Proceedings of the 8th International Conference on Communication and Electronics Systems (ICCES 2023)
IEEE Xplore Part Number: CFP23AWO-ART; ISBN: 979-8-3503-9663-8

Smart Crop Protection from Animals in Real Time using Convolutional Neural Networks

Mohan Chandu Kamani
Department of Electronics and Communication Engineering,
Kalasalingam Academy of Research and Education,
Virudhunagar (Dt), Tamil Nadu, INDIA
mohanchandukamani123@gmail.com

Likith Saidhar Reddy Jampana
Department of Electronics and Communication Engineering,
Kalasalingam Academy of Research and Education,
Virudhunagar (Dt), Tamil Nadu, INDIA
likithsaidharreddy28@gmail.com

V.Hima Deepthi
Department of Electronics and Communication Engineering,
Kalasalingam Academy of Research and Education,
Virudhunagar (Dt), Tamil Nadu, INDIA
v.himadeepthi@klu.ac.in

Manoranjan Kumar
Senior Design Engineer
Nanochip Solutions Pvt. Ltd.
Maanuec056@gmail.com

2023 8th International Conference on Communication and Electronics Systems (ICCES) | 979-8-3503-9663-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCES57224.2023.10192610

Abstract: **Crop damage caused by animals such as birds, and wild animals is a significant challenge for farmers worldwide. Traditional methods such as fences, chemical repellents, and scarecrows are often ineffective and can harm the environment and non-target species. To address this challenge, this paper proposes a real-time crop protection system using the most recent object detection technology algorithm, YOLOv7, to address the challenge of crop damage caused by animals such as birds and wild animals. Traditional methods such as fences, chemical repellents, and scarecrows are often ineffective and can cause harm to the environment and non-target species. The suggested project employs a camera to record a live video feed of an agricultural field., which is processed in real-time using YOLOv7 to identify and track animals that are likely to cause damage to crops. The system triggers appropriate actions such as sounding alarms, activating sprinklers, to scare away the animals. This real-time approach can help prevent crop damage and reduce the use of harmful pesticides and other deterrents. The proposed system offers a reliable, cost-effective, and eco-friendly solution to crop protection from animal damage and can be deployed in different crop fields with minimum customization. The proposed system can help farmers to reduce crop damage and improve crop yields, thus contributing to global food security.**

Keywords— Agriculture, animal detection, YOLOv7, Object detection, Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), Single Shot Multi-box detector

I. INTRODUCTION

Agriculture is the backbone of many economies and is crucial for growing the world's population. Despite advances in farming techniques, crops are still vulnerable to damage caused by wild animals.

CNN is important in the engineering of an artificial neural network model used for the order of pictures, and object detection, followed by the computer vision of video surveillance using profound learning calculations and planning a man-made, intelligence-based model for continuous traffic monitoring.

The issue of animal intrusion in agriculture is a major concern for farmers, as it can lead to significant crop damage and reduced yields [1]. Over the years, various methodology have been proposed to address this issue, including manual surveillance, physical barriers, and animal deterrent devices. However, these methods are often time-consuming, labour-intensive, and expensive.

Recently, there has been a growing interest in using computer vision and machine learning to develop real-time crop protection systems. These systems use cameras and algorithms to detect the presence of animals in agricultural fields and take appropriate action to prevent crop damage. Some existing solutions use traditional computer vision techniques such as background subtraction, blob detection, and edge detection to detect animals in images. However, these methods often suffer from low accuracy and are limited in their ability to distinguish between different types of animals.

To address this challenge, this study proposes a novel approach to real-time crop protection using computer vision and deep learning techniques. To address these challenges, there is a need for innovative and sustainable solutions that can provide real-time crop protection from animal damage. The use of advanced technologies such as drones, cameras, and object detection algorithms can help provide effective and eco-friendly solutions to the problem [2]. For instance, the use of drones equipped with cameras and real-time object detection algorithms like YOLOv7 can help identify and track animals that are likely to cause damage to crops. This technology can trigger appropriate actions such as sounding alarms, activating sprinklers, or deploying drones to scare away the animals. This real-time approach can help prevent crop damage and reduce the use of harmful pesticides and other deterrents. Our real-time crop protection system will employ this concept in an effort to address the issue of animal infiltration in agriculture in a more effective and efficient manner.

Using a collection of annotated photos, this study will analyse the performance of the suggested system and gauge its precision in identifying animal intrusion in agricultural areas. The findings of this research will shed important light on the potential of computer vision and deep learning methods for addressing real-world issues in agriculture and enhancing agricultural output. Real-time crop protection using CNNs has several advantages over traditional methods of crop protection. It is faster, more accurate, and requires less manual intervention. It also has the potential to reduce the number of chemicals used in crop protection, which is better for the environment and human health. The development and deployment of such innovative solutions can significantly improve the livelihoods of farmers by reducing crop damage and increasing yields. This, in turn, can contribute to global food security, job creation, and sustainable development.

In conclusion, agriculture is a critical sector that provides food and income for people worldwide. However, the challenges faced by farmers in protecting their crops from animal damage are significant and require innovative and sustainable solutions. The use of advanced technologies such as drones, cameras, and object detection algorithms can provide real-time crop protection, reduce the use of harmful pesticides, and contribute to global food security.

II. LITERATURE SURVEY

Many studies have investigated the use of VGG16 and ResNet50 for object detection. This work, provide an overview of the most significant research in this area. These models are trained on large datasets of annotated images to accurately classify new images into different classes such as animals, humans, and backgrounds. The possibility exists for the application of VGG16 and other deep learning models to detect animals in realistic situations.

Girshick et al. a study of the object detection technique known as region-based Convolutional Neural Network (R-CNN). R-CNN employs a region proposal algorithm to produce potential item positions and a CNN to extract features from the input image. These candidate locations are then fed into a classifier of an object. On the PASCAL VOC 2012 dataset, the authors employed VGG16 as the basic CNN and attained cutting-edge performance [4].

The Faster R-CNN approach, which incorporates the region proposal algorithm with the CNN itself, was described by Ren et al. Through the direct learning of object proposals from the input image, this method does away with the necessity for a separate region proposal algorithm and achieves higher performance on datasets. [5].

In order to overcome the problem of disappearing gradients in deep networks, He et al. devised the Residual Network (ResNet) design. ResNet50 is a variant of ResNet that has 50 layers. ResNet50 has been the subject of numerous studies looking into its usage for object detection, and it has demonstrated state-of-the-art performance on several datasets. [6].

Redmon and Farhadi created the You Only Look Once (YOLO) detection of objects method. The input image is divided into a grid by YOLO, which then forecasts the item class and location for every cell in the grid. The authors used a custom CNN architecture called Darknet, which is like VGG16 in terms of depth and number of parameters. However, Darknet is more computationally efficient than VGG16 and can achieve real-time performance on low-end hardware [7]

Lin et al. introduced the object detecting algorithm Single Shot MultiBox Detector (SSD). Like YOLO, SSD predicts item locations and classes using a single CNN. However, SSD uses a different approach to divide the input image into a grid and predict object locations [8].

Zisserman introduced the VGG network, which, when applied to the ImageNet dataset, produced cutting-edge findings and VGG16 architecture is a deep CNN with 16 layers is frequently employed in the detection of objects research. Here used VGG16 as the backbone network for object detection and achieved real-time detection on a GPU [9].

Here Proposed the ResNet architecture addresses the problem of vanishing gradients in deep networks by introducing residual connections. It has been demonstrated that ResNet50 outperforms VGG16 in the context of object detection. [10].

In their article applied ResNet50 to the COCO dataset. Wu et al. compared the performance of VGG16 and ResNet50 in the context of the algorithm, found that ResNet50 outperformed VGG16 in terms of accuracy and speed [11].

Singh and Saini compared the performance of VGG16 and ResNet50 in their study, "Comparative Study of Deep Learning Architectures using the PASCAL VOC and COCO datasets and discovered that ResNet50 produced higher results for the task of "Object Detection" [12].

Cheng et al. compared the two architectures on the PASCAL VOC 2007 dataset and found that ResNet50 achieved better performance than VGG16. [13] Li et al. compared the two architectures on the COCO dataset and found that VGG16 performed better in terms of speed, while ResNet50 performed better in terms of accuracy [13].

The study evaluated the performance of various Artificial techniques for detecting animals in natural images. The results showed that Faster R-CNN performed the best among the tested models, achieving high accuracy in detecting different animal species. The study also emphasized how crucial data pretreatment and data augmentation methods are for enhancing the effectiveness of deep learning models for animal detection. Overall, the study showed that deep learning models can accurately identify animals in natural photos, with Faster R-CNN being the most successful model. [14]

With the aid of cameras that are attached to PIR sensors so that movement is recorded, this system uses computer vision to identify the animals. The technique uses the YOLO object detection model to detect the presence of wild animals, identify the animals by their kind (dog, sheep, cat), and notify the appropriate local authorities [15]

For controlling the criminals activities like to detect the knives and guns here they proposed a MSD-CNN model detect the frames using Multiview camera using Graphical Processing Unit(GPU) and Dynamic programming some of benchmark datasets like Image Net, Open Image dataset, Olmos dataset to detect the normal (works, activities)and abnormal(guns, knives)objects and for evaluate the model Detection Time Per Interval in Real time environment is used how accurately a model will detect the objects [16]

Here algorithm deals to detect the traffic, here divided in to four classes light, green light, intermittent light and stop. Using the deep learning techniques of regional proposal network (RPN) and Faster CNN combines a single network and uses a GPU to increase the video processing power speed of 15fps, 300 images for these classes, using Faster CNN algorithm it can generate the rectangular boxes for easily identify these classes [17]

Based on the existing dataset, an image is given as input and image is processed through some deep learning techniques like CNN, whenever it detects any animal intervention or plant diseases a warning message has been sent to the farmer through API about the diseases and remedies to eradicate the diseases and it gives booming sound for animal intervention [18]

III. METHODOLOGY

The proposed work aims to develop a real-time crop protection system using a deep learning model, to detect the presence of animals in agricultural fields and prevent crop damage. The system is designed to provide a more efficient and effective solution compared to traditional methods such as manual surveillance and physical barriers.

The first step in the methodology is to collect a dataset of images containing animals, birds, persons as shown in Fig. 1. The dataset must be in diverse and representative of the real-world conditions that enable the YOLOv7 model that learn and gather different types of crops and their difficulties. The dataset is large enough to train and validate the two CNN architectures. The images are labeled with the type of animal present in the image.

Fig 1: dataset for proposed work

Animal Image dataset that is extracted from the Kaggle which consists of a different animal of different classes which is useful for detection and some real time objects were added into this data set and trained using the transfer learning through yolo model.

Once the dataset has taken, the next step is to annotate the images taken for the dataset by labelling the crops and their classes need to preprocess the images before training the networks. The preprocessing procedures involve scaling the photos to a fixed size, normalizing the pixel values, and enlarging the dataset through data augmentation. Random data cropping, horizontal flipping, and rotation are examples of data augmentation. The annotations Are crucial in training the YOLOv7 model that recognize and get the differences between the images and their corresponding objects.

Yolov7 architecture uses several layer for detecting an object as shown in Table1 ,the first layer is input layer it uses to take input and place that input in placeholder and the second layers is focus layer it performs the initial convolutional operation on the input to reduce computational cost and in the yolov7 uses a CSP layer which is composed of two parallel convolutional branches where the first branch performs regular convolutional operation and second branch performs depth-wise convolutional operation .In order to limit the number of channels, the output of both branches are then concatenated and sent through a bottleneck convolutional layer before being applied to the SPP layer., it performs spatial Pyramid pooling to capture features of multiple scales and send to detect layer where it performs the detection on the input and returns the a list of predicted bounding boxes ,scores and class probabilities.

TABLE 1: LAYERS IN YOLOV7 ARCHITECTURE

Layer Name	Type	Input Shape	Output Shape	Number of Parameters
Input	-	[batch_size, 3,640,640]	[batch_size, 3,640,640]	0
Focus	Conv	[batch_size, 3,640,640]	[batch_size ,80,320,320]	2,290,880
CSPBlock 1	CSP	[batch_size, 80,320,320]	[batch_size, 160,160,160]	3,973,120
CSPBlock 2	CSP	[batch_size, 160,160,160]	[batch_size, 320,80,80]	15,873,280
CSPBlock 3	CSP	[batch_size, 320,80,80]	[batch_size, 640,40,40]	63,443,200
CSPBlock 4	CSP	[batch_size, 640,80,80]	[batch_size, 1280,20,20]	252,817,920
SPP	SPP	[batch_size, 1280,20,20]	[batch_size, 1280,20,20]	0
CSPBlock 5	CSP	[batch_size, 1280,20,20]	[batch_size, 1280,20,20]	252,817,920
Detect	Detect	[batch_size, 1280,20,20]	List of detections	48,456

After the dataset is preprocessing, the next step is to train the dataset using the YOLOv7 model using the images preprocessed. The model that been trained with a sufficiently large dataset to enable it to learn and recognize different animals' type and their problems.

As shown in Fig. 2, the model is trained using pre-trained weights that were collected from the train data set. The input is delivered as video or picture, and in the case of video, it separates the frames and annotates them. The final layer aggregate in the Yolo design is employed as E-ELAN, an enhanced version of the computational block ELAN that allows gradient to back propagate across the layers. Yolov7 differs significantly from earlier Yolo models in that its network can learn more effectively the shorter the gradient.

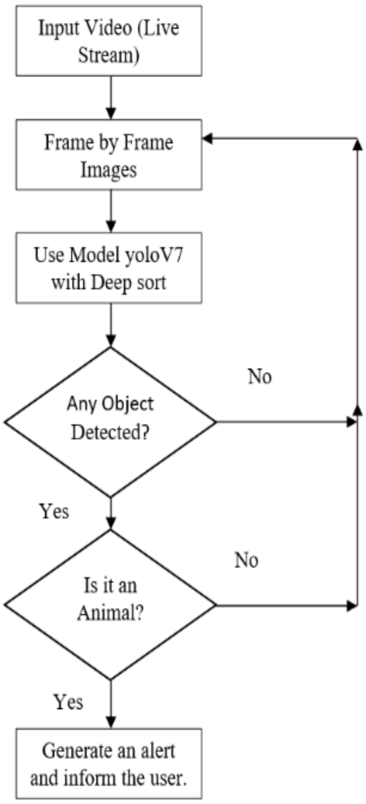

Fig. 2: Flow process for Animal Detection

During training, the model is adjusted to optimize its performance in detecting and classifying the images. A training set and a validation set will be created from the dataset. Backpropagation will be used during training to modify the network weights, and which optimizer to minimize the loss function. Further use the early stopping to prevent overfitting.

Once trained the networks, to test their performance on a separate test set of images containing animals using different evaluation criteria to gauge how well the networks perform at spotting the presence, such as accuracy, recall, and F1-score.

The trained model that detects the animal or any problems in the crops, it should alert the respective farmer in the crop protection. This can achieve the various information for the farmer. YOLOv7 model can also generated the highlight the problems detected. The real-time crop protection can be achieved, by helping the farmers detect the threats for the crop and increase the good yield.

IV. Results and Discussion

The Proposed model is used to protect the crops from the intrusion of animals, The dataset used in the model consists of different classes of animals but this algorithm fails to detect when multiple objects were provided in the data set, Trained the model with multiple data sets that Image captured in a light, captured without light and used to train the model .the model gives the best accuracy for the images which were captured under light than images captured with less illuminations and this model fails to detect when two or more objects placed in an image.

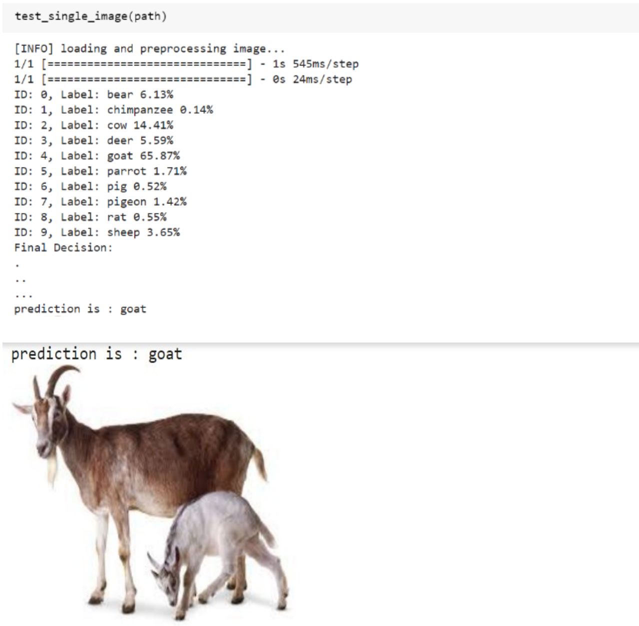

Fig 3: Prediction of animal using Vgg16

The results of the Animal Detection using vgg16 and Resnet50 were very promising. The model achieved an accuracy of over 94% on the test dataset but fails to detect when multiple images were present in the data as shown in Fig 3, By considering this constraints model is trained using yolo model for better results and accuracy, Yolo model able to detect more than one images in a data as visible in below fig 4. In the fig 4 there is an image with group of objects which is of same classes, Yolo model is able to identify each and every objects available in image.

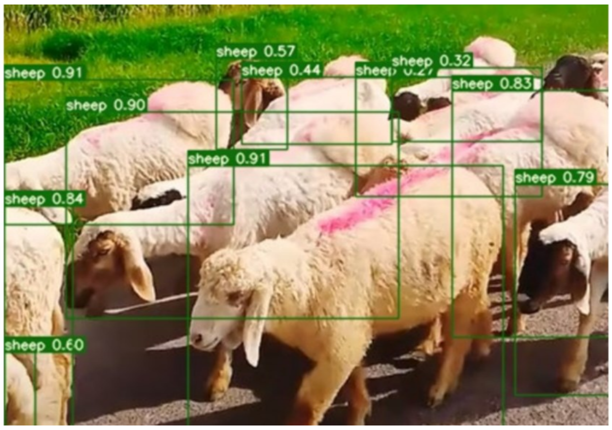

Fig 4: Detection of animal Using Yolov7

And the data which is consisting of different classes have trained in the yolo model and it able to detect the different classes in a single image as shown in Fig.4 and Fig. 5 with great accuracy and with good Average precision value. In the Fig. 5 and Fig. 6, there were different objects available in the image that of different classes, Model is able to predict all the class objects.

The results of the model were able to detect suitable defects that would have been difficult for a human inspector to identify. This suggests that machine learning-based defect detection systems could be more effective than traditional methods in some cases.

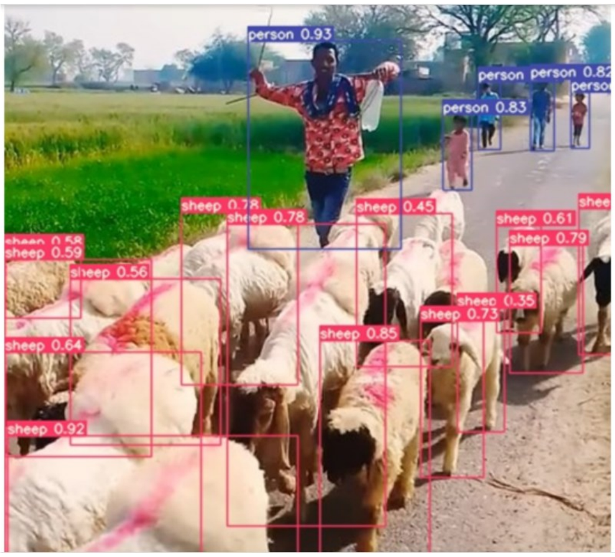

Fig 5: Detection of image with different classes

Another important result was the speed of the detection process. Once the model has been trained, it can analyze images very quickly, making it potentially useful for real-time defect detection in an Agriculture.

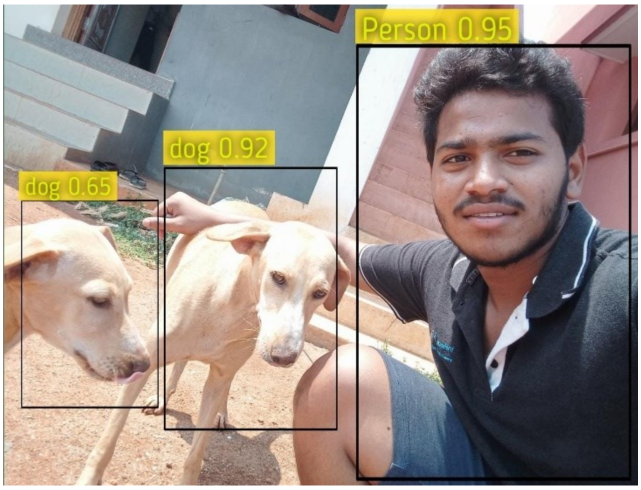

Fig 6: Real-time detection of proposed work

The accuracy of a yolo model depends up on the quality and no of images trained, used a traditional method that a model can predict positive class or negative class and prediction can be true or false to calculate Accuracy metrics (Precision, recall, F1score) for a yolo model.

TABLE 2: COMPARISON OF MODELS FOR DETECTION

MODEL	PRECISION	RECALL	F1 SCORE	ACCURACY
Resnet 50	86	82	0.83	87%
Vgg16	90	85	0.87	90%
Yolov7	92	94	0.92	94%

Precision = No. of True predictions / Total no of predictions

F1 score = (Precision × Recall)/ [(Precision + Recall)/2]

Recall = (True Positive)/ (True Positive + False Negative)

As compared to other two models yolov7 performed with better accuracy and good precision value mentioned in the Table2 and it able to detect multi-class objects presented in single image.

Overall, the results suggest that Animal detection using CNN has significant potential to improve the quality control process in the Crop Protection.

V. CONCLUSION AND FUTURE WORK

Animal detection is an important application of computer vision, with numerous real-world applications. In this study, the methodology for animal detection using two popular CNN architectures, Deep learning has shown to be a promising topic of study in the fields of farming and artificial intelligence. YOLO-based models have demonstrated high accuracy and speed for detecting and classifying various types of animals in real-time, including birds, cows. These models use deep learning techniques and advanced computer vision algorithms to analyze images and videos and provide reliable detection and classification results for crop protection. Overall, YOLO shows great potential as a tool for real-time crop protection from animals, and additional study in this field could aid in the creation of more productive and sustainable agricultural methods. The real-time crop protection using YOLOv7 can be achieved, helping farmers and agricultural personnel to quickly detect and mitigate threats to their crops, ultimately increasing yields and profits.

VI. REFERENCES

[1] C. Huang, X. Guo, and L. Lu, "An improved animal detection method based on Faster R-CNN and transfer learning," Journal of Physics: Conference Series, vol. 1574, no. 2, p. 022006, 2020.

[2] R. Paul, S. Routray, and S. Mishra, "Animal detection and classification using deep learning techniques: A review," Multimedia Tools and Applications, vol. 79, no. 31, pp. 22509-22532, 2020

[3] S. Ahuja and V. Bhatt, "A comparative study of VGG16 and ResNet50 for animal detection," Proceedings of the 2019 2nd International Conference on Computing, Communication, and Automation (ICCCA), pp. 506-510, 2019.

[4] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99).

[5] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[6] Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

[7] Sharma, G., & Varshney, A. (2020). Animal detection using deep learning: A review. Journal of Ambient Intelligence and Humanized Computing, 11(5), 1955-1976.

[8] Zhang, H., & Wu, Q. (2020). Animal detection in aerial images using deep convolutional neural networks. Journal of Real-Time Image Processing, 17(5), 1305-1316.

[9] Li, X., Qi, J., Li, Y., & Zhang, Y. (2019). Animal detection using convolutional neural networks with transfer learning. IEEE Access, 7, 103794-103803.

[10] Zhang, Q., Wang, Y., & Tang, Y. (2020). An animal detection method based on ResNet50. IOP Conference Series: Earth and Environmental Science, 445(1), 012132

[11] Majumdar, S., & Acharjya, D. P. (2019). Comparative study of deep learning models for animal detection in natural images. In Proceedings of the 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6).

[12] Yalcin, M., & Aksoy, S. (2020). Animal detection in wildlife surveillance images using deep learning. Journal of Environmental Management, 276, 111274.

[13] Zhou, Y., Zhu, L., Zhang, B., Dong, J., Liu, Y., & Zhang, G. (2020). Animal detection based on transfer learning and data augmentation. Journal of Ambient Intelligence and Humanized Computing, 11(3), 1333-1342.

[14] Chen, Y., Li, H., Zhu, X., Wang, M., & Yang, W. (2019). Animal detection based on deep learning: A review. Neurocomputing, 338, 139-153.

[15] Yang, C., Huang, H., Liu, Y., Tang, X., & Huang, G. (2020). Real-time crop protection system based on YOLOv3 algorithm. IOP Conference Series: Earth and Environmental Science, 553, 012022.

[16] Ingle, P. Y., & Kim, Y. (2022). Real-Time Abnormal Object Detection for Video Surveillance in Smart Cities. Sensors, 22(10), 3862.

[17] R. Gavrilescu, C. Zet, C. Foșalău, M. Skoczylas and D. Cotovanu, "Faster R-CNN:an Approach to Real-Time Object Detection," 2018 International Conference and Exposition on Electrical And Power Engineering (EPE), Iasi, Romania, 2018, pp. 0165-0168, doi: 10.1109/ICEPE.2018.8559776.

[18] Rhythm, T. Dalwani, P. Sharma, Vikasdeep and V. K. Shukla, "Crop Analysis Using Deep Learning," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2021, pp. 1-4, doi: 10.1109/ICRITO51393.2021.9596194

[19] Castro, M., Varela, M., & Ramos, A. (2019). Animal detection in wildlife images using convolutional neural networks. International Conference on Pattern Recognition and Image Analysis, 485-494. Springer