Grape Farm Pest Bird Sound Detection System using Embedded Systems

Euhid Aman[1], Wen-Ren Yang[2]

[1]*Presidency University, Bangalore, India*
[2]*National Changhua University of Education, Changhua City, Changhua County - Taiwan*
**euhidaman@gmail.com*

Keywords: Artificial Neural Networks, Embedded Systems, Deep Learning, Sound Classification

Abstract

Crops are essential for fuelling economies, and agricultural fields require proper care to ensure food production. This project focuses on addressing the presence of pest birds, specifically mynah birds, in vineyards. The proposed approach utilizes sound signals to detect and identify the mynah birds in the vineyard. By analysing the surrounding environment's audio inputs, the system can accurately identify distinct sound patterns of mynah birds, using Deep Learning. Additionally, the project incorporates distance estimation techniques using ultrasonic sensors to calculate the bird's proximity from a fixed position within the vineyard. All the detection and estimation processes are performed on the RP2040 microcontroller, specifically the Cortex-M0+ 133MHz variant. After detection has been done, a robotic vehicle outfitted with red diode lasers can be dispatched to the desired location to chase the pest birds away while protecting the vineyards from undesired exposures and crop losses.

1. Introduction

Crop damage caused by birds as illustrated in Fig. 1, is a persistent challenge in agriculture, with traditional methods like nets and loud noises becoming ineffective due to bird habituation. To address this issue, it is crucial to develop flexible and adaptable bird expulsion strategies, including real-time bird detection. This research proposes a cost-effective and weather-resistant technology for real-time bird detection in vineyards, utilizing microcontrollers.

The primary objective of this project is to create a real-time bird detection system, specifically designed to locate mynah birds in grape fields. The proposed system utilizes Raspberry Pi – Pico microcontroller [1], as a cheaper and resilient solution for implementation on a large scale, considering the extreme temperatures and environmental conditions prevalent in the field.

The system incorporates various sensors, such as PDM-MEMS sensors [2] for sound detection, and HC-SR04 ultrasonic sensors [3] for accurate distance estimation. By analysing the sound signals in real-time from the surrounding environment, the system can identify the unique sound patterns of mynah birds, using a neural network model. The detected bird's location is then determined, and this information can be communicated to a central server system for implementing appropriate measures to repel the birds effectively.

Fig. 1. Damage caused by Mynah pest-birds in a vineyard.

2. Methodology

2.1 SOUND DATA INFERENCE

2.1.1 Sound Data Collection and Splitting: The baseline neural network model for sound classification will be trained using the ESC-50 dataset, comprising 2,000 environmental sound recordings across 50 classes. The 5-second audio files are sliced into 16,000-sample segments and filtered to remove silent portions. Additionally, to augment the dataset and enhance training, the original audio samples are strided every 4,000 samples. The ESC-50 dataset is divided into training, validation, and test sets based on the fold column, following the method outlined in TensorFlow's Transfer learning with YAMNet for environmental sound classification guide.

2..1.2 Baseline Model Creation: After separating the features from the audio data, the Keras API of TensorFlow will be used to build the model. The model is compiled with accuracy as its metrics, an Adam optimizer, and a sparse categorical crossentropy loss function. Additionally, early stopping, and dynamic learning rate scheduler callbacks were defined for training. The model calculated a loss of 39% and an accuracy of 24.44% under those circumstances. Finally, the baseline model must be preserved for additional sound categorizations and model optimizations.

2.1.3 Transfer Learning: To accurately detect mynah bird sounds, a custom dataset comprising mynah sounds must be utilized. This dataset is mixed with background noises from TensorFlow's speech commands directory (Fig. 2). Data augmentation techniques can be employed to enhance the existing mynah sounds dataset, such as adding white noises, incorporating random silences, and combining multiple audio signals. Segmenting the mynah sounds dataset into 1-second

https://doi.org/10.1049/icp.2023.3167

soundbites and merging it with the background noises improves dataset consistency. Once the dataset is augmented and expanded, it should be divided again for training, validation, and testing purposes. Subsequently, the head and tail of the baseline model need to be replaced to create a binary classifier specifically designed to detect mynah sounds. The model must then be retrained using the new and augmented dataset.

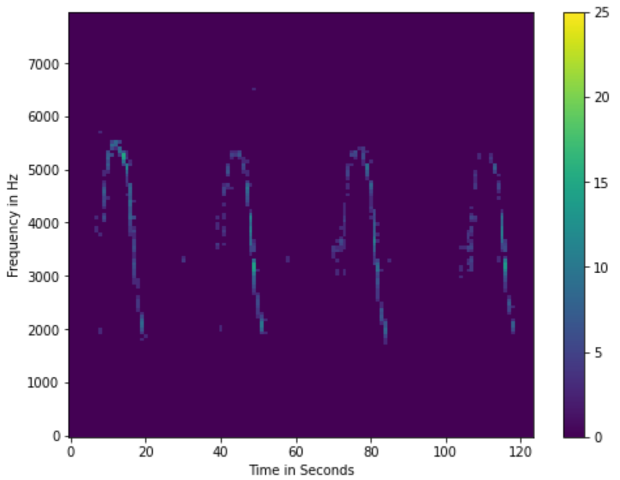

Fig. 2. Spectrogram of pest bird audio, augmented with white noise for training enhanced baseline model.

2.1.4 Quantization Aware Training and Model Compression: After training the mynah sound detection model, the subsequent phase involves quantizing the model for RP2040 microcontroller deployment. Quantization-aware training [4] is employed to reduce model weight size and unnecessary parameters, enhancing compatibility with lower-precision computations during actual use. In our study, the quantized model's total parameters reduced to 812 as shown in Fig. 3, from the original 14,973.

```
 Layer (type)                Output Shape              Param #
=================================================================
 input_1 (InputLayer)        [(None, 124, 129, 1)]     0

 resizing (Resizing)         (None, 32, 32, 1)         0

 normalization (Normalizatio  (None, 32, 32, 1)        3
 n)

 conv2d (Conv2D)             (None, 13, 13, 8)         520

 max_pooling2d (MaxPooling2D  (None, 6, 6, 8)          0
 )

 flatten (Flatten)           (None, 288)               0

 dropout (Dropout)           (None, 288)               0

 dense_1 (Dense)             (None, 1)                 289

=================================================================
Total params: 812
Trainable params: 809
Non-trainable params: 3
```

Fig. 3. Quantized model for deployment on Rpi-Pico

2.1.5 Model Deployment and Inference on RP2040: Upon compiling and building the model, it is flashed onto the Rpi-2040 microcontroller. This involves configuring the board's LEDs, setting up the TFLite library, and establishing the CMSIS-DSP based signal processing pipeline. The PDM-MEMS microphone is activated for real-time audio input.

The deployed model achieves an accuracy of 69.07% on the test data, effectively discerning the mynah bird's shrill sound amidst surrounding noises. Its robust filtering mechanism enables the model to differentiate the target sound from unwanted background noise, exhibiting its focused detection capabilities.

2.2 DISTANCE ESTIMATION

The system utilizes multiple sound sensors kept at various angles, to detect mynah or pest birds. When a bird is detected, past a certain threshold, an array of ultrasonic sensors estimate the bird's distance from the transmitter. The HC-SR04 ultrasonic sensor measures the time for sound to travel and calculates distance using the formula:

$$s = \frac{t}{2} \times \mathrm{c} \qquad (1)$$

where t is the round-trip time and c is sound speed (approx. 343m/s at room temperature). Then a vehicle fitted with red-diode lasers is sent to chase the birds away.

3. Conclusion

The Grape Farm Pest Bird Sound Detection System effectively identifies pest birds within grape fields using a Raspberry Pi Pico microcontroller, PDM MEMS microphone, and HC-SR04 ultrasonic sensor. This embedded project employs continuous field sounds for swift responsiveness, even with limited RAM (264kb). Real-time monitoring ensures prompt actions against bird infestations, diminishing the need for harmful pesticides. The proposed system was able to accurately distinguish mynah bird sounds from background noises with 69.07% precision amidst a bustling environment of other birds and environmental sounds. This illustrates the potential of embedded systems to tackle agricultural challenges, scalable to larger farms and adaptable for diverse pest bird detection. This paper offers insights into creating and deploying an inventive agricultural solution, underscoring embedded system’s effectiveness in real-world problem-solving.

4. Acknowledgements

This research is supported by Artificial Neural Network Applications, TEEP (Taiwan Experience Education Program), Taiwan Ministry of Education, project no. 1112506208.

4. References

[1] ‘Raspberry Pi Pico RP2040 Microcontroller Board’, https://www.okdo.com/p/raspberry-pi-pico/, accessed June 2023

[2] Da Silva, B., Braeken, A., Touhafi, A., "FPGA-Based Architectures for Acoustic Beamforming with Microphone Arrays: Trends, Challenges and Research Opportunities," Computers, vol. 7, no. 3, p. 41, 2018.

[3] Morgan, E. J., "HC-SR04 ultrasonic sensor," Nov 16, 2014.

[4] Y. Zhou, et al., "Adaptive quantization for deep neural network," Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, no. 1, 2018.