Automatically identifying animal using single shot Detection

[1]Mr.Yashraj Pandey, [2]Mr.Srikanth.M, [3]Mr.S Suranpal singh Bali,
[4]Mrs.O.Vasantha kumari [5]Mr.P.Dinesh Kumar, [6] Dr.G.Victo Sudha George

[1][2][3]Student, B.tech CSE, Department of CSE, Dr.M.G.R. Educational and Research Institute
[4]Assistant Professor, Department of CSE, Dr.M.G.R Educational and Research institute. .
[5]Professor, Department of CSE, Dr.M.G.R. Educational and Research Institute.
[6]Professor, Department of CSE, Dr.M.G.R. Educational and Research Institute.

Abstract: Crop damage caused by animal attacks is one of the major threats in reducing the crop yield. Due to the expansion of cultivated land into previous wildlife habitat, crop raiding is becoming one of the most antagonizing human wildlife conflicts. Farmers in India face serious threats from pests, natural calamities & damage by animals resulting in lower yields. Traditional methods followed by farmers are not that effective and it is not feasible to hire guards to keep an eye on crops and prevent wild animals. Since safety of both hum an and animal is equally vital, it is important to protect the crops from damage caused by animal as well as divert the animal without any harm. Thus, in order to overcome above problems and to reach our aim, we use machine learning to detect animals, entering into our farm by using deep neural network concept, a division in computer vision. In this project, we will monitor the entire farm at regular intervals through a camera which will be recording the surrounding throughout the day. With the help of a machine learning model, we detect the entry of animals and we play appropriate sounds to drive the animal away. This report specifies various libraries and concepts of convolutional neural networks used to create the model.

Keywords: Convolutional Neural Network, Deep learning, Prediction, Training and Validation, Play Sound.

I. INTRODUCTION

Agriculture meets the food demands of the population and also provides various raw materials for different industries. Interference of animals in agricultural lands causes a huge loss of crops. Crop damage due to raiding wild animals has become a major issue of concern these days. Animals like wild boars, macaques, porcupines, deer, monkeys and bears are extremely destructive and have also caused human casualties in certain occasions. Small farmers can even lose up to half of their yield to animals and they cannot take any harsh measures due to the strict wildlife laws. Human-elephant conflict is rising intensely as elephants are a highly conflict prone wildlife species, especially in India. Thus, there is need for a system to detect any intrusion which can help the farmers to drive away these animals as soon as they learn about their intrusion.

Computer vision is applicable to many fields like medical field, robotics, remote sensing, machine vision, content-based image retrieval. Computer vision solves many problems in different disciplines. Computer

vision also applied in the security field to perform automatic surveillance and access control and attendance management. The computer vision can be applied in agriculture field in many ways like disease detection of a tree by examining leaves or flowers or fruits and quality control of agricultural products.

The computer vision techniques can be applied in order to provide security from wild animals in agriculture. In agriculture fields near to forest areas have a severe threat from wild animals, which attacks regularly on farms. These attacks causing huge damage to agricultural crops subsequently causes significant financial losses to farmers.

Some measures are taken by the farmers by installing electrical fences to the farms, big flood lights in the farm. Some even resort to hiring guards. Installing an electrical fence is much costlier to equip huge farms and kills so many animals, which is even illegal in certain places and affects the biodiversity. Other existing techniques also are not effective due to several reasons, cost being one of them.

In this project, we proposed a new and cost effective solution for agriculture security from animals. It is a proactive solution which gives alerts to the farmers when animals come near to the farms. It also causes certain siren to be played whenever any animals are detected and is directed towards the animal in an attempt to scare them away. Here, we are implementing a solution that recognizes animals when it is captured on camera.

II. RELATED WORK

[1] Computer Vision is applied in agriculture field for food grading, disease identification of the plants and agro-farms security. Huge crop damage is caused by the wild animal attacks on the agriculture farms. Here are some traditional techniques followed by the local farmers, but which are not effective. This problem can be solved using computer vision techniques. In this paper, we proposed an algorithm to detect animals in a given image. WCoHOG is a Histogram oriented gradients based feature vector with better accuracy. It is an extension of Co-occurrence Histograms of Oriented Gradients (CoHOG). In this paper LIBLINEAR classifier is used in order to get better accuracy for high dimensional data. The experiments were conducted on two benchmark datasets called Wild-Anim and CamaraTrap dataset. Experimental results prove that W-CoHOG performs better than existing state of the art methods.

Summary: This journal discusses about the Computer Vision and its application in the field of agricultural safety.

[2] Agriculture is still one of the most crucial sectors of the Indian economy. It is important for human survival as well as economic growth. Traditional systems like humanoid scarecrows are used even today in an agricultural field to stop birds and animals from disturbing and feeding on growing crops. There are many loopholes in such ideas and so enhancing agricultural security has become a major issue these days. Thus, this paper focuses on proposing a system which detects the intruders, monitors any malicious activity and then reports it to the owner of the system. It acts as an adaptable system which provides a practicable system to the farmers for ensuring complete safety of their farmlands from any attacks or trespassing activities.

[3] We present a class of efficient models called MobileNets for mobile and embedded vision applications. MobileNets are based on a streamlined architecture that uses depthwise separable convolutions to build light weight deep neural networks. We introduce two simple global hyperparameters that efficiently tradeoff between latency and accuracy. These hyper-parameters allow the model builder to choose the right sized model for their application based on the constraints of the problem. We present extensive experiments on resource and accuracy tradeoffs and show strong performance compared to other popular models on ImageNet

classification. We then demonstrate the effectiveness of MobileNets across a wide range of applications and use cases including object detection,finegrain classification, face attributes and large scale geo-localization.

[4] Crops are vulnerable to wild animals. Therefore, it is very important to monitor the nearby presence of animals. Then the actuation of various devices should follow to repel the hazardous animals. Traditional methods have been widely applied depending on the kinds of produce and imperiling animals. In this paper, we propose a method to protect farms from wild animals via ubiquitous wired network devices, which is applied to farm along with traditional methods to improve the protection performance. Operational amplifier circuits are utilized mainly for the detection of animal intrusion from the outside of farms. The proposed monitoring scheme is to provide an early warning about possible intrusion and damage by wild animals.

III. EXISTING SYSTEM

Since most of the farms in India are small, most farmers relies on medieval techniques like using a scare crow, or relying on guards to monitor crops. More recently, crops are also being protected using electric fencing but it can be highly cost inefficient which a small farmer cannot afford. Even if they can afford it, in most cases it is illegal to use such fences which governments uses as a measure to conserve the wildlife populations.

Also, in busy seasons like the harvesting time, it can get difficult to have a guard, guarding and monitoring the crops from animals.

IV. PROPOSED SYSTEM

We propose an AI based surveillance system to detect and monitor the presence of any animal. A camera can be placed conveniently at location(s) where any possible animal might enter from. The system uses computer vision using OpenCV to process the feed from the camera. Pre-trained model Mobile Net SSD (Single Shot Detector) is used to detect the animals in the farms. The model is trained on MS COCO image dataset. A siren is fired on detecting an animal which can act as a deterrent to the animal. It can also notify the farmer so that he/she can take the concerned action as required in time.

V. METHODOLOGY

i. Import all the Libraries/packages.
ii. Load the pretrained ImageNet SSD model.
iii. Load MobileNetSSD_deploy.prototxt.txt which defines all the layers in the model.
iv. Start the live video feed.
v. The images are then preprocessed.
vi. We use OpenCV's blobFromImage which performs certain preprocessing to convert it as a 4 dimensional blob.
vii. All the objects (if any) are detected in that blob using the model.
viii. If the probability of those detected exceeds certain threshold (which is 0.2 in our case), only then we will consider that the object is present.
ix. The objects is then bounded in a bounding box.
x. A label along with the probability of success is displayed to the user in the live feed itself.
xi. A siren is played if an animal is detected for a while.

FLOWCHART

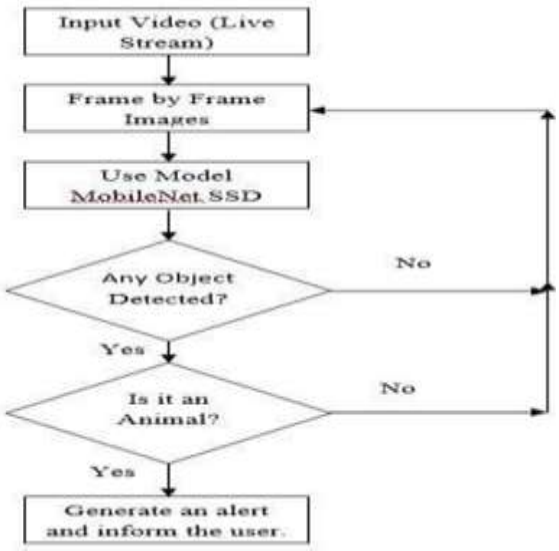

Fig 5.1 Flowchart

VI. MODULES

Upload (Live): Upload a video as a live feed using a webcam (or any camera attached in a farm).

View: Video can be viewed live in a dialog box.

Pre-processing: Data Preprocessing is a technique that is used to convert the raw data into a clean data set. Cleaning the data refers to removing the null values, filling the null values with meaningful value, removing duplicate values, removing outliers, removing unwanted attributes. If dataset contains any categorical records means convert those categorical variables to numerical values. In this case, we are taking a live video feed in the form of images and resizing them to a standard size.

Identifying Features: We use MobileNet SSD pretrained model which identifies features in any image using a Convolution Neural Network (CNN) model.

Model: SSD (Single Shot Detector) is a popular algorithm in object detection. It's generally faster than RCNN. SSD has two components: a backbone model and SSD head. *Backbone* model usually is a pretrained image classification network as a feature extractor. Here, we will use MobileNet SSD model to detect the objects. Here, VGG Net is used as a backbone model to extract the features from the images. Convolution layers (CNN) are then used for object detection in the images using the feature map generated by VGG net layer. The model is able to detect multiple objects in any given image. For the purpose of classification, the model uses softmax in the last layer. Softmax takes in a vector of numbers and converts them to probabilities which are then used for image generating results. Softmax converts logits into probabilities by taking the exponents from every output and then normalize each of these numbers by the sum of such exponents, such that the entire output vector adds up to one.

Prediction: A live video feed is taken in frame by frame as individual images. These images the then fed into the model after preprocessing to detect animals (if any exists).

User Interface: A dialog box opens up while taking in the live video feed. The frames or images from the video are used to detect objects.

The objects are then bounded in a bounding box along with a label and the probability of success in also displayed in there. A siren is then played if any animal is detected for a while.

VII. ALGORITHM

. Convolutional Neural Network

A convolutional neural network(CNN) uses a variation of the multilayer perceptrons. A CNN contains one or more than one convolutional layers. These layers can either be completely interconnected or pooled.Before passing the result to the next layer, the convolutional layer uses a convolutional operation on the input. Due to this

convolutional operation, the network can be much deeper but with much fewer parameters.Due to this ability, convolutional neural networks show very effective results in image and video recognition, natural language processing, and recommender systems.Convolutional neural networks also show great results in semantic parsing and paraphrase detection. They are also applied in signal processing and image classification.

Image Input Layer:

Create an image input layer using image input layer. An image input layer inputs images to a network and applies data normalization. Specify the image size using the input Size argument. The size of an image corresponds to the height, width, and the number of color channels of that image. For example, for a grayscale image, the number of channels is 1, and for a color image it is 3.

Convolution Layer:

Convolutional layers are the major building blocks used in convolutional neural networks.A convolution is the simple application of a filter to an input that results in an activation. Repeated application of the same filter to an input results in a map of activations called a feature map, indicating the locations and strength of a detected feature in an input, such as an image.The innovation of convolutional neural networks is the ability to automatically learn a large number of filters in parallel specific to a training dataset under the constraints of a specific predictive modeling problem, such as image classification. The result is highly specific features that can be detected anywhere on input images.

Pooling Layer:

It is common to periodically insert a Pooling layer in-between successive Convolution layers in a CNN architecture. Its function is to progressively reduce the spatial size of the representation to reduce the amount of parameters and computation in the network, and hence to also control overfitting. The Pooling Layer operates independently on every depth slice of the input and resizes it spatially, using the MAX operation. The most common form is a pooling layer with filters of size 2x2 applied with a stride of 2 downsamples every depth slice in the input by 2 along both width and height, discarding 75% of the activations. Every MAX operation would in this case be taking a max over 4 numbers (little 2x2 region in some depth slice). The depth dimension remains unchanged.

Input Sequence Layer:

A sequence input layer inputs sequence data to a network. In the general case, input sequences and output sequences have different lengths (e.g. machine translation) and the entire input sequence is required in order to start predicting the target. This requires a more advanced setup, which is what people commonly refer to when mentioning "sequence to sequence models" with no further context. Here's how it works:

In the general case, input sequences and output sequences have different lengths (e.g. machine translation) and the entire input sequence is required in order to start predicting the target. This requires a more advanced setup, which is what people commonly refer to when mentioning "sequence to sequence models". Here's how it works:

- A RNN layer (or stack thereof) acts as "encoder": it processes the input sequence and returns its own internal state. Note that we discard the outputs of the encoder RNN, only recovering the state.
- Another RNN layer (or stack thereof) acts as "decoder": it is trained to predict the

next characters of the target sequence, given previous characters of the target sequence.

Output Layers:

Softmax and Classification Layers:

Softmax converts logits into probabilities by taking the exponents from every output and then norms each of these numbers by the sum of such exponents, such that the entire output vector adds up to one – every probability should be one. Generally, cross-entropy loss is the loss of such a problem in several classes. In the last layer of an image classification network such as CNN (e.g. VGG16) used in ImageNet competitions, softmax is also applied.

A softmax layer applies a softmax function to the input. A classification layer computes the cross-entropy loss for multi-class classification problems with mutually exclusive classes. Create a classification layer using classification Layer. For classification problems, a softmax layer and then a classification layer must follow the final fully connected layer. The softmax function is also known as the normalized exponential and can be considered the multi-class generalization of the logistic sigmoid function. For typical classification networks, the classification layer must follow the softmax layer. In the classification layer, train Network takes the values from the Softmax function.

VIII. RESULTS

The proposed scheme presents a novel approach to detect for any intrusion in farms. This approach has been implemented by the using neural networks.
We have successfully developed a deep learning model using the deep neural network architecture to detect the presence of any animal in the farm. We were able to get a net frame rate of 18 frames per second (approximately).

Fig.7.1 Animals Detected

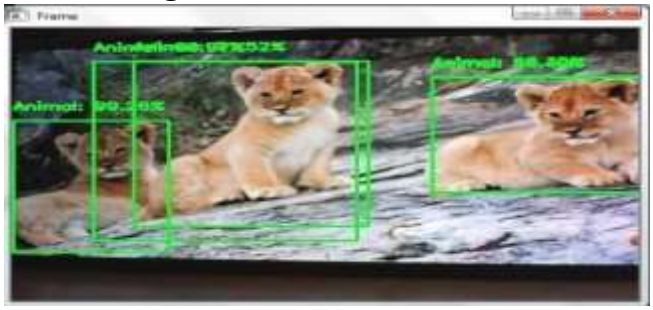

Fig 7.2 Animals Detected

All the objects (if any) are detected in that blob using the model. If the probability of those detected exceeds certain threshold (which is 0.2 in our case), only then we will consider that the object is present. The objects is then bounded in a bounding box. A label along with the probability of success is displayed to the user in the live feed itself. A siren is played if an animal is detected for a while. The live video feed can be closed using the key 'q‘.

Fig 2: Comparison between accuracies of algorithms

DATASET	TYPE	MOBILENETv2-SSD		Tiny-YOLO	
		ACC.(%)	FPR (%)	ACC.(%)	FPR(%)
Tiger	Detection	80.80	1.76	**89.00**	**0.00**
	Re-Id	**99.79**	**0.21**	85.00	0.08
Elephant	Detection/Re-Id	92.56	0.1	**97.00**	**0.09**
		89.47	**4.06**	74.00	0.00
Jaguar	Detection Re-Id	97.05	**3.04**	72.00	0.00

IX. CONCLUSION

The problem of damaging crops by wild animals has become a major social problem in the current time. It requires urgent attention and an effective solution. The proposed method allows us to detect any animal

presence or intrusion in farms using video from any camera device placed in the farms. The object detection model worked almost consistently at 18 frames per second.

It is a cheap and robust system. The siren scares the intruders away as well as it can alert the farmer to take action. Thus, this application can be used to protect crops in the farm. It might be very useful for agricultural purposes instead of traditional methods used today.

REFERENCES

[1] Parikh, M. et al. “Wild-Animal Recognition in Agriculture Farms Using W-COHOG for Agro-Security.” (2017).

[4] Neural Networks for Mobile Vision Applications."

[5] Deshpande, Abhinav. (2016). Design and Implementation of an Intelligent Security System for Farm Protection from Wild Animals. 5. pp.2319-7064.

[6] M. Gogoi and S.R. Philip, “Protection of Crops from Animals using Intelligent Surveillance System,” Journal of Applied and Fundamental Sciences, vol.1, no.2, pp.200- 206, 2015.

[7] S. Pandey and S. B. Bajracharya, “Crop protection and its effectiveness against wildlife: A case study of two villages of shivapuri national park, nepal,” Nepal Journal of Science and Technology, vol. 16, no. 1, pp. 1– 10, 2015.

[2] S. Yadahalli, A. Parmar and A. Deshpande, "Smart Intrusion Detection System for Crop Protection by using Arduino," *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, 2020, pp. 405-408, doi: 10.1109/ICIRCA48905.2020.9182868.

[3] Howard, Andrew & Zhu, Menglong & Chen, Bo & Kalenichenko, Dmitry & Wang, Weijun & Weyand, Tobias & Andreetto, Marco & Adam, Hartwig. (2017)."MobileNets: Efficient Convolutional

[8] V. Bavane, A. Raut, S. Sonune, A. Bawane, and P. Jawandhiya, “Protection of crops from wild animals using intelligent surveillance system.”

[9] R. Bhardwaj, K. Bera, O. Jadhav, P. Gaikwad, and T. Gupta, “Intrusion detection through image processing and getting notified via sms and image,” 2018.

[10] R. M. Antunes and F. L. Grilo, “Intruder alarm systems-the road ahead,” in Advanced Technologies. IntechOpen, 2009.