[ 1/14] Uploading 20210710-130535.wav OK
[ 2/14] Uploading 20210710-130603.wav OK
[ 3/14] Uploading 20210710-130544.wav OK
[ 4/14] Uploading 20210710-130553.wav OK
[ 5/14] Uploading 20210710-130738.wav OK
[ 6/14] Uploading 20210710-130718.wav OK
[ 7/14] Uploading 20210710-130649.wav OK
[ 8/14] Uploading 20210710-130700.wav OK
[ 9/14] Uploading 20210710-130630.wav OK
[10/14] Uploading 20210710-130621.wav OK
[11/14] Uploading 20210710-130709.wav OK
[12/14] Uploading 20210710-130611.wav OK
[13/14] Uploading 20210710-130728.wav OK
[14/14] Uploading 20210710-130639.wav OK
```

Done. Files uploaded successful: 14. Files that failed to upload: 0.

Job completed

Raw Data uploaded to Edge Impulse Studio as .wav



Did you know? You can capture data from any device or development board, or upload your existing datasets - [Show options](#)



DATA COLLECTED
2m 49s



LABELS
2



Collected data



SAMPLE NAME	LABEL	ADDED	LENGTH	⋮
20210710-130621.wav.2a5e0...	hot	Today, 14:02:55	4s	⋮
20210710-130630.wav.2a5e0...	hot	Today, 14:02:54	3s	⋮
20210710-130700.wav.2a5e0...	hot	Today, 14:02:52	4s	⋮
20210710-130649.wav.2a5e0...	hot	Today, 14:02:52	5s	⋮
20210710-130718.wav.2a5e0...	hot	Today, 14:02:52	5s	⋮
20210710-130738.wav.2a5e0...	hot	Today, 14:02:51	5s	⋮
20210710-130553.wav.2a5e0...	hot	Today, 14:02:51	4s	⋮
20210710-130544.wav.2a5e0...	hot	Today, 14:02:51	4s	⋮
20210710-130603.wav.2a5e0...	hot	Today, 14:02:51	3s	⋮
20210710-130535.wav.2a5e0...	hot	Today, 14:02:48	5s	⋮
20210710-130416.wav.2a5dv...	cool	Today, 14:02:12	4s	⋮
20210710-130408.wav.2a5dv...	cool	Today, 14:02:11	4s	⋮

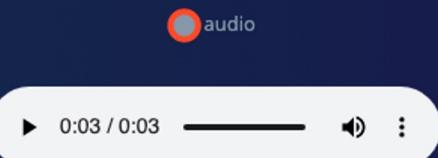
Record new data

[Connect using WebUSB](#)

No devices connected to the remote management API.

RAW DATA

20210710-130621.wav.2a5e0r33



Raw Data cleaned as split in 1 second samples

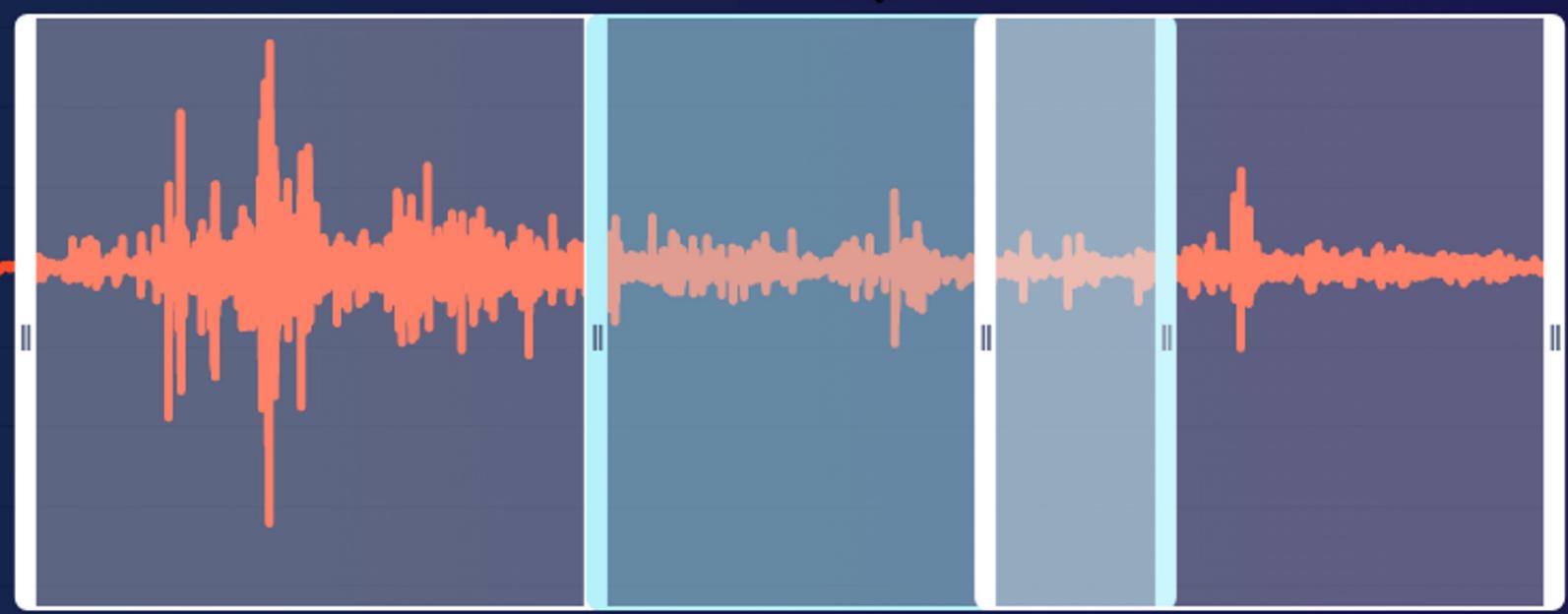
Split sample '20210710-130621.wav.2a5e0r33'

4

+ Add Segment

Set segment length (ms.): **Remove segment**

Apply



0

379

758

1138

1517

1896

2276

2655

3035

3414

A small circular icon with a blue center and an orange ring around it, representing audio content.

▶ 0:01 / 0:01 ━━ 🔍 ⋮

Cancel

Raw Data cleaned as split in 1 second samples Shift samples [?](#)

Split



⚡ An impulse takes raw data, uses signal processing to extract features, and then uses a learning block to classify new data.

Time series data

Axes
audio

Window size 1000 ms.

Window increase 500 ms.

Zero-pad data

Audio (MFE)

Name

Input axes audio

Neural Network (Keras)

Name

Input features MFE

Output features
3 (cool, hot, noise)

Output features

3 (cool, hot, noise)

Save Impulse

Audio (MFE)
Extracts a spectrogram from audio signals using **Mel-filterbank energy features**, great for non-voice audio.

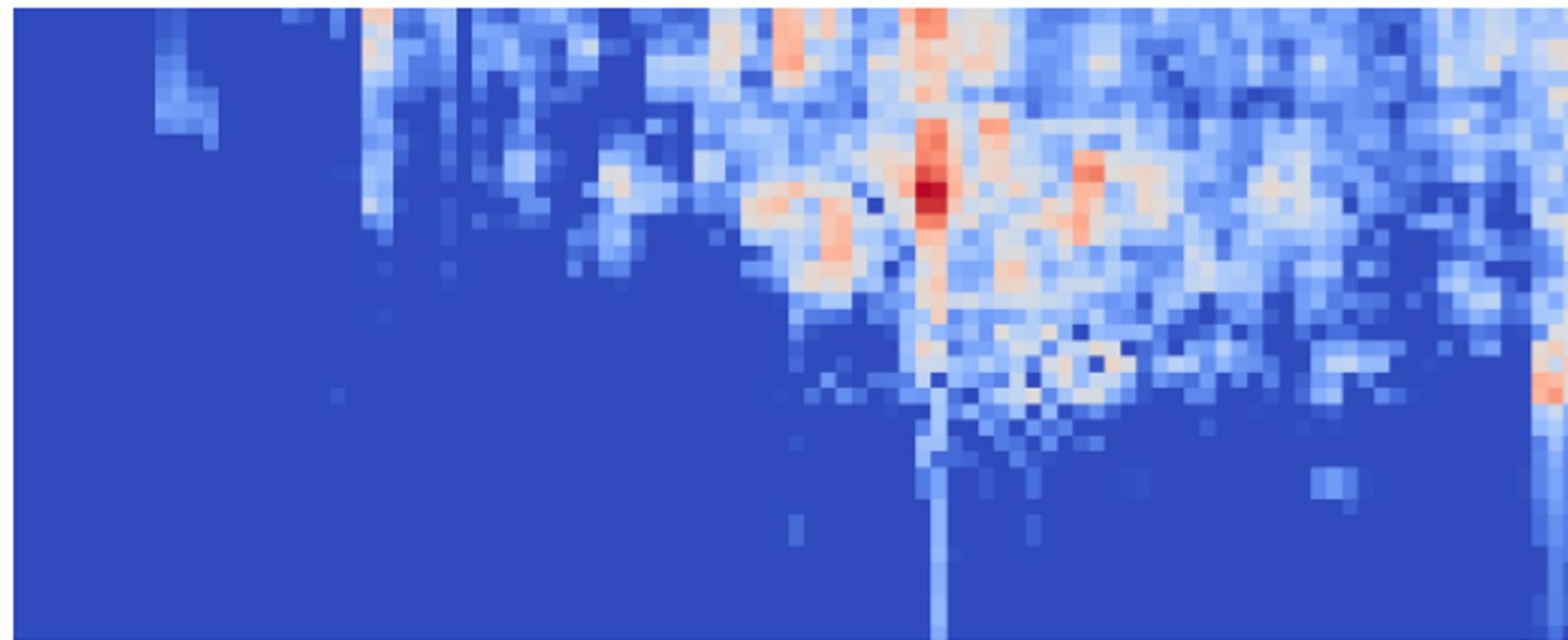
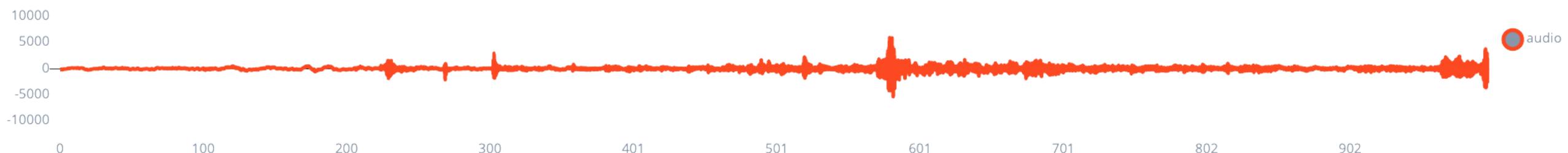


Add a processing block

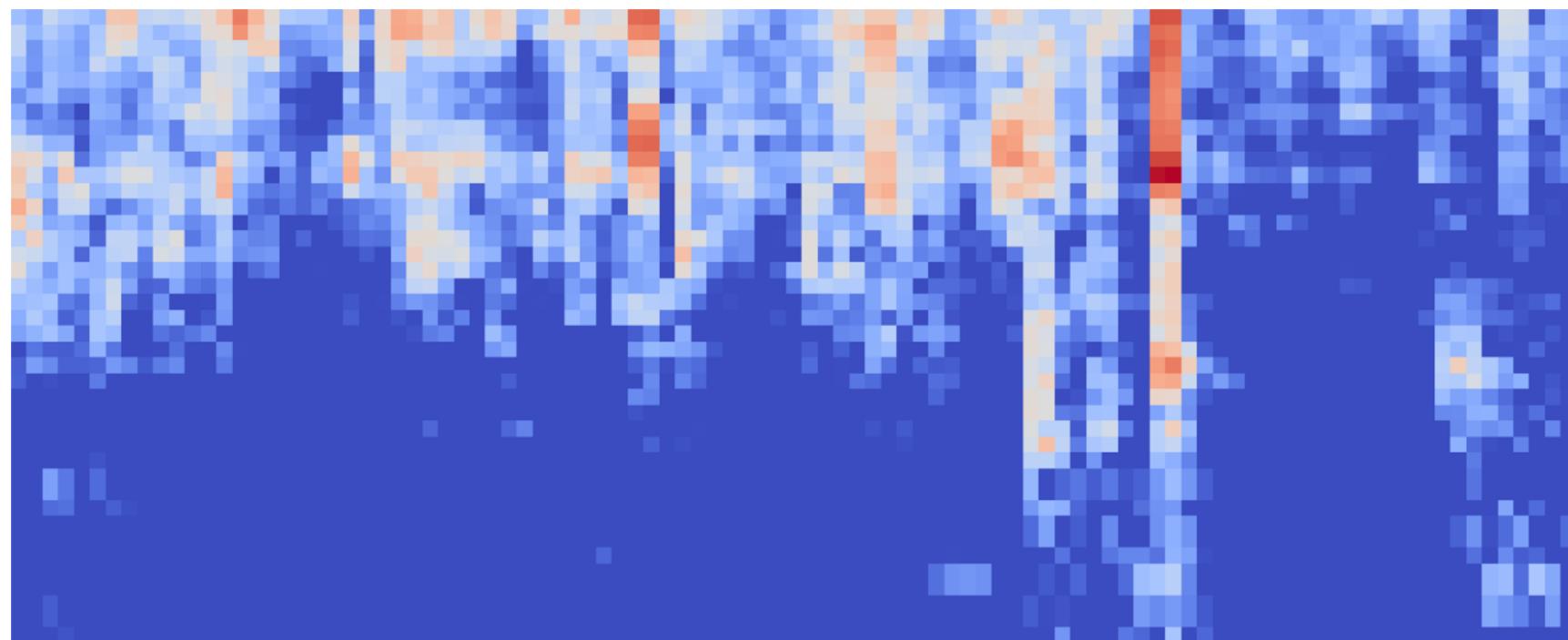
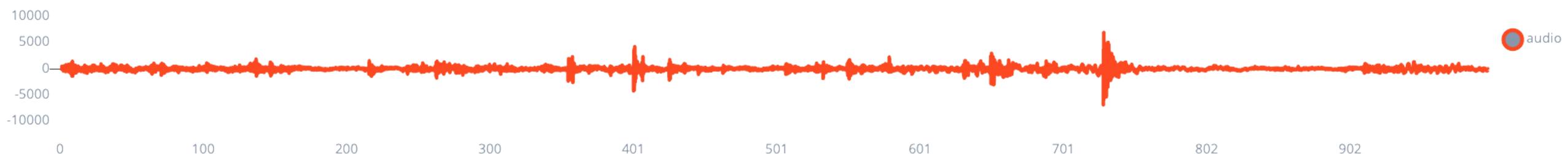


Add a learning block

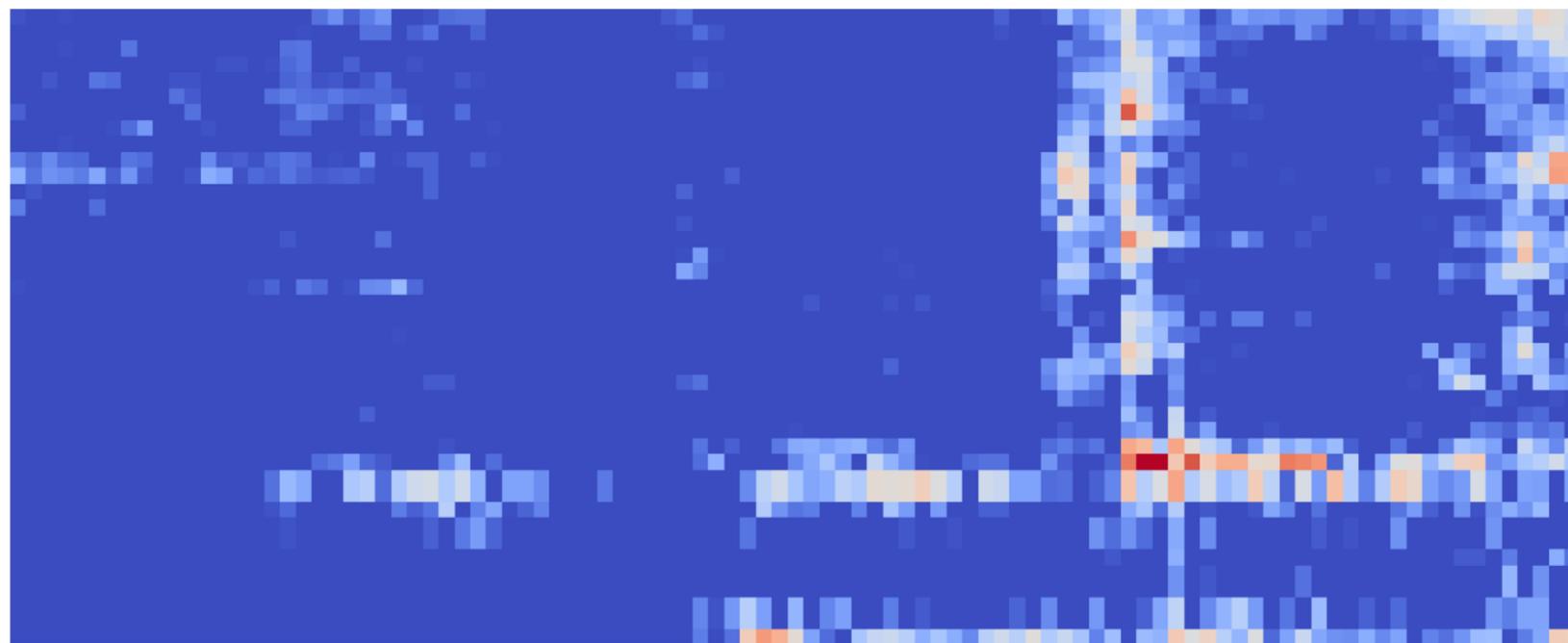
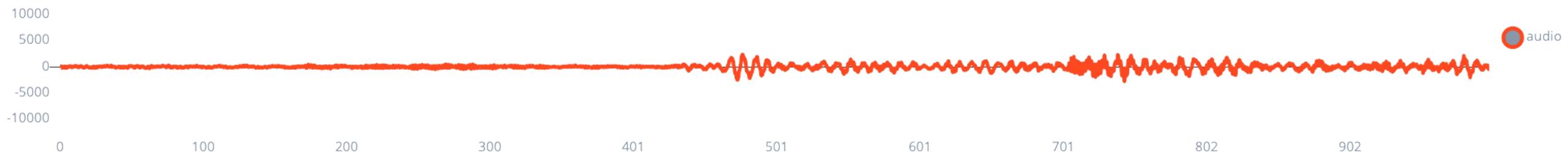
Cool Water 1 second sample



Hot Water 1 second sample



Noise 1 second sample



Parameters Generate features

Training set

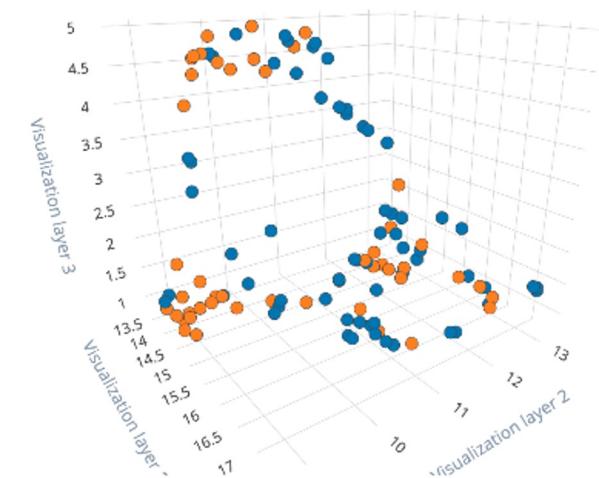
Data in training set	2m 35s
Classes	3 (cool, hot, noise)
Window length	1000 ms.
Window increase	500 ms.
Training windows	155

Generate featuresFeature explorer (155 samples) ?

X Axis Y Axis Z Axis

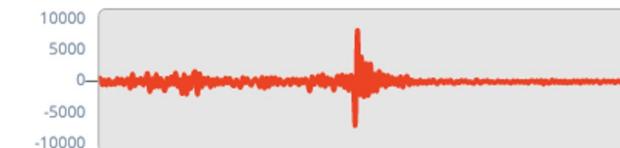
Visualization layer 1 Visualization layer 2 Visualization layer 3

- cool
- hot
- noise



20210710-125930.wav.2a5dv05j.s2

Label: cool

[View sample](#)[View features](#)▶ 0:00 / 0:01 ◀ ⋮On-device performance ?

PROCESSING TIME
250 ms.



PEAK RAM USAGE
25 KB

#1 ▾ Click to set a description for this version

Neural Network settings

Training settings

Number of training cycles ②

Learning rate ②

Minimum confidence rating ②

Audio training options

Data augmentation ②

Neural network architecture

Architecture presets ② [1D Convolutional \(Default\)](#) [2D Convolutional](#)

Input layer (3,960 features)

Reshape layer (40 columns)

1D conv / pool layer (8 neurons, 3 kernel size, 1 layer)

Dropout (rate 0.25)

1D conv / pool layer (16 neurons, 3 kernel size, 1 layer)

Dropout (rate 0.25)

Flatten layer

Add an extra layer

Output layer (3 features)

Training output

Model

Model version: ② [Quantized \(int8\)](#)

Last training performance (validation set)

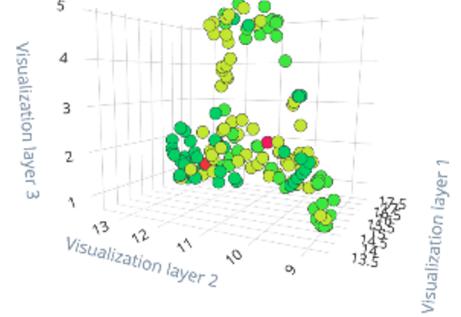
 ACCURACY	93.5%	 LOSS	0.13
---	--------------	---	-------------

Confusion matrix (validation set)

	COOL	HOT	NOISE
COOL	92.9%	7.1%	0%
HOT	16.7%	83.3%	0%
NOISE	0%	0%	100%
F1 SCORE	0.93	0.83	1.00

Feature explorer (full training set) ②

- cool - correct
- hot - correct
- noise - correct
- cool - incorrect
- hot - incorrect



Visualization layer 3

Visualization layer 2

Visualization layer 1

On-device performance ②

 INFERRING TIME	17 ms.
 PEAK RAM USAGE	10.9K
 FLASH USAGE	31.4K

This lists all test data. You can manage this data through [Data acquisition](#).

Test data

[Classify all](#)

Set the 'expected outcome' for each sample to the desired outcome to automatically score the impulse.

SAMPLE NAME	EXPECTED OUTCOME	LENGTH	ACCURACY	RESULT	...
20210710-130728.wav	hot	1s	100%	1 hot	...
20210710-130639.wav	hot	1s	100%	1 hot	...
20210710-130553.wav	hot	1s	100%	1 hot	...
20210710-130535.wav	hot	1s	100%	1 hot	...
20210710-130535.wav	hot	1s	100%	1 hot	...
20210710-130224.wav	cool	1s	0%	1 noise	...
20210710-130304.wav	cool	1s	100%	1 cool	...
20210710-130236.wav	cool	1s	100%	1 cool	...
20210710-130256.wav	cool	1s	100%	1 cool	...
20210710-130224.wav	cool	1s	100%	1 cool	...
20210710-130142.wav	cool	1s	0%	1 noise	...
20210710-130100.wav	noise	1s	100%	1 noise	...
20210710-130041.wav	noise	1s	100%	1 noise	...
20210710-125854.wav	noise	1s	100%	1 noise	...
20210710-125854.wav	noise	1s	100%	1 noise	...

Model testing output

Model testing results

ACCURACY

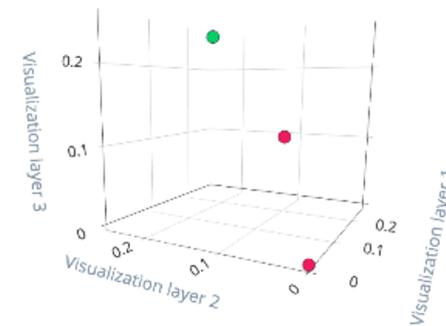
86.67%

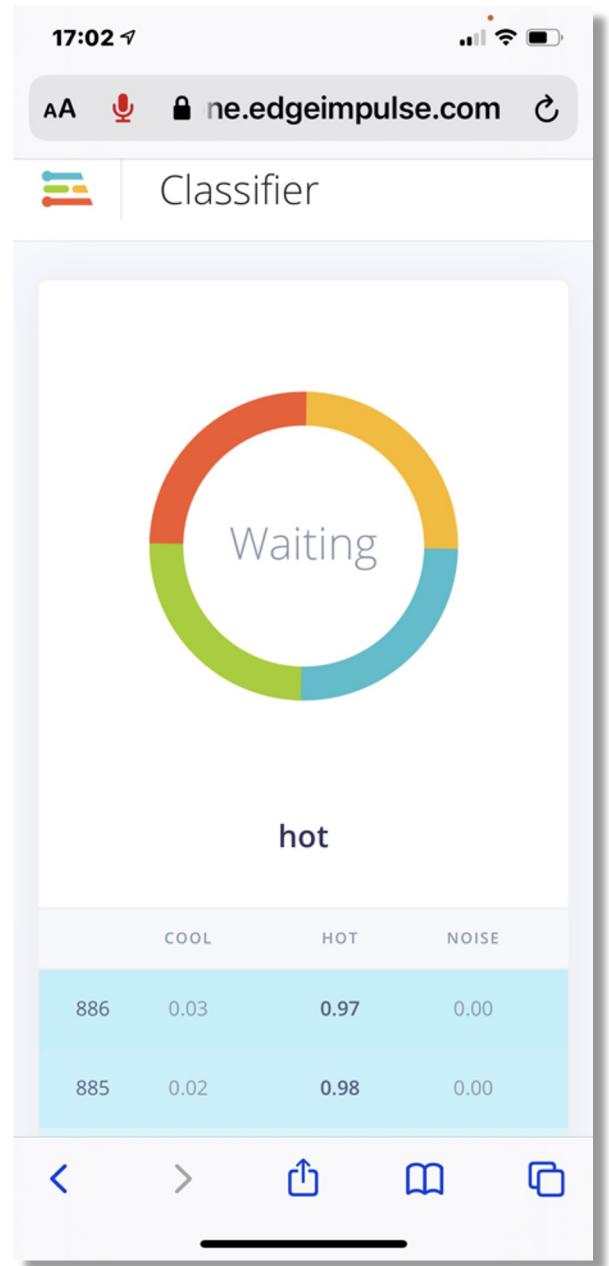
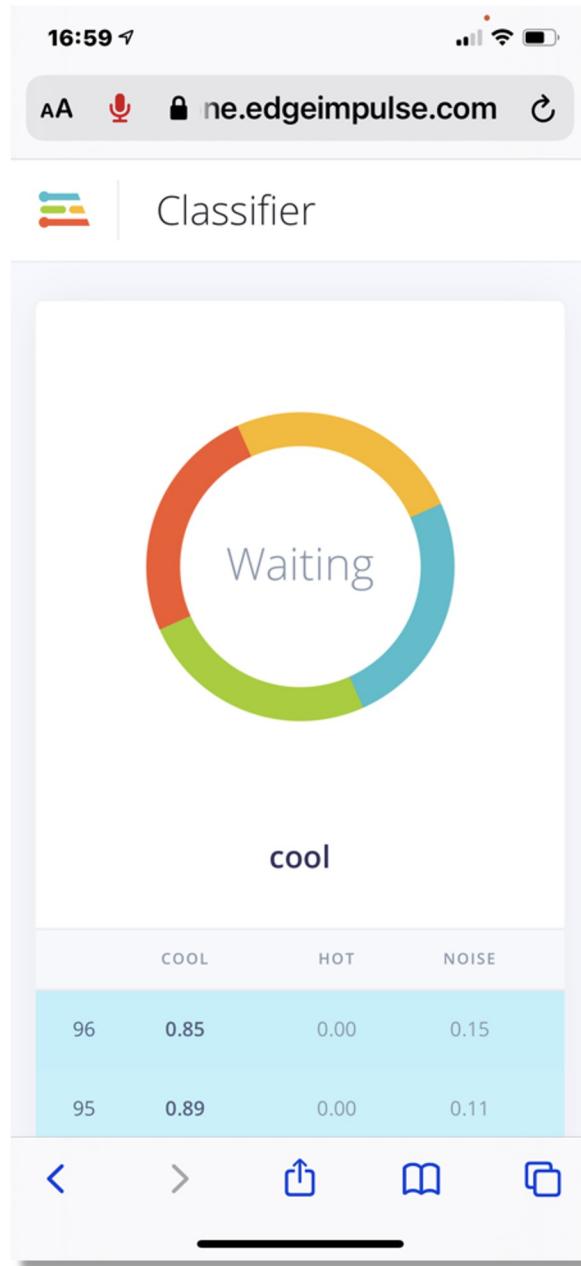
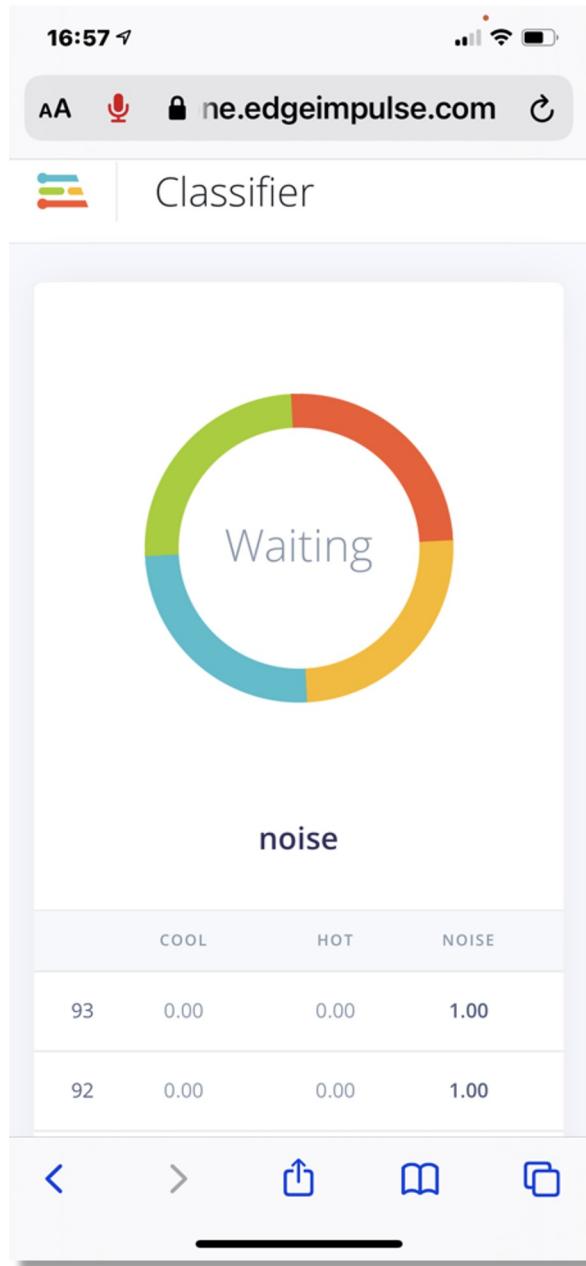


	COOL	HOT	NOISE	UNCERTAIN
COOL	66.7%	0%	33.3%	0%
HOT	0%	100%	0%	0%
NOISE	0%	0%	100%	0%

Feature explorer

- cool - correct
- hot - correct
- noise - correct
- cool - incorrect





Live Classifier (Off line) using iphone

Select optimizations (optional)

Model optimizations can increase on-device performance but may reduce accuracy. Click below to analyze optimizations and see the recommended choices for your target. Or, just click Build to use the currently selected options.



Enable EON™ Compiler

Same accuracy, up to 50% less memory. Open source.



Available optimizations for NN Classifier

Quantized (int8)	RAM USAGE	LATENCY	CONFUSION MATRIX			
	10.9K	17 ms	66.7	0	33.3	0
Currently selected	FLASH USAGE	ACCURACY	0	100	0	0
	31.4K	86.67%	0	0	100	0
Unoptimized (float32)	RAM USAGE	LATENCY	CONFUSION MATRIX			
	33.9K	78 ms	66.7	0	33.3	0
Click to select	FLASH USAGE	ACCURACY	0	100	0	0
	38.0K	86.67%	0	0	100	0

Estimate for Cortex-M4F 80MHz (ST IoT Discovery Kit)

Arduino File Edit Sketch Tools Help

/dev/cu.usbmodem144301

ICTP - PSYCOACOUSTICS TEMPERATURE Project

Inferencing settings:

- Interval: 0.06 ms.
- Frame size: 16000
- Sample length: 1000 ms.
- No. of classes: 3

Predictions (DSP: 126 ms., Classification: 21 ms., Anomaly: 0 ms.):

- PREDICTION: ==> noise with probability 1.00
- Predictions (DSP: 126 ms., Classification: 21 ms., Anomaly: 0 ms.):
- PREDICTION: ==> noise with probability 1.00
- Predictions (DSP: 126 ms., Classification: 20 ms., Anomaly: 0 ms.):
- PREDICTION: ==> noise with probability 1.00

Autoscroll Show timestamp Both NL & CR 115200 baud Clear output

Preview

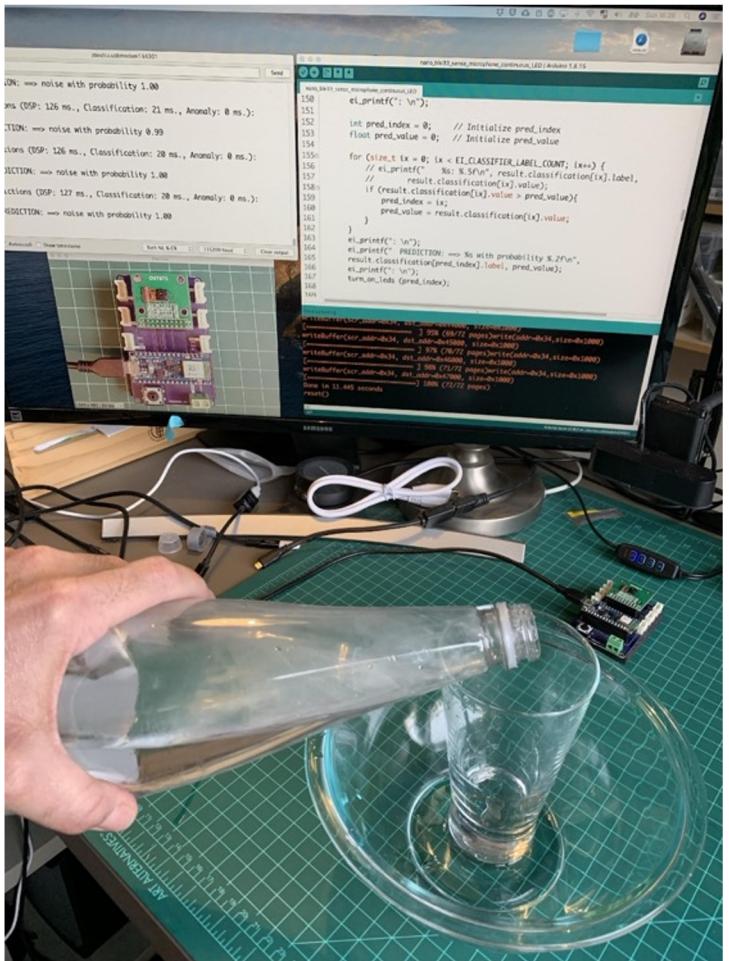
640 x 480 : 60 fps

88 */**
89 * @brief Special Postprocess function for RGB LEDs
90 */
91
92 void turn_off_leds(){
93 digitalWrite(LED_R, HIGH);
94 digitalWrite(LED_G, HIGH);
95 digitalWrite(LED_B, HIGH);
96 }
97
98 /*
99 * cool: [0] ==> Blue ON
100 * hot: [1] ==> Red ON
101 * noise: [2] ==> ALL OFF
102 */
103
104 void turn_on_leds(int pred_index) {
105 switch (pred_index)
106 {
107 case 0:
108 turn_off_leds();
109 digitalWrite(LED_B, LOW);
110 break;
111
112 case 1:
113 turn_off_leds();
114 digitalWrite(LED_R, LOW);
115 break;
116
117 case 2:
118 turn_off_leds();
119 break;
120 }
121 }
122 }

Done uploading.
writeBuffer(scr_addr=0x34, dst_addr=0x44000, size=0x1000)
[=====] 95% (69/72 pages) write(addr=0x34, size=0x1000)
writeBuffer(scr_addr=0x34, dst_addr=0x45000, size=0x1000)
[=====] 97% (70/72 pages) write(addr=0x34, size=0x1000)

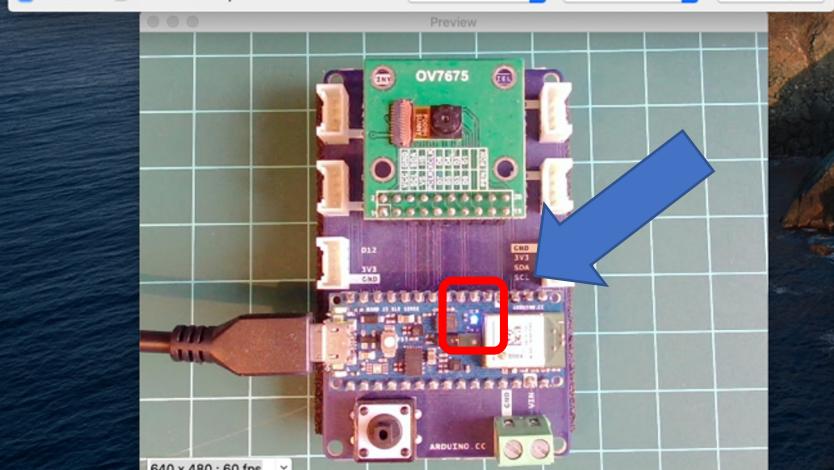
87

Arduino Nano 33 BLE on /dev/cu.usbmodem144301



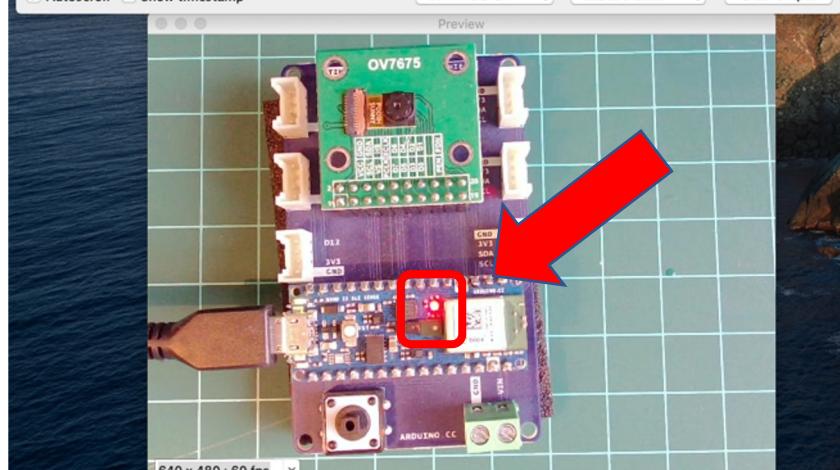
```
/dev/cu.usbmodem144301 Send

PREDICTION: ==> noise with probability 1.00
:
Predictions (DSP: 129 ms., Classification: 20 ms., Anomaly: 0 ms.):
:
PREDICTION: ==> noise with probability 1.00
:
Predictions (DSP: 129 ms., Classification: 21 ms., Anomaly: 0 ms.):
:
PREDICTION: ==> noise with probability 1.00
:
Predictions (DSP: 129 ms., Classification: 20 ms., Anomaly: 0 ms.):
:
PREDICTION: ==> cool with probability 0.96
```



```
/dev/cu.usbmodem144301
Send

PREDICTION: ==> noise with probability 0.99
:
Predictions (DSP: 129 ms., Classification: 21 ms., Anomaly: 0 ms.):
:
PREDICTION: ==> cool with probability 0.68
:
Predictions (DSP: 129 ms., Classification: 20 ms., Anomaly: 0 ms.):
:
PREDICTION: ==> cool with probability 1.00
:
Predictions (DSP: 129 ms., Classification: 20 ms., Anomaly: 0 ms.):
:
PREDICTION: ==> hot with probability 0.99
:
```



Thanks



TINYML4D