Estimation of UAV Count Using Thermal Imaging and Lightweight CNN

W. A. Nelson, A. Jha, A. Kumar, L. R. Cenkeramaddi

Abstract: **Illegal and improper use of UAVs has damaged public property and challenged the safety and security of the civilian population. Due to their small form factor, UAVs are undetectable using conventional aircraft detecting methods. In this work, we have addressed this issue by utilizing thermal images to detect and estimate the UAV count in a multi-UAV setting. Thermal imaging-based detection provides a number of advantages, including night vision, temperature sensitivity, low visibility, camouflage penetration, and non-invasiveness. It is a non-contact and non-intrusive detecting method that can detect hidden objects or people even in low-visibility environments such as smoke and fog. Experiments were carried out by capturing thermal images of a multi-UAV setting, where an arbitrary number of UAVs are flying in a random manner. Further, a UAV-thermal dataset is also developed so as to facilitate further research. Extensive experiments were carried out and the reported results show that the proposed model accurately estimates the number of UAVs with an accuracy** 99.9%.

C.1 Introduction

Advancements in unmanned aerial vehicles (UAV) technology have sparked interest in utilizing UAVs for various applications such as package delivery, maintaining law and order, surveying, disaster management, defense, etc [1]. However, in recent times, the illegal and improper use of UAVs has elevated a growing concern to detect, monitor, and track UAV-related activities [2]. In this context, techniques to identify and detect illegal and intruder UAVs are of prime significance. There are a lot of works in the literature which address this issue using various sensors and techniques. Some of the notable ones include using acoustic sensors [3], RGB images [4], radar modules [5], etc. However, in this work, we focus our attention on using thermal imaging for UAV detection. The primary motivation for using a thermal imaging-based solution can be attributed to the improved visibility that thermal sensors offer in extreme light and dark conditions. Subsequently, thermal sensors are invariant to background noise as compared to acoustic sensors [6], [7].

There have been multiple attempts throughout the literature that use thermal imaging for UAV detection. One such notable work is in [8], where the authors develop a UAV detection and tracking system using deep learning techniques. A data set containing UAVs in visual and thermal modes was generated using adversarial data augmentation methodologies. Additionally, extensive experiments were conducted that revealed that the proposed model performed well on real-world UAV images with complex environments despite being trained on synthetic data. In [9] and [10], the authors develop UAV detection systems to detect obstacles using thermal sensors during the night. Particularly, in [9], the proposed method uses RGB camera and ADS-B signals for validating the obtained results. Results also showed that the detection accuracy was 100% for extreme illumination conditions and during the night in all cases. The authors in [11] used faster-RCNN, a saliency map, and a magnifying small objects (MSO) module to develop a UAV detection system. The thermal saliency map helps to extract meaningful features from the thermal images. Additionally, the MSO module enhances the resolution of small objects before feeding the thermal image to the model. The MSO module thereby increases the accuracy of the model. Reported results indicate more than 93% accuracy for the proposed approach. In [12], the authors investigate the effectiveness of training a UAV detection system using limited thermal images. In order to improve the accuracy of the model, the RGB images containing UAVs were preprocessed to retain characteristics found in thermal imagery. The model was hence trained on preprocessed RGB and limited thermal images. It was reported that the accuracy of the trained model improved significantly based on the type of preprocessing performed on the RGB images.

It can be seen from the above paragraph that most of the literature pertaining to UAV detection using thermal imagery lacks sufficient real-world datasets. Synthetic thermal images are generated from RGB images or other sources to mimic thermal image characteristics. Additionally, the literature focuses on the detection of just a single UAV. In real-world intruder and trespassing scenarios, simultaneous detection of multiple UAVs is crucial to ensure safety and security.

Our work aims to bridge this gap by providing a method to accurately estimate the number of UAVs in a multi-UAV scenario. Additionally, this work will act as a reference for future detection systems to embark on the challenge of detecting multiple UAVs simultaneously. The main contributions of this paper are:

- Proposed a simple cost-effective approach to estimate the count of UAVs in multi-UAV setting by utilizing thermal imaging.
- Developed a lightweight machine learning model that can be used on the edge to estimate the UAV count in a multi-UAV scenario.
- Developed a thermal dataset that contains a total of 10 UAVs flying simultaneously in all directions in a random manner.

The remainder of the paper is organized as follows. Section C.2 provides the problem formulation. Section C.3 explains the proposed method that includes the

Table C.1: Experimental parameters.

Parameters	**Details**
Semicircular area	Radius: 5 meters
Measurement duration	5 minutes
UAV model	**Count**
DJI Mavic 2 Enterprise	1
DJI Mini 2	1
DJI Mini SE	2
DJI Mini 3 Pro	1
Tello EDU	4
SYMA X30	1

Table C.2: Thermal camera specifications.

Parameters	**Value**
Thermal sensitivity (mK)	< 50
Resolution (pixels)	160×120
Spectral Range (microns)	$8 - 14$ (nominal)
Frame Rate (Hz)	8.7 (effective)
Horizontal Field of View (°)	57
Lens Type	f/1.1
Size (mm)	$10.5 \times 12.7 \times 7.14$

experiments, measurement setup, dataset details, and proposed convolutional neural network (CNN) architecture. Section C.4 provides the results obtained after using the proposed approach. Finally, Section C.5 concludes the paper by providing a summary and possible future extensions.

C.2 Problem formulation

UAVs are powered with battery sources to conduct their various operations such as flight, landing, and hovering. These operations dissipate energy. Additionally, the rotation of propellers and other aerodynamic moving parts also results in energy dissipation. A part of the energy dissipated is in the form of heat or thermal energy. The dissipated thermal energy is characteristic of each UAV and depends on various factors such as shape, size, internal components, and battery type of the UAV. Hence, these thermal signatures can be utilized to detect and estimate the UAV count in the captured frame. Hence in this work, the primary objective is to estimate the UAV count from a captured frame by utilizing thermal imaging and machine learning techniques.

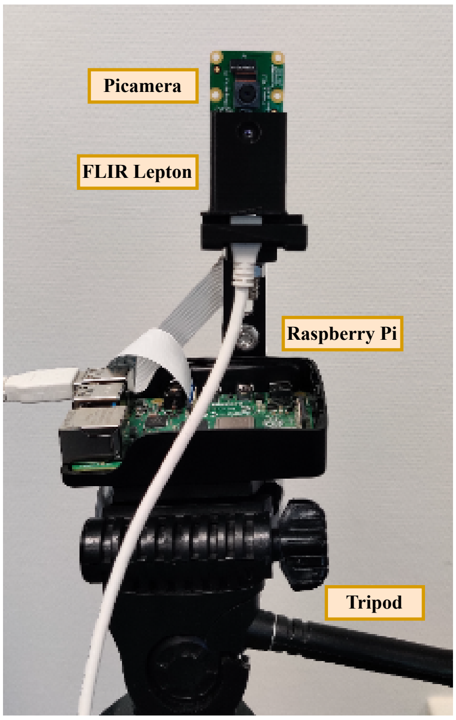


Figure C.1: Experimental setup used for collecting thermal images of UAVs.

C.3 Method

Each UAV has its own characteristic thermal signature. The proposed method exploits these thermal signatures to detect UAVs. The thermal signatures are captured by using a thermal camera, in this case, the FLIR Lepton thermal camera [13]. The captured thermal images are then provided to our custom lightweight CNN model to estimate the UAV count in the captured frame. The precise steps are outlined below:

Step 1: Collect thermal images of various types of UAVs using the FLIR Lepton thermal camera.
Step 2: Preprocess the obtained raw thermal images by resizing, normalizing, grayscale conversion, etc. before feeding to the proposed CNN architecture.
Step 3: Train the processed images with our custom lightweight CNN model to obtain an estimated UAV count in the scene.

C.3.1 Experiment

The experiment was conducted primarily in an indoor lab environment that consisted of a semicircular area of radius 5 meters. Each measurement of the experiment involved capturing thermal images of UAVs flying in the area. The measurements consisted of an arbitrary number of UAVs flying in all directions in a random manner. The measurements were taken in the order of increasing number of UAVs in the scene. Different UAV models such as DJI series [14], Tello EDU series [15], and SYMA series [16] were employed for the measurements. Details regarding the UAV

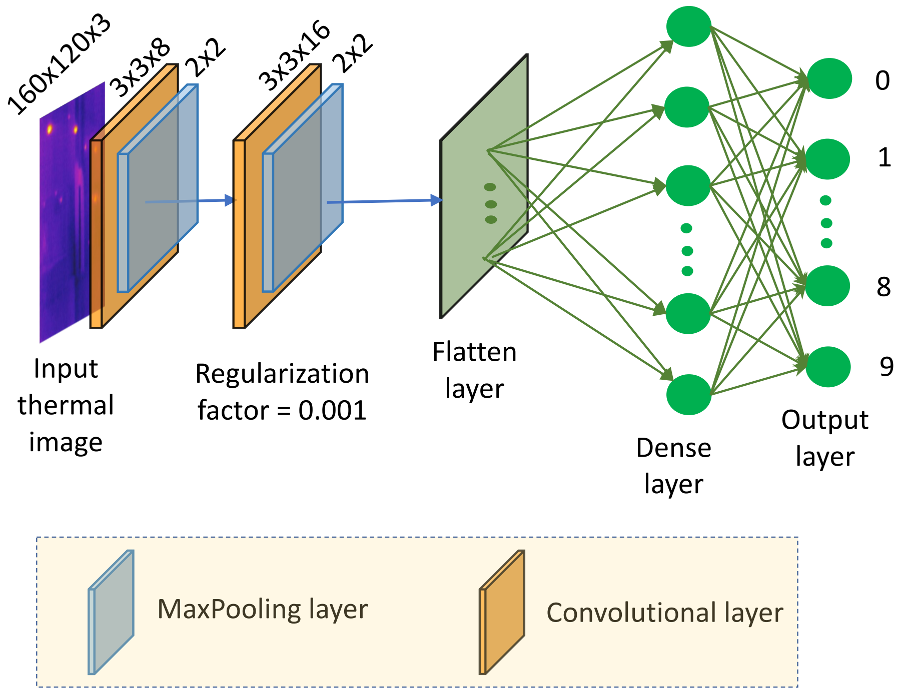


Figure C.2: Proposed CNN architecture for estimating the UAV count from thermal images.

model types can be found in Table C.1. As detecting small-sized UAVs is more challenging, the designed experiments also used Tello EDU UAVs for measurement. All the Tello EDU UAVs were programmed to fly in predefined trajectories. The remaining UAVs however, were operated manually by human control. Additional details regarding the experiment can be found in Table C.1.

The thermal images are captured using the FLIR Lepton 3.5 camera [13]. The FLIR Lepton 3.5 outputs a thermal image that has a resolution of 160×120 pixels with an effective frame rate of 8.7 Hz. The FLIR Lepton offers a nominal spectral range of $8 - 14$ microns with a thermal sensitivity of less than 50 mK. Due to its small size and excellent performance, it serves to be a good candidate to be used in the proposed method. Additional details regarding the thermal camera can be obtained from Table C.2.

The thermal camera is mounted on top of a static tripod as shown in Fig. C.1. The power to the thermal camera is supplied with the help of a USB A connection from the Raspberry Pi Model 4B [17]. A Picamera [18] is also fitted just below the FLIR Lepton. The purpose of the Picamera is to capture RGB images and provide a ground truth for the thermal images. The Picamera is also powered by the Raspberry Pi and operates synchronously with the thermal camera. In addition to powering the two cameras, the Raspberry Pi also serves the purpose of capturing and storing the thermal and RGB images obtained from the FLIR Lepton and Picamera respectively. The Raspberry Pi is installed with Ubuntu 18.04 [19] as the operating system. Ubuntu 18.04 was selected so as to facilitate seamless integration and installation of ROS Melodic[20] and other dependent software packages. ROS Melodic helps to capture data simultaneously from both the FLIR Lepton and Picamera. A portable

Table C.3: Layerwise architecture details of the proposed CNN model

No.	Layer	Output Size	Parameter
1	Input	[(None, 160, 120, 3)]	0
2	Conv2D (1)	[(None, 160, 120, 8)]	224
3	MaxPooling2D (1)	[(None, 80, 60, 8)]	0
4	Conv2D (2)	[(None, 80, 60, 16)]	1168
5	MaxPooling2D (1)	[(None, 40, 30, 16)]	0
6	Flatten	[(None, 19200)]	0
7	Dense	[(None, 10)]	192010

power bank is used to supply power to the Raspberry Pi board.

As mentioned above, each measurement consisted of flying a fixed number of UAVs in all directions. Thermal and RGB images for each of these flight trajectories are captured and stored. A total of 10 measurements are taken where each measurement consisted of 1, 2, 3, etc. up to 10 UAVs being flown in random directions. Each measurement lasted for a total of 5 minutes so as to capture at least 2500 thermal images. Different UAV models were used in these measurements to improve variability. Further, one measurement was conducted in an outdoor environment to introduce noise and additional variability to the data set. This in turn provides a way to understand how the proposed method performs in outdoor scenarios.

C.3.2 Dataset Details

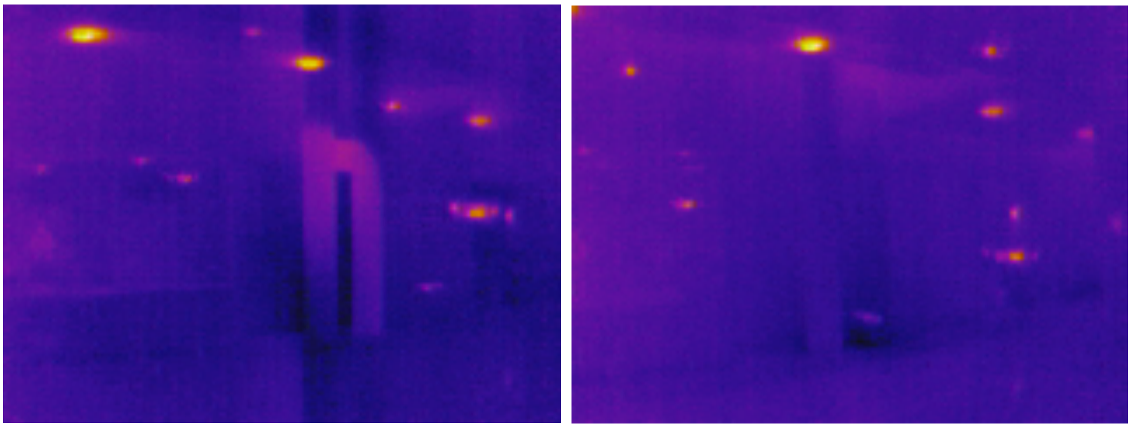
The thermal images captured using the FLIR Lepton camera have an image resolution of $n_{row} \times n_{col} = 160 \times 120$ pixels. Using the setup described in the above section, we have created a data set $D_j = \{y_j^i\}_{i=1}^{N=1000}$ of thermal images where y represents each thermal image in a set of 1000 of images from the class j. Here class j denotes the images that have j UAVs present in them. For example, class 2 denotes thermal images with 2 UAVs. A subset from this data set is provided to the proposed CNN model as input for training.

C.3.3 Convolution neural network (CNN)

The proposed CNN model is shown in Fig. C.2. The input to the CNN is the thermal image of size 160×120. This is followed by convolution (Conv2D) and max pooling layers. After the input layer, the image is fed into a convolutional layer with 8 filters and a kernel size of 3×3. Features are extracted in this layer and the reduced image is fed to a max pooling layer of dimension 2×2. The reduced image then undergoes a second convolutional operation with a L2 regularizer that has a regularization value of 0.001. This is further followed by max pooling, flatten, and dense layers. All the convolutional layers are activated using the ReLu non-linear activation function. Finally, the output layer uses softmax activation to provide outputs in terms of the probability of the number of UAVs present in the captured image. Additional architectural details are outlined in Table C.3.

(a) (b)

Figure C.3: RGB images from Picamera for (a) 6 UAV measurement scenario, (b) 8 UAV measurement scenario.



(a) (b)

Figure C.4: Thermal image from FLIR Lepton for (a) 6 UAV measurement scenario, (b) 8 UAV measurement scenario.

The UAV-thermal dataset that is developed for this work contains 10,000 thermal images. The images are segregated into folders ranging from 1 to 10, where each folder represents a class depicting the number of UAVs in the scene. The dataset is divided into the train, test, and validation sets in the ratio 80%, 10%, and 10% respectively. The sparse categorical cross-entropy function is used to minimize the loss. Further, the model uses an adaptive momentum (Adam) optimizer with 10^{-4} as the learning rate during training.

C.4 Results

In this paper, we propose a low-cost approach to estimate the UAV count in a scene using lightweight CNN. Fig. C.3 shows the visual RGB images of the measurement scenarios with 6 and 8 UAVs flown in a random fashion. Fig. C.4 depict the

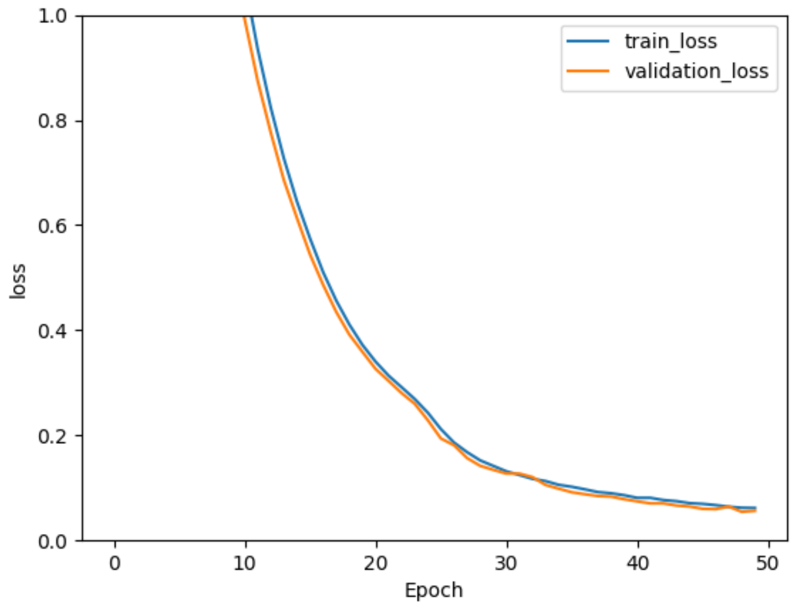


Figure C.5: Training and validation loss curve.

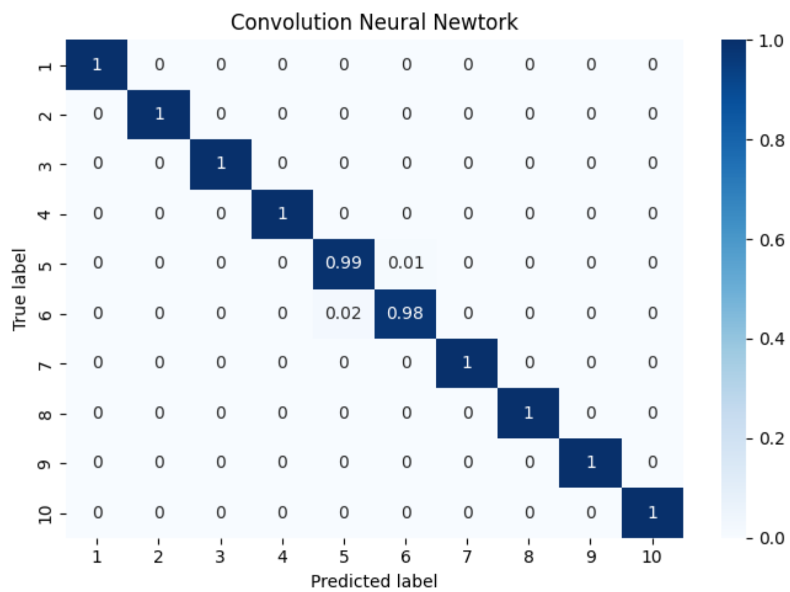


Figure C.6: Confusion matrix obtained on the test data set.

corresponding thermal images for the 6 and 8 UAV scenarios respectively. The proposed CNN model is trained on the thermal images from the created dataset. We use 10-fold cross-validation to train the proposed CNN model. The dataset is split with $(80, 10, 10)\%$ for train, testing, and validation respectively. Fig. C.5 shows the loss curve obtained during the training. It can be observed from Fig. C.5 that the training and validation loss converge to a minimum after 40 epochs. The model attains 99.91% validation accuracy from the 10-fold hold-out validation. Further, after the training, the model is tested with previously unseen data (10% from the dataset) to evaluate the performance. We obtained an average of 99.9% accuracy on the test set. The confusion matrix depicting the model performance is provided in Fig. C.6.

C.5 Conclusion and Discussion

In this paper, we proposed a method to estimate the number of UAVs in a scene. An end-to-end technique is developed that captures thermal images using embedded hardware and later processes and estimates the number of UAVs. Additionally, a UAV-thermal dataset is also developed that consists of a combination of 10 UAVs flying in all directions in a random manner. Extensive experiments demonstrate that the proposed CNN architecture was able to provide 99.9% accuracy in estimating the exact number of UAVs present. Future extensions to this work include UAV model identification in a multi-UAV setting that can help address UAV security threats.

References

[1] A. N. Wilson, Abhinav Kumar, Ajit Jha, and Linga Reddy Cenkeramaddi. Embedded sensors, communication technologies, computing platforms and machine learning for uavs: A review. *IEEE Sensors Journal*, 22(3):1807–1826, 2022.

[2] Matthew Ritchie, Francesco Fioranelli, and Hervé Borrion. Micro uav crime prevention: Can we help princess leia? *Crime Prevention in the 21st Century: Insightful Approaches for Crime Prevention Initiatives*, pages 359–376, 2017.

[3] Wilson AN, Ajit Jha, Abhinav Kumar, and Linga Reddy Cenkeramaddi. Estimation of number of unmanned aerial vehicles in a scene utilizing acoustic signatures and machine learning. *The Journal of the Acoustical Society of America*, 154(1):533–546, 2023.

[4] Jie Zhao, Jingshu Zhang, Dongdong Li, and Dong Wang. Vision-based anti-uav detection and tracking. *IEEE Transactions on Intelligent Transportation Systems*, 23(12):25323–25334, 2022.

[5] A. N. Wilson, Abhinav Kumar, Ajit Jha, and Linga Reddy Cenkeramaddi. Multitarget angle of arrival estimation using rotating mmwave fmcw radar and yolov3. *IEEE Sensors Journal*, 23(3):3173–3182, 2023.

[6] Linga Reddy Cenkeramaddi, Jyoti Bhatia, Ajit Jha, Santosh Kumar Vishkarma, and J. Soumya. A survey on sensors for autonomous systems. In *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pages 1182–1187, 2020.

[7] A. N. Wilson, Khushi Anil Gupta, Balu Harshavardan Koduru, Abhinav Kumar, Ajit Jha, and Linga Reddy Cenkeramaddi. Recent advances in thermal imaging and its applications using machine learning: A review. *IEEE Sensors Journal*, 23(4):3395–3407, 2023.

[8] Ye Wang, Yueru Chen, Jongmoo Choi, and C-C Jay Kuo. Towards visible and thermal drone monitoring with convolutional neural networks. *APSIPA Transactions on Signal and Information Processing*, 8:e5, 2019.

[9] Adrian Carrio, Yucong Lin, Srikanth Saripalli, and Pascual Campoy. Obstacle detection system for small uavs using ads-b and thermal imaging. *Journal of intelligent & robotic systems*, 88:583–595, 2017.

[10] Petar Andraši, Tomislav Radišić, Mario Muštra, and Jurica Ivošević. Night-time detection of uavs using thermal infrared camera. *Transportation Research Procedia*, 28:183–190, 2017.

[11] Naoual El-Djouher Mebtouche and Nadia Baha. Robust uav detection based on saliency cues and magnified features on thermal images. *Multimedia Tools and Applications*, 82(13):20039–20058, 2023.

[12] Lars Sommer and Arne Schumann. Deep learning-based drone detection in infrared imagery with limited training data. In *Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies IV*, volume 11542, page 1154204. SPIE, 2020.

[13] TELEDYNE FLIR. TELEDYNE FLIR Lepton. [Online]. Available: https://www.flir.eu/products/lepton/?model=500-0771-01&vertical=microcam&segment=oem.

[14] DJI. DJI Camera Drones. [Online]. Available: https://www.dji.com/no/camera-drones.

[15] Ryze Robotics. Tello EDU. [Online]. Available: https://www.ryzerobotics.com/tello-edu.

[16] SYMA. SYMA X30 Foldable Drone. [Online]. Available: https://www.symatoys.com/goodshow/x30-syma-x30-foldable-drone.html.

[17] Raspberry Pi. Raspberry Pi 4. [Online]. Available: https://www.raspberrypi.com/products/raspberry-pi-4-model-b/.

[18] Picamera. Picamera Documentation. [Online]. Available: https://picamera.readthedocs.io/en/release-1.13/.

[19] Ubuntu. Ubuntu 18.04. [Online]. Available: https://releases.ubuntu.com/18.04/.

[20] ROS Wiki. ROS Melodic. [Online]. Available: http://wiki.ros.org/melodic.