https://doi.org/10.31871/IJNTR.7.12.13 International Journal of New Technology and Research (IJNTR)
ISSN: 2454-4116, Volume-7, Issue-12, December 2021 Pages 18-22

An Agricultural AIoT Bird Repellent System with Machine-Learning based Moving Object Detection

Guan-Hsiung Liaw, Chia-He Li

***Abstract*—In recent years, with the maturity of deep learning technology, the problem of image recognition has become easier to solve, and it has begun to be applied to various industries. We try to apply this technology to solve the problem of bird damage in traditional agriculture, because in addition to natural disasters, the largest case of agricultural damage should be the gnawing disaster of birds or insects. In this paper, we propose an artificial intelligence of things (AIoT) bird repellent system architecture and make its prototype. In this system, we trained a neural network model that can identify bird flocks, and developed moving bird flock identification technology based on this, which will actually be implemented on a small single-board computer equipped with GPU (such as NVIDIA Jetson Nano) . The system can be set up in farmland to capture real-time video through a camera and identify the moving bird flocks by a small single-board computer equipped with GPU. When a flock of birds appears, the small single-board computer transmits a message through the LoRa low-power long-distance wireless communication interface to the ultrasonic bird repellent devices deployed in the farmland, and emits ultrasonic sound to drive the bird flock away.**

***Index Terms*—AIoT (Artificial Intelligence of Things), Bird Repellent System, Deep Learning, YOLO (You Only Look Once).**

I. Introduction

In recent years, due to the impact of declining birthrate, the rural population has been reduced and aging, resulting in a shortage of manpower in agriculture. Therefore, the development of many agricultural automation machines has emerged to reduce the manpower spent in farming. On the other hand, the annual loss of agricultural products caused by the gnawing of birds or pests is also a common and serious problem. However, there are not many effective and environmentally friendly prevention methods. These issues are urgently needed for contemporary agriculture.

At present, the most commonly used in farmland is the scarecrow as the main device for frightening and repelling the birds. However, the scarecrow is an immobile device. The birds will also accumulate experience and learning. Over time, they will know that it is a harmless device. Not only do birds no longer fear, they even stop on top of the scarecrow to rest. Therefore, many farmers will switch to using firecrackers to drive away birds. This method requires farmers to go to a certain place to release firecrackers at a certain point in time, but this approach will increase the labor cost, and also generate noise and air pollution problems.

Although the birds in the farmland will cause the loss of crops, the birds themselves are harmless to the farmers, so the best way is to drive them away instead of culling them. However, although the sound of firecrackers or metal percussion can scare away birds, it also interferes with the daily life of human beings, and can even cause fright and become an alternative pollution. The method of repelling birds and birds should be friendly to humans and the environment in order to achieve a win-win result.

The main principle of using firecrackers to drive away flocks of birds in farmland is that birds are afraid of strong flashing light and sound waves of certain frequency bands will interfere with the birds' nervous and physiological systems, causing their uncomfortable reactions. The audio frequency that birds and rodents feel uncomfortable is 25k~34.5kHz, but the human hearing range is 20~20kHz, so the audio frequency that is uncomfortable to birds is in the ultrasonic range that humans can't hear, and will not cause human hearing discomfort. , So ultrasound is very suitable as a friendly bird repellent tool.

In this paper, we propose an Artificial Intelligence of Things (AIoT) bird repellent system architecture and its prototype implementation. First, we train a neural network model that can identify bird flocks, and use this model to develop an identification method for moving bird flocks. In the prototype implementation, the model and the identification method of moving birds are implemented in a single-board computer with a camera. This single-board computer will be installed in the farmland for real-time bird flock identification. Based on the data we have collected and the actual experience provided by farmers, we know that most poultry will forage from early morning to morning and afternoon to evening. Therefore, during these periods, the single-board computer is activated and uses the camera to capture video and then perform real-time moving bird flock identification. Once the flock of birds flying into the farmland is detected, the single-board computer will remotely start the ultrasonic bird repellent device deployed in the farmland via the LoRa low-power wireless communication interface to emit ultrasonic sound for repelling the bird. Strong random flashing light can also emit at the same time to enhance the effect of driving away the birds.

The rest of the paper is organized as follows. In Section 2 the related works are introduced. In Section 3, the proposed system architecture and prototype implementation are described. In Section 4, some issues about detection of moving bird flocks are discussed. Finally, the conclusion is

Guan-Hsiung Liaw, Department of Information Engineering, I-Shou University, Kaohsiung City, Taiwan, R.O.C.

Chia-He Li, Intelligent Multimedia System Technology Dept., Division for Video Processing & Long Range Communications Technology, Information and Communications Research Laboratories, Industrial Technology Research Institute (ITRI), Chutung, Hsinchu, Taiwan, R.O.C.

given in Section 5.

II. Related Works

A. Deep Learning – YOLOv3

Redmo et al. proposed a deep learning technique for rapid object detection in images—YOLO (You Look Only Once)[1]. The principle is to select the area of the object to be identified by framing the bounding box in the image, and then perform feature extraction and classification of the objects in the pixels in the bounding box.

YOLO is based on the deep learning architecture developed by Convolutional Neural Network (CNN) [2]. The conceptual diagram of the operation of CNN is shown in Figure 1. In the operation of CNN, an input image (Input Image) will be input into the Convolution Layer for convolution operation, as shown in Figure 2. Some filters used to extract features, called Feature Detectors, are used to perform convolution operations with the input image. The size and number of feature detectors can be customized.

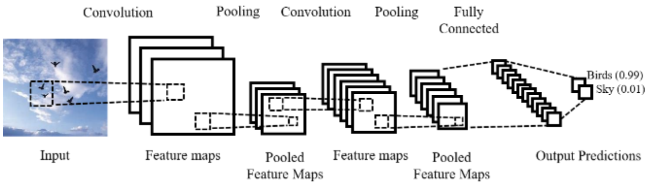

Figure 1: Convolutional Neural Network (CNN).

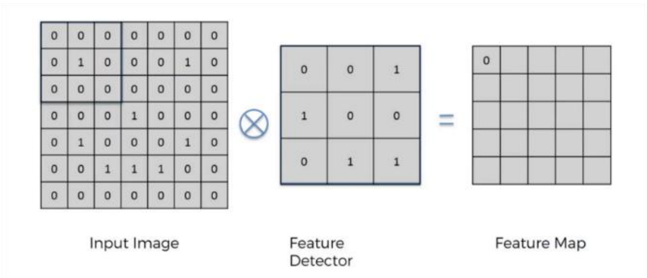

Figure 2: Convolution Layer

The convolution operation is to perform the matrix inner product sum operation on a block of the original image and the randomly generated feature detector to generate each grid value in the Feature Map. By analogy to complete the convolution operation of the entire input picture, a complete Feature Map will be generated. The number of Feature Maps depends on the number of masks originally preset, so the output of the convolutional layer will be multiple Feature Maps. The purpose of Feature Detector is to help extract some feature values in the image, that is, to find the outline of the object in the image.

Figure 3: Pooling Layer

Next, perform the Pooling process for all feature maps, as shown in Figure 3. Currently Max Pooling is the most commonly used method of Pooling. Pick a custom-sized matrix from the feature map and keep its maximum value, while discarding the rest of the values. Finally, multiple Pooled Feature Maps are generated. The purpose of this is to select out blocks with obvious characteristics.

Next, perform the Flattening operation on several pooled feature maps, that is, convert the two-dimensional array of the matrix into a one-dimensional array, as shown in Figure 4. After entering the Fully Connected Layer, the flattened one-dimensional array can be converted into a primary building block, then the most basic neural network can be constructed.

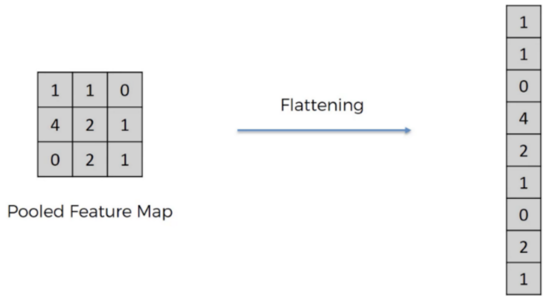

Figure 4: Fully Connected Layer.

Nowadays, many neural model architectures are defined according to CNN. YOLOv3 [3] uses the Darknet-53 neural network model framework extended from the Residual Network model [4], which contains 53 layers of convolutional layers. The object detection method of YOLOv3 is based on the improved method of YOLOv2 [5]. It uses Feature Pyramid Networks (FPN) technology. From the original feature maps with a single layer of 13*13, it has been changed to feature maps with multiple layers of 13*13, 26*26, and 52*52. Using the anchor box algorithm, each grid area in each layer of feature maps is matched with 3 anchor boxes to predict the bounding boxes. In this way, the ability to detect small objects can be improved. The method of object detection is to calculate a confidence score to evaluate the accuracy of the bounding box to predict the object.

In this paper, we choose YOLOv3-tiny as the architecture for real-time bird flock detection, as shown in Figure 5. This architecture is based on the YOLOv3 architecture and removes some feature layers, making YOLOv3-tiny a lightweight network framework.

Layer	Type	Filters	Size/Stride	Input	Output
0	Convolutional	16	3 × 3/1	416 × 416 × 3	416 × 416 × 16
1	Maxpool		2 × 2/2	416 × 416 × 16	208 × 208 × 16
2	Convolutional	32	3 × 3/1	208 × 208 × 16	208 × 208 × 32
3	Maxpool		2 × 2/2	208 × 208 × 32	104 × 104 × 32
4	Convolutional	64	3 × 3/1	104 × 104 × 32	104 × 104 × 64
5	Maxpool		2 × 2/2	104 × 104 × 64	52 × 52 × 64
6	Convolutional	128	3 × 3/1	52 × 52 × 64	52 × 52 × 128
7	Maxpool		2 × 2/2	52 × 52 × 128	26 × 26 × 128
8	Convolutional	256	3 × 3/1	26 × 26 × 128	26 × 26 × 256
9	Maxpool		2 × 2/2	26 × 26 × 256	13 × 13 × 256
10	Convolutional	512	3 × 3/1	13 × 13 × 256	13 × 13 × 512
11	Maxpool		2 × 2/1	13 × 13 × 512	13 × 13 × 512
12	Convolutional	1024	3 × 3/1	13 × 13 × 512	13 × 13 × 1024
13	Convolutional	256	1 × 1/1	13 × 13 × 1024	13 × 13 × 256
14	Convolutional	512	3 × 3/1	13 × 13 × 256	13 × 13 × 512
15	Convolutional	255	1 × 1/1	13 × 13 × 512	13 × 13 × 255
16	YOLO				
17	**Route 13**				
18	Convolutional	128	1 × 1/1	13 × 13 × 256	13 × 13 × 128
19	Up-sampling		2 × 2/1	13 × 13 × 128	26 × 26 × 128
20	**Route 19 8**				
21	Convolutional	256	3 × 3/1	13 × 13 × 384	13 × 13 × 256
22	Convolutional	255	1 × 1/1	13 × 13 × 256	13 × 13 × 256
23	YOLO				

Figure 5: YOLOv3-tiny.

https://doi.org/10.31871/IJNTR.7.12.13
International Journal of New Technology and Research (IJNTR)
ISSN: 2454-4116, Volume-7, Issue-12, December 2021 Pages 18-22

B. NVIDIA Jetson Nano

NVIDIA Jetson Nano[6] is an embedded single-board computer equipped with both CPU and GPU. The default operating system executed by the CPU is Ubuntu Linux. There are 4 USB 3.0 ports in the hardware architecture, which can meet the requirements of various USB devices on the market. Jetson Nano is equipped with a GPU produced by NVIDIA, which can conveniently do high-performance visual processing. In addition, it also contains 40 GPIO pins that can be used for external peripheral module control. It is a development board specially designed for the AIoT industry. Its product specifications are shown in Figure 6.

GPU	NVIDIA Maxwell architecture with 128 NVIDIA CUDA® cores
CPU	Quad-core ARM Cortex-A57 MPCore processor
Memory	4 GB 64-bit LPDDR4, 1600MHz 25.6 GB/s
Storage	16 GB eMMC 5.1
Video Encode	250MP/sec 1x 4K @ 30 (HEVC) 2x 1080p @ 60 (HEVC) 4x 1080p @ 30 (HEVC) 4x 720p @ 60 (HEVC) 9x 720p @ 30 (HEVC)
Video Decode	500MP/sec 1x 4K @ 60 (HEVC) 2x 4K @ 30 (HEVC) 4x 1080p @ 60 (HEVC) 8x 1080p @ 30 (HEVC) 9x 720p @ 60 (HEVC)
Camera	12 lanes (3x4 or 4x2) MIPI CSI-2 D-PHY 1.1 (1.5 Gb/s per pair)
Connectivity	Gigabit Ethernet, M.2 Key E
Display	HDMI 2.0 and eDP 1.4
USB	4x USB 3.0, USB 2.0 Micro-B
Others	GPIO, I²C, I²S, SPI, UART
Mechanical	69.6 mm x 45 mm 260-pin edge connector

Figure 6: The specification of NVIDIA Jetson Nano.

C. LoRaWAN

LoRaWAN is a low-power long-distance wireless transmission wide area network (Lower-Power Wide Area Network, LPWAN) [7]. Its underlying modulation uses LoRa low-power long-distance wireless modulation technology. The foundation of its network architecture is based on a star topology, as shown in Figure 7. The gateway is an intermediate bridge between the terminal equipment and the back-end server, and is responsible for transferring information between them. LoRaWAN uses a lightweight protocol design to reduce equipment costs. It also uses license-free wireless frequency bands to reduce network construction costs. LoRaWAN is very suitable for communication in a wide range of fields, with a range of up to 10 kilometers, and can even be used in indoor and underground environments. Such a protocol is very suitable for the development of the Internet of Things.

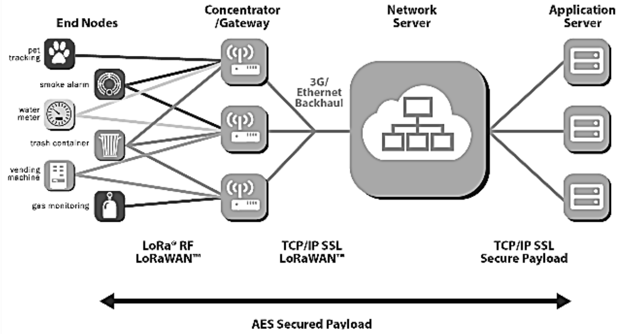

Figure 7: Network architecture of LoRaWAN.

The LoRaWAN network is a star topology structure, including End Nodes, Concentrators/Gateways, Network Server and Application Server. End nodes send messages to the gateway through the LoRa wireless channel, and the gateway then forwards the message to the network server through the back-end line (such as 3G, Ethernet, etc.)

D. MQTT

MQTT (Message Queuing Telemetry Transport) [8] is a widely used IoT message exchange protocol. MQTT was originally designed to be able to send and receive and process messages under the conditions of narrow bandwidth and low power consumption. The Broker is used as a bridge for message communication, and the Publish/Subscribe mode is adopted to allow clients to forward the Topic through the Broker. The transmission layer of the message uses the TCP protocol.

E. Node-Red

Node-RED was originally developed by IBM [9] as a graphical web service editor that connects hardware devices, APIs, and online services. Its characteristic is that programmers can connect various "nodes" in the editor to form desired application services. After the deployment process, the Node.js interpreter engine is used to execute the JavaScript code in each node in order to achieve the purpose of quickly realizing Web application services. Node-RED also has built-in rich dashboard components, which can generate visual charts. In addition, Node-RED also supports various Internet data transmission protocols, such as HTTP, MQTT, etc., and supports multiple database manipulation functions, such as MariaDB, PostgresSQL, and so on. It is quite suitable for development and use as a back-end application service of the Internet of Things. The prototype system in this paper uses Node-Red as the back-end development tool.

III. System Architecture and Prototype Implementation

A. System Architecture

The architecture of the AIoT intelligent bird repellent system proposed in this paper is shown in Figure 8. It contains the following components:

- <u>Smart Bird Flock Detection Camera</u>: It is installed in an appropriate position on the farmland, monitors a specific area, automatically recognizes the harmful birds that appear, and directly controls the ultrasonic bird repellent within the range to send out ultrasonic waves to drive away the birds. This camera contains an embedded artificial intelligence bird flock identification system. Through the collection of pictures of various bird flocks, a neural network model that can identify the flocks of birds is trained, and then converted into model files for inference, and imported into this system for operation.
- <u>Ultrasonic Bird Repeller</u>: It can be remotely controlled to continuously emit ultrasound waves of a specific frequency that cause discomfort for birds to drive away harmful birds on site. This device is also powered by solar energy so that it can be installed in any place freely.

- LoRaWAN Gateway: This paper adopts low-power long-distance wireless network, LoRaWAN, as the wireless communication channel between smart bird detection cameras and ultrasonic bird repellers on farmland. It can also be used to exchange various realtime/non-realtime data or control messages between smart bird detection camera and the backend application server.
- Machine Learning Model Construction Server: Since the construction of the bird flock identification model requires a lot of calculations, it is necessary to rely on a server with GPUs to build a suitable model in a short time.
- Bird Repellent Application Server: This server will perform status monitoring, data collection, and statistical analysis of the entire bird repellent system. It also provides external APIs, which can be used to develop user-side APPs or web pages.

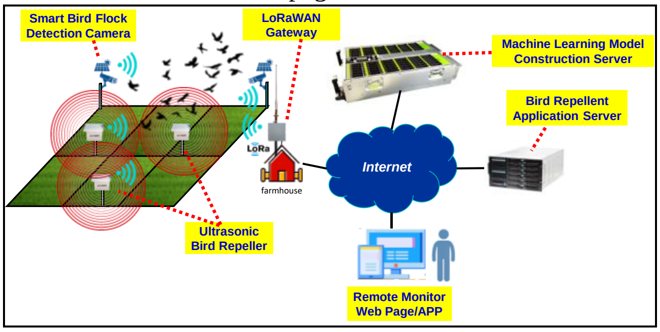

Figure 8: The proposed system architecture.

This system will be set up in the farmland, through the camera to capture the image to identify whether there are birds moving to the farmland, and then send messages through the LoRaWAN gateway to control the ultrasonic bird repellent to generate ultrasonic waves to friendly repel the birds. While repelling birds, the gateway can send relevant information about system operation to the back-end bird repellent application server through the MQTT protocol and store it in the database. Back-end users can interact with the server through remote monitoring pages or mobile apps to obtain information about system operation, and can adjust the configuration settings of the system to make the operation of the system more in line with the needs of users.

B. Prototype Implementation

The GPU server used to train the bird detection model in this paper consists of an E5-2620v4 CPU, 96GB RAM, and three TX 2080 TI GPUs, and uses Ubuntu Linux as the operating system. On this server, the YOLOv3-tiny network architecture is used for migration learning, and a model that can detect bird objects is trained. The model was then run on NVIDIA Jetson Nano. The Smart Bird Flock Detection Camera in Figure 8 is composed of a Jetson Nano equipped with a camera, a LoRaWAN wireless communication module, and a solar power kit, as shown in Figure 9.

On NVIDIA Jetson Nano, we use the Python programming language to write code. Keras package [12] is used to import the trained model, and a camera is installed on the Jetson Nano to continuously capture images in an uninterrupted manner. Each image is input to the trained model for bird flock detection. When Jetson Nano continuously recognizes that there are birds in the three images and the center of gravity of the birds is moving downward, it remotely controls the ultrasonic bird repellent device in the farmland through the LoRaWAN wireless network to make it emit ultrasonic waves. At the same time, useful information is transmitted to the back-end application server through the MQTT protocol. In the application server, Node-RED is used to develop various user application services. Figure 10 shows the actual prototype system installed in farmland. Note that the Ultrasonic Bird Repeller is also powered by solar power kit.

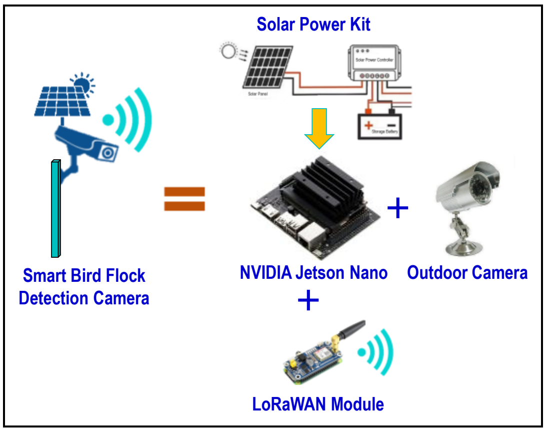

Figure 9: Implementation of Smart Bird Flock Detection Camera.

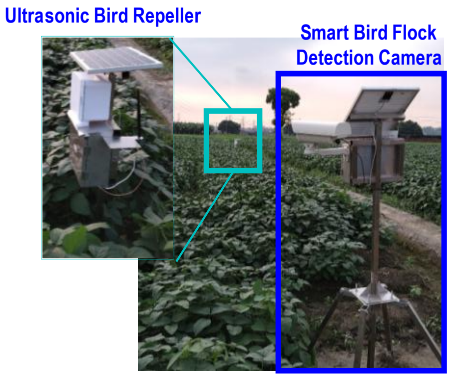

Figure 10: The actual prototype installed in farmland.

IV. ISSUES ABOUT BIRD FLOCK DETECTION AND REPELLING

This paper uses the YOLO real-time object detection technology to train the model, following the steps in [10] for experiments. We obtained 2240 images containing birds and their annotation files from COCO dataset (2014) [11], and discarded some bird images that are rarely seen in farmland. Finally, 810 pictures were retained as a data set for training. In addition, we also added 200 bird flock flying images taken from the Internet and on-site to the training data set. Next, we run the code provided by the official website on the GPU server to train the model and generate the final model file. Then we use the trained model to make inferences on the test pictures to verify the effectiveness. The achieved mean average precision is about 78% with 6000 epochs.

After the model is verified, we tested the model on the NVIDIA Jetson Nano. Image file of Jetpack SDK 4.2 is installed on Jetson Nano. The system contains Ubuntu (version 18.04) operating system, CUDA (version 10.0), and python (version 3.6.9) software. The main packages required to execute image inference code are Keras (version 2.2.4) and tensorflow-gpu (version 1.13.1).

An AVC test movie with length of five minutes and seven seconds and screen width of 1280 (px) * 720 (px) was used for verification. We use the OpenCV package in the code to import this video, and the number of frames extracted from the video is 9230. Then use the Keras package to import the trained model into the program, and make inferences on the images in sequence. The fps (Frame per Second) observed during the experiment is about 4-6 frames. The experimental result screen is shown in Figure 11. It can be found that the recognition effect of birds closer to the camera is better than that of distant birds. This is because the contours of distant birds are less clear such that it is more difficult to recognize. Due to the insufficient feature description capabilities of small objects, the detection performance is poor. In the future, a neural network framework with a larger number of layers can be selected for model training, which may solve such problems.

As for the actual effect of the ultrasonic bird repellent on repelling the birds, it needs a long period of field experiment to collect the complete response of the bird flock to the ultrasonic wave. Only after this can the actual modification and optimization of the bird repelling system be made.

Figure 11: Screenshot of bird flock detection experiment.

V. Conclusion

In recent years, with the maturity of deep learning technology, the problem of image recognition has become easier to solve, and it has begun to be applied to the intelligence of various industries. This paper attempts to combine artificial intelligence with Internet of Things technology, and proposes A system that automatically recognizes the flocks of birds flying into the farmland and activates the ultrasonic bird repellent by remote control to drive away the birds that are harmful to crops. This system will increase crop yields, increase farmers' incomes, and attract young people to return to their hometowns to engage in agriculture. It creates an easy, convenient, and systematic intelligent agricultural bird repelling system.

In the future, the bird group recognition model can be optimized, such as adding pictures of birds appearing in the actual farmland to the data set, or adding pictures of other categories, such as airplanes, kites, etc., for training. In this way, the misjudgment during identification may be resolved, and the effect of bird detection may be improved.

In addition to the increase in the amount of data and the number of categories, different types of neural network frameworks can also be imported to generate new training models. Then evaluate what kind of model framework is more suitable for this system, so that the detection on small objects will be more accurate.

Due to advances in hardware, artificial intelligence technology will be able to conduct experiments and analysis more quickly in the development. The development of artificial intelligence combined with cross-domain technology will be a major trend in the future.

References

[1] [1] J. Redmon, S. Divvala, and R. Girshick, "You only look once: Unified, real-time object detection," *IEEE conference on computer vision and pattern recognition (CVPR)*, 2016.

[2] H.-C. Shin, et al.," Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, vo. 35, no. 5, 2016.

[3] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *IEEE conference on computer vision and pattern recognition (CVPR)*, 2016.

[5] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," *IEEE conference on computer vision and pattern recognition (CVPR)*, 2017.

[6] Nvidia Jetson Nano, Accessed on Dec. 5, 2021 [Online]. Available: https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-nano/

[7] A. Lavric and V. Popa, "Internet of things and LoRa™ low-power wide-area networks: a survey," *2017 International Symposium on Signals, Circuits and Systems (ISSCS)*, 2017.

[8] MQTT, Accessed on Dec. 5, 2021 [Online]. Available: https://mqtt.org/.

[9] Node-RED, Accessed on Dec. 5, 2021 [Online]. Available: https://nodered.org/.

[10] YOLO v3: Real-Time Object Detection. Accessed on Dec. 5, 2021 [Online]. Available: https://pjreddie.com/darknet/yolo/.

[11] COCO Dataset. Accessed on Dec 5, 2021 [Online].Available: https://cocodataset.org/#home.

[12] Kares. Accessed on Dec 5, 2021 [Online].Available: https://keras.io/.