

# Robust UAV vs Bird Classification using Ensemble Deep Learning with Confidence-Based Decision and Explainable AI

Dr. Sundaragiri Dheeraj  
Associate Professor

Department of Computer Science and Engineering  
Sreenidhi Institute of Science and Technology  
Hyderabad, India  
[dheeraj.s@sreenidhi.edu.in](mailto:dheeraj.s@sreenidhi.edu.in)

A.Jayasri  
Student,

Department of Computer science and Engineering  
Sreenidhi institute of science and technology,  
Hyderabad,India  
[jayaenigathala@gmail.com](mailto:jayaenigathala@gmail.com)

P.Shivani  
Student,

Department of Computer science and Engineering  
Sreenidhi institute of science and technology,  
Hyderabad,India  
[peesarishivani@gmail.com](mailto:peesarishivani@gmail.com)

Yenugula Sai Pragnya  
Student,

Department of Computer Science and Engineering  
Sreenidhi Institute of Science and Technology  
Hyderabad, India  
[spyenugula05@gmail.com](mailto:spyenugula05@gmail.com)

**Abstract**— Unmanned Aerial Vehicle (UAV) detection has become increasingly critical in surveillance, airspace management, and security systems, where distinguishing UAVs from visually similar objects such as birds remains a significant challenge. This work presents a robust ensemble deep learning framework for UAV versus bird classification, integrating multiple convolutional neural network architectures with confidence-based decision strategies and explainable artificial intelligence. The methodology incorporates comprehensive image preprocessing techniques including resizing, augmentation, and normalization to enhance generalization. Multiple deep learning models—Custom CNN, ResNet50, MobileNet, EfficientNet, and ConvNeXt—are employed to extract discriminative features. A confidence-based ensemble mechanism is utilized to improve classification reliability, while Gradient-weighted Class Activation Mapping (Grad-CAM) is integrated to provide visual explanations of model predictions, enhancing interpretability. Performance evaluation demonstrates that the ResNet50 model achieves superior results with an accuracy of 99.8%, precision of 99.8%, recall of 99.8%, and F1-score of 99.8%, indicating its effectiveness in handling complex visual patterns. The system is deployed using a Flask-based web interface with SQLite integration, enabling user authentication, image upload, and real-time prediction.

**Keywords**— UAV Classification, Ensemble Deep Learning, ResNet50, Grad-CAM, Explainable AI, Flask Framework, Image Processing, Confidence-Based Decision.

## I. INTRODUCTION

The rapid expansion of unmanned aerial vehicles has significantly transformed modern aerial applications, including surveillance, environmental monitoring, logistics, and security operations. However, the increasing presence of UAVs in shared airspace has introduced critical challenges in distinguishing them from naturally occurring flying objects such as birds. Due to similarities in shape, motion patterns,

and size, accurate identification becomes difficult under varying environmental conditions. These challenges have been widely discussed in UAV detection studies where visual ambiguity remains a primary concern in automated recognition systems [1].

Advancements in artificial intelligence and computer vision have enabled the development of automated detection approaches for aerial object classification. Deep learning methods have shown strong capabilities in extracting hierarchical features from complex image data, improving recognition accuracy in challenging scenarios [2]. Research in multi-modal perception further highlights the importance of combining diverse feature representations to enhance robustness in dynamic environments [3]. Active learning strategies have also been explored to improve model adaptability in visual inspection tasks [4].

However, limited supervision and insufficient labeled data remain significant barriers in training highly generalized models for aerial classification tasks [5]. Intelligent monitoring systems in related domains such as traffic management demonstrate the importance of reliable classification under real-world variability [6]. Additionally, validation strategies for machine learning models emphasize the need for structured evaluation frameworks to ensure dependable performance in safety-critical applications [7].

Recent advancements in perception systems for navigation and object recognition highlight the importance of accurate detection mechanisms in complex visual environments [8]. Fairness and decision-making challenges in automated systems further underline the necessity for transparent and reliable classification approaches [9]. Studies on bird and drone detection using both classical and deep learning methods demonstrate persistent challenges in reducing misclassification errors under diverse conditions [10].

Benchmark studies in drone versus bird detection further confirm the difficulty of achieving consistent performance across different datasets and environmental variations [11].

The objective is to develop an intelligent classification framework capable of accurately distinguishing UAVs from birds using advanced deep learning techniques. The system integrates multiple convolutional architectures including Custom CNN, ResNet50, MobileNet, EfficientNet, and ConvNeXt, along with preprocessing strategies, confidence-based decision mechanisms, Grad-CAM-based explainability, and a Flask-based interface for real-time interaction and interpretable output generation.

## II. RELATED WORK

Research on aerial object recognition has gained significant attention due to the increasing use of unmanned aerial vehicles in civilian and surveillance environments. Studies focusing on large-scale benchmark evaluations highlight the persistent difficulty in distinguishing drones from birds due to overlapping visual and motion characteristics under real-world conditions [12]. These investigations emphasize the importance of standardized evaluation frameworks and diverse datasets to improve robustness in detection systems.

Efforts to enhance detection precision while reducing computational complexity have been explored using both traditional machine learning and deep learning-based approaches. Research demonstrates that lightweight detection mechanisms combined with feature optimization can improve efficiency in real-time aerial monitoring scenarios [13]. Comparative analyses of convolutional neural network models further indicate that architectural depth and design variations significantly influence classification performance in UAV identification tasks [14].

Deep learning-based classification approaches have been widely adopted for distinguishing drones and birds, showing strong capability in learning hierarchical visual features from complex image data [15]. Additionally, object detection frameworks such as real-time single-stage detectors have been applied to aerial surveillance, offering faster inference speeds while maintaining competitive accuracy in dynamic environments [16].

Improvements in deep residual learning architectures have also been explored for fine-grained bird classification tasks, demonstrating enhanced feature extraction and improved recognition accuracy in visually similar categories [17]. Transfer learning strategies have further contributed to improving classification performance by adapting pre-trained models to specialized datasets, reducing training complexity and improving generalization ability [18].

Research on UAV classification using aerial imagery datasets highlights the effectiveness of transfer learning in adapting deep learning models to domain-specific aerial recognition tasks [19]. Beyond traditional classification approaches, advancements in optimization-based robotic perception and self-supervised learning emphasize the

growing importance of adaptive learning systems capable of operating in uncertain environments [20].

Broader perspectives in intelligent systems research highlight the role of human-centered design and emerging technologies in shaping automated decision-making systems, particularly in safety-critical domains [21]. Additionally, developments in edge-cloud computing environments focus on optimizing data processing and deployment efficiency, enabling scalable and responsive intelligent monitoring systems [22].

Overall, existing studies collectively demonstrate that while significant progress has been made in aerial object classification, challenges such as visual similarity, environmental variability, computational constraints, and model interpretability still persist. These limitations motivate the need for more robust, explainable, and scalable solutions capable of handling complex real-world scenarios with higher reliability and transparency.

## III. MATERIALS AND METHODS

The system proposes an advanced intelligent framework for accurate discrimination between UAVs and birds using multiple deep learning architectures combined with adaptive decision strategies. It integrates Custom CNN, ResNet50, MobileNet, EfficientNet, and ConvNeXt to extract rich hierarchical features from aerial imagery. A confidence-based fusion mechanism is applied to combine outputs from different models, ensuring improved reliability in classification decisions. The preprocessing pipeline includes resizing, augmentation, and normalization to enhance data consistency and robustness under varying environmental conditions. To improve interpretability, Grad-CAM is incorporated to visualize critical regions influencing predictions, increasing transparency in automated decision-making. The framework also emphasizes efficient deployment considerations for real-time inference in constrained environments [23]. It leverages structured data representation strategies to improve generalization across diverse visual scenarios [24]. Additionally, insights from environmental monitoring systems guide the adaptation of robust feature learning techniques for dynamic aerial conditions [25]. A Flask-based interface with SQLite integration is included to support secure authentication, image upload, and real-time classification output. This ensures smooth interaction between users and the system, enabling practical deployment in surveillance and monitoring applications where accurate and interpretable aerial object recognition is essential.

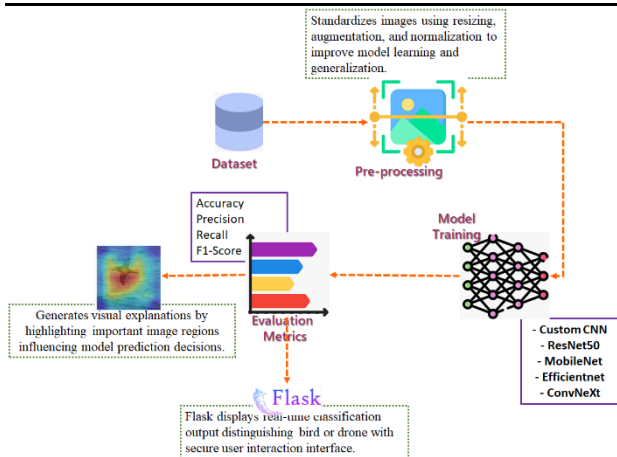


Fig.1 System Architecture

Figure 1 illustrates the comprehensive machine learning architecture designed for bird and drone classification. The workflow begins with a dataset that undergoes pre-processing, including resizing and augmentation, to enhance generalization. Various deep learning models, such as ResNet50 and ConvNeXt, are trained and subsequently assessed using evaluation metrics like accuracy and F1-score. To ensure transparency, visual explanations highlight influential image regions. Finally, the system is deployed via a Flask web interface, providing real-time, secure classification results for end-user interaction.

### A) Dataset Collection:

The dataset collection process, illustrated in Fig. 2, focuses on constructing a diverse and balanced set of aerial images categorized into birds and drones to support reliable visual discrimination under real-world conditions. The images are collected from varied environments, including urban areas, open skies, forests, and mountainous regions, ensuring strong variability in background complexity, lighting intensity, and object scale. This diversity helps the system learn robust feature representations that are less sensitive to environmental changes. Both close-range and long-distance views are included to capture differences in object clarity and resolution, improving adaptability to real surveillance scenarios.

To further enhance dataset richness, augmentation techniques such as rotation, flipping, and shearing are applied, as shown in the transformed samples in Fig. 2. These transformations simulate different viewing angles and orientations, enabling improved spatial invariance and reducing model overfitting. The inclusion of both clear silhouettes and partially occluded objects increases the difficulty level of classification, strengthening feature learning capability.

Additionally, motion variations and background clutter are incorporated to replicate real-world aerial monitoring challenges. This comprehensive dataset construction ensures that the model is exposed to a wide range of visual patterns, ultimately improving its ability to distinguish between natural

bird movement and structured drone flight behavior in complex environments.

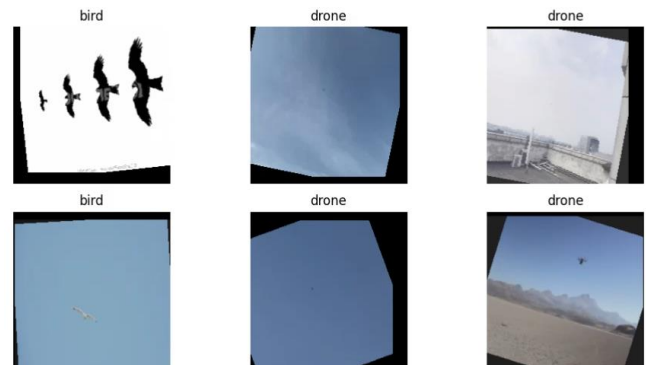


Fig.2 Dataset

### B) Image Processing:

Image processing plays a crucial role in preparing aerial images for effective learning by ensuring uniformity, enhancing variability, and improving feature quality before classification. The process begins with resizing, where all input images are scaled to a fixed dimension. This step is essential because aerial images collected from different sources often vary in resolution and aspect ratio. By standardizing image size, the system ensures consistent input format, reduces computational complexity, and allows efficient batch processing during model training and inference.

Random horizontal flip is then applied as a data augmentation technique to simulate different viewing angles. Since aerial objects such as birds and drones can appear from multiple directions in real-world scenarios, this transformation helps the system learn orientation-invariant features. It improves generalization by exposing the model to mirrored versions of the same image, reducing sensitivity to directional bias.

Random vertical flip further enhances variability by flipping images along the vertical axis. This is particularly useful in aerial surveillance scenarios where objects may be captured from unusual perspectives due to camera tilt or movement. It increases the diversity of training samples and strengthens the model's ability to recognize objects regardless of vertical orientation.

Random rotation is used to rotate images by different angles, simulating real-world camera movements and object orientations. Aerial images are rarely captured in perfectly aligned positions, and rotation helps the model become robust to such variations. This technique improves spatial understanding and ensures that classification remains accurate even when objects are tilted or rotated in unpredictable directions.

Color jitter is applied to adjust brightness, contrast, saturation, and hue of images. Environmental factors such as sunlight, shadows, fog, and atmospheric conditions can significantly affect image appearance. By introducing

controlled variations in color properties, the system learns to focus on structural and shape-based features rather than being influenced by lighting inconsistencies. This enhances robustness in diverse environmental conditions.

Normalization is the final preprocessing step, where pixel values are scaled to a standardized range. This ensures numerical stability during model training and accelerates convergence. By reducing variations in pixel intensity distributions, normalization allows the model to learn more efficiently and prevents dominance of large-scale features over smaller yet important details.

Overall, the combination of resizing, flipping, rotation, color jitter, and normalization creates a strong preprocessing pipeline that significantly improves dataset quality. It enhances variability, reduces overfitting, and ensures that the learning system is exposed to diverse and realistic aerial conditions, ultimately leading to more accurate and reliable classification performance in distinguishing between birds and drones.

### C) Algorithms:

**Custom CNN:** Custom Convolutional Neural Network is a deep learning architecture manually designed for image classification tasks using stacked convolutional, pooling, and fully connected layers. It learns hierarchical representations from input images by progressively extracting low-level to high-level visual patterns such as edges, textures, shapes, and object structures. In UAV versus bird classification, it processes preprocessed aerial images to identify distinguishing visual characteristics between natural and mechanical flying objects. The architecture is flexible, allowing customization of layers, filter sizes, activation functions, and depth based on dataset complexity. This adaptability makes it suitable for experimental analysis and baseline performance comparison. It is computationally efficient compared to very deep architectures and can be trained with moderate resources. The model plays an important role in understanding fundamental feature learning behavior in image classification tasks and provides interpretability due to its straightforward structure, making it useful for controlled experimentation and performance benchmarking in aerial object recognition scenarios.

**ResNet50:** ResNet50 is a deep convolutional neural network built using residual learning with skip connections that allow direct flow of information across layers. This design solves the vanishing gradient problem and enables effective training of very deep networks. It extracts hierarchical and complex features from images by passing input through multiple residual blocks, each learning refinement over previous representations. In UAV versus bird classification, it captures fine-grained differences in shape, texture, and structure, making it highly effective in distinguishing visually similar aerial objects. The residual connections preserve essential information while enabling deeper feature extraction without degradation in performance. Its strong representational power and stability

make it suitable for complex image recognition tasks. The model is widely used for high-accuracy visual classification due to its ability to generalize well across varied datasets and environmental conditions, ensuring reliable performance in challenging aerial image analysis scenarios.

**MobileNet:** MobileNet is a lightweight convolutional neural network designed for efficient computation using depthwise separable convolutions. This structure splits standard convolution into depthwise and pointwise operations, significantly reducing computational cost and number of parameters while maintaining effective feature extraction. It processes images in a resource-efficient manner, making it suitable for environments with limited processing power. In UAV versus bird classification, it identifies essential visual features such as shape and motion-related patterns while ensuring fast inference. Its compact architecture enables deployment in real-time systems and mobile or edge-based applications. Despite being lightweight, it maintains competitive performance by focusing on efficient representation learning. The model is particularly useful where speed and efficiency are more critical than extremely deep feature extraction. It balances accuracy and computational cost, making it a practical choice for real-time aerial object classification tasks requiring quick response and minimal hardware requirements.

**EfficientNet:** EfficientNet is a convolutional neural network that introduces compound scaling to uniformly balance depth, width, and resolution of the model. This balanced scaling strategy allows it to achieve high performance without excessive computational complexity. It processes images at optimized resolutions, enabling extraction of both fine-grained details and global structural patterns. In UAV versus bird classification, it effectively captures subtle visual differences between similar aerial objects under diverse environmental conditions. The architecture improves efficiency by scaling model dimensions in a controlled manner rather than arbitrarily increasing size. This leads to better accuracy with fewer parameters compared to traditional deep networks. EfficientNet demonstrates strong generalization capability across varied datasets and image conditions. Its design ensures optimal utilization of computational resources while maintaining high representational power, making it suitable for complex image classification tasks where both accuracy and efficiency are critical requirements in real-world aerial monitoring scenarios.

**ConvNeXt:** ConvNeXt is a modern convolutional neural network architecture inspired by transformer-based design improvements while retaining convolutional operations. It integrates advanced architectural enhancements such as large kernel convolutions, layer normalization, and simplified network design to improve learning efficiency and performance. It captures both local and global contextual information from images through optimized feature extraction mechanisms. In UAV versus bird classification, it enhances discrimination by focusing on spatial relationships,

object structure, and contextual cues within aerial images. The architecture is designed for scalability and improved training stability, allowing it to handle complex visual variations effectively. ConvNeXt bridges the gap between traditional convolutional networks and modern transformer-inspired models, offering strong representation learning capabilities. Its optimized structure improves accuracy and robustness across diverse image conditions, making it highly effective for challenging classification tasks involving visually similar aerial objects in dynamic and real-world environments.

#### D) Integration of XAI & Flask:

The integration of explainable artificial intelligence and a web deployment framework enhances both transparency and usability in the aerial image classification system. Explainable artificial intelligence is incorporated using Gradient-weighted Class Activation Mapping, which visually highlights the regions of an image that contribute most to the prediction outcome. This allows users to understand why a particular image is classified as a bird or a drone by focusing on discriminative features such as shape, texture, and movement-related patterns. By providing heatmap-based visual explanations, the system improves interpretability and builds trust in automated decision-making processes, especially in safety-critical environments where understanding model behavior is essential.

The Flask framework is used to create a lightweight and interactive web interface that enables real-time communication between users and the classification system. It handles user authentication, image upload, and result display in an efficient manner. SQLite is integrated to securely manage user credentials and maintain session data. Once an image is uploaded through the interface, it is processed and passed through the trained models, and the prediction along with the explainable visualization is returned to the user.

The combination of explainable AI and Flask ensures that the system is not only accurate but also transparent and user-friendly. Users can interact seamlessly with the application while gaining insights into how decisions are made, making the overall system practical for real-world aerial monitoring and surveillance applications.

#### IV. EXPERIMENTAL RESULTS

**Accuracy:** The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

**Precision:** Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

**Recall:** Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

**F1-Score:** F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

$$F1\ Score = 2 * \frac{Recall * Precision}{Recall + Precision} * 100 \quad (1)$$

Table.1 Performance Evaluation

Model	Accuracy	Precision	Recall	F1-score
Custom CNN	0.950	0.952	0.950	0.950
ResNet50	0.998	0.998	0.998	0.998
MobileNet	0.973	0.973	0.973	0.973
EfficientNet	0.995	0.995	0.995	0.995
ConvNeXt	0.984	0.984	0.984	0.984

Table 1 presents the performance evaluation of different deep learning models used for classification, comparing accuracy, precision, recall, and F1-score values to analyze their effectiveness in distinguishing aerial objects.

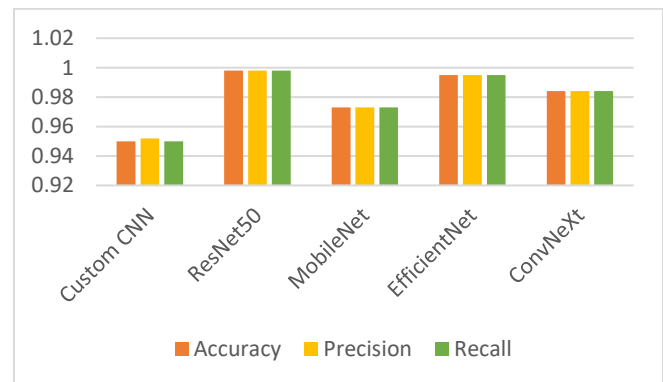


Fig.3 Comparison Graph

Figure 3 displays a performance comparison of several deep learning models based on accuracy, precision, and recall metrics. ResNet50 and EfficientNet emerge as the top performers, maintaining near-perfect scores, while the Custom CNN shows slightly lower but still competitive results.

## V. CONCLUSION

Accurate differentiation between UAVs and birds has been achieved through a robust ensemble deep learning framework that integrates multiple convolutional architectures with confidence-based decision making. The system effectively handles visual similarity challenges by leveraging advanced preprocessing techniques and diverse feature extraction capabilities, resulting in highly reliable classification performance. Among the evaluated models, ResNet50 demonstrated superior performance with an accuracy of 99.8%, precision of 99.8%, recall of 99.8%, and F1-score of 99.8%, confirming its effectiveness in capturing fine-grained visual patterns. The incorporation of Gradient-weighted Class Activation Mapping (Grad-CAM) enhances transparency by providing visual explanations, allowing users to interpret model decisions and verify the regions influencing predictions. This significantly improves trust and usability in real-world surveillance and monitoring applications. Furthermore, the deployment through a Flask-based web interface with SQLite support enables seamless user interaction, including secure authentication, image uploading, and real-time prediction, making the system practical and accessible. The overall outcome demonstrates a high-performance, interpretable, and deployable solution for UAV versus bird classification, contributing to improved safety, monitoring efficiency, and intelligent airspace management systems.

Future enhancements can focus on extending the classification framework to multi-class aerial object detection, including helicopters, airplanes, and unknown flying objects, to improve real-world applicability. Integration of temporal information from video streams using architectures such as LSTM or 3D CNNs can further enhance detection accuracy by capturing motion patterns. Incorporating lightweight transformer-based models may improve performance on edge devices with limited computational resources. Advanced explainability methods beyond Grad-CAM, such as SHAP or LIME, can provide deeper interpretability. Additionally, deploying the system on cloud and IoT-enabled platforms with real-time alert mechanisms and continuous learning pipelines can ensure scalability, adaptability, and improved robustness in dynamic environments.

## REFERENCES

- [1] Shafiq, K. (2023). *Drone Detection and Classification using Machine Learning* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).
- [2] Valle, M. (2022). Automated generation of diverse test cases for autonomous systems using evolutionary search and explainable AI.
- [3] Wang, C., Xu, Y., & Zhang, D. (2025). Multi-Modal Perception and Fusion for Maritime Autonomy: A Survey.
- [4] Bruce, D., Sanghvi, H. A., Behara, R., & Agarwal, A. An Active Learning Artificial Intelligence Driven Framework for Building Collaborative Visual Inspection Models. Available at SSRN 4617143.
- [5] Albert, P. (2023). *Deep learning for computer vision constrained by limited supervision* (Doctoral dissertation).
- [6] Nigam, N., Singh, D. P., & Choudhary, J. (2023). A review of different components of the intelligent traffic management system (ITMS). *Symmetry*, 15(3), 583.
- [7] Gallitz, O. C. (2023). *Validation of Machine Learning Algorithms by Design with Applications for Automated Driving* (Doctoral dissertation, Technische Universität München).
- [8] Susmit, T. A., Mehejabin, M., Hasan, I., Kausar, A. I., & Akbar, S. B. (2024). *Pathway to perception: a smart navigation approach for visually impaired individuals leveraging YOLO, faster R-CNN, and LLAMA* (Doctoral dissertation, Brac University).
- [9] Nashed, S. B., Svegliato, J., & Blodgett, S. L. (2023). Fairness and sequential decision making: Limits, lessons, and opportunities. *arXiv preprint arXiv:2301.05753*.
- [10] R. Abu Zitar, M. Kassab, A. El Fallah, and F. Barbaresco, "Bird/drone detection and classification using classical and deep learning methods," 2023. DOI: 10.22541/au.168075364.45332093/v1
- [11] A. Coluccia, A. Fascista, A. Schumann, et al., "Drone vs. bird detection: Deep learning algorithms and results from a grand challenge," 8, vol. 21, MDPI, 2021, p. 2824. DOI: 10.3390/s21082824. [Online]. Available: <https://www.mdpi.com/1424-8220/21/8/2824>.
- [12] A. Coluccia, A. Fascista, L. W. Sommer, et al., "The drone-vs-bird detection grand challenge at icassp 2023: A review of methods and results," *IEEE Open Journal of Signal Processing*, vol. 99, pp. 1–15, 2024. DOI: 10.1109/OJSP.2024.3379073. [Online]. Available: [https://www.researchgate.net/publication/379099583\\_The\\_Drone-vs-Bird\\_Detection\\_Grand\\_Challenge\\_at\\_ICASSP\\_2023\\_A\\_Review\\_of\\_Methods\\_and\\_Results](https://www.researchgate.net/publication/379099583_The_Drone-vs-Bird_Detection_Grand_Challenge_at_ICASSP_2023_A_Review_of_Methods_and_Results).
- [13] M.Kassab, R. AbuZitar, F. Barbaresco, and A. E. F. Seghrouchin, "Drone detection with improved precision in traditional machine learning and less complexity in single-shot detectors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 4, pp. 3847–3859, 2024. DOI: 10.1109/TAES.2024.3368991.
- [14] H. M. Oh, H. Lee, and M. Y. Kim, "Comparing convolutional neural network (cnn) models for machine learning-based drone and bird classification of anti drone system," in 2019 19th International Conference on Control, Automation and Systems (ICCAS), IEEE, 2019, pp. 87–90. DOI: 10.23919/ICCAS47443.2019.8971699. [Online]. Available: <https://ieeexplore.ieee.org/document/8971699>.
- [15] S. S. Selvi, S. Pavithra, R. Dharini, and E. Chaitra, "A deep learning approach to classify drones and birds," in 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), IEEE, 2022, pp. 1–5. DOI: 10.1109/MysuruCon55714.2022.9993706. [Online]. Available: <https://ieeexplore.ieee.org/document/9993706>.
- [16] S. Singha and B. Aydin, "Automated drone detection using yolov4," *Drones*, vol. 5, no. 3, p. 95, 2021. DOI: 10.3390/drones5030095. [Online]. Available: <https://www.mdpi.com/2504-446X/5/3/95>.
- [17] H. Song, "Bird image classification based on improved resnet-152 image classification model," *Applied and Computational Engineering*, vol. 54, pp. 206–212, 2024. DOI: 10.54254/2755-2721/54/20241530. [Online]. Available: <https://www.ewadirect.com/proceedings/ace/article/view/11068>.
- [18] C.-L. Yang, Y. Harjoseputro, Y.-C. Hu, and Y.-Y. Chen, "An improved transfer learning for image-based species classification of protected Indonesian birds," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 1057–1072, 2022. DOI: 10.32604/cmc.2022.031305. [Online]. Available: <https://www.techscience.com/cmc/v73n1/48224>.
- [19] O.S.Albahri, A. Ruskic, A. Badnjevic, D. Cosic, and B. Gavrilovic, "Uav classification based on transfer learning using aerial imagery datasets," *Drones*, vol. 9, no. 2, p. 110, 2023. DOI: 10.3390/drones9020110. [Online]. Available: <https://www.mdpi.com/2504-446X/9/2/110>.
- [20] Avigal, Y. (2024). *Enabling Efficient and Reliable Robot Manipulation Through Optimization, Interactive Perception and Self-Supervised Learning* (Doctoral dissertation, University of California, Berkeley).
- [21] Daugherty, P., & Wilson, H. J. (2022). *Radically human: How new technology is transforming business and shaping our future*. Harvard Business Press.
- [22] Withana, S. D. (2024). *Cyberinfrastructure Knowledge Network: Optimizing Edge-Cloud-Continuum for Efficient Data Capture, Modeling, and Deployment*. Indiana University.



# International Journal of DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X

[www.ijdim.com](http://www.ijdim.com)

Original Research Paper

---

- [23] Goldstein, O. (2021). *Real-time cost-aware machine learning at the edge*. University of California, Los Angeles.
- [24] Leung, B. (2022). *Understanding Learned Visual Invariances through Hierarchical Dataset Design and Collection*. University of California, San Diego.
- [25] Dupuis, C. (2024). *Monitoring canopy disturbances in Central Africa: from the ground to the sky*. Universite de Liege (Belgium).