Appendix A. Example Data Sheet - Open-Source Sensor

OVERVIEW **15**
Compliance and Certifications 15
Description 15
Features 15
Use Cases 15
MODEL CHARACTERISTICS **16**
Software Flow Diagram 17
Dataset Nutrition Label 18
IoT Security and Privacy Label 21
Machine Learning Model Specification 23
Performance Analysis 25
Environmental Sensitivity 26
Demographic biases 26
HARDWARE CHARACTERISTICS **27**
Hardware Details 28
Device Diagrams 29
Bill of Materials 30
Environmental Impact 31
Acronyms 32
Glossary 33

OVERVIEW

PA1 Person Detection Module

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. As of the time of writing, the sensor does not have any certifications from third-party organizations.

Description

The PA1 Person Detection Module is a cost-effective device that uses a machine learning (ML) algorithm to detect the presence of a person within its range. The sensor is equipped with cameras and sensors that capture images and data from the surrounding environment. These images and data are then processed by the on-device ML algorithm to identify people. When a person is detected, the sensor sends an alert or trigger to connected devices or systems, allowing them to perform specific actions such as activating security cameras, turning on lights, or opening doors. The person detection sensor is ideal for use in security, home automation, and other applications that require quick and accurate detection of people.

The sensor has a small form factor and utilizes a monochrome camera with a field of view of 320 x 320 (QVGA). The sensor is equipped with an onboard 3.3V regulator, which enables it to operate with an input voltage range of 3.5V - 5.5V when enabled, or 3.0V - 3.6V when disabled. 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. The sensor has a maximum data rate of 100 kb/s and a wide sensitivity coverage of 0.1 - 10 klux.

Features

- Real-time person detection with on-device ML
- Indoor and outdoor use
- Low power consumption
- Onboard camera
- Small form factor: 10 x 10 x 2 mm
- I2C serial communication
- Wide sensitivity coverage: 0.1 - 10 klux

Use Cases

- Security
- Home automation
- Consumer appliances

MODEL CHARACTERISTICS

Software Flow Diagram

Grayscale images (320x320) are collected and resized to 96x96 via bilinear interpolation. Images are fed into a MobileNetV1 architecture trained and optimized through Edge Impulse. The output probability is communicated via Qwiic interface to the application processor.

Start
320x320 raw grayscale image from the camera

↓

Preprocessing: Bilinear Interpolation
96x96 grayscale image

↓

Model Architecture	MobileNetV1 96x96 with no final dense layer
framework	TensorFlow Lite Micro
training cycles	20
training learning rate	0.0005
validation set	20%
dropout	0.1
Quantization	int8
Compiler	EON Tuner
Inference time	~ 2200 ms
Peak RAM usage	131 Kb
Flash Usage	309.5 Kb

↓

Single Output

Person Detection Probability

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	**Yes**	**Yes**	**Unsure**	**No**
	COCO Dataset	https://arxiv.org/pdf/1906.05721.pdf	Not Applicable	Not Applicable

Do Not Use

- **Domain.** Military or weaponized applications
- **Image Detection for hi-res images.** The model is designed for lo-fi uses, and other models exist for hi-res images that are fine-tuned to that purpose
- **Object Identification more specific than person/not-person.** The data was cleaned and labeled specifically for person/not-person. Re-labeling the dataset for other purposes does not ensure proper diversity of data for another purpose.

Intended Use

- **Intended Domain.** Internet of Things
- **Intended Domain.** Image Recognition
- **Intended Domain.** On-Device Intelligence
- **Intended Domain.** Person Detection
- **Intended Use.** Train neural network models to detect the presence of a person in images when deployed on resource-constrained microcontrollers.
- **Other Responsible Uses.** Object Detection and Recognition
- **Other Responsible Uses.** Scene Understanding
- **Other Responsible Uses.** Image Captioning

Collection process

The MS-COCO dataset was collected through sourcing diverse images from Flickr and using Amazon Mechanical Turk for human annotators to draw polygons around object instances and provide descriptive captions for each image, followed by quality control measures to ensure annotation consistency. The Visual Wake Words dataset was derived from this by selecting the subject of images containing "person" and "non-person" labels.

General risks

Any additional risks?

Individual Information

yes

Consent

Consent was not given.

Generalized Inferences

The original source material, from COCO, is mainly made up of photographs from Flickr, and it's not clear to what extent the users of Flickr are representative of the population at large outside the U.S., for instance.

Generalized Inferences - Mitigation

Identifying a specific use case for models made using this dataset, creating a list of situations in which people would be found for that use case, and then reviewing the base dataset to ensure it has a diversity of images related to the situations you identify (this may be a somewhat manual process).

Sensitive Content

Not Applicable

Documented Known Issues

https://medium.com/@jamie_34747/how-i-found-nearly-300-000-errors-in-ms-coco-79d382edf22b

Other Known Issues

Some items in both the person and non-person categories are known to be mislabeled.

Number of issues

Risky	2
Safe	1
Unknown	4

Feature selection

Which columns were chosen and why?

Cultural or Domain Assumptions

Proxy Characteristics

Planning Representation

Domain Knowledge

Some familiarity with the style of how images are labeled in the COCO datasets would be helpful

Number of issues

Risky	1
Safe	1
Unknown	2

Representation

Which rows were included and why?

Subpopulation Information

Not Applicable

Representation

Unknown

Individual Inferences

Decisions or predictions based on the dataset may not accurately account for individual variations, such as clothing and accessories worn by an individual, and could result in overgeneralized outcomes that don't consider unique circumstances or factors. Additionally, the data may include bias due to its data collection practices which may lead to unfair or discriminatory decisions.

Individual Inferences - Mitigation

Collection Representation

Other Representation Issues

Number of issues	
Risky	1
Safe	0
Unknown	5

Data values

What values are in each column?

Collection and Labeling Protocols

The data was generalized from its original description to be that of a "person" or "not person", which required scraping of the original dataset based on search parameters entered by the authors of the dataset. The upstream dataset used Amazon Mechanical Turk workers to label pictures as well, on a custom interface created by the upstream dataset authors.

Data Imputation Protocols

Data Manipulation Protocols

Missing Data

The dataset is derived from MS-COCO and thus contains all items within that dataset that include person and non-person tags.

Raw Data

Number of issues	
Risky	1
Safe	0
Unknown	4

IoT Security and Privacy Label

This device contains a camera that takes pictures at 1 s intervals. No other sensory data is collected. Raw data is contained solely within the ML module, with only high-level features transmitted to the main processor (i.e., no image data is accessible by the main processor). This module has no internet connectivity or data storage capacity outside the model and software.

Security & Privacy Overview

Harvard University

Person Detection Module PA1
Firmware version: 0.1 **- updated on:** 2023-02-20
The device was manufactured in: United States

Security Mechanisms

Security updates	No security updates
Access control	No user account is allowed

Data Practices

	Visual	Audio	Physiological	Location
Sensor data collection	Visual			
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

PUBLIC DOMAIN

Security & Privacy Details

Harvard University

Person Detection Module PA1
Firmware version: 0.1 - **updated on:** 2023-02-20
The device was manufactured in: United States

Security Mechanisms

Security updates	No security updates
Access control	No user account is allowed.
Security oversight	No security audits
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	Data will not be linked with other data sources
What will be Inferred from User's Data	Presence of a human
Special data handling practices for children	No
In Compliance with	GDPR
Privacy policy	Not disclosed

More Information

Call Harvard University with your questions at	Not disclosed
Email Harvard University with your questions at	ml-sensors@googlegroups.com
Functionality when offline	Full functionality on offline mode
Functionality with no data processing	Not disclosed
Physical actuations and triggers	Device performs customized actions when person is detected.
Compatible platforms	Not disclosed

CMU IoT Security and Privacy Label CISPL 1.0 iotsecurityprivacy.org

PUBLIC DOMAIN

Machine Learning Model Specification

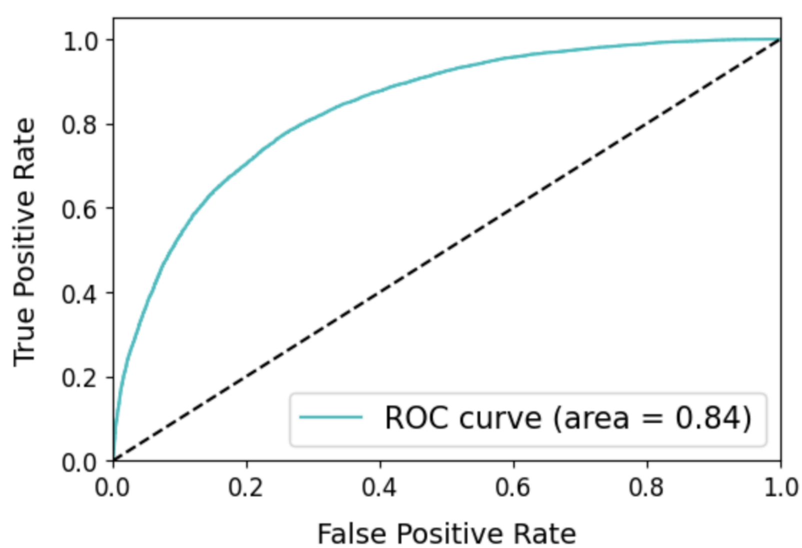
The person detection model was created using transfer learning with the MobileNetV1 neural network (see architecture here) on Edge Impulse. The training and testing of the model were done using a subset of images from the MS-COCO 2017 dataset, which is widely used for image recognition. Only images containing humans were selected from the dataset, totaling 109,604 images. The derived dataset is equivalent to the Visual Wake Words dataset. A train/validation split ratio of 0.8 was used.

The input to the model is a 96x96 raw image in 8-bit grayscale format, equivalent to 9,216 features. The training process was carried out over 20 cycles with a learning rate of 0.0005 and a test set of 20% on MobileNetV1 with a dropout of 0.1 and no final dense layer. The output layer of the model produces a two-class vector of results, indicating the probability of a person being present in the image. The unoptimized (float32) model has an accuracy of 76.3%, with a false positive (FP) and false negative (FN) rate of 20.7% and 26.8%, respectively. The model was quantized to int8 and deployed on Edge Impulse using the integrated EON-Compiler to produce a C++ library. The quantized model has an accuracy of 75.5%, with an FP and FN rate of 23.9% and 25.1%.

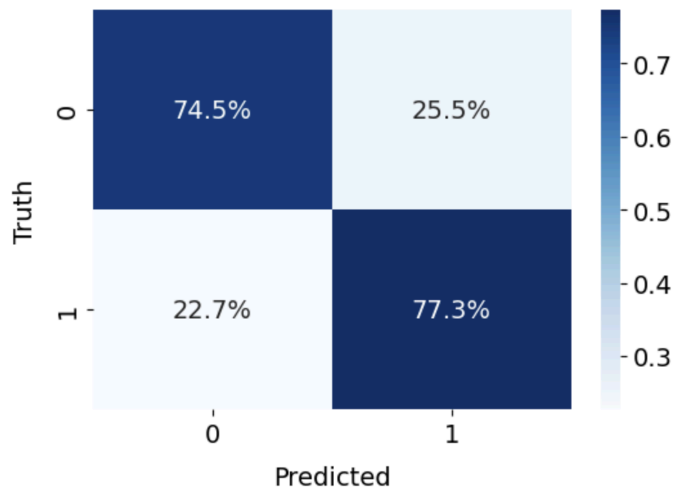
To enable live person detection, a set of image provision scripts was added to the software pipeline. The scripts continuously capture data from the onboard camera and pass it to the model in the appropriate scale and format. Using the Arm GNU Toolchain, the Pico-SDK, and the resulting C++ library, the model was built and compiled into a binary file that can be flashed to the ML board [See README/GitHub Repo]. The output of the model is an output vector consisting of a non-person score and a person score, which is communicated through a serial connection and can be viewed on a serial monitor.

Model workflow and characteristics can be viewed through the public Edge Impulse project version here.

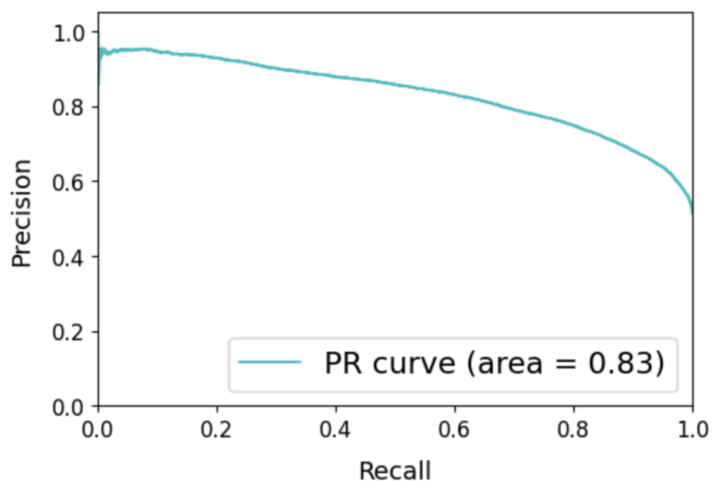
(a) Receiver Operating Characteristic Curve



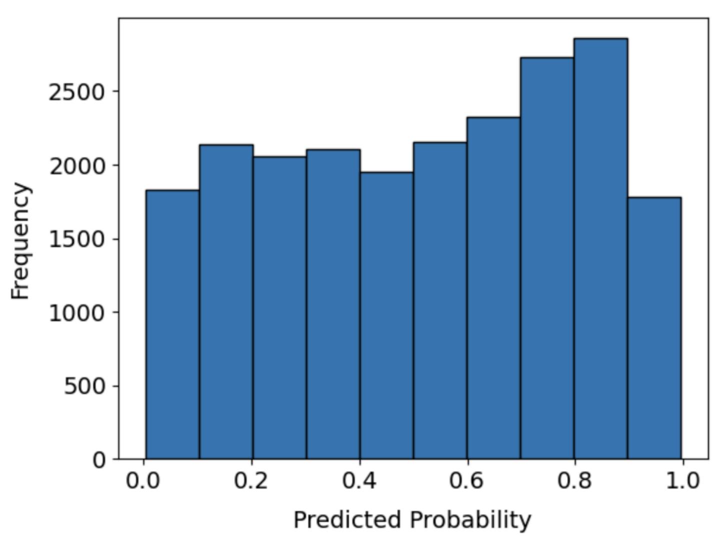
(b) Confusion Matrix



(c) Precision-Recall Curve



(d) Histogram of Predicted Probabilities



Performance Analysis

The end-to-end performance of the person detection sensor model was tested through an experimental study. 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.

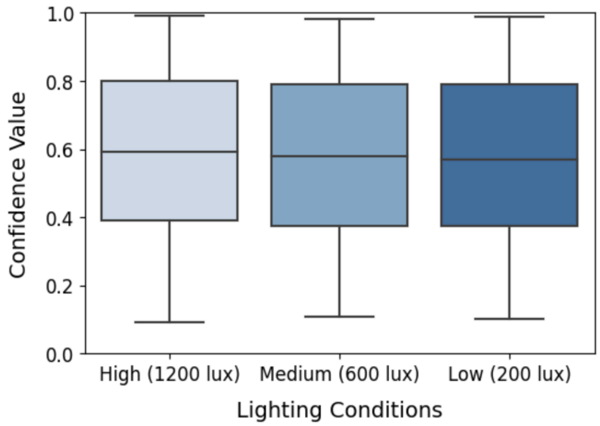
The lighting levels were controlled using a dimmer switch that had three levels of operation, corresponding to 208±31, 584±51, and 1149±59 lux, respectively. When the lights were turned off, the illuminance meter gave a reading of zero lux. When all the lights were turned on at full strength, the sensor 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).

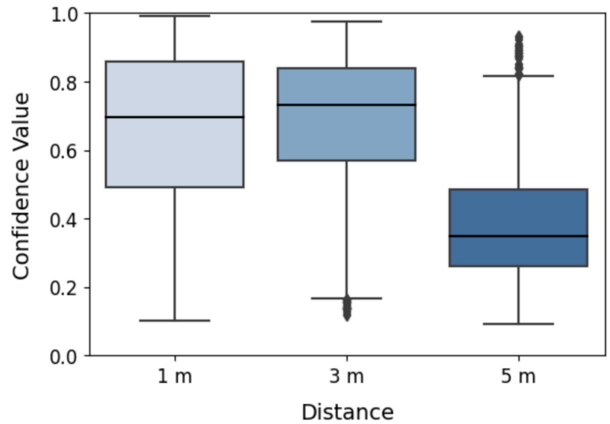
Before entering the study environment, the participants were asked to provide their gender identity and evaluate their skin tone according to the Monk Skin Tone (MST) Scale to evaluate algorithmic bias. The study evaluated algorithmic bias by bucketing skin tone into three categories: light (MST 0-4), 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.

Environmental Sensitivity

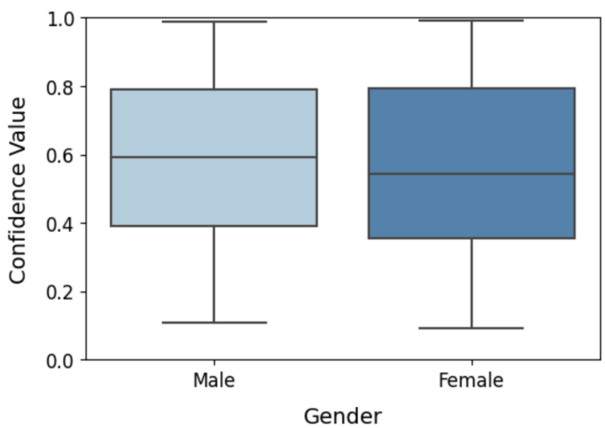
The device shows a marginal decrease in performance under decreased lighting conditions. A marked drop off in performance is observed at distances 3-5 meters from the sensor.

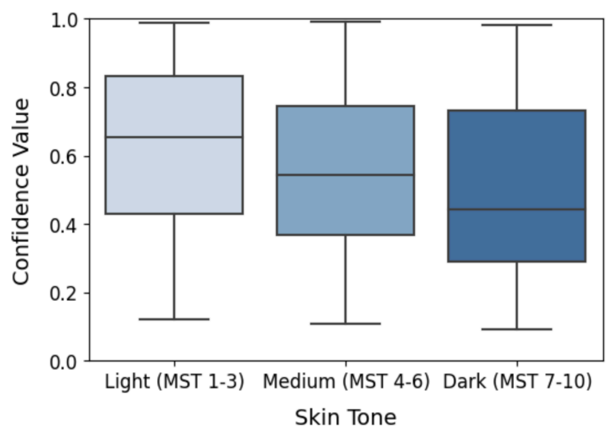




Demographic biases

A small gender bias is observed in model performance. A large skin tone bias was observed, showing approximately a 20% decrease in the confidence value for individuals with a darker skin tone.





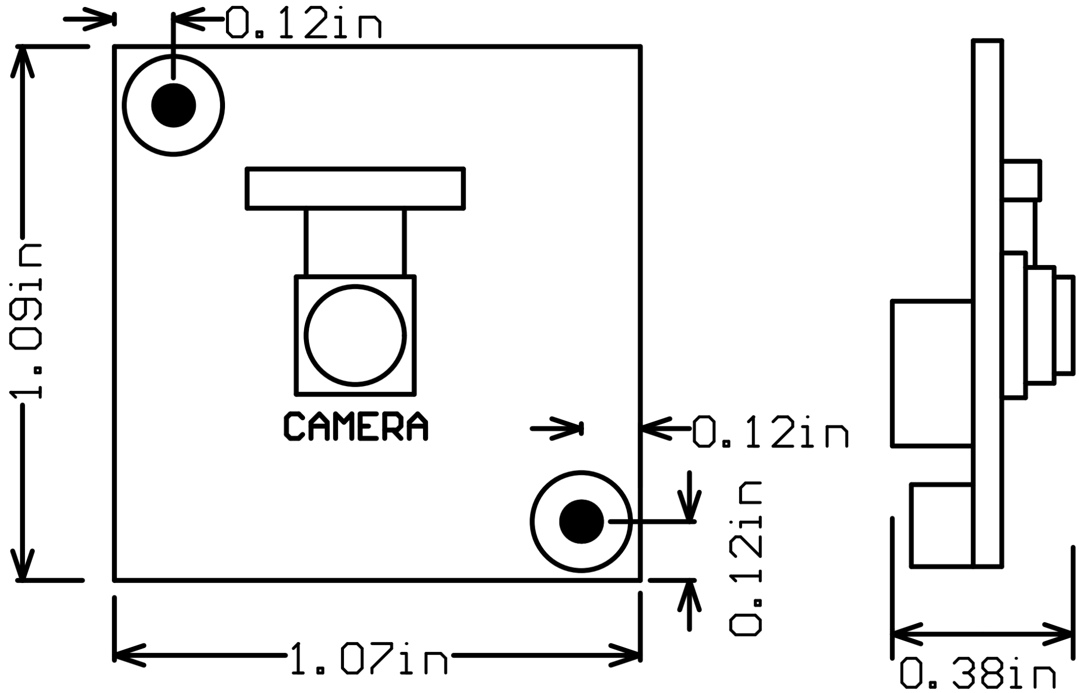
HARDWARE CHARACTERISTICS

Hardware Details

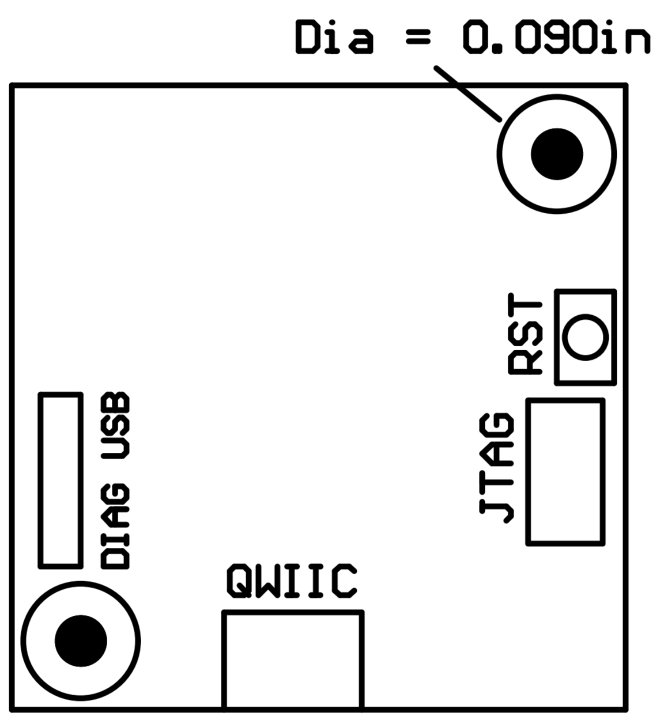
Camera Specifications (see here)	
Field of view (horizontal)	87º
Color Filter Array	Bayer, Monochrome
Frame Rate	45FPS @ 6MHz
Pixel Array (Active/ Effective)	324 x 324 / 320 x 320
Electrical Specifications	
Operating Voltage Range (regulator enabled)	3.5V to 5.5V
Operating Voltage Range (regulator disabled)	3.0V to 3.6V
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	Red arrow on sticker points up.
GPIO mode	SCL/SDA lines can be customized to make programmable flag lines (I_{out} max = 12 mA)
Diagnostic LED	Default behavior of green LED on board: illuminates for one second on power-up, then illuminates when person detected.
Data Transfer and Format	Single byte: number from 0-255 representing confidence score
I2C Address	0x22

Device Diagrams

Front and side view of sensor.



Back view of sensor.



Bill of Materials

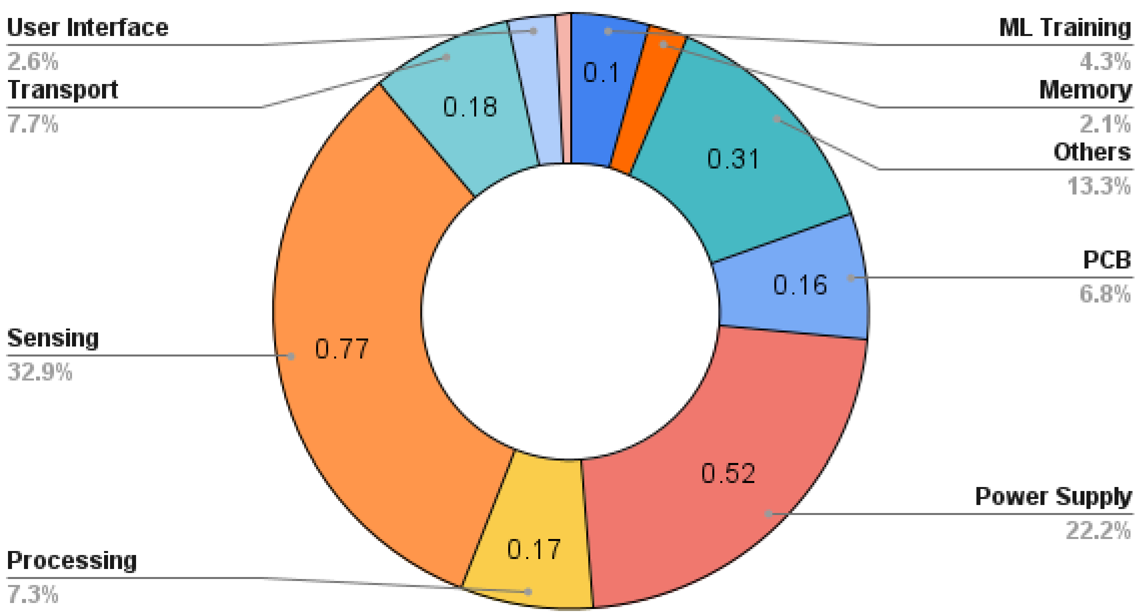
The following is a comprehensive list of materials required to assemble the PA1 person detection module, commonly referred to as the bill of materials. All unit cost values quoted in minimum order quantity of one.

Category In TinyML Calculator	Component	Unit Cost ($)	Quantity	Manufacturer	Link to Datasheet (if available)
Functional Components					
✔	RP2040 Microcontroller	1.00	1	Raspberry Pi	https://datasheets.raspberrypi.com/rp2040/rp2040-datasheet.pdf
✔	QVGA Camera Module HM01B0	8.90	1	HiMax	https://cdn.sparkfun.com/assets/7/f/c/8/3/HM01B0-MNA-Datasheet.pdf
✔	Flash Memory W25Q16JVSNIQ	0.36	1	Winbond Electronics	https://www.winbond.com/resource-files/w25q16jv%20spi%20revg%2003222018%20plus.pdf
✔	12 MHz Crystal Oscillator 445C25D12M00000	0.42	1	CTS-Frequency Controls	https://www.mouser.com/datasheet/2/96/008-0360-0-786290.pdf
Power Circuitry					
	Voltage Regulator TLV70228 2.8V	0.69	1	Texas Instruments	https://www.digchip.com/datasheets/download_datasheet.php?id=3747267&part-number=TLV70228
Indication					
✔	LTST-C190KGKT LED	0.05	1	Lite-On Inc.	https://www.digikey.com/htmldatasheets/production/37809/0/0/1/ltst-c190kgkt.pdf
Connectors					
	FFC connector FH26W-31S-0	1.28	1	Hirose Electric Co Ltd	https://www.hirose.com/product/download/?distributor=digikey&type=specSheet&lang=en&num=FH26W-31S-0.3SHW(60)
	Qwiic connector PRT-14417	0.57	1	SparkFun Electronics	https://www.mouser.com/datasheet/2/813/Qwiic_Connector_Datasheet-1223982.pdf
Passive Components					
✔	Resistors	0.01	10	-	N/A
✔	Capacitors (low value)	0.01	15	-	N/A
✔	Capacitors (high value)	0.05	7	-	N/A
✔	Ferrite bead 600Ω	0.07	2	-	N/A
✔	Printed circuit board	0.50	1	-	N/A
	Total	**14.51**			

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 10 out of 13 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 **2.34 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
SNR	Signal-to-noise ratio
COCO	Common Objects in Context
FFC	Flexible Flat Cable
GDPR	General Data Protection Regulation
ML	Machine Learning
I2C	Inter-Integrated Circuit
LED	Light-Emitting Diode
MCU	Microcontroller Unit
SCL	Serial Clock
SDA	Serial Data
GPIO	General Purpose Input Output
SDK	Software Development Kit
QVGA	Quarter Video Graphics Array

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.
SNR	Signal-to-noise ratio. This number characterizes the ratio of the fundamental signal to the noise spectrum up to half the Nyquist frequency.
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.
False Positive	A situation in person detection where the system incorrectly identifies an object or pattern as a person when it is not.
False Negative	A situation in person detection where the system fails to identify a person when one 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.
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.
Precision	A performance metric that measures the proportion of correctly identified persons among all the instances identified as persons by the system. It quantifies the system's ability to avoid false positives.
Recall (Sensitivity)	A performance metric that measures the proportion of correctly identified persons among all the actual persons present in the data. It quantifies the system's ability to avoid false negatives.
Threshold	A predefined value used to determine whether the output of a person detection system indicates the presence or absence of a person. Adjusting the threshold affects the balance between false positives and false negatives.
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.