A Novel Resource-Constrained Insect Monitoring System based on Machine Vision with Edge AI

Amin Kargar, Walsh Scholar
Tyndall National Institute
University College Cork
Cork, Ireland
amin.kargar@tyndall.ie

Mariusz P. Wilk
Tyndall National Institute
University College Cork
Cork, Ireland
mariusz.wilk@tyndall.ie

Dimitrios Zorbas
Department of Computer Science
Nazarbayev University
Nur-Sultan, Kazakhstan
dimzorbas@ieee.org

Michael T. Gaffney
Horticulture Development Department
Teagasc Ashtown Food Research Centre
Dublin, Ireland
Michael.gaffney@teagasc.ie

Brendan O’Flynn
Tyndall National Institute
University College Cork
Cork, Ireland
brendan.oflynn@tyndall.ie

Abstract— Effective insect pest monitoring is a vital component of Integrated Pest Management (IPM) strategies. It helps to support crop productivity while minimising the need for plant protection products. In recent years, many researchers have considered the integration of intelligence into such systems in the context of the Smart Agriculture research agenda. This paper describes the development of a smart pest monitoring system, developed in accordance with specific requirements associated with the agricultural sector. The proposed system is a low-cost smart insect trap, for use in orchards, that detects specific insect species that are detrimental to fruit quality. The system helps to identify the invasive insect, Brown Marmorated Stink Bug (BMSB) or Halyomorpha halys (HH) using a Microcontroller Unit-based edge device comprising of an Internet of Things enabled, resource-constrained image acquisition and processing system. It is used to execute our proposed lightweight image analysis algorithm and Convolutional Neural Network (CNN) model for insect detection and classification, respectively. The prototype device is currently deployed in an orchard in Italy. The preliminary experimental results show over 70 percent of accuracy in BMSB classification on our custom-built dataset, demonstrating the proposed system feasibility and effectiveness in monitoring this invasive insect species.

Keywords—Machine Vision, Image processing, Deep Learning, Edge AI, Integrated Pest Monitoring, Food Security

I. INTRODUCTION

The global food supply is greatly affected by pest insects, which contribute to food shortages and reduced quality of fruit and vegetables. Pest Insect monitoring is one of the critical aspects of an IPM strategy that helps growers to control the pest population in the fields and develop strategies to prevent and reduce crop damage [1]. At present, trap monitoring mainly involves the use of commercial traps to attract insects to land on their sticky surface; through the use of pheromones. A human operator manually counts the number and variety of insects captured on the trap at specific time intervals. This then helps the farmers make informed decisions as part of an integrated IPM strategy and mitigate the negative effect of the pest insects’ presence through an optimal selection of pest control techniques.

The farming community faces the dangers of the rapid spread of destructive non-native pest insect species in different parts of the world. BMSB is one such insect. It is an invasive shield bug native to East Asia (China and Japan) that was first seen in Italy in 2012 and just two years later, increased damage to fruit crops was observed [2]. It continues to spread northwards and the presence of two adult males were reported in the UK in 2020 [2]. Therefore, increased monitoring approaches, particularly automated monitoring approaches are needed to deal with pest insects, in particular invasive species such as the BMSB. In this context, a collaborative EU research project titled, HALY.ID, is aimed to address this challenge through the development of new technologies focused on the monitoring of the BMSB insect [3].

Currently, farmers use commercial, pheromone loaded, sticky traps in their crops to attract target insect pest species. These traps are periodically checked manually for insects to establish the type and estimated quantity of the insect present in the field or orchard [4]. This is a time consuming task which can require a lot of labour and often requires a high level of entomological expertise for correct identification that could and should be automated. The research community has recognised this challenge. To this end, researchers have proposed several camera-based monitoring systems. These systems take images of a trap and then carry out image analysis to detect the presence of pests investigated [5]. These systems improve pest monitoring performance, but they have limitations as follows. Generally, only one side of the trap is monitored whereas commercial insect traps are generally two-sided [6][7]. The high cost of such systems can be a significant drawback in agricultural deployments where scalability is critical [5]. Image processing and Machine Learning (ML) algorithms are computationally intensive. For this reason, some of the proposed methods require either high processing power for local computing or cloud-based solutions which lead to higher cost and power consumption.

In this paper, a novel stationary smart pest trap system for detecting the BMSB is proposed. It is a low-cost, automated, edge device that consists of a camera for taking images from the trap, and a servo motor for rotating the trap so that images from both sides of the trap can be captured and analysed separately. It uses a low-cost and low-power Microcontroller Unit (MCU) that performs all data processing on the edge device itself, including image processing and ML algorithms. The edge device detects and identifies the BMSB in images taken of the traps, and only the relevant results are sent to the cloud system. In fact, in the image processing phase, just the suspected areas that are potentially filled with a BMSB are extracted and cropped for classification and transmission. Thus, instead of the entire image, some small sub-images are utilized for classification and transmission. This significantly
reduces the bandwidth and storage requirements in what are typically resource-constrained systems. This device also uses standard double-sided commercial sticky traps that growers normally use for manual assessment of insect populations [8]. This device was developed as part of the HALY.ID research project. Fig. 1 shows the proposed system operational within a fruit orchard in the Emilia-Romagna region in Italy.

This paper is organized as follows. Related works are discussed in Section II. In Section III, the proposed method is described. Experimental results are shown and discussed in Section VI. Section V concludes this research study.

II. RELATED WORK

Remote insect pest monitoring is an active research area. The research community recognises the potential of technology in this application space. For example, in [9] the authors proposed a remote greenhouse pest monitoring system. Their system is based on Wireless imaging and Sensor Nodes (WiSN) which are composed of a Raspberry Pi (RPi) 3 and a RPi camera that are placed in front of a yellow sticky trap. In this configuration, the captured image is transferred to a server for analysis via a 4G modem communication link. The authors in [10] proposed a remote insect trap monitoring approach using four-layer Internet of Things (IoT) to construct the remote trap insect monitoring system, as well as Deep Learning frameworks for classification. This is a server- or cloud-based system with four IoT layers. This approach increases the cost and bandwidth requirements. The system described is capable of monitoring only single-sided traps.

Recently, edge-based systems have attracted much attention because of their advantages over cloud- or server-based systems, including lower latency, lower data transfer volumes and network traffic, and the ability to deploy in remote locations with limited connectivity [11]. Many of the recent smart pest monitoring systems found in the literature are using a RPi as the main data processing unit, i.e. the control and data processing functions are performed by it. In [7], the authors proposed an automated light trap to monitor moths. This system uses a RPi 4 to execute for computing, an Ultra HD web camera, a light ring, and a UV light source. In the study described, the trap, a white sheet with a sugar coating, is placed in front of the camera. The UV light is used to attract insects at night. The camera captures images from just one side of the trap. The authors in [6] proposed a vision-based insect counting and identification system and implemented it on a RPi 2 Model B with a camera mounted in front of the trap. Their proposed detection algorithm is complex as it needs approximately 5 minutes for one cycle of insect detection and identification. A similar approach was described in [12]. A RPi 3 with an Intel Neural Compute Stick was used to detect Codling Moths with the VGG16 Deep Neural Network (DNN) model. Both of these systems used a RPi single-board computer with a general-purpose microprocessor that consumes multiple times the power of an MCU-based board.

In addition, there are a number of studies such as [13] and [14] which attempt to improve the detection and recognition algorithms and as such are not concerned about the hardware, cost and other factors that are important in the considered application space, i.e. agricultural settings.

Current literature shows a limited number of studies that focus on edge devices, i.e. computing is carried out near the

<table>
<thead>
<tr>
<th>Study</th>
<th>Title</th>
<th>Trap type</th>
<th>Edge Computing</th>
<th>Image processing unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nam et al. [14]</td>
<td>Pest detection on Traps using Deep Convolutional Neural Networks</td>
<td>Single sided</td>
<td>No</td>
<td>CPU/GPU</td>
</tr>
<tr>
<td>Bjerke et al. [7]</td>
<td>An Automated Light Trap to Monitor Moths (Lepidoptera) Using Computer Vision-Based Tracking and Deep Learning</td>
<td>Single sided</td>
<td>Yes</td>
<td>CPU (RPi)</td>
</tr>
<tr>
<td>Zhong et al. [6]</td>
<td>A Vision-Based Counting and Recognition System for Flying Insects in Intelligent Agriculture</td>
<td>Single sided</td>
<td>Yes</td>
<td>CPU (RPi)</td>
</tr>
<tr>
<td>Brunelli et al. [12]</td>
<td>Energy Neutral Machine Learning Based IoT Device for Pest Detection in Precision Agriculture</td>
<td>Single sided</td>
<td>Yes</td>
<td>CPU (RPi)</td>
</tr>
<tr>
<td>Proposed system</td>
<td></td>
<td>Double Sided</td>
<td>Yes</td>
<td>MCU (OpenMV - STM32)</td>
</tr>
</tbody>
</table>
source of data (e.g. fruit orchards). Those works that do perform the computing function at the edge do not focus on low-cost, resource-constrained, MCU-based architectures. Furthermore, existing literature reports on works that monitor only single-sided traps while double-sided traps are also mainly used. Table I shows the comparison of mentioned related studies with this work. This paper proposes a novel edge device with Edge Artificial Intelligence (AI) that uses a low-cost MCU for automated insect monitoring with double-sided commercial sticky traps.

III. PROPOSED METHOD

This work is focused on the development of a stationary edge device, also referred to as an end-device or sensor unit. The edge device is a low-cost, low-power, MCU-based system that incorporates a low-cost camera. It is capable of taking images and processing them. The system can be IoT enabled so that the output could also be sent to the cloud via a telecommunications link; to provide data for the growers’ IPM strategy. The proposed general system block diagram is shown in Fig. 2.

In order to meet the requirement of being able to automatically identify and count the insect pest of interest, i.e. the BMSB, the proposed device should be integrated with a standard commercially available insect trap that the growers normally use in manual monitoring of their orchards. The proposed system must meet certain requirements to be attractive to users in the agricultural sector, such as low cost, low power consumption, and effectiveness at detecting the target insects [5].

To this end, an IoT sensor system that incorporates both a trap and a camera was developed. The camera can take images of both sides of the trap at regular time intervals using a servo mechanism. The MCU can process the images to detect and identify the insects. The proposed system details are elaborated in the following sub-sections.

A. Hardware platform

One of the main factors that drive many of the system requirements for the final prototype is the camera specification. The balance between camera resolution and energy consumption of the overall system (often battery powered for remote deployment) should be considered. Similarly, the MCU must be powerful enough to process the images but not too powerful so as to avoid unnecessarily increasing the cost.

Based on the identified requirements, the OpenMV Cam STM32H7 Plus module was selected for this study [15].

OpenMV is an MCU-based image acquisition board with an STM32H743II ARM Cortex M7 processor with 32 MB of SDRAM, 1MB of SRAM, 32 MB of external flash and 2 MB of internal flash running at 480 MHz. This is a small, inexpensive, and low-power system that incorporates an OV5640 imaging sensor that can acquire a 5MP resolution image. In the system developed, images of 2MP resolution are used to achieve lower power consumption while still being able to capture and identify the smallest features of the BMSB, such as antennas and legs.

B. System Design

The device must not have an adverse impact on the trap’s fundamental requirement, its attractiveness to insects. The trap attractiveness to insects could be reduced if the device is not designed appropriately. For example, the effectiveness of the trap can be impaired based on its positioning or the amount of light it is exposed to and this needs to be one of the system-level design parameters considered and discussed below [5].

1) Illumination

Vision sensors are designed to capture optical data that their sensor arrays are exposed to. Illumination plays a critical role in these sensors’ effectiveness. The less controlled the environment is, the more difficult it is to extract the desired information from the images. One way to control the environment is to introduce artificial illumination to the scene, thus, introducing a level of control over the environment in which the vision sensors operate. This approach is also widely used in machine vision applications with an associated power drain which needs to be considered [16][17]. The impact of illumination is clearly visible in Fig. 3. As is clear from this figure, illumination has a significant impact on the captured image quality and can directly affect image processing and deep learning performance. This allows for image acquisition at night, i.e. minimising other sources of illumination.

2) Servo Motor Integration

Commercial insect traps are generally double-sided, i.e. insects are captured on both sides. This can pose a significant challenge for a camera-based sensor. Firstly, both sides of the trap must be captured and analysed. Secondly, the camera must not be occluding either of the trap’s two sides or else its effectiveness could be adversely affected.

The above problems can be solved by introducing a novel approach that utilises a low-cost servo motor. The servo motor can hold and rotate the trap. When the camera is inactive, the trap’s surface is parallel to the camera’s optical axis, thus avoiding occlusions, as shown in Fig. 1. When the pictures are to be captured, the servo motor rotates the trap so that the camera can capture images from both sides of the trap.
C. Image Processing System Algorithm

The proposed system algorithm is shown in Fig. 4. This algorithm first attempts to extract the Region of Interest (ROI) or multiple ROIs from the captured image of the trap. It then runs a CNN for insect classification. The area of the ROI is such that the insect of interest almost entirely fills its area. In this way, only small portions of captured images (ROIs with suspected BMSBs) are used for further processing and transferring, which reduces system requirements significantly. The steps of the proposed algorithm are also shown in Fig. 5 which depicts the image processing pipeline using an example input image; from the raw input image to the end result.

It must be noted that the algorithm under development must be as “light” as possible so as to run on the edge device which is resource-constrained in terms of processing power, memory, and energy. In this context, in the first phase (pre-processing), two simple image processing techniques are used to prepare the raw input image (see Fig. 4). These techniques are as follows:

1- Gaussian Blur: The Gaussian blur/smoothing filter with the 3x3 kernel is applied to the input image to reduce the high-frequency noise component and details [18].

2- Since there is a strong difference in colour and intensity between black (insects) and the white background colour (trap), histogram-based thresholding is utilized to divide the image into two classes, i.e. background and foreground. To this end, Otsu’s method [19] is used to find the optimal threshold value.

Then, a simple blob detector (available in OpenMV libraries [20]) is utilized to detect and extract possible ROIs from the image. The blob detector finds a blob based on two main parameters including Pixel Area and Circularity. Pixel Area is set based on the minimum and maximum size of the studied insect which is approximately 12-17 mm in length [21] and Circularity is adjusted so as to approximate the insect’s shape to an Elliptical Shape.

Finally, just the extracted ROIs are passed to a lightweight deep learning model for insect classification. In this phase, the extracted small images are fed to the deep learning model to label with the input ROI with either a Yes or No depending on whether there is a BMSB present in the extracted image or not. As mentioned before, the deep learning model should be as small as possible so that it can fit in the MCU’s memory. For this reason, parameter quantisation was performed on the model which converts floating-point numbers to other data types. In this study, all the weights and biases were converted to 8-bit integers [-128,127] instead of float32 which significantly reduces memory and computing requirements. The (int8) quantised CNN model is proposed with the model size of just 57kB to show the feasibility of such systems.

The model architecture is shown in Fig. 6. This is a CNN-based model comprising two convolutional layers with 32 and 16 channels and a kernel size of 3x3. Each convolutional layer is followed by a max-pooling layer. In the final part, there is a fully-connected layer connected to the output layer (softmax). This model totally has 51,730 parameters and the TensorFlow framework was used to design, train and quantise the model.

IV. EXPERIMENTAL RESULTS

In this section, the performance of the proposed prototype is evaluated based on preliminary results. Note, that the proposed algorithm was implemented on the OpenMV MCU
platform, and all the computation and analysis were carried out on this board. Moreover, it must be noted that this task was carried out before the field deployment. Thus, we did not have access to real-world images of the BMSB in the orchards. For this reason, the initial dataset was manually created in laboratory conditions using dead (deep-frozen) BMSB specimens that had been collected in the field by our project collaborators. The specimens were manually glued to the trap instead, as shown in Fig. 5, simulating a deployment capture scenario.

As mentioned in the previous section, in the first phase, the proposed system finds and extracts ROIs from the whole input image. Our laboratory experiments (see Fig. 7) showed that the proposed method for ROI extraction was sufficient for successfully detecting all desired blobs based on Pixel Area and Circularity parameters.

Regarding the deep learning part, Adam and Binary Cross-entropy are chosen as the optimizer and loss function, respectively. The kernel size of 3x3, same padding with strides 1 are set for 2-D convolutional layers, and their activation functions are Rectified Linear Unit (ReLU). The output layer activation function is Softmax and a dropout layer with the rate of 0.25 is used to tackle the overfitting during the training phase. Fig. 8 shows the model accuracy and loss curves. Based on this figure, model accuracy becomes steady at around 97% after 15 epochs.

The model performance is reported in Table II, i.e. the binary classification. The model with parameters in float32 data type achieved 96.7% accuracy and over 97% for recall, precision and F-score with the size of 210 kB. On the other hand, the quantised model with parameters in the int8 data type achieved lower accuracy, i.e. 88.3%. However, the model size was significantly reduced; by a factor of approximately 3.7, and its size in memory was reduced to 57 kB.

The quantised model, i.e. that with parameters in the int8 data type, was embedded in the MCU’s firmware. The proposed prototype system was (and still is – September 2022) deployed in an orchard in the Emilia-Romagna region in Italy during the summer months. Table III reports the model classification performance from this deployment. The data collection was carried out over a period of two months during which the BMSB was present in the orchard. Table III shows that the classification accuracy in real-world field deployment was lower as compared to the corresponding experiments in the lab, i.e. shown in Table II. This result was expected as the images acquired in the orchard were contaminated noise and

<table>
<thead>
<tr>
<th>Table II</th>
<th>PROPOSED CNN MODEL LABORATORY RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy (%)</td>
</tr>
<tr>
<td>Float32</td>
<td>96.7</td>
</tr>
<tr>
<td>Int8</td>
<td>88.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table III</th>
<th>PROPOSED CNN MODEL FIELD DEPLOYMENT RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy (%)</td>
</tr>
<tr>
<td>Int8</td>
<td>70.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P = 1, T = 1</th>
<th>P = 1, T = 0</th>
<th>P = 1, T = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>P = 0, T = 0</td>
<td>P = 0, T = 0</td>
<td>P = 0, T = 0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P = 1, T = 0</th>
<th>P = 1, T = 1</th>
<th>P = 1, T = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P = 1, T = 0</th>
<th>P = 1, T = 1</th>
<th>P = 1, T = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

Fig. 7. a) Original input image from artificial insects, b) Image pre-processing, c) Blob detection on input image in greyscale.

Fig. 8. a) Training and validation Accuracy, b) Training and validation Loss.

Fig. 9. Some results of deployed prototype.
by-catch (random non-BMSB insects on the trap) to a greater extent. In this regard, some of the classified images captured at the deployment site in Italy are shown in Fig. 9. There are some false positives or negatives in the field deployment results. For example, regarding the second image in the first row of Fig. 9, it was predicted as BMSB while it is not. This is because it has the same shape as the BMSB but it has a different colour and it is not in the used dataset for training. Also, a similar result occurred when BMSB lands unevenly on the trap. Hence, further improvement of the deep learning model and training it with images captured in real-world conditions can further increase the accuracy.

These results show the output of the early stage, edge-compatible, image processing algorithms. The proposed method has shown promising results. It demonstrates the potential that this system has in the considered application space, i.e. automated pest insect monitoring as part of an IPM strategy.

V. CONCLUSIONS AND FUTURE WORK

In this study, a novel prototype system for automated insect pest monitoring with a commercial trap was proposed. It is a low-cost edge device with Edge AI for detecting and identifying insects of interest, in this case the BMSB. An OpenMV Cam H7 Plus was selected for the proof-of-concept prototype in this study to satisfy the balance between cost, energy consumption and processing power. As this is a resource constrained MCU based board, a lightweight image processing algorithm is proposed to extract only ROIs with candidate blobs (suspected BMSB) for further processing instead of processing whole images, thus reducing system requirements. Then the extracted ROIs are fed to a lightweight CNN model for classification and BMSB inference. The proposed system was validated in field deployment in an orchard in Italy during the summer months and an evaluation of performance is underway. The preliminary results show promising results with BMSB classification accuracy of over 70 percent. It is an ongoing study. The accuracy of the Edge AI is expected to increase once the CNN updated model is retrained with a dataset containing images from real-world field deployment conditions. Also, it must be noted that as an edge-based system, the power consumption aspect is the next stage of this study, even though it is expected the energy consumption of our MCU-based system to be multiple times lower than that of the RPI-based solutions.

ACKNOWLEDGMENTS

This project is co-funded by the European Regional Development Fund (ERDF) under Ireland’s European Structural and Investment Funds Programmes 2014–2020. This work was carried out as part of the Haly.1D project – 2020EN508 funded by Ireland’s Department of Agriculture, Food and the Marine under Grant: 2020 Trans National ERA-Net. The first author is supported by a Walsh Scholarship funded by Teagasc, The Irish Food and Agriculture Authority. Aspects of this work have been supported by Science Foundation Ireland under Grant 12/RC/2289-P2-INSIGHT, 13/RC/2077-CONNECT, 16/RC/3835-VISTAMILK and 16/RC/3918-CONFIRM which are co-funded by the European Regional Development Fund (ERDF) under Ireland’s European Structural and Investment Funds Programmes 2014–2020.

REFERENCES


