Appendix B. Example Data Sheet - Commercial Sensor

OVERVIEW

Person Sensor V1.0
SEN-21231

Compliance and Certifications

The person detection sensor complies with essential industry standards and regulations, including RoHS for environmental safety and GDPR for protecting individual privacy. The sensor has been audited by Kodelski Security for security and privacy implications.

Description

The Person Sensor is a small, low-cost hardware module that detects nearby peoples' faces, and returns information about how many there are, where they are relative to the device, and performs facial recognition. It is designed to be used as an input to a larger system, for example to wake up a kiosk display from sleep mode when somebody approaches, mute a microphone when nobody is present, or orient a fan so it's always pointing at the nearest person.

The sensor has a small form factor and utilizes a monochrome camera with a field of view of 640 x 480 (VGA). The input voltage for the sensor is 3.3V and the typical operating current for the sensor is 40 mA. The sensor communicates via I2C/Qwiic mode, conforming to SparkFun Qwiic electrical/mechanical specifications, and has a maximum cable length of 1 m at 400 kb/s. Longer cables can be used at lower data rates. The sensor has a maximum data rate of 400 kb/s.

Features

- Real-time person + head pose tracking with on-device ML
- Real-time person identification with on-devicce ML
- Low power consumption
- Onboard camera
- Small form factor: 22 x 20 x 10 mm
- I2C serial communication
- Lead-free

Use Cases

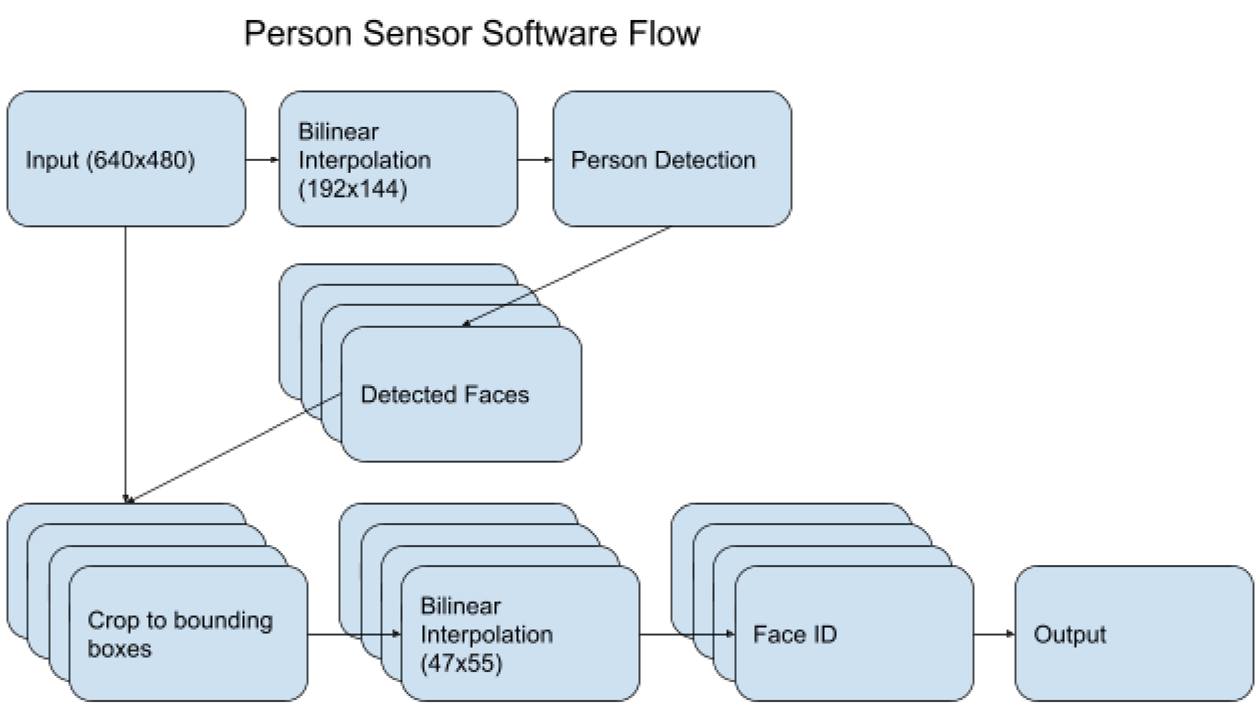
- Security
- Home automation
- Consumer appliances

MODEL CHARACTERISTICS

Software Flow Diagram

8-bit grayscale images (640x480) are resized to 192x144 and passed into a RetinaFace model trained to detect human faces. This model outputs a list of faces with coordinates for a bounding box around each face as well as five key facial landmarks. If identity is enabled, bounded faces are cropped out of the original image and rescaled to 47x55 and passed into a DeepID model to generate an embedding. This embedding is compared with saved facial IDs, and the nearest ID is returned along with the bounding box and information about whether the person is facing the sensor. Output information is communicated via Qwiic interface to the application processor.

	Person Detection Model	**Person ID Model**
Architecture	RetinaFace	DeepID
Framework	TFLite Micro	TFLite Micro
Validation Set	20%	-
Quantization	int8	int8
Inference Time	140 ms	125 ms
Peak RAM Usage	442.6 kB	189 kB
Flash Usage	449 kB	397 kB



Dataset Nutrition Label

The data nutrition label is publicly available here, with some important features outlined below.

At a Glance

About humans	Upstream sources	Technical review	Ethical review	Update frequency
Yes	**No**	**Unsure**	**No**	**No**

Intended Use

- **Intended Domain.** Face Detection and Landmark Detection
- **Intended Use.** Face Detection and Landmark Detection

Restrictions on Use

- no

Known Uses

Do Not Use

- **Domain.** Military

General risks

Any additional risks?

Individual Information

no

Consent

Yes,

Generalized Inferences

Most face image sources and existing datasets over-represent people in developed countries. Since this dataset contains images available on the internet, it probably suffers a similar bias.

Generalized Inferences - Mitigation

Sensitive Content

Not Applicable

Documented Known Issues

Other Known Issues

Number of issues

Risky	0
Safe	3
Unknown	4

IoT Security and Privacy Label

Security & Privacy Overview

Useful Sensors

Person Sensor V1.0
Firmware version: Not disclosed **- updated on:** 2023-05-03
The device was manufactured in: China

Security Mechanisms

Security updates	No security updates
Access control	Not disclosed

Data Practices

Sensor data collection	Visual	Audio	Physiological	Location
Sensor type	Camera			
Purpose	Providing and improving device functions			
Data stored on the device	No device storage			
Data stored in the cloud	No cloud storage			
Data shared with	Not shared			
Data sold to	Not sold			
Other collected data				

Privacy policy	Not disclosed

More Information

Detailed Security & Privacy Label:
Not disclosed

CMU IoT Security and Privacy Label CISPL 1.0 iotsecurityprivacy.org

Security & Privacy Details

Useful Sensors

Person Sensor V1.0
Firmware version: Not disclosed **- updated on:** 2023-05-03
The device was manufactured in: China

Security Mechanisms

Security updates	No security updates *Sensor is a standalone unit. For security and privacy, the firmware cannot be changed and only sensor outputs are available.*
Access control	Not disclosed
Security oversight	Audits performed by third-party security auditors *Third party security audit performed by Kudelski Security*
Ports and protocols	Not disclosed
Hardware safety	Not disclosed
Software safety	Not disclosed
Personal safety	Not disclosed
Vulnerability disclosure and management	Not disclosed
Software and hardware composition list	Not disclosed
Encryption and key management	Not disclosed

Data Practices

Sensor data collection	**Visual**
Sensor type	Camera
Data collection frequency	Continuous
Purpose	Providing and improving device functions
Data stored on the device	No device storage
Local data retention time	No retention
Data stored in the cloud	No cloud storage
Cloud data retention time	No retention
Data shared with	Not shared
Data sharing frequency	Not shared
Data sold to	Not sold
Other collected data	None

Data linkage	Not disclosed
What will be Inferred from User's Data	Not disclosed
Special data handling practices for children	Not disclosed
In Compliance with	Not disclosed
Privacy policy	Not disclosed

More Information

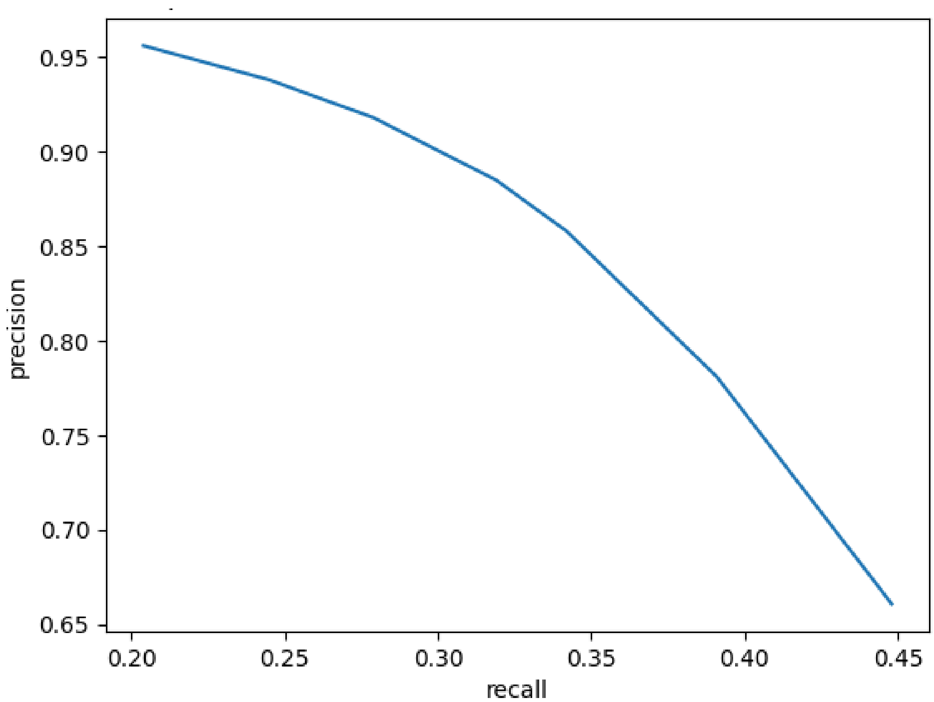
Call Useful Sensors with your questions at	1 805 813 7571
Email Useful Sensors with your questions at	contact@usefulsensors.com
Functionality when offline	Full functionality on offline mode
Functionality with no data processing	Not applicable
Physical actuations and triggers	Not disclosed
Compatible platforms	Not disclosed

CMU IoT Security and Privacy Label CISPL 1.0 iotsecurityprivacy.org

Machine Learning Model Specification

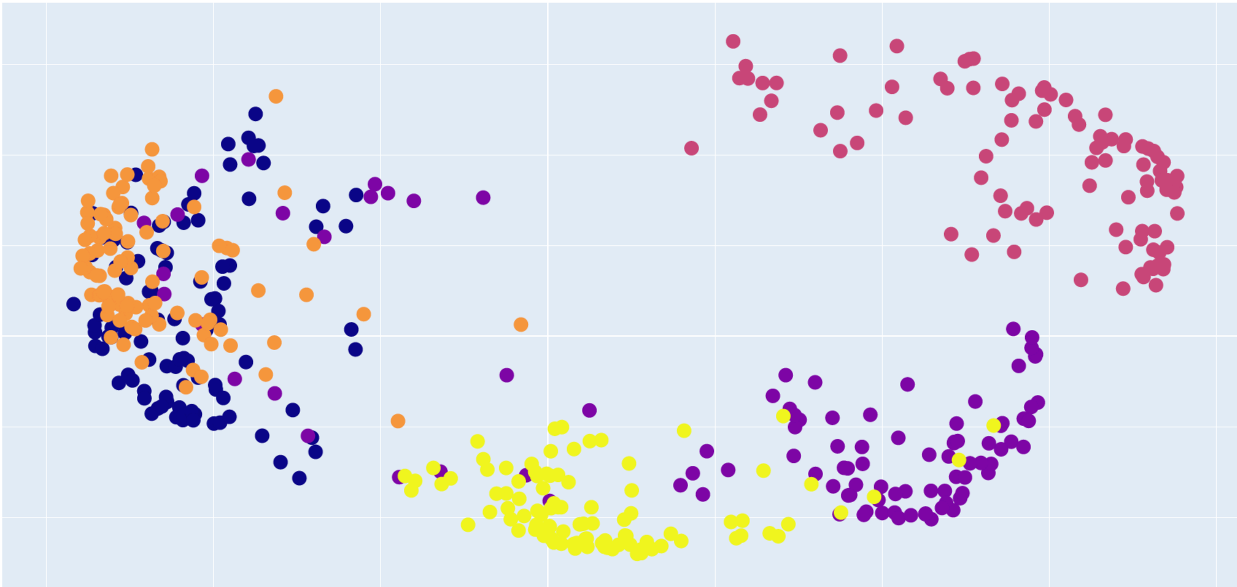
Person Detection Model

The person detection model was trained on a proprietary dataset of ~30,000 images with 300k labeled faces and five facial landmarks per face. The model input is a 192x144 raw image in 8-bit grayscale format, equivalent to 27,648 features. The training process was performed until the model accuracy ceased to improve. Final model performance achieved a precision of 91.8% on the test set, using a threshold of 0.7. The precision-recall curve of the model on the test set is shown below. The model was quantized to 8-bit integer using post training quantization using the Tensorflow Lite converter and is deployed using the Tensorflow Lite Micro runtime.



Face Identification Model

The face identification model was fine-tuned on a proprietary dataset encompassing ~4000 images across 5 identities captured using the sensor camera module. The input to the model is a 47x55 raw image in 8-big grayscale format, equivalent to 2,585 features. To produce the best separation between faces a dense classification layer was added to the model, and several iterations of freezing either the classification layer or the model was used to achieve a higher accuracy on the fine-tuning dataset. Finally, the classification layer was removed and embedding separation was evaluated using Principal Component Analysis (PCA) in three dimensions.



Each color represents one of five unique identities in the validation dataset. Distances between points indicate approximate distances between embeddings simplified to 2-D space.

The model was quantized to 8-bit integer using post training quantization using the Tensorflow Lite converter and is deployed using the Tensorflow Lite Micro runtime. On the sensor, the embedding generated by the Face ID model is compared against registered faces, and if a face with similar enough features is found, that identity is used. Otherwise, an identity of -1 is returned to indicate that no registered identity was found.

Performance Analysis

The end-to-end performance of the person detection sensor model was tested through an experimental study conducted in the Science and Engineering Complex (SEC) at Harvard University. The study involved 40 participants and evaluated the accuracy of the model under different lighting conditions using three identical sensors.

The study room measured 25 x 31 x 10 ft and contained 32 ceiling lights that were uniformly distributed in a 4 x 8 grid. The lighting conditions were captured quantitatively for each participant using a Lux LCD Illuminance Meter (Precision Vision, Inc.) and a C-800-U Spectrometer (Sekonic Corporation).

The sensors were mounted on a wooden board affixed to the wall at a height of 1.5 m above the ground. The participants were evaluated at three different distances (1.5 m, 4.5 m, and 7.5 m) from the sensors under each lighting condition. The ambient lighting in the room was provided by artificial lights, and blackout curtains were used to block the ambient lighting from outside.

The lighting levels were controlled using a dimmer switch that had three levels of operation, corresponding to 208±31, 584±51, and 1149±59 lux, respectively. When the lights were turned off, the illuminance meter gave a reading of zero lux. When all the lights were turned on at full strength, the sensor gave an average reading of 1149 lux. The color temperature of the lighting was measured to be 5600 K, corresponding to white light. Colored tape was placed on the ground to demarcate the locations where participants should stand during the experiment (i.e., 1.5, 4.5, and 7.5 m from the sensor array).

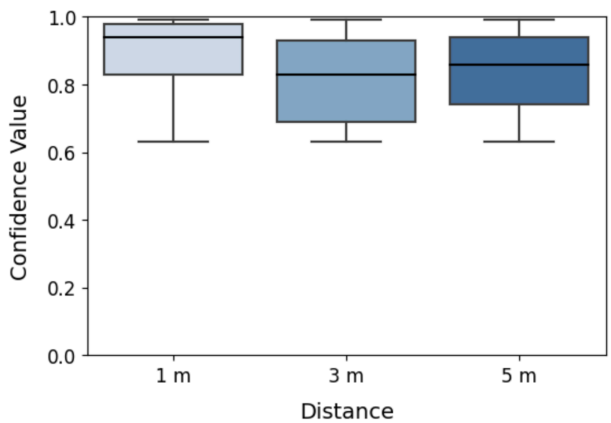
Before entering the study environment, the participants were asked to provide their gender identity and evaluate their skin tone according to the Monk Skin Tone (MST) Scale to evaluate algorithmic bias. The study evaluated algorithmic bias by bucketing skin tone into three categories: light (MST 0-4), medium (MST 5-7), and dark (MST 8-10). At each location and lighting condition, ten readings were taken from each sensor and averaged.

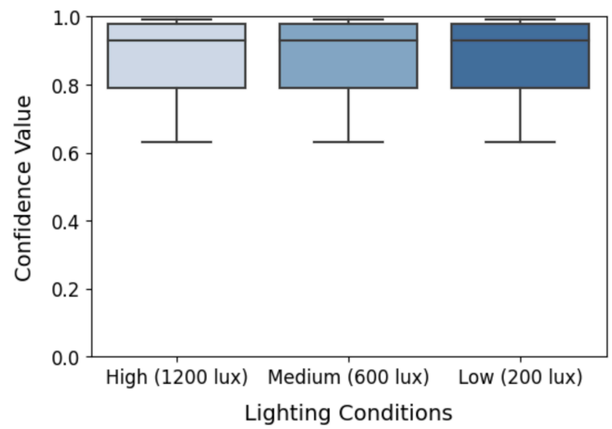
Participants were recruited using flyers, and those interested filled out a Study Interest Form. Upon arrival, participants signed a Consent Form indicating their willingness to participate in the study. The accuracy of the model is provided in the following graphs as a function of lighting condition, distance, gender identity, and skin tone. Overall, 63.2% of the participants were male, and 36.8% were female; the percentage of participants corresponding to each skin tone group was: 47.4% light, 39.4% medium, and 13.2% dark.

This anonymous study was approved by the Institutional Review Board of Harvard University on 6 April 2023 (Project Code: IRB23-0136).

Environmental Sensitivity

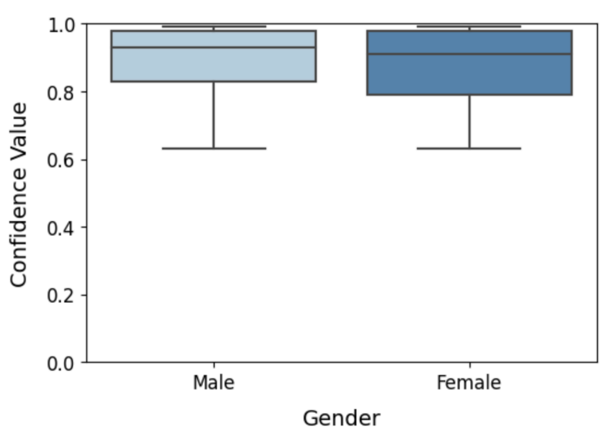
The device showed no decrease in performance under decreased lighting conditions. A moderate drop off in performance of around 10% is observed at distances 3-5 meters from the sensor.

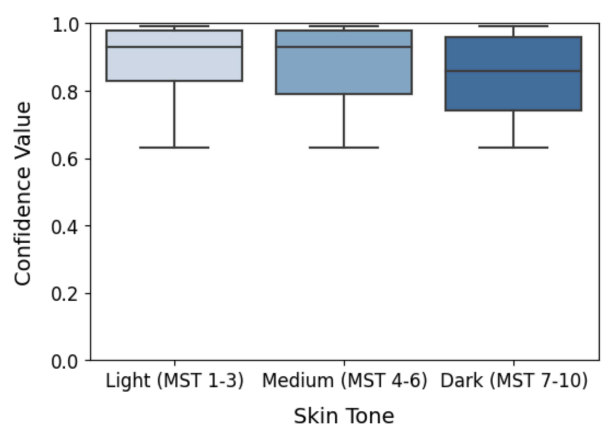




Demographic biases

A small gender bias is observed in model performance. A moderate skin tone bias was observed, showing approximately a 10% decrease in the confidence value for individuals with a darker skin tone.





HARDWARE

Hardware Details

Camera Specifications (see here)	
Field of view (horizontal)	110º
Color Filter Array	Bayer, Monochrome
Frame Rate	60FPS @ 48MHz
Pixel Array (Active/ Effective)	644 x 484 / 640x480
Electrical Specifications	
Operating Voltage Range (regulator enabled)	3.1V to 3.5V
Operating Current	40 mA
Operating Temperature	-20 ºC to 85 ºC
Communication Specifications	
I2C/Qwiic mode	Conforms with SparkFun Qwiic electrical/mechanical specifications. https://www.sparkfun.com/qwiic
Max cable length	1 m
Max data rate	100 kb/s
Module Orientation	Text on sensor is upright, up arrow points upwards
GPIO mode	INT pin is high when person is detected
Diagnostic LED	Default behavior of green LED on board: illuminates when person detected.
Data Transfer and Format	See I2C Protocol Table
I2C Address	0x63

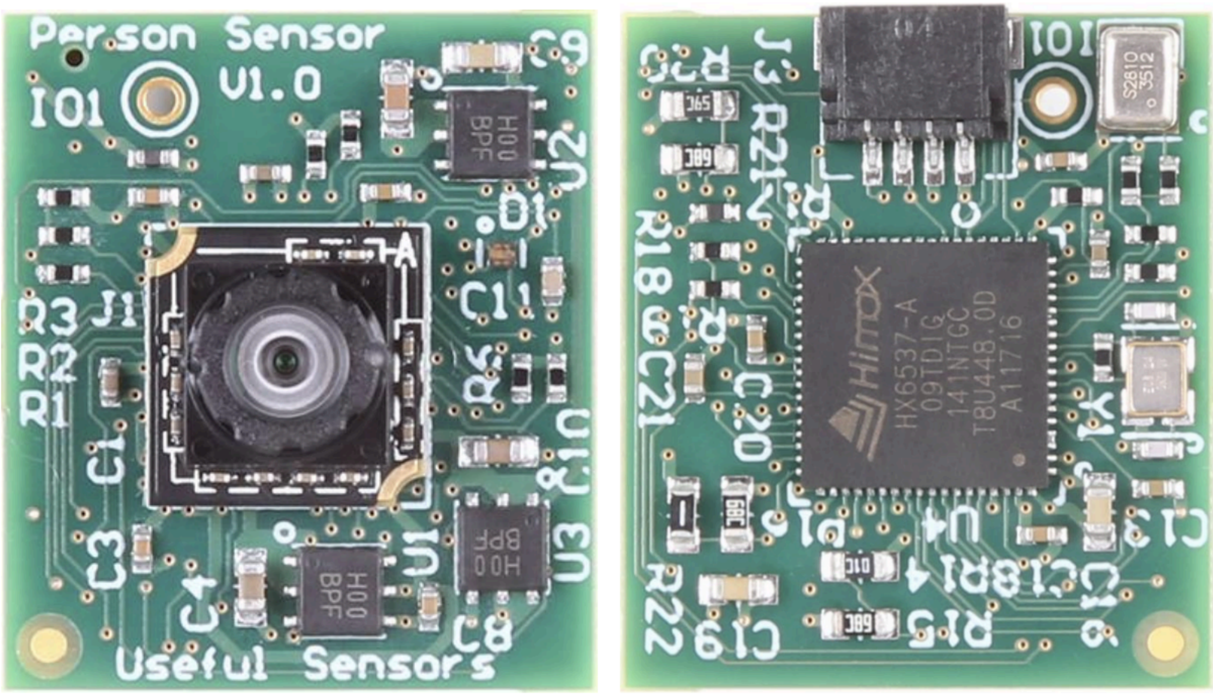
I2C Protocol

Address	Name	Default	Description
0x01	Mode	0x01 (continuous)	Mode. See mode table below.
0x02	Enable ID	0x00 (False)	Enable / Disable the ID model. With this flag set to False, only capture bounding boxes.
0x03	Single shot	0x00	Trigger a single-shot inference. Only works if the sensor is in standby mode.
0x04	Label next ID	0x00	Calibrate the next identified frame as person N, from 0 to 7. If two frames pass with no person, this label is discarded.
0x05	Persist IDs	0x01 (True)	Store any recognized IDs even when unpowered.
0x06	Erase IDs	0x0	Wipe any recognized IDs from storage.
0x07	Debug Mode	0x01 (True)	Whether to enable the LED indicator on the sensor.

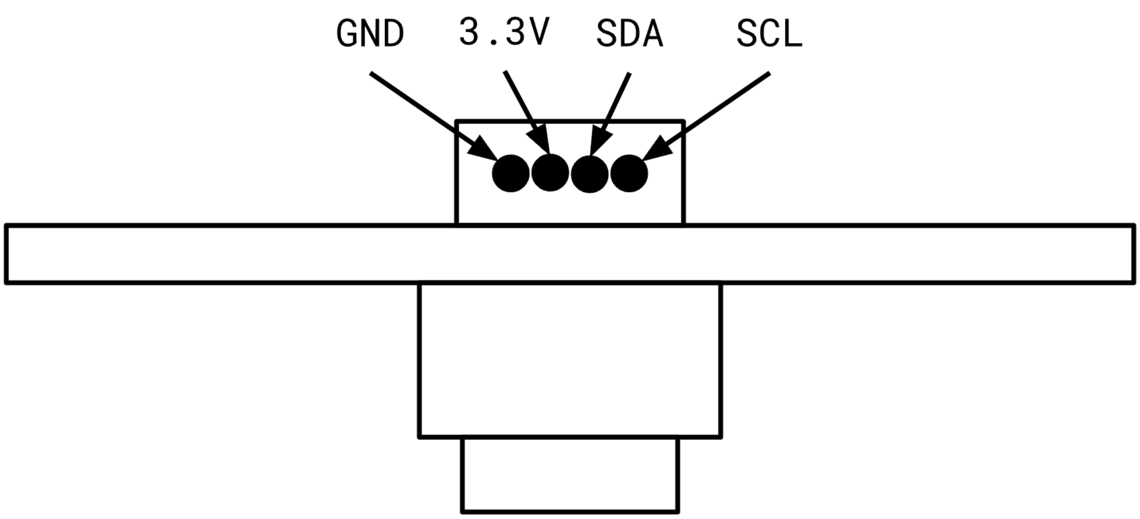
Mode	Name	Description
0x00	Standby	Lowest power mode, sensor is in standby and not capturing.
0x01	Continuous	Capture continuously, setting the GPIO trigger pin to high if a face is detected.

Device Diagrams

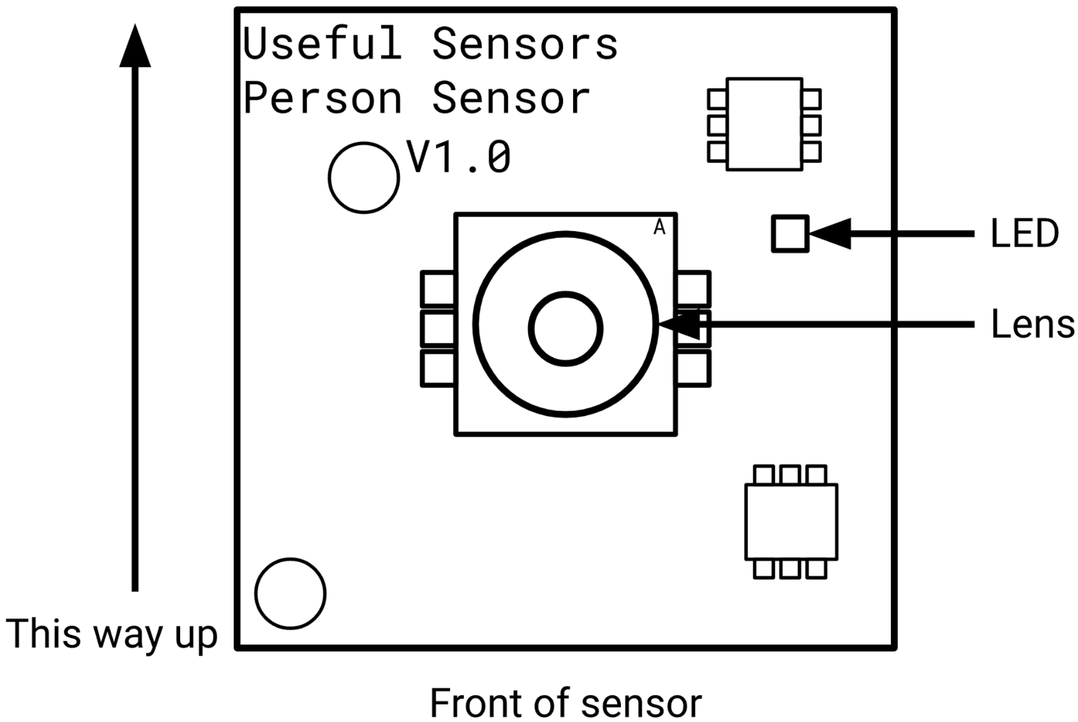
Front and backside of the sensor module.



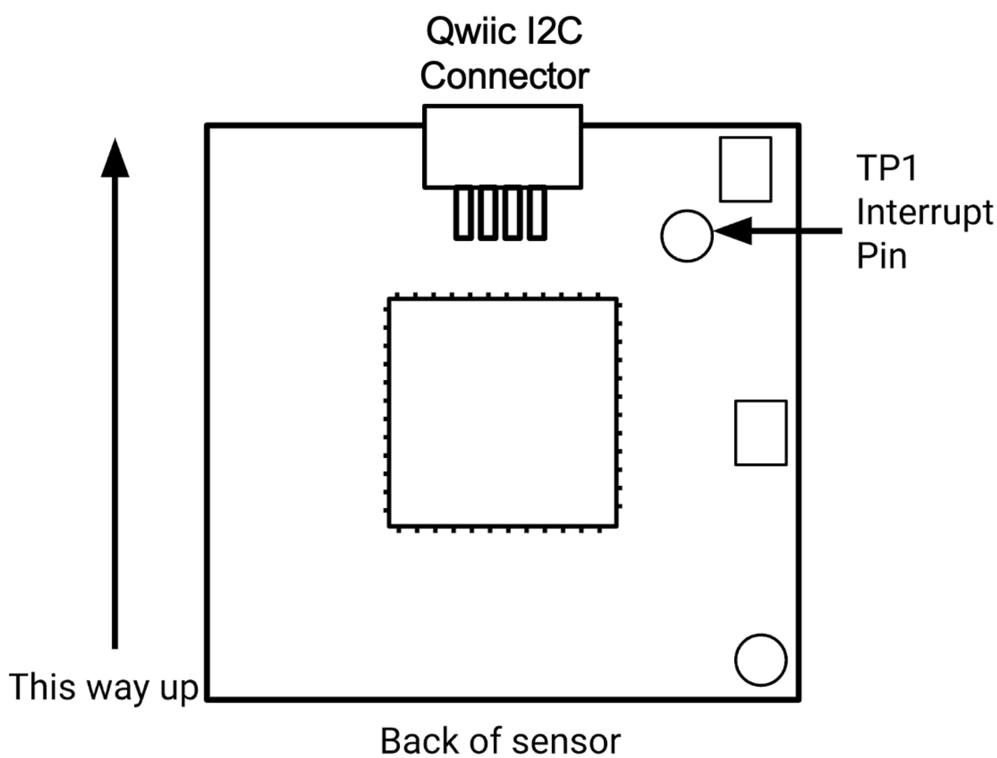
Qwiic connector interface.



Frontside schematic of sensor.



Backside schematic of sensor.



Bill of Materials

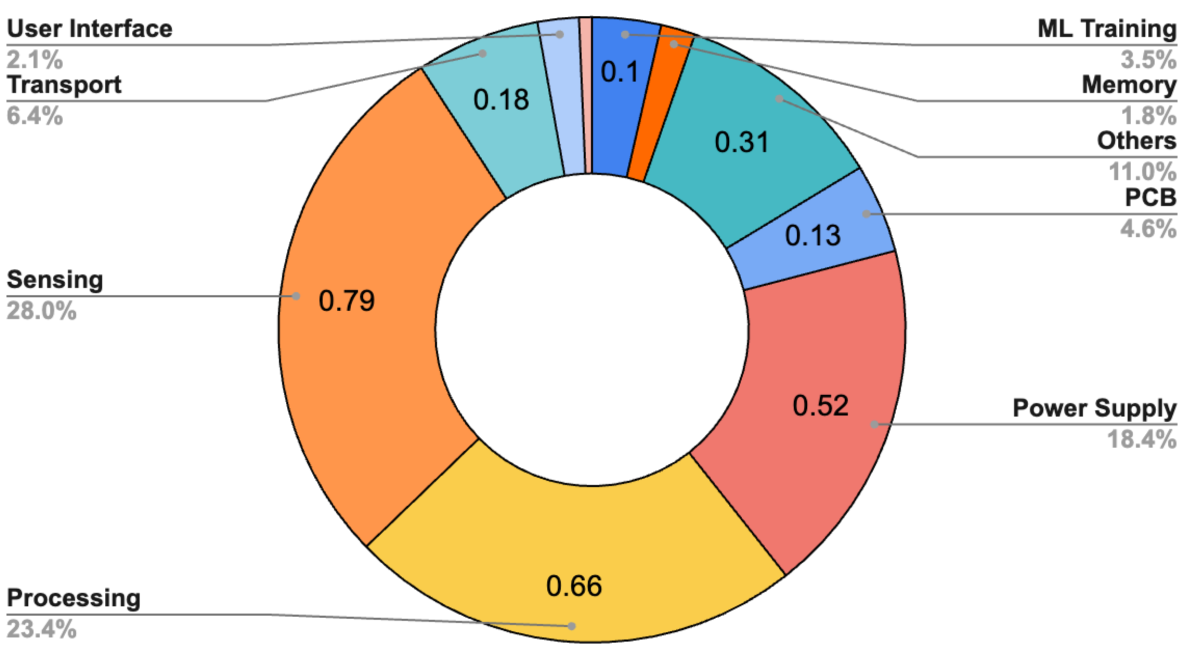
The following is a comprehensive list of materials required to assemble the Person Sensor V1.0. All unit cost values quoted in minimum order quantity of one.

Category In TinyML Calculator	Component	Unit Cost ($)	Quantity	Manufacturer	Link to Datasheet (if available)
Functional Components					
✔	Himax MCU HX6537/39/40-A	14.50	1	HiMax	https://cdn.sparkfun.com/assets/6/6/7/e/8/HX6537-A_DS_public_v01_1_.pdf
✔	Camera Module HM0360-MWA	8.81	1	HiMax	https://cdn.sparkfun.com/assets/d/2/9/9/7/Pre-HM0360_DS_preliminary_v04__Ltd._-1.pdf
✔	MEMS Microphone SPH0641LM4H-1	1.05	1	Knowles	https://media.digikey.com/pdf/Data%20Sheets/Knowles%20Acoustics%20PDFs/SPH0641LM4H-1.pdf
✔	Crystal Oscillator ECS-240-10-36-CKM-TR	0.57	1	ECS Inc.	https://ecsxtal.com/store/pdf/ECX-2236.pdf
Power Circuitry					
	Adjustable Linear Voltage Regulator R1173D001B-TR-FE	1.33	3	Nisshinbo Micro Devices Inc.	https://www.nisshinbo-microdevices.co.jp/en/pdf/datasheet/r1173-ea.pdf
Indication					
✔	RGB LED	0.1	1	Harvatek Corporation	https://media.digikey.com/pdf/Data%20Sheets/Harvatek%20PDFs/B39D3RGB-F6C0001HOU1930.pdf
Connectors					
	Board to Camera OK-10F030-04	1.22	1	AliExpress	N/A
	Qwiic JST SH 4-pin Right Angle Connector	0.40	1	Adafruit	N/A
Passive Components					
✔	Misc resistors	0.01	15	-	N/A
✔	Misc capacitors	0.01	17	-	N/A
✔	Misc inductors	0.01	1	-	N/A
	Total	**30.97**			

Environmental Impact

With the widespread deployment of smart sensors, it is essential to consider and be conscious of the environmental impact such ubiquitous computing may have. Thus another component we advocate to be included in the datasheet is an “environmental impact" section that outlines the device footprint. Using the methodology of [9], we generated a sample of what this section might look like as part of the datasheet for our sensor specifically. We capture the carbon footprint (CO2-eq.) of our ML sensor in the chart below. Due to the limited amount of data available on electronic device footprint we were not able to capture every single component. We were able to account for 8 out of 11 components from our bill of materials, though, which we feel captures the concept sufficiently for the sake of demonstration. We were unable to find data for the connectors and voltage regulator. However, in addition to the bill of materials, we capture the carbon footprint for the ML sensor’s model training, transport, and three-year use.

The total carbon footprint, including embodied and operational footprint, of our ML Sensor is approximately **2.82 kg CO2-eq**. The chart below shows how the footprint is broken down. The majority of the footprint can be attributed to the power supply and camera sensor.



We note that we do not claim that this is 100% accurate but rather a representative approximation of the sensor’s environmental impact and what other future datasheet should aim to include.

Acronyms

Acronym	Description
GDPR	General Data Protection Regulation
GPIO	General Purpose Input Output
ML	Machine Learning
I2C	Inter-Integrated Circuit
ID	Identifier
IoU	Intersection Over Union
LED	Light-Emitting Diode
MCU	Microcontroller Unit
MEMS	Microelectromechanical System
MST	Monk Skin Tone Scale
PCA	Principal Component Analysis
RGB	Red Blue Green

Glossary

Lux	Photometric unit of luminance (at 550 nm, 1 lux = 1 lumen/m^2 = 1/683 W/m^2)
Sensitivity	A measure of pixel performance that characterizes the rise of the photodiode or sense node signal in Volts upon illumination with light. Units are typically V/(W/m^2)/sec and are dependent on the incident light wavelength. Sensitivity measurements are often taken with 550 nm incident light. At this wavelength, 683 lux is equal to 1 W/m^2; the units of sensitivity are quoted in V/lux/sec. Note that responsivity and sensitivity are used interchangeably in image sensor characterization literature so it is best to check the units.
IoU	Intersection Over Union (IoU) is a metric used to evaluate the accuracy of an object detector on a specific dataset. It measures the overlap between the predicted bounding box (from the detector) and the ground truth bounding box. Values range between 0 and 1, where a higher value indicates better prediction accuracy. A value of 0 indicates no overlap, while a value of 1 indicates perfect overlap (the predicted box matches the ground truth exactly).
Inference	The process of applying a trained machine learning model to unseen data for making predictions or classifications. In the context of person detection, it involves analyzing images or video frames to determine if a person is present.
Accuracy	A performance metric that measures the overall correctness of a person detection system, indicating the percentage of correctly identified persons in the total number of instances.
Principal Component Analysis	A statistical procedure that transforms a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. These components are orthogonal to each other and capture the variance in the data in decreasing order.
Monk Skin Tone Scale	A 10-shade system, developed by Google, designed to provide a more inclusive representation of diverse skin tones in image-based technologies to address the challenges of representation in image-based technologies, especially for people of color.
Training Set	Labeled examples or samples used to teach a machine learning model to recognize and classify objects accurately. In the case of person detection, it comprises images or videos with annotated information about the presence or absence of people.
Test Set	A subset of the dataset that is strictly used to evaluate the performance of a model after it has been trained. The test set provides an unbiased evaluation of a model's generalization to new, unseen data. It should never be used during training or hyperparameter tuning.
Validation Set	A subset of the dataset, separate from the training set, used to evaluate a model during training. It provides an intermittent check on the model's performance, allowing for hyperparameter tuning and model selection. By evaluating model performance on a validation set, one can detect issues like overfitting (where the model performs exceptionally well on the training set but poorly on new, unseen data). Once the model is optimized using the validation set, its final performance is then assessed on the test set.
Person Detection	The process of identifying the presence and location of a person within an image or video stream.
Sensor	A device that detects and measures physical or environmental properties, such as the presence of a person, and converts them into electrical signals.

USER STUDY FORMS

Study Flyer


Harvard Edge Computing Group
MACHINE LEARNING SENSORS
EXPERIMENTAL STUDY
INFO & SIGN-UP
PARTICIPANTS NEEDED!
Come help us test the first machine
learning sensor!
CONTACT :
matthew_stewart@g.harvard.edu
yasmineomri@college.harvard.edu
tinyurl.com/mlsensors

Machine Learning Sensors Experimental Study Interest Form

The Edge Computing Group is seeking participants for an experimental study evaluating a new paradigm of machine learning sensors that we are designing.

We are looking into the next generation of sensors, ML sensors, which use on-device machine learning to extract useful information from the raw data before reporting it to the outer system. The ML sensor paradigm comes with significant benefits when it comes to modularity, composability, power efficiency, privacy and security, and more! Part of establishing the ML sensor paradigm is reimagining what the conventional sensor data sheet. More particularly, we want to test the end-to-end performance of the sensors through a representative study to see how it performs in the real world on a set of people that it was not trained on. In order to investigate potential algorithmic bias, we will be collecting data on participants' sex and skin tone throughout the study.

The study will take place in the SEC, where a room will be set up with a series of sensors that will output the probability of seeing a person. The study will involve varying distances and lighting settings and is estimated to take approximately 2 minutes per person. Participants will be provided with snacks and required to fill out a brief consent form prior to participation. The sensors will take instantaneous images for processing, but none of the information will be saved to ensure privacy preservation.

If you have any questions or are interested in helping out with this study, please contact matthew_stewart@g.harvard.edu or yasmineomri@college.harvard.edu. For more information about our project, please see our high-level blog or this paper.

Name:

____

Email:

____

Which of these weeks would you be available to take part in this experiment? We will be reaching out by email closer to the date to organize concrete times.

- ☐ April 3-7
- ☐ April 10-14
- ☐ April 17-21
- ☐ April 24-28

Machine Learning Sensors Experimental Study Consent Form

The Edge Computing Group is looking for participants to take part in an experimental study we are devising to evaluate a new paradigm of intelligent sensor that we are developing.

We are looking into the next generation of sensors, ML sensors, which use on-device machine learning to extract useful information from the raw data before reporting it to the outer system. The ML sensor paradigm comes with significant benefits when it comes to modularity, composability, power efficiency, privacy and security, and more. Part of establishing the ML sensor paradigm is reimagining the conventional sensor data sheet. More particularly, we want to test the end-to-end performance of the sensors through a representative study to see how it performs in the real world on a set of people that it was not trained on.

Please review and confirm your agreement to the following:

What You Will Do in this Study: Participants will be requested to select their gender identity and skin tone (using the Monk Skin Tone Scale). You will be asked to input this information at the end of this consent form and prior to entering the study room. This information will remain anonymous, as no name or additional information about the participant will be recorded.

The participant will then enter the study room and stand in front of a group of six sensors at three varying distances. At each distance, three different light settings will be tested. One individual in the room will be altering the lighting conditions and another taking note of the sensor outputs. The sensor predictions (detection of a person) are recorded and coupled with the skin tone and gender identity to investigate potential algorithmic bias. We expect the experiment to take about 2-3 minutes per person.

Participation in this study is voluntary. You do not have to be in this study if you do not want to and you can quit the study at any time.

Risks and Benefits: There are no known risks to participants. Participants will be compensated with snacks for their time. The results of this study intend to be published and could lead to the development of more accurate and effective person detection technology, which could have a positive impact on various fields, such as security, healthcare, and transportation.

Confidentiality: No privacy-sensitive information will be recorded, such as images or personal details. Skin tone and gender identity will be deidentified prior to data analysis, ensuring participants remain entirely anonymous.

We will only report a summarized analysis of the variability of the sensors' prediction accuracies with skin tone and gender.

If you have questions about the survey or study, contact Matthew Stewart (matthew_stewart@g.harvard.edu), Yasmine Omri (yasmineomri@college.harvard.edu) or Professor Vijay Janapa Reddi (vj@eecs.harvard.edu).

I confirm that I am at least eighteen years of age, I understand and have read the points above, and consent to the collection, use, and sharing of your anonymous responses.

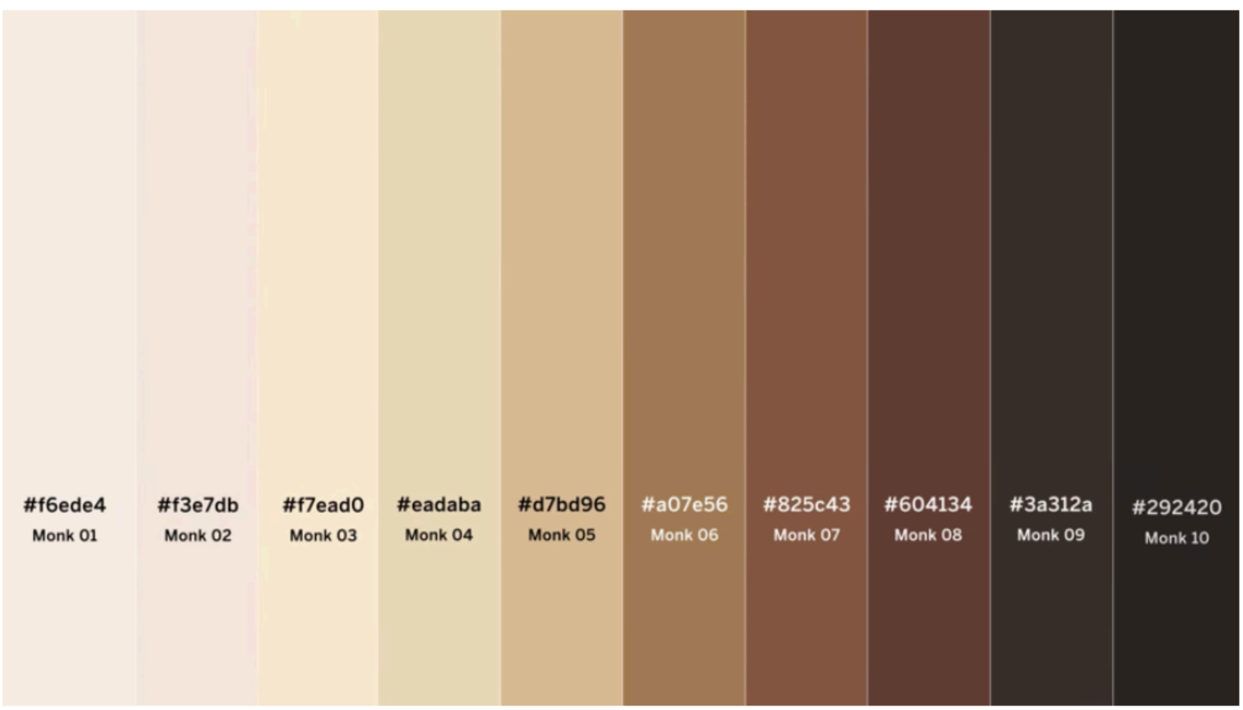
- ☐ I consent

Please select your gender identity:

- ☐ Male
- ☐ Female
- ☐ Transgender Male
- ☐ Transgender Female
- ☐ Non-binary
- ☐ Other (Please Specify)

Which Monk Scale skin tone most closely matches your skin color?

Please refer to the Monk Scale skin tone chart below or on the table. Ask a team member if you are having trouble identifying the appropriate value.



- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7
- ☐ 8
- ☐ 9
- ☐ 10