Proceedings of the 6th International Conference on Data Intelligence and Cognitive Informatics (ICDICI-2025)
DVD Part Number: CFP25VL6-DVD; ISBN: 979-8-3315-0312-3

IoT-BASED ANIMAL SPECIES RECOGNITION AND BEHAVIOR ANALYSIS USING DEEP LEARNING TECHNIQUE

V.Janani, Research Scholar,
Department of Computer Science
School of Computing Sciences
Vels Institute of Science, Technology & Advanced Studies (VISTAS),
Pallavaram,Chennai, India
samibps17@gmail.com

Dr.S.Kamalakkannan, Professor,
Department of Computer Applications
School of Computing Sciences
Vels Institute of Science, Technology & Advanced Studies (VISTAS),
Pallavaram,Chennai, India
kannan.scs@vistas.ac.in

Dr.P.Kavitha, Assistant Professor
Department of Computer Applications
School of Computing Sciences
Vels Institute of Science, Technology & Advanced Studies (VISTAS),
Pallavaram,Chennai, India
pkavikamal@gmail.com

***Abstract*— Developing tools to analyze and identify wild animal behavior is vital for wildlife management, as it helps monitor stress and well-being and informs conservators. As a result, the development of apps can help breeders make decisions to increase production performance. Wild animal behavior, such as aggression, sitting, standing, and walking, is investigated to minimize its impact through classification and identification. However, image analysis seldom attempts to directly determine "behavioral states" (e.g., activities or facial emotions) from the image; instead, it often depends on identifiable body components. The proposed Region-based CNN with a bio-inspired optimization model as fish swarm optimization (FSO) enhances wild animal identification and behavior recognition through spatial and temporal features. The proposed model achieves precision, recall, accuracy, and f1 score like 0.85, 0.84, 0.94, and 0.84. Additionally, cloud-based data logging and monitoring is essential for sending mail notifications to the breeder. The animal behavior status is updated and monitored in the Adafruit IoT cloud.**

Keywords— Region-based Convolutional Neural Network, Fish Swarm Optimization, Internet of Things cloud

I. INTRODUCTION

Real-time animal behavior analysis in Adafruit Cloud is made possible by IoT, which uses sensors to collect data on posture and movement. Data from these sensors is sent to the dashboard's "animal" or "animal-name" feeds. The meal, for example, adapts to the animal's sitting habit. This makes it possible to continuously monitor operations without requiring human inspections. IoT is crucial for intelligent animal monitoring systems because it enables users to examine these feeds, spot behavioral patterns, spot problems early, and respond quickly. The study of animal behavior involves observing, describing, and comprehending how animals engage with one another and their surroundings. The field of animal behavior research is changing quickly because to the development of advanced behavior detection technologies and the ongoing introduction of novel experimental techniques. This development is especially important for improving our knowledge of neuroethological issues, as demonstrated by the use of mice to study conditions such as Alzheimer's and improve animal care in agriculture. The incorporation of state-of-the-art technology is the driving factor behind this acceleration of advancement, with deep learning particularly noteworthy as a disruptive force that changes the methods used by researchers to study and analyse animal behavior.

Observing wild or semi-wild animals can teach biologists and ethnologists about behavior, habits, possible diseases, relationships with other animals and the environment, and movement patterns. Learning and understanding such facts about animals requires knowledge about behavior, movement patterns, and other biometric data, such as the animal's body temperature and heart rate over extended periods. Consequently, the development and application of animal monitoring systems have been the subject of several studies. Capturing the animal to attach the tracking device and ensure it stays on it for the necessary time is one of the most difficult and costly aspects of animal monitoring. The intricacy of the task demands that these devices be fault-tolerant, lightweight, and have low power consumption; yet, this poses a technological challenge to their design and execution. Giving biologists a small, long-lasting device that can process large quantities of data and be accessed online is difficult.

The proposed system aims to detect and classify animal behavior settings using deep learning and IoT technology. The system collects data through IoT devices and processes wild animal images using data augmentation techniques to enhance the dataset. This data is then used to classify animal behavior status using a Region-Based Convolutional Neural Network (RCNN). A bio-inspired optimization algorithm, specifically the Fish Swarm Optimization Algorithm, is employed to improve prediction accuracy further. The processed behavior status is subsequently updated in the Adafruit cloud. Additionally, this status is sent via email notifications to inform relevant parties in real-time.

II. RELATED WORK

To improve fish detection in commercial trawls and solve unsustainable fishing bycatch and discards, the Coordinate-Aware Mask R-CNN (CAM-RCNN). The approach employs a multiscale retinex, a compound dice and cross-entropy loss function, group normalization, and coordconv for picture improvement. CAM-RCNN achieved superior accuracy in the North Sea datasets, highlighting its efficacy. Future work will expand aquatic animal classes and develop a lightweight version for real-time deployment, promoting sustainable fishing and biodiversity [1].

A version of the MS R-CNN model, the PigMS R-CNN framework, is presented in this study to segment sticky pig images in group-housed situations. PigMS R-CNN head network is used for regression, classification, and segmentation, and a residual network with 101 layers and FPN for feature extraction and ROI generation. With an F1 score increase from 0.9228 to 0.9374, Soft-NMS improves precision and supports welfare assessment and pig behavior monitoring applications. [2].The study enhances Mask R-CNN for precise cattle contour extraction using Generalized Color Fourier Descriptors for image pre-enhancement, an optimally smaller filter size, multiscale semantic features, and a specialized sub-network for improved accuracy. Post-enhancement with the Grabcut algorithm achieved a 0.93 mean Average Precision, outperforming current methods [3].

The study uses sophisticated object detection algorithms to increase the speed and accuracy of animal species recognition. We investigate a deformable convolutional neural network (D-CNN) and four R-CNN models. Steps include model review, dataset preparation, training, enhancement, performance evaluation, system proposals based on superior models, and future work to address limitations [4]. The paper introduces a precise chicken image segmentation methodology to improve animal welfare monitoring. It entails constructing a specialized dataset and employing MSAnet, an end-to-end network integrating multiscale information and attention mechanisms for feature extraction. A combined loss strategy enables deep supervision. Despite challenges, the framework demonstrates promising results, with potential applications beyond poultry farming [5].

The study introduces a methodology to enhance wildlife instance segmentation in camera trap images, emphasizing efficiency and accuracy with limited data. It employs Few-Shot Object Detection (FSOD) for species recognition and initial bounding boxes, followed by a deep snake model for precise segmentation within these boxes. Experimental results validate its effectiveness, suggesting future improvements in dataset diversity and classification accuracy [6]. The research employs Mask R-CNN to precisely segment green sea turtle images, which is crucial for future identification tasks. Contrast-limited Adaptive Histogram Equalization (CLAHE) addresses challenges like image quality and contrast. Combining Mask R-CNN with CLAHE increases Intersection over Union (IoU) by 1.55%, showing promise for improved segmentation. Future work aims to refine segmentation smoothness and reduce errors using techniques like the Laplacian filter [7].

The study enhances fish segmentation in sonar images, which is crucial for monitoring fish behavior in challenging environments. Mask R-CNN is the primary tool, but variability in sonar images between shallow waters poses challenges. A preprocessing CNN standardizes feature maps for Mask R-CNN, improving adaptability. Experimental results demonstrate the superiority of Mask R-CNN when applied to PreCNN's outputs, especially in new fish-farming environments, streamlining segmentation accuracy and transferability across varied contexts. Future work aims to refine the framework with a lightweight fish-instance segmentation network integrated with PreCNN for optimized efficiency [8].

The study introduces a methodology to streamline manual image segmentation in digitized specimen analysis. Researchers trained a specialized deep learning model for fish segmentation in varied backgrounds, validating its accuracy via visual and quantitative assessment. They developed Sashimi, a user-friendly toolkit for automating segmentation accessible to non-programmers with features for model sharing and performance evaluation [9]. The study employs precision livestock farming to detect cattle estrus for breeding management. It utilizes video inputs and a Mask R-CNN framework for cattle segmentation, followed by lightweight tracking. Surveillance cameras capture data for training, achieving 95.5% accuracy in estrus detection. Future work includes improving segmentation and analyzing mounting behavior for optimal artificial insemination timing [10].

An automated body weight estimation system for meat rabbits was developed using a hardware system with a rabbit image acquisition robot and a software system with a weight estimation model. The robot captured top-view images processed through a segmentation and weight-fitting network, achieving a 4.3% relative error in weight estimation. Challenges include camera stability and occlusion in segmenting multiple rabbits, requiring future optimization and integration of edge computing for wider applications in commercial meat rabbit farming [11].

The study addresses biodiversity decline by analyzing animal distribution and behavior using camera traps. A deep learning approach detects and classifies animals, focusing on nighttime activity. Evaluation of methods like Mask R-CNN determines effective detection techniques. Future efforts aim to improve action detection for comprehensive biodiversity monitoring and population estimation [12]. A study proposes enhancing deep learning models for accurately segmenting multiple cow objects in images. It introduces a framework based on Mask R-CNN and Generalized Color Fourier Descriptors to improve efficiency. Evaluation using pre-trained weights shows a mean Average Precision of 0.93, outperforming existing models. Future work includes integrating tracking algorithms for real-time monitoring [13].

The study addresses the computer vision challenge of identifying individual cattle in open-range environments, which is crucial for agriculture. It uses the Mask R-CNN framework for cattle segmentation, achieving 95% accuracy. Results suggest its competitiveness with other models and potential for broader applications in animal farming, including inventory management and health monitoring. Future work aims to refine the algorithm for animal tracking and counting [14]. The MaskDis R-CNN methodology tackles complex pig segmentation issues by combining MS R-CNN with an adversarial network, MaskDis, improving pixel-level feature learning. Achieving 92.03% precision, 92.18% recall, and an F1 score of 0.9210 surpasses YOLACT and Swin Transformer. Challenges include computation and occlusion limitations, driving future optimizations for faster inference and enhanced farming practices [15].

The study uses deep learning to explore segmenting laboratory rats in thermal images for social behavior tests. It evaluates U-Net and V-Net architectures in various temperature ranges. Results show better segmentation with narrower temperature ranges. Specific ranges improve outcomes, and 16-bit resolution doesn't consistently enhance results. The study underscores the potential of machine learning in biomedical tasks, proposing further exploration of instance segmentation algorithms with larger datasets [16]. Camera traps, vital in biology for biodiversity study, have hesitated towards depth estimation. The study advocates for an automated camera trap method with depth estimation, introducing D-Mask R-CNN. Plans include expanding datasets and integrating them into ecological modeling, promising automation in camera trap studies [17].

The project aimed to enhance the detection of tiny deer in Nepal's Chitwan National Park using thermal UAV photos. Variations in thermal images presented challenges for traditional models. Detection was improved by an upgraded Faster R-CNN model that used a ResNet-based Feature Pyramid Network (FPN)—tested five models with various ResNet versions on a dataset of 13,509 annotated deer instances and 2278 images. With 91.6% Average Precision (AP) for all deer and 73.6%, 93.4%, and 94.3% AP scores for small, medium, and giant deer, respectively, the integrated Faster R-CNN, FPN, and ResNet18 models performed exceptionally well [18]. The study uses UAVs and thermal imaging to estimate wild turkey populations in Texas. Automating airborne survey video processing provides a fast, reproducible counting method essential for wildlife management. The approach employs a deep learning semantic segmentation pipeline with the Mask R-CNN algorithm, followed by Data Association and Filtering (DAF) to count roosting turkeys, removing false positives accurately. Drones fly at night to gather data, with promising preliminary results from 280 video frames. Future work will refine accuracy by examining altitude, speed, and nighttime timing [19].

The study explores deep instance segmentation algorithms for analyzing laboratory rodents in thermal images, using the non-intrusive nature of thermal imaging to monitor activity and physiological changes. The system employed Mask R-CNN and TensorMask, finding that pre-training TensorMask on visible light images yielded the highest precision. The contributions confirm the effectiveness of these algorithms for rodent detection and segmentation in thermal images, showing that pre-training on visible light enhances performance. Future work will further focus on expanding the thermal image database to improve Tensor Mask [20]. Innovative picture fusion and a new learning method are used in this work to build a UGV-based livestock detection and counting system. To increase accuracy with the fewest possible annotations, it uses a Dual-scale image Decomposition Fusion and a Seed Label focused Object Detector. The Restricted Supervised Learning (RSL) approach improves precision significantly, achieving up to 91.56% accuracy with YOLOv2. Extensive testing shows 98.3% accuracy on benchmark datasets. Future work will focus on detecting animal health status and integrating additional sensors for broader agricultural use [21].

The study used a non-invasive camera trap network to monitor ungulate populations in Latvia. Images of wild boar and deer were collected across various locations and times. Two object detection frameworks, RetinaNet and Faster R-CNN, automate animal identification and classification. Data augmentation techniques optimized model performance, and the models' effectiveness was evaluated on the collected dataset, enhancing wildlife population management and conservation strategies [22]. This article presents SAWIT, a dataset for research on tiny animals gathered in 2021 over seven months with video traps. Thirty-four thousand four hundred thirty-four photos were annotated by experts, who identified 34,820 animals in seven categories. The Faster RCNN and YOLO object identification algorithms were benchmarked and obtained mean Average Precision (mAP) scores ranging from 58.5% to 62.6%. Results suggest incorporating temporal data to improve detection, especially for hidden species. The framework advances computer vision-based wildlife monitoring, balancing accuracy and efficiency [23]. The study utilized Mask R-CNN, a deep learning method, to automate Holstein Friesian dairy cow detection and segmentation in surveillance videos. Eight cameras were set up in a barn in northern Germany, capturing 36 cows. The model, trained on Microsoft's dataset, achieved high precision (91% for detection, 85% for segmentation) and emphasized the potential of automated monitoring in cattle farming [24].

The study introduces an improved Mask R-CNN model to tackle livestock image segmentation challenges in precision farming. It overcomes the limitations of traditional methods by optimizing filter sizes, leveraging multiscale features, and integrating a sub-network. Experiments achieved a mean average precision of 0.93, surpassing existing models, with a future focus on handling overlapping regions and differentiating cattle body parts [25]. The study innovatively tackles challenges in wild animal surveys via unmanned aircraft systems (UASs) and deep learning. Tactics like shortening feature stride, optimizing anchor sizes and introducing a challenging negative class improve detection, notably for small animals. Tests on a kiang survey in the Tibetan Plateau show substantial enhancement, making surveys 25 times faster or achieving an F1 score of around 0.90. The research highlights the effectiveness of UAS and deep learning fusion in wildlife surveys, offering practical solutions for field challenges [26].

The study compares animal detection methods in thermal images due to nocturnal animal activity. Recent technology enables thermal sensing for wildlife research. It evaluates classical HOG/SVM and deep neural network approaches such as Faster RCNN and YOLO using mAP and training time metrics. YOLOv3 achieves >90% mAP, outperforming others. The study advances autonomous wildlife observation systems, emphasizing neural networks' superiority and suggesting future research on species differentiation and broader applications in nocturnal activities and security [27]. The study used aerial images to test Mask R-CNN for counting cattle in diverse environments. Optimal parameters were set and performance evaluated, achieving 94% accuracy in pastures and 92% in feedlots. Mask R-CNN outperformed other methods, especially in challenging scenarios. The manual assessment showed high precision and counting accuracy. Integration with central processing was recommended for efficiency [28].

The method for animal abundance estimation in dense habitats or underground leverages thermal imagers, but manual review of lengthy footage is costly and inefficient. Off-the-shelf systems yield low-res imagery, complicating detection. The proposed solution uses cost-effective thermal imagers and a Distant-YOLO algorithm for automated detection, trained on diverse Australian data, offering a robust tool for wildlife monitoring [29]. The study assessed cattle breathing using infrared thermography and Mask R-CNN. Mask R-CNN detected cattle noses in RGB images, aligning them with temperature data. Breathing patterns were derived by averaging temperatures in the region of interest. Results showed 76% nose detection accuracy and a high correlation ($R2 = 0.91$) with thermal imaging. The method is efficient, scalable, and accurate for analyzing breathing in cattle [30].

Insufficient research has been done on combining deep learning and thermal imaging for temporal behavior tracking and pattern recognition across time and environments; little research has been done on domain adaptation or transfer learning to apply trained models across varied geographical locations or climates without retraining; and there is a need for super-resolution techniques, denoising algorithms, or attention mechanisms to improve deep learning performance on degraded inputs. These models are unable to accurately classify species using low-resolution or high-noise thermal imagery, especially in dense or mixed-species habitats.

III. Proposed Methodology

The proposed methodology for animal behavior detection integrates region-based convolutional neural networks (RCNN) with the Fish Swarm Optimization Algorithm, utilizing Internet of Things (IoT) technology. This comprehensive system consists of several key components: image data augmentation, classification, and optimization, as shown in Figure 1.

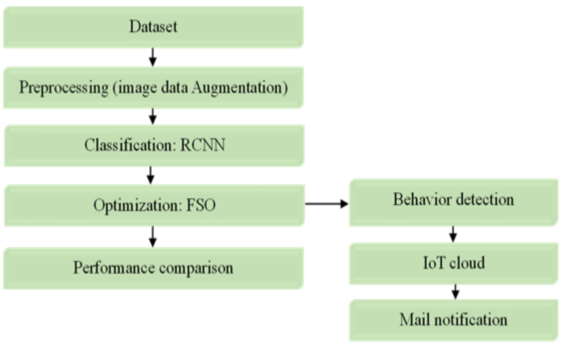

Fig 1. The architecture of the proposed methodology

A. *Preprocessing: Image data augmentation*

The initial phase involves collecting and altering photos of wild creatures. Data augmentation techniques like as rotation, scaling, and flipping are used to these images in order to expand the dataset and improve the robustness of the model. This step ensures the neural network can generalize well to various scenarios and lighting conditions. Data augmentation is the process of artificially producing new data from preexisting data. It is mainly utilized to train new machine learning (ML) models. The first training of machine learning models requires large and diverse datasets. Still, it can be challenging to locate sufficiently different real-world datasets because of data silos, legal constraints, and other problems. Data augmentation involves minor adjustments to the original data to enlarge the dataset artificially. Many sectors increasingly use generative artificial intelligence (AI) technologies for quick, high-quality data augmentation. Some operations of data augmentation

- Rotation: As the name implies, a rotation procedure rotates the image by a predetermined amount.
- Sharing: Shearing can also be used to change the image's orientation.
- Zooming: The zooming action provides the ability to zoom in or out.
- Cropping: You can resize the image by cropping it or focusing on a particular portion.
- Flipping: Flipping can change the image's orientation. Flipping can be done vertically or horizontally.

B. *Proposed algorithm RCNN (Region convolution neural network)*

The first effort to develop an object identification model that requires a CNN that has already been trained is R-CNN. Next, a quick evaluation of the Fast R-CNN is conducted. The process of producing region ideas must be taken into account, even if it is faster than the R-CNN.After a binary classification task, the RPN generates a score that indicates whether the anchor or the region of the image under evaluation includes a background or foreground object. There are two possible scores: "object" and "no object." The objectless score is the name given to this number. The classifier network generates and displays the score for each class to which the objects would be assigned. The architecture of RCNN shown in figure 2.

C. *Algorithm 1 (RCNN)*

- Initialize the parameters
- Initialize and process every frame
- Reading and returning every frame with its coordinate points
- Creating anchor box from every frame
- Passing anchor box frame to convert as input
- Extracting anchor and bounding box to detect an object
- For i=1 (No. Of object detections in a range):
  - Anchor box – bounding box (0,0, i)
  - Bounding masks – masks [i]

Left box, right box, top box, and bottom box – int[box*masks];

  - Square box- int[start(x,y)], [end(x,y)]
  - Detecting object in boundary box[i]

Thus far, method one has been devised, whereby the region proposal component and the object detection component are considered distinct entities. The efficacy of the region proposal method determined the classifier's performance, which fed the outputs of the area proposal approaches. The study on RCNN suggested applying a single strategy to both tasks, which meant that the classifier and the region suggestion used the same convolutional characteristics.

An updated bounding box Ri close to the controlled box (Ri, f (Ri)) that was there before. It is midway between the neural network's convolution and ROI pooling layers. If the divide and conquer model cannot find the required bounding box, it will replace the slower R-CNN by repeatedly improving its architecture. An area of interest is first searched over the preferred bounding box. Subsequently, the interest region is divided recursively into searches over progressively smaller subregions.

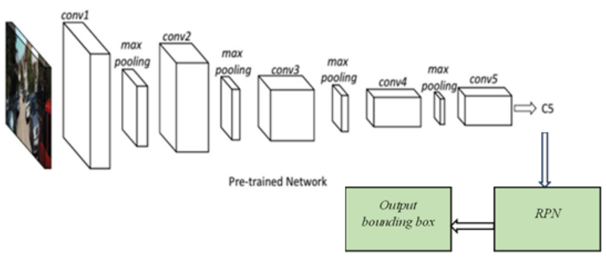

Fig 2. Architecture of RCNN

Allow Ri to serve as the initial bounding box for each iteration. The probability that the i[th] region of interest is situated in the region Ri is indicated by the score f (Ri). The rim vector is another name for the vector [Ti, Bi, Li, Ri]. The same search methodology is used, but fewer and fewer subregions are covered. The recursive tree ends when region R (the leaf node) has fewer bounding boxes than that. The probability P (Ri) is then calculated using the responses to the fragments. This process is summarized in Algorithm 2.

D. Algorithm 2(RCNN)

- Describe the search region encompassing Ri
- Evaluate the p(Ri)
- Evaluate the f(Ri)
- Continue steps 1-4 until the amount of P(Ri) reaches 0 or even the amount of iterations T exceeds a specified threshold.
- Change the coordinates to Ri.

The initial bounding box score is evaluated using a classifier and ROI pooling. The region suggestions are then extracted using a selective search technique, and the area proposals with the highest score are selected as the starting bounding box. The non-maximal suppression (NMS) approach is the name of it.

The embedded design of the proposed algorithm. The existing two-stage detection pipeline's architecture allows the context refinement technique to be used in both the final refinement phase and the region proposal generating stage. The leaf node is the region in the recurrence tree, Ri. The area Ri of the node and the consistent region of the node's ascendants are where the Ri samples are favored in the recurrence tree. They come from the root node Ri. The root node is indicated by equation (1), and it can be found by using the formula to find the bounding box solution, f (Ri).

$$f(Ri) \sim maximum\{Ri|R'\epsilon Ri'\} \tag{1}$$

Were,

- Ri is a particle's position or solution in a search space.
- $f(Ri)$ fitness function value of that solution.
- Ri' neighboring solutions around RiR_iRi (based on mutation, crossover, or heuristic rules).
- R' a candidate solution from the neighborhood.

The joint scoring approach aims to reason about the section's distinctive regions of interest collaboratively. The paired potential model's spatial connection between candidates is distinct from the prior research's unary potential score functions, which are well-defined by the local object detector at a matching position. The notation kij 1, k. 1,..., k. represents the edge cluster index Ij. Together with the θ_i^U and θ_(ij,)^P, the joint score function S (y, w) of unary and paired potentials, which connects all of the trainee labels in a single image and is the trainable parameter, is dependent on y.

$$S(y,w) = \sum_{i\epsilon S} y_i \theta_i^U(\omega) + \sum_{(i,j)\epsilon E} y_i y_j \theta_{ij,k_{ij}}^P(\omega) \tag{2}$$

Where y_i is the label/decision variable for node I, ω is the uncertainty parameter, $\theta_i^U(\omega)$ is the unary potential $y_j\theta_{ij,k_{ij}}^P(\omega)$ is the pairwise potential S is the set of nodes, and E is the set of edges representing relationships between nodes.

Image and potential correlation are achieved by applying the score function Eq. (2) to a feedforward neural network. A region of interest's feature vector can be mapped using the extra feedforward networks called pairwise and unary networks. The pairwise potential in the pairwise potential maps the concatenated pairs of a region of interest, whereas the bounding box in the unary potential maps an area of interest. An excellent style manual for science writers is [7].

IV. Fish swarm algorithm (FSA)

The Fish Swarm Optimization Algorithm is incorporated to enhance the RCNN's prediction accuracy. This bio-inspired optimization algorithm mimics the natural behavior of fish swarms to find optimal solutions. By optimizing the RCNN's weights and parameters, this algorithm helps achieve higher accuracy and better performance of the behavior classification model.

An algorithm for fish swarming that was motivated by the way aquatic fish move in liquid media. The target is repeatedly approached after being chosen at random. Being short-sighted is regarded as the first stage, with consequences at the end. The parameters are taken into consideration, and the initial values are fixed. The best optimal solution is reached by appropriate initial value selection. Fish can travel great distances in pursuit of a more expansive habitat that suits them. As a result, the fish can endure any unfavorable circumstance. However, some faults may result in poorer stability at larger values. One possible component of local search with a bigger FSA display location is global search. To identify the parameters that will make the algorithm accurate and reliable, the best fitness can be found. Fish are swift navigators that can move from local best search results to a target. The FSA method has undergone numerous design modifications to accommodate various problems. The differences in the algorithm can be classified as FSA solutions for discrete and continuous, composite and binary, multi-parameter, and hybrid FSA. The flow diagram of FSA shown in figure 3.

The four primary behaviors observed in the FSM are following, swarming, preying, and random behaviors. Fish use preying to travel into areas with high food concentrations. The mathematical representation is given in Eq. (3), considering the fish's optical distance.

$$X_j = X_i^{(t)} + V * rand() \tag{3}$$

Were,

- X_j This is the position after applying the velocity update with randomness, $X_i^{(t)}$ Denotes the current position of the particle i at iteration t.
- V This is a velocity or a step-size scaling factor.
- $Rand\ ()$ Represents a random number drawn from a uniform distribution, usually in the range [0,1].

If the fish prey is unsatisfied after a certain number of attempts, it will be computed randomly using Equation (4).

$$X_i^{(t+1)} = X_i^{(t)} + V * rand() \tag{4}$$

Where,

- $X_i^{(t+1)}$ The updated position.

- $X_i^{(t)}$ The current position of the $t+1$ individual at iteration t.
- $rand()$ A random number generated from a uniform distribution, typically in the range [0,1].

Fish cooperate to protect themselves from danger. Equation (5) can mathematically calculate a fish swarm's central position.

$$x_{cd} = \frac{\sum_{j=1}^{nf} x_{jd}}{nf} \tag{5}$$

Where x_{cd} is The centroid or mean coordinate of all nf individuals in the d dimension, x_{jd} is the coordinate value of the j, and nf is The number of individuals.

Fish when they find a good food concentration. Equation (6) represents the preying movement for fish in step movement.

$$X_i^{(t+1)} = X_i^{(t)} + S * rand() * \left(\frac{X_j - X_i^{(t)}}{\left|\left|X_j - X_i^{(t)}\right|(t)\right|}\right) \tag{6}$$

Where

- $X_i^{(t+1)}$ Direction vector from x_i to X_j

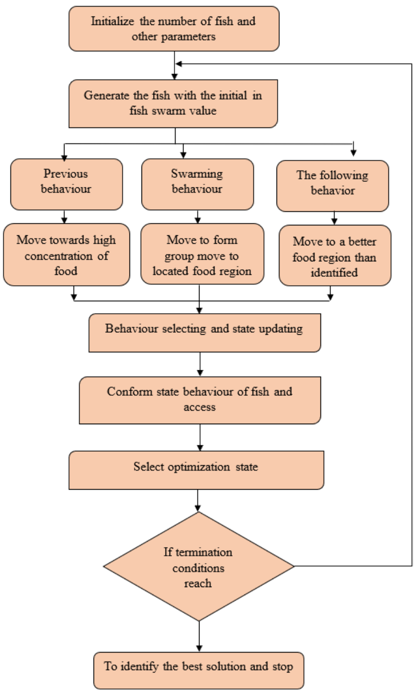

Fig 3. Flow diagram of fish swarm optimization

V. IoT Monitoring and Notification

A. IoT Integration

IoT devices are deployed within the wildlife coastal area to capture real-time data on animal behavior continuously. These devices have cameras and sensors that monitor the animals and feed data into the system. The integration of IoT ensures a constant stream of data, enabling real-time monitoring and analysis. IoT makes it possible to analyze animal behavior in real-time in Adafruit Cloud by employing sensors to gather information about posture and movement. The "animal" or "animal-name" feeds on the dashboard receive data from these sensors. When an animal sits, for instance, the feed adjusts to reflect this behavior. This facilitates ongoing activity monitoring without the need for human inspections. IoT is essential for intelligent animal monitoring systems because it allows users to analyze these feeds, identify trends in behavior, identify issues early, and take prompt action.

B. Cloud and Notification

The processed behavior status is updated in the Adafruit cloud, which provides a centralized data storage and access platform. This cloud-based system allows wildlife conservators and researchers to access the information remotely. Additionally, the system is configured to send email notifications with updated behavior status, ensuring that relevant parties are promptly informed about any significant changes in animal behavior.

VI. Result and Discussion

A. Experiments setup

This study was implemented using an anaconda spyder tool to identify the behavior of wild animals. The table 1 shows the detailed experiment specifications of this study.

TABLE 1 EXPERIMENTS SPECIFICATION

Metrics	Description
Software Tool	Anaconda Spyder (CPU) or Google colab (T4 GPU)
RAM	16 GB
Python version	3.11
Tensorflow version	2.13.0
Keras version	2.13.1

B. Data collection process and dataset

The dataset is collected from an animal image dataset with 90 pieces. The dataset contains five folders of the behavior of images like aggressive, calm, laydown, sitting, and standing of wild animals. A total of 643 images with the shape of 243, 243,3 from the five categories. The number of images in five individual categories is graphed in Figure 4.

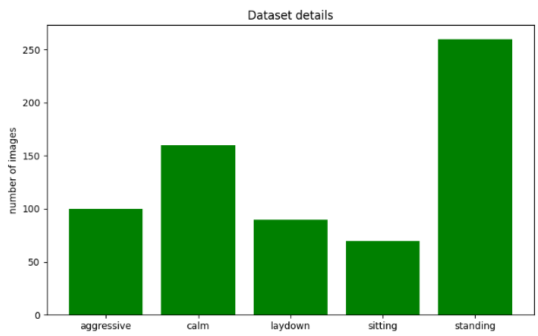

Fig 4. Details of the proposed dataset

C. Performance Metrics

A confusion matrix is a fundamental tool in evaluating the performance of a classification model, particularly in binary classification tasks. It displays the predicted and actual value counts, providing a detailed breakdown of the model's performance.

TABLE 2: PERFORMANCE ANALYSIS

Methods	Precision	Recall	Accuracy	F1 score
RCNN	0.92	0.88	0.91	0.90
RCNN -TL	0.83	0.83	0.93	0.83
RCNN+FSO	0.85	0.84	0.94	0.84

Techniques like feature selection optimization (FSO) aid in improving the precision of models like RCNN by removing superfluous or unnecessary features, which helps the model concentrate on important patterns that differentiate across classes. Precision is increased because fewer false positives result from this noise reduction. The model's comprehension and prediction confidence may also be increased by adjusting hyperparameters, implementing more reliable data augmentation, and utilizing transfer learning with domain-specific data. Precision is successfully increased by the improvement of learning procedures and input data.

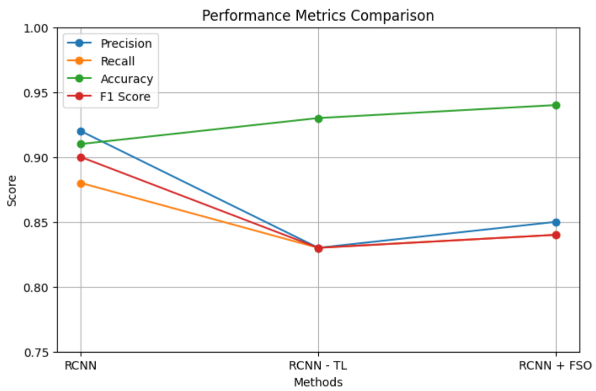

Fig 5. Performance analysis of existing and proposed system

The proposed system analysis of wild animal behavior using RCNN with the FSO algorithm produced performance metrics compared with the other algorithms, such as RCNN and RCNN-TL, as shown in Table 2 and Figure 5.

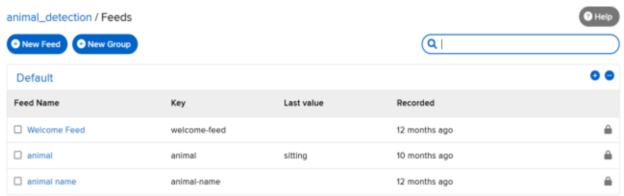

Fig 6. Animal behavior analysis in Adafruit cloud

IoT is essential to animal behavior analysis using Adafruit Cloud because it makes it possible to collect data in real time and monitor it remotely linked to sensors. When a device detects movement or sitting, it may transmit updates to the cloud, which stores, visualizes, and analyses data. This makes it possible for farmers or researchers to effectively monitor animals' health and activity patterns through readily available cloud platforms like Adafruit. The integration of IoT guarantees ongoing, automated monitoring without human involvement, improving decision-making in animal care and administration, as shown in Figure 6. Additionally, it sends periodic email notifications with the latest data to designated recipients, as shown in Figure 7.

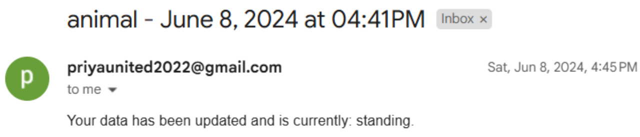

Fig 7. Mail notification of animal behavior

VII. CONCLUSION

The region-based convolution neural network, or RCNN, was developed in this suggested system to classify animal behavior. The performance metrics of proposed animal behavior are precision, recall, and accuracy. The F1 score values calculated depend on the RCNN model with bio-inspired optimization algorithms, such as fish swarm optimization, to enhance the model's performance. The designed IoT device's embedded system may use the suggested model to train it end-to-end and use it to classify animal behavior in real-time and across data collection. By combining RCNN with Fish Swarm Optimization and IoT technology, this methodology offers a robust and efficient solution for monitoring and classifying animal behavior. Data augmentation, feature extraction and selection, and advanced classification and optimization techniques ensure high accuracy and real-time updates, ultimately enhancing animal welfare and management. The prediction results as precision, recall, accuracy, and F1 score of 0.85, 0.84, 0.94 and 0.84.

REFERENCES

[1] Yi, Dewei, Hasan Bayarov Ahmedov, Shouyong Jiang, Yiren Li, Sean Joseph Flinn, and Paul G. Fernandes. "Coordinate-Aware Mask R-CNN with Group Normalization: An underwater marine animal instance segmentation framework." Neurocomputing 583, pp. 127488, 2024.

[2] Tu, Shuqin, Weijun Yuan, Yun Liang, Fan Wang, and Hua Wan. "Automatic detection and segmentation for group-housed pigs based on PigMS R-CNN." Sensors 21, no. 9, pp. 3251, 2021.

[3] Bello, Rotimi-Williams, Ahmad Sufril Azlan Mohamed, and Abdullah Zawawi Talib. "Contour extraction of individual cattle from an image using enhanced Mask R-CNN instance segmentation method." Ieee Access 9, pp. 56984-57000, 2021.

[4] Ibraheam, Mai, Kin Fun Li, Fayez Gebali, and Leonard E. Sielecki. "A performance comparison and enhancement of animal species detection in images with various r-cnn models." AI 2, no. 4, pp. 552-577, 2021.

[5] Li, Wei, Yang Xiao, Xibin Song, Na Lv, Xinbo Jiang, Yan Huang, and Jingliang Peng. "Chicken image segmentation via multiscale attention-based deep convolutional neural network." IEEE Access 9, pp. 61398-61407, 2021.

[6] Tang, Jiaxi, Yaqin Zhao, Liqi Feng, and Wenxuan Zhao. "Contour-Based Wild Animal Instance Segmentation Using a Few-Shot Detector." Animals 12, no. 15, pp. 1980, 2022.

[7] Syahiran Soria, Mohamad, Khalif Amir Zakry, Irwandi Hipiny, Hamimah Ujir, Ruhana Hassan, and Alphonsus Ligori Jerry. "An Instance Segmentation Method for Nesting Green Sea Turtle's Carapace using Mask R-CNN." International Journal of Computing and Digital Systems 16, no. 1, pp. 201-211, 2024.

[8] Chang, Chin-Chun, Yen-Po Wang, and Shyi-Chyi Cheng. "Fish segmentation in sonar images by mask r-cnn on feature maps of conditional random fields." Sensors 21, no. 22, pp. 7625, 2021.

[9] Schwartz, Shawn T., and Michael E. Alfaro. "Sashimi: A toolkit for facilitating high‐throughput organismal image segmentation using

deep learning." Methods in Ecology and Evolution 12, no. 12, pp. 2341-2354, 2021.

[10] Noe, Su Myat, Thi Thi Zin, Pyke Tin, and Ikuo Kobayashi. "Automatic detection and tracking of mounting behavior in cattle using a deep learning-based instance segmentation model." Int. J. Innov. Comput. Inf. Control 18, no. 1, pp. 211-220, 2022.

[11] Duan, Enze, Hongyun Hao, Shida Zhao, Hongying Wang, and Zongchun Bai. "Estimating body weight in captive rabbits based on improved mask RCNN." Agriculture 13, no. 4, pp. 791, 2023.

[12] Schindler, Frank, and Volker Steinhage. "Identification of animals and recognition of their actions in wildlife videos using deep learning techniques." Ecological Informatics 61, pp.101215, 2021.

[13] Bello, Rotimi-Williams, Ahmad Sufril Azlan Mohamed, Abdullah Zawawi Talib, Daniel A. Olubummo, and O. Charles Enuma. "Enhanced deep learning framework for cow image segmentation." IAENG International Journal of Computer Science 48, no. 4, pp.1-10, 2021.

[14] Bello, Rotimi-Williams, and Mojisola Abosede Oladipo. "Mask the Target: Mask R-CNN-Based Approach for Open-Range Individual Cattle Segmentation." Rotimi-Williams Bello, Mojisola Abosede Oladipo," Mask the Target: Mask R-CNN-Based Approach for Open-Range Individual Cattle Segmentation 2023.

[15] Tu, Shuqin, Qiantao Zeng, Haofeng Liu, Yun Liang, Xiaolong Liu, Lei Huang, and Zhengxin Huang. "MaskDis R - CNN: An instance segmentation algorithm with adversarial network for herd pigs." IET Image Processing 17, no. 12, pp. 3488-3499, 2023.

[16] Mazur-Milecka, Magdalena, and Jacek Ruminski. "Deep learning based thermal image segmentation for laboratory animals tracking." Quantitative InfraRed Thermography Journal 18, no. 3, pp. 159-176, 2021.

[17] Haucke, Timm, and Volker Steinhage. "Exploiting depth information for wildlife monitoring." arXiv preprint arXiv: 2102.05607, 2021.

[18] Lyu, Haitao, Fang Qiu, Li An, Douglas Stow, Rebecca Lewison, and Eve Bohnett. "Deer survey from drone thermal imagery using enhanced faster R-CNN based on ResNets and FPN." Ecological Informatics 79, pp. 102383, 2024.

[19] Kassim, Yasmin M., Michael E. Byrne, Cristy Burch, Kevin Mote, Jason Hardin, David R. Larsen, and Kannappan Palaniappan. "Small object bird detection in infrared drone videos using mask R-CNN deep learning." Electronic Imaging 32, pp. 1-8, 2020.

[20] Mazur-Milecka, Magdalena, Tomasz Kocejko, and Jacek Ruminski. "Deep instance segmentation of laboratory animals in thermal images." Applied Sciences 10, no. 17, pp. 5979, 2020.

[21] Meena, S. Divya, and Loganathan Agilandeeswari. "Smart animal detection and counting framework for monitoring livestock in an autonomous unmanned ground vehicle using restricted supervised learning and image fusion." Neural Processing Letters 53, no. 2, pp. 1253-1285, 2021.

[22] Vecvanags, Alekss, Kadir Aktas, Ilja Pavlovs, Egils Avots, Jevgenijs Filipovs, Agris Brauns, Gundega Done, Dainis Jakovels, and Gholamreza Anbarjafari. "Ungulate detection and species classification from camera trap images using RetinaNet and faster R-CNN." Entropy 24, no. 3, pp. 353, 2022.

[23] Nguyen, Thi Thu Thuy, Anne C. Eichholtzer, Don A. Driscoll, Nathan I. Semianiw, Dean M. Corva, Abbas Z. Kouzani, Thanh Thi Nguyen, and Duc Thanh Nguyen. "SAWIT: A small-sized animal wild image dataset with annotations." Multimedia Tools and Applications, pp. 1-26, 2023.

[24] Salau, Jennifer, and Joachim Krieter. "Instance segmentation with Mask R-CNN applied to loose-housed dairy cows in a multi-camera setting." Animals 10, no. 12, pp. 2402, 2020.

[25] Bello, Rotimi-Williams, Ahmad Sufril Azlan Mohamed, and Abdullah Zawawi Talib. "Enhanced Mask R-CNN for herd segmentation." International Journal of Agricultural and Biological Engineering 14, no. 4, pp.238-24, 2021.

[26] Peng, Jinbang, Dongliang Wang, Xiaohan Liao, Quanqin Shao, Zhigang Sun, Huanyin Yue, and Huping Ye. "Wild animal survey using UAS imagery and deep learning: modified Faster R-CNN for kiang detection in Tibetan Plateau." ISPRS Journal of Photogrammetry and Remote Sensing 169, pp.364-376, 2020.

[27] Popek, Łukasz, Rafał Perz, and Grzegorz Galiński. "Comparison of different methods of animal detection and recognition on thermal camera images." Electronics 12, no. 2, pp. 270, 2023.

[28] Xu, Beibei, Wensheng Wang, Greg Falzon, Paul Kwan, Leifeng Guo, Guipeng Chen, Amy Tait, and Derek Schneider. "Automated cattle counting using Mask R-CNN in quadcopter vision system." Computers and Electronics in Agriculture 171, pp. 105300, 2020.

[29] Ulhaq, Anwaar, Peter Adams, Tarnya E. Cox, Asim Khan, Tom Low, and Manoranjan Paul. "Automated detection of animals in low-resolution airborne thermal imagery." Remote Sensing 13, no. 16, pp. 3276, 2021.

[30] Kim, Sueun, and Yuichi Hidaka. "Breathing pattern analysis in cattle using infrared thermography and computer vision." Animals 11, no. 1, pp. 207 2021.