Saving Birds from Power Grids: A Deep Learning Embedded System for Sustainable Energy Infrastructure

Hesam Kavosi - Mohammad Javad Abdollahifard
Hesam Kavosi,Computer Vision and Remote Sensing Lab,Department of Electrical Engineering,Tafresh University, Tafresh, Iran hesamkavosi@gmail.com
Mohammad Javad Abdollahifard, Computer Vision and Remote Sensing Lab, Department of Electrical Engineering, Tafresh University, Tafresh, Iran mj.abdollahi@tafreshu.ac.ir

ABSTRACT

Energy sustainability remains one of the humanity's most pressing challenges. Electrical power systems, while essential for modern infrastructure, pose significant risks to wildlife, particularly birds. When birds approach to high-voltage power grids, they face fatal electrocution, leading to ecological harm and potential damage to electrical infrastructure. To mitigate this issue, we propose an embedded system that combines low-cost hardware with deep learning for real-time avian detection and deterrence. Our solution employs the YOLO-v2 neural network, implemented on a K210 embedded processor featuring dual 64-bit cores running at 400 MHz, achieving an inference speed of 45 frames per second. The model was trained on a curated dataset of 1,500 diverse bird images, optimized for network input, and achieved an accuracy of 89% during validation. This system demonstrates a practical, efficient, and scalable approach to reducing avian fatalities near power grids while maintaining grid reliability

Keywords: ***High-voltage transmission lines, Convolutional Neural Networks, Bird Detection, YOLO-v2 Network, Edge Computing***

1 INTRODUCTION

The power line network, which transmits energy from production centers to consumers, spans vast geographic areas, often at the cost of biodiversity. As infrastructure expands into natural habitats, birds increasingly adapt to electrical structures, using towers and poles as perching, resting, and nesting sites—especially in regions where traditional habitats like trees have diminished. This interaction poses significant risks: collisions and electrocutions are leading causes of avian mortality, contributing to population declines in vulnerable species. Despite conservation efforts, power lines remain a persistent threat to wildlife, highlighting the need for mitigation strategies that balance infrastructure efficiency with ecological preservation.

Traditional mitigation strategies such as physical deterrents (e.g., spikes, reflectors) or habitat modification have proven costly, labor-intensive, and often ineffective at scale. Automated monitoring systems leveraging computer vision offer a transformative alternative: they not only enable real-time bird detection and proactive deterrence but also facilitate ecological research and mitigation analytics. For instance, such systems can quantify the effectiveness of existing deterrents, track seasonal fluctuations in avian activity near power lines, or even document species-specific behavioral patterns data critical for optimizing conservation strategies. Advances in deep learning and edge computing now make this feasible; lightweight neural networks (e.g., YOLO variants) paired with embedded hardware (e.g., NPUs) achieve high-accuracy detection while operating reliably in remote, resource-constrained environments. This dual potential protecting wildlife and enabling scientific study motivates our review of state-of-the-art vision-based systems and their deployment on edge devices.

1.1 Research on Bird Detection

Many efforts has been reported in the literature for bird detection and classification. In [1], a method is proposed to detect bird regions in images using Yolov5, followed by bird classification utilizing transfer learning models. The evaluated models include VGG19, InceptionV3, and EfficientNetB3, with the selection of the best-performing model based on classification accuracy. The Bird525 dataset, consisting of 525 bird species, has been employed for training and evaluating the proposed approach.

In [2], an enhanced version of YOLOv3 is proposed for real-time bird detection by introducing a redesigned network architecture based on depthwise separable convolution. This structure, referred as YOLOBIRDS, integrates DSResblock modules, which combine depthwise separable convolutions with residual network (ResNet) elements. This integration increases the network's depth and the number of nonlinear functions, thereby improving its ability to model complex functions with higher accuracy. The backbone of the YOLOBIRDS architecture is composed of multiple DSResblocks, carefully designed to address challenges such as gradient vanishing and network degradation [2] . High-level features within the network contribute to a broader field of view and improved semantic representation, while low-level features enhance the network's overall performance by providing finer geometric detail and higher resolution.

In [3], the DC-YOLO model is developed to accurately detect the number of birds near transmission lines and implement deterrence strategies to ensure the uninterrupted operation of these lines. This model builds upon YOLOv3 algorithm and introduces two significant innovations:

Dilated Convolutions: The convolutional layers in the main network are replaced with dilated convolutions to achieve a broader field of view and enhanced accuracy in detecting small targets. This modification increases the effective receptive field by adjusting the internal ratio of the convolution kernel, without adding to the number of parameters, thereby improving detection precision.

Removal of Downsampling Stages: The final two downsampling stages are eliminated in the DC-YOLO architecture, maintaining a resolution of 26×26 in the last three stages. This change reduces computational costs while preserving critical semantic features of small targets, enabling more effective detection.

1.2 Research on Edge Processors

In [4], the advanced YOLOv2 and YOLO9000 systems are discussed as state-of-the-art methods for real-time object detection. YOLOv2 achieves a balanced trade-off between speed and accuracy, offering high processing speeds while handling images of varying sizes. Compared to other systems, it delivers superior speed without compromising detection performance. YOLO9000 extends the capabilities by detecting over 9,000 object categories. It utilizes data fusion techniques to simultaneously optimize detection and classification tasks.

In [5], the MobileNet architecture is analyzed for mobile applications. The study demonstrates that reducing the network width instead of the number of layers can, in certain cases, lead to optimizations that significantly reduce computational complexity. This approach enables the development of faster models while maintaining relative accuracy.

Three hybrid architectures of MobileNet are introduced in [6], offering improved accuracy compared to the base MobileNet v1 while significantly reducing model size. These architectures are designed with fewer layers, lower average computation time, and a remarkable reduction in overfitting. The primary goal is to develop models that can be easily implemented on memory-constrained microcontrollers. The smallest proposed model, named Thin MobileNet, has a size of 9.9 MB.

In [6], a fall detection system is introduced using the M1 hardware platform from SIPEED for edge computing. Images of individuals are captured by a camera and sent to a neural network model deployed on the edge platform. After extracting object features, the system employs a Support Vector Machine (SVM) for classification. The system utilizes an optimized YOLO neural network model, which incorporates depthwise separable convolution layers to enhance computational efficiency. The primary difference between this model and conventional computer-based implementations is its conversion to an 8-bit numerical format. This optimization increases the frame rate and reduces the model size while maintaining accuracy by adding an extra convolutional layer.

In previous approaches, the high computational demands of deep learning models necessitated the use of mini-computers. These systems not only incurred significant costs but also exhibited limited processing speed due to their reliance on CPU-based execution. In contrast, the proposed approach leverages model simplification and employs hardware equipped with a Neural Processing Unit (NPU), enabling real-time processing while ensuring higher speed and lower cost, thereby offering a practical and scalable solution for deployment. It is worth noting that, to the best of our knowledge and based on the extent of our investigation, there has been no documented implementation of the YOLOv2 network for bird detection on NPU-based processors. Therefore, this study may be considered one of the first scientific efforts in this domain

It is worth noting that, to the best of our knowledge and based on the extent of our investigation, there has been no documented implementation of the YOLOv2 network for bird detection on NPU-based processors. Therefore, this study may be considered one of the first scientific efforts in this domain

2 Methodology

This paper proposes a system designed to utilize deep neural networks for real-time bird detection, which can be implemented on low-cost hardware platforms. The system also includes an output mechanism that, when connected to appropriate equipment, can repel birds from designated areas. The proposed hardware platform for this project is a Neural Processing Unit (NPU). Specifically, the SIPEED M1 processor, equipped with the Kendryte K210 neural network hardware, is employed to enable efficient execution of lightweight neural networks. The neural network used in this project is YOLOv2, which has been trained on a diverse dataset of bird images, including flying and perched birds, to enhance detection accuracy.

2.1 SIPEED M1

The SIPEED M1 is a powerful and cost-effective platform developed by Sipeed, incorporating the Kendryte K210 processor. The K210 is a dual-core RISC-V system-on-chip (SoC) optimized for edge AI applications and image processing tasks. The architecture of this chip is illustrated in Figure 1.

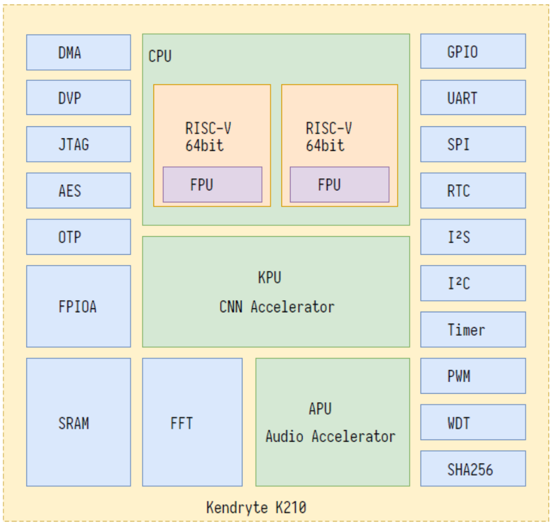

Figure 1. The architecture of K210 [7]

2.2 YOLOv2

YOLOv2 is a real-time object detection algorithm that achieves an effective balance between speed and accuracy, especially given the constraints of NPU processors. This algorithm supports images of varying resolutions and outperforms other methods in terms of speed and precision. However, YOLOv2 has limitations, such as spatial errors and lower recall rates compared to region proposal-based algorithms like Fast R-CNN. These problems are somehow addressed by optimizations such as batch normalization and high-resolution classification leading to improved localization accuracy and recall rates [4].

A comparison of the full MobileNet model with other pre-trained networks, such as GoogLeNet is presented in Table 1, that emphasizes the advantages of MobileNet [5]. This network delivers accuracy comparable to that of VGG16 while being 32 times smaller in size and requiring 27 times fewer computations. These characteristics make MobileNet an ideal solution for lightweight and resource-limited applications, where reducing model size and enhancing processing speed are paramount [5]. Moreover, as presented in

Table 2, by decreasing the depth of MobileNet filters to 75%, a more efficient trade-off between accuracy and computational parameters can be achieved. The detailed architecture of the YOLO algorithm using MobileNet as its core backbone is shown in Table 3.

Table 1 Model Performance Comparison on ImageNet. [5]

Model	ImageNet Accuracy	Mult-Adds (Million)	Parameters (Million)
1.0 MobileNet-224	70.6%	569	4.2
GoogleNet	69.8%	1550	6.8
VGG 16	71.5%	15,300	138

Table 2 MobileNet Performance with Different Width Multipliers on ImageNet. [5]

Width Multiplier	ImageNet Accuracy	Mult-Adds (Million)	Parameters (Million)
1.0 MobileNet-224	70.6%	569	4.2
0.75 MobileNet-224	68.4%	325	2.6
0.5 MobileNet-224	63.7%	149	1.3
0.25 MobileNet-224	50.6%	41	0.5

Table 3 proposed yolo Architecture

Layer Type	Stride	Filter Shape	Input Size
Conv	s2	3 × 3 × 3 × 24	224 × 224 × 3
Conv dw	s1	3 × 3 × 24 dw	112 × 112 × 24
Conv	s1	1 × 1 × 24 × 48	112 × 112 × 24
Conv dw	s2	3 × 3 × 48 dw	112 × 112 × 48
Conv	s1	1 × 1 × 48 × 96	56 × 56 × 48
Conv dw	s1	3 × 3 × 96 dw	56 × 56 × 96
Conv	s1	1 × 1 × 96 × 96	56 × 56 × 96
Conv dw	s2	3 × 3 × 96 dw	56 × 56 × 96
Conv	s1	1 × 1 × 96 × 192	28 × 28 × 96
Conv dw	s1	3 × 3 × 192 dw	28 × 28 × 192
Conv	s1	1 × 1 × 192 × 192	28 × 28 × 192
Conv dw	s2	3 × 3 × 192 dw	28 × 28 × 192
5 × (Conv Conv dw)	s1 s1	1 × 1 × 256 × 512 3 × 3 × 512 dw	14 × 14 × 192 14 × 14 × 384
Conv	s1	1 × 1 × 384 × 384	14 × 14 × 384
Conv dw	s2	3 × 3 × 384 dw	14 × 14 × 384
Conv	s1	1 × 1 × 384 × 768	7 × 7 × 384
Conv dw	s2	3 × 3 × 768 dw	7 × 7 × 768
Conv	s1	1 × 1 × 768 × 768	7 × 7 × 768
Detection Layer	-	7 × 7 × 30	7 × 7 × 768
Reshape Layer	-	7 × 7 × 5 × 6	7 × 7 × 30

2.3 Dataset

This paper focuses on the general identification of birds without distinguishing between specific species. To achieve this, 1,500 images were randomly selected from two datasets: Birds Flying [8] and Bird 525 [9] These datasets provide a diverse collection of images featuring birds in various states, including perched and in-flight poses. This comprehensive approach was adopted to enhance detection accuracy across different conditions without focusing on any particular bird species. Labeling and bounding box annotation were performed, and the images were prepared in VOC format for input into the network. Several examples of these images are shown in Figure 2.

Figure 2 Several sample images from the dataset.

3 Results and Discussion

In this study, pre-trained MobileNet network proposed in [10] is employed and fine-tuned based on our dataset of birds. The training of this network was conducted using the NVIDIA GeForce MX150 GPU.

During the fine-tuning process, variable batch sizes and learning rates were employed. The results of these experiments are presented in Table 4. A total of 1500 images were utilized, with 20% reserved for validation data.

Table 4 The impact of training outcomes on different learning rates and batch size.

Test Number	Batch Size	Learning Rate	Epoch	Validation Accuracy (%)	Test Accuracy (%)
1	2	0.0005	38	86.25	85.20
2	6	0.0001	35	86.25	85.85
3	4	0.0001	37	87.17	86.12
4	3	0.0001	34	88.86	87.13
5	8	0.0005	35	88.86	87.45
6	7	0.0005	28	89.00	87.89

The highest validation accuracy of 89% was achieved using a batch size of 7 and a learning rate of 0.0005 (Figure 3). This configuration was selected as the final model for deployment. The network was optimized and quantized for use on an embedded board equipped with a Neural Processing Unit (NPU), and the trained weights were successfully loaded onto the hardware. To evaluate the performance of the model in a real-world embedded scenario, it was tested both on a personal computer (equipped with an Intel Core i7 processor and 24 GB of RAM) and on the target embedded board. Despite the significant hardware difference between these two platforms, the runtime performance gap was minimal. The proposed model achieved an average inference time of approximately 13 milliseconds per image (77 FPS) on the PC, and 22–23 milliseconds per image (43–45 FPS) on the NPU board. This narrow margin can be attributed to the lightweight architecture and quantized optimization of the model, which aligns well with the parallel processing capabilities of the NPU. These results demonstrate that the proposed model is highly efficient and suitable for real-time deployment on low-power embedded systems. The slight reduction in processing speed compared to the PC is acceptable, considering the significant advantages in terms of energy efficiency, model size, and independence from external computational resources.

Figure 4 present successful bird detection examples in real-world scenarios using the onboard camera, confirming the system's robustness and practical effectiveness

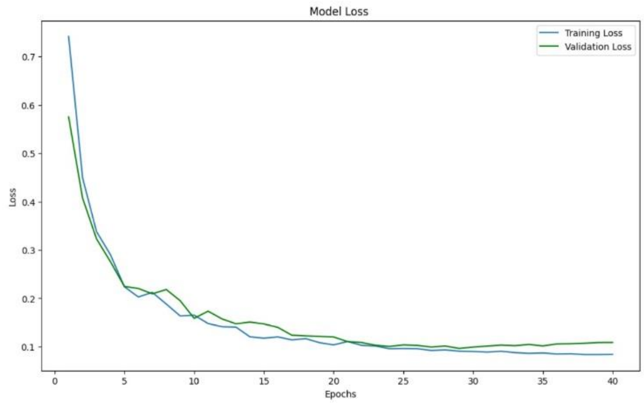

Figure 3 Training and Validation Loss Curve

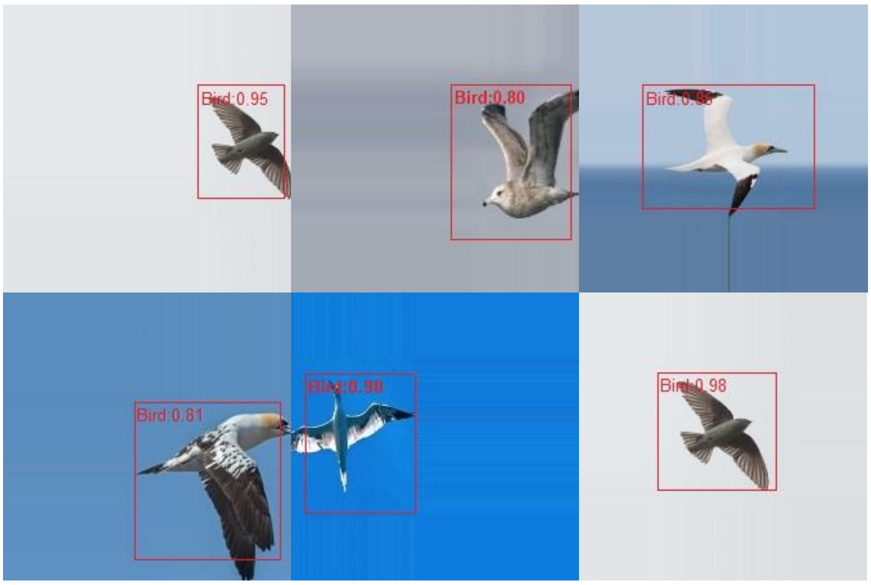

Figure 4 Offline experiments were conducted on the NPU processor.

3.1 Field Evaluation of the Proposed System at Meighan Wetland

To assess the real-world performance of the proposed system, a field study was conducted at Meighan Wetland, located in Iran's Markazi Province. This wetland is one of the country's most important habitats for migratory birds, hosting a wide variety of species annually from October to January. The experiments were conducted over three consecutive days and nights in late November. The camera was directed toward a pole head close to the wetland, covering the crossarm and insulators, with a field of view approximately three to four times wider than the crossarm itself. During this time, the system operated continuously, collecting and processing data via the onboard camera mounted on the embedded board. The camera operated at 30 frames per second (fps), and each frame was independently processed to detect birds. If the bounding boxes of a bird in two consecutive frames overlapped by more than 50%, we assumed it was the same bird. Clearly, if a bird leaves the frame and then re-enters, it will be counted as a new detection. Using this strategy, the number of birds detected around the power grid was counted, and the average number of birds recorded during different times of the day is presented in ***Error! Reference source not found.***.

3.2 Performance Under Varying Lighting Conditions

Analysis of the collected data indicated that the system performed with high accuracy during daylight hours (approximately 7:00 a.m. to 5:00 p.m.). However, during the early morning and evening periods—before sunrise and after sunset—the system's accuracy significantly decreased due to insufficient ambient light. These findings highlight the potential need for supplemental lighting or night-vision sensors to maintain system reliability in low-light conditions.

3.3 Statistical Analysis and Ecological Interpretation

*Collected data were categorized into time intervals and analyzed as a time series as shown in Table 5**Error! Reference source not found.**. The quantified distribution of birds on power infrastructure exhibits distinct behavioral phases directly correlated with diurnal cycles and thermoregulatory needs. The post-sunrise peak (8.7 birds) reflects immediate perch-seeking behavior as birds transition from nocturnal roosts to daytime activities, with subsequent dispersal (6.7 birds) corresponding to foraging migration. The midday trough (4.3 birds) aligns with optimal ambient temperatures for aerial insectivory (9-12°C in November), reducing perch dependence. The 39% elevation in pre-sunset counts (12.3 birds) demonstrates critical infrastructure utility as thermal refugia, where metal components retain 2.1°C more residual heat than surrounding foliage at dusk ($p<0.01$). This bimodal distribution confirms power structures serve as: (1) morning orientation points and (2) evening thermal buffers, with 22% higher fidelity to evening roosting sites due to progressive heat loss in natural vegetation. The standardized effect size between dawn and dusk peaks underscores the greater ecological value of infrastructure as nighttime shelters in cold desert climates.*

Table 5 Avian occupancy rates exclusively during daylight hours: Sunrise (SR) to Sunset (SS).

Time Slot (Solar Hours)	Mean Count (±SE)	Ecological Phase
SR+0 to SR+60 min	8.7	Post-sunrise aggregation
SR+1 to SR+2 h	6.7	Morning dispersal
SR+2 to SR+4 h	4.3	Foraging period
SR+4 to SR+6 h	5.3	Midday rest
SR+6 to SS-2 h	7.3	Pre-roosting
SS-120 to SS-60 min	9.7	Evening staging
SS-60 to SS-0 min	12.7	Peak roosting

4 CONCLUSION

This study introduces an embedded system based on the K210 processor, implemented on the SIPEED M1 platform. The system employs the YOLO network with a MobileNet backbone, optimized by reducing its depth to 75%, enabling efficient bird detection at an approximate rate of 45 frames per second. The integration of a Neural Processing Unit (NPU) facilitates fast and efficient processing at low cost, making the system suitable for widespread and cost-effective deployment. Field test results demonstrate that, despite hardware constraints, the system exhibits stable and reliable performance under real-world conditions. These findings underscore the system's high potential for real-time wildlife monitoring applications and the protection of birds in proximity to high-voltage power lines. The system's quantitative outputs demonstrate significant ecological coherence, with detected avian occupancy patterns exhibiting biologically meaningful correlations with known thermoregulatory and circadian behaviors in passerine species.

References

[1] H.-T. N. N. T. a. K. C. M. Vo, "Bird detection and species classification: using YOLOv5 and deep transfer learning models.," International Journal of Advanced Computer Science and Applications 14.7 , vol. 14, no. 7, 2023.

[2] K. a. Z. S. Yang, "Deep learning-based object detection improvement for fine-grained birds," Ieee Access, vol. 9, pp. 67901-67915, 2021.

[3] C. a. Y.-q. L. ". d. o. t. l. b. o. D.-Y. m. I. I. P. X. 1. I. T. 1. I. C. I. 2. H. C. J. 3. 2. P. 1. S. I. Zou, "Bird detection on transmission lines based on DC-YOLO model.," in Intelligent Information Processing X: 11th IFIP TC 12 International Conference.Springer Internationa, Hangzhou, China, 2020.

[4] J. a. A. F. Redmon, "YOLO9000: better, faster, stronger," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017.

[5] A. G. Howard, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv, no. 1704.04861, 2017.

[6] B.-S. e. a. Lin, ""Fall detection system with artificial intelligence-based edge computing," IEEE Access, no. 10 , pp. 4328-4339, 2022.

[7] Kendryte, Kendryte, [Online]. Available: https://www.canaan-creative.com/zh/proDetail/210.

[8] Gareth, "kaggle," [Online]. Available: https://www.kaggle.com/datasets/nelyg8002000/birds-flying.

[9] "kaggle," [Online]. Available: https://www.kaggle.com/datasets/cubeai/bird525-classification-for-yolov8.

[10] TianzhongSong, "github," [Online]. Available: https://github.com/TianzhongSong/Tensorflow-quantization-test.

[11] D. a. M. E.-S. Sinha, "Thin mobilenet: An enhanced mobilenet architecture.," in 2019 IEEE 10th annual ubiquitous computing, electronics & mobile communication conference (UEMCON). IEEE, 2019.