IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

STOCHASTIC MODEL FOR RECOGNITION OF THE OBJECTS FROM THE REAL-TIME VIDEO STREAM AND DETERMINATION OF ANIMAL IN AGRICULTURAL LAND

[1]Y.Ezhil Dharani,[2]S.Ramalakshmi,[3] Dr.M.Malini

[1]Student,[2] Assistant Professor,[3]Director

[1] Master of Computer Application,

[1]Sri Sarada College For Women, Ariyakulam,Tirunelveli

Abstract: The study carries the stochastic model for recognition of intrusion of animal into human communities in hunt of food, and destroy crops which leads to a major threat in agricultural field covered in Western Ghats for benefits of farmers' income due to loss in crop yield. The input data are collected as real-time video stream of animal roaming in Western Ghats available from the source of net which includes 1265 training dataset. The framework of this study detect the animals when entered into farmland. The proposed method of this study is ResNet50 model used to detect animals. The training data are collected as images and video stream are gathered from forest department. The intrusion of animal was detected using object detection methodology that is Deep Neural Network module in deep learning. Classification of the animals are done by utilizing ResNet50 Architecture of Convolutional Neural Network. ResNet50 is a special architecture of deep learning model which was mainly used for classification of images. The result of the study achieves detection accuracy of 80.47 percent. Thus this study gives high animal detection accuracy, cost-effectiveness, efficient animal surveillance through re-identification.

***Index Terms* – ResNet50 model, Convolutional Neural Network, Deep Neural Network.**

I. INTRODUCTION

This In India, intrusion of animal in farmland is an undesirable but typical event, particularly in villages and towns near forest areas. Reported causes for such conflict include dwindling forest cover and expanding human population, causing animals to be displaced from their natural habitat and forcing them to stray close to human settlements in search of food. Intruding animals usually attack and steal cattle or raid farms, however there are times when humans are caught in the crossfire. A separate, nation-wide study on human animal conflict in vulnerable areas conducted by the Centre for Wildlife Studies, Bengaluru found that 71% of the households surveyed had suffered crop loss, and 17% had suffered livestock loss. A recent report published by the Centre for Wildlife Studies estimates that there are approximately 80,000 recorded human-wildlife conflicts every year in India, with elephant, tiger, and leopard and rhesus macaque being the top conflict prone species. The problem is acute as well as widespread, and an effective solution can reduce difficulties of the affected place. Food needs to contain a high nutritional value and its security must be guaranteed around the world. Crop damage caused by animal invasion is one of the leading causes of crop yield reduction. The farm areas near the forest boundaries are apparently affected by the wild animal attacks. Agriculture plays a crucial role in the nation's progress. Organizational issues have long been a source of concern for the country's development. Farmers face a slew of problems, including a lack of water for irrigation, crop withering due to climate changes, nutrient-depleted soils, and pest and wildlife damage to crops. Untamed animals walk over harvests and consume them, lowering yield. In such scenarios we require smart vision which could efficiently predict and notify the farmer at earliest. Further a visual system is used to solve this problem by predicting the animal before it enters the field. This whole new part of the visual system comes under the concept of Deep learning in artificial intelligence. By automatically detecting animal invasions, this study presents a method that leverages deep learning and computer vision to prevent animal-human conflicts. The main aim of the study is to prevent crops from destroying by untamed animals by accurate animal detection and identification of the type of animal, the system's ability to support massive implementations and a fail-safe user alert mode.

3.1 Existing Model

Saieshwar Radhakrishnan (2018) depicted an image processing and machine learning based approach to classify the animal as threat and hence alert the farmer. The sample data consists of Nilgai and dog. The image is segmented into parts using Watershed algorithm. The features are extracted from the training set by using 2D Gabor filter bank. Classification is done using Support Vector Machines algorithm. The result of this study gives the overall accuracy of 54.32%. Vikas Bavane (2018) said that farmlands or agricultural lands surveillance is very important to protect the area from animals. This problem is so pronounced that sometimes the farmers decide to leave the areas barren due to such frequent animal attacks. This system helps us to keep away such wild animals from the farmlands as well as provides surveillance functionality. Cheng et al. (2017) developed the system uses discriminative features for classifying the bird species based on parts of birds that uses a support vector machine along with Normal Bayes classifier. Fine-Grained Image Categorization – Technique for discriminating fine-grained classes which can be divided into two very important main groups useful for future work as well. Mohamed Elsayed Abd Elaziz (2019) presented a new multi-objective metaheuristic based on a multiverse optimization algorithm on segment grayscale images via multi-level thresholding. The result showed that the presented method provides better approximation to the optimal than the other algorithms in terms of hypervolume and spacing plus the quality of the segmented image is better than those of the other methods in terms of uniformity measures. Ashwini L. Kadam (2021) described the system that attempts to address the wild animal intrusions patterns analysis in the farm field. The data is collected from the farm field such as motion detection data, time, temperature, humidity, and GPS data. The intrusion patterns were analyzed in kinds of graphs date versus time and temperature versus humidity to make it easy to understand the farmer, when intrusion is occurring. Yamanaka et al (2018) made a tensor flow implementation of last and accurate image super resolution by CNN with skip connection and network in network a deep learning based single image super resolution model

3.2 Objectives

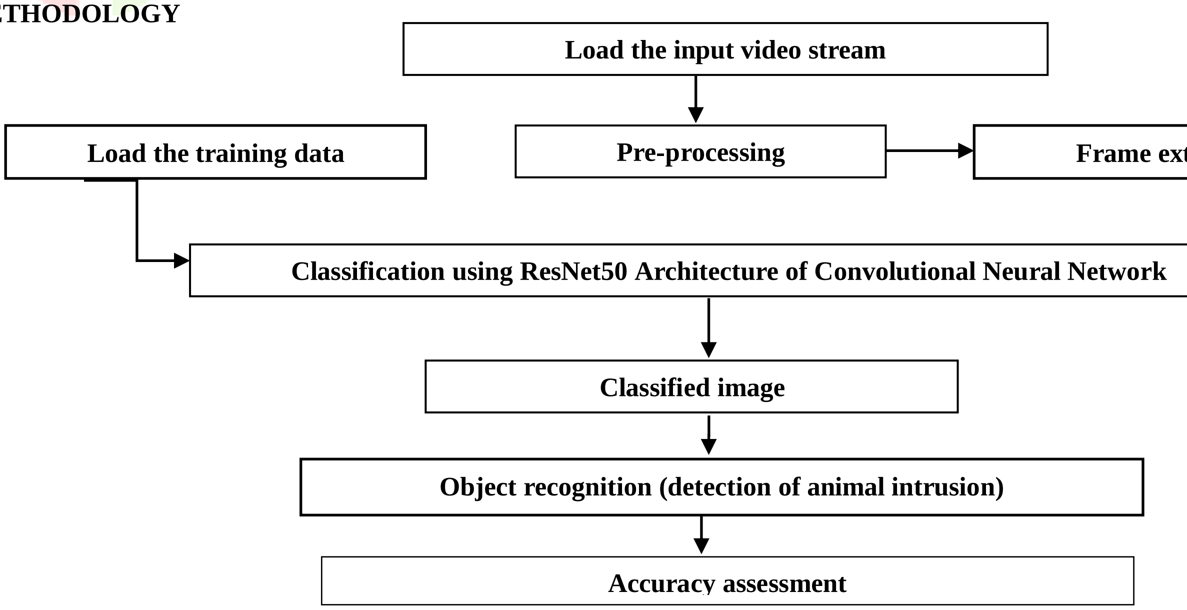
- To develop model for determination of intrusion of animal which destroy farmland in Western Ghats.
- To recognize the objects from real-time video stream.
- To identify the animal using feature extraction.
- To classify the identified animals using ResNet50 Architecture of Convolutional Neural Network

The accuracy assessed

3.3 Data for Study

The input data are collected as real-time video stream of animal roaming in Western Ghats available from the source of net which includes 1265images. In this framework, video stream is captured when there is an intrusion of animal in farmland and then image was classified as person or animal using Convolution Neural Network (CNN) and deep learning technique. The smart farm protection system gives reliable security and safety to crops. This study was being selected for assuring the wellbeing of creatures while warding them off and diminishes the exertion made by man in securing the field. Classification of the intrusion of animal was done by utilizing Convolutional Neural Network. CNN a special architecture of artificial neural networks model was mainly used for classification of images. Convolutional neural Network takes image as an input. Accordingly, a filter is applied to the image in ConvNet. This filter consists of weights when multiplied with pixel values in image to reduce the features, the clarity of original image is reduced. The data regarding these intrusions of animals were sent to the cloud by means of the web. This way it is easy to arrive at useful information regarding the intrusions and take measures against it. The study provides a solution without hurting creatures or setting human life at stake. Hence, this approach is helpful to the farmers in protecting fields, saving them from financial losses.

I. RESEARCH METHODOLOGY



3.1 General

The aim is to develop a detector that analyses the video images from camera traps in real-time. Classes to identify are zebra, humans and wild horse, wild sheep and elephant. The scope of this study also includes to extract events of importance.

3.2 Pre-Processing

Image pre-processing are the steps taken to format images before used by model training and inference. This includes, but is not limited to, resizing, orienting, and color corrections. Image augmentation manipulations are forms of image pre-processing. The steps are applied to training and test sets, image augmentation is only applied to the training data. Thus, a transformation that could be an augmentation in some situations may best be a pre-processing step in others.Pre-processing is required to clean image data for model input. Image pre-processing may also decrease model training time and increase model inference speed. If input images are particularly large, reducing the size of these images will dramatically improve model training time without significantly reducing model performance. The steps involved in Pre-processing are resize which is used in changing the size of an image, by preserving scale is not always required, filling in dead pixels with reflected image content is often best, and down sampling large images to smaller images is often safest, orientation: is used when an image is captured, it contains metadata that tells our machines the orientation by which to display that input image relative to how it is stored on disk, metadata from images and ensure pixels are all ordered the same way and grayscale is used when color changes are an example of image transformations that may be applied to all images (train and test) or randomly altered in training only as augmentations.

3.3 Convolutional Neural Network

In deep learning, a convolutional neural network (CNN/ConvNet) is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other.Convolutional neural networks are composed of multiple layers of artificial neurons. Artificial neurons, a rough imitation of their biological counterparts, are mathematical functions that calculate the weighted sum of multiple inputs and outputs an activation value. When you input an image in a ConvNet, each layer generates several activation functions that are passed on to the next layer.

Convolutional: Convolutional layers consist of a rectangular grid of neurons. It requires that the previous layer also be a rectangular grid of neurons. Each neuron takes inputs from a rectangular section of the previous layer; the weights for this rectangular section are the same for each neuron in the convolutional layer. Thus, the convolutional layer is just an image convolution of the previous layer, where the weights specify the convolution filter.

Max-Pooling: After each convolutional layer, there may be a pooling layer. The pooling layer takes small rectangular blocks from the convolutional layer and subsamples it to produce a single output from that block. There are several ways to do this pooling, such as taking the average or the maximum, or a learned linear combination of the neurons in the block. Our pooling layers will always be max-pooling layers; that is, they take the maximum of the block they are pooling.

Fully-Connected: Finally, after several convolutional and max pooling layers, the high-level reasoning in the neural network is done via fully connected layers. A fully connected layer takes all neurons in the previous layer (be it fully connected, pooling, or convolutional) and connects it to every single neuron it has. Fully connected layers are not spatially located anymore (you can visualize them as one-dimensional), so there can be no convolutional layers after a fully connected layer.

$z=x*f$

X represent input image ‘f’ represent the filter

$$W_{out} = \frac{w - f + 2p}{s}$$

$$f(x) = sin(2\sigma sigma(w_2 \sigma relu(w_1 x + b_1) + b_2) + b_3) - 1$$

Pooling Layer

Pooling layers, also known as down sampling, conducts dimensionality reduction, reducing the number of parameters in the input. Similar to the convolutional layer, the pooling operation sweeps a filter across the entire input, but the difference is that this filter does not have any weights. Instead, the kernel applies an aggregation function to the values within the receptive field, populating the output array. There are two main types of pooling:

Max pooling: As the filter moves across the input, it selects the pixel with the maximum value to send to the output array. As an aside, this approach tends to be used more often compared to average pooling.

Average pooling: As the filter moves across the input, it calculates the average value within the receptive field to send to the output array.

3.4 Feature Extraction

Extracting features from the output of video segmentation. Feature extraction is the time consuming task in CBVR. This can be overcome by using the multi core architecture. These mainly include features of key frames, objects, motions and audio/text features.

Features of Key Frames are classified as color based, texture based and shape based features. Color-based features include color histograms, color moments, color correlograms, a mixture of Gaussian models, etc. split the image into 5×5 blocks to capture local color information.Texture-based features are object surface-owned intrinsic visual features that are independent of color or intensity and reflect homogenous phenomena in images. Features based on colour: Colour histograms, a mixture of Gaussian models, colour moments, colour coral grams etc. are in the features of colour based. Colour based feature extraction are dependent on the spaces of colour for example, the HSV, RGB, YCBCR, HVC and normalized r-g and YUV.

3.5 Object Recognition

One of the applications of the deep learning technique called Convolutional Neural Network used for object recognition. Observing wild animals in their natural environments is a central task in ecology. The fast growth of human population and the endless pursuit of economic development are making over-exploitation of natural resources, causing rapid, novel and substantial changes to Earth's ecosystems. An increasing area of land surface has been transformed by human action, altering wildlife population, habitat and behaviour. More seriously, many wild species on Earth have been driven to extinction, and many species are introduced into new areas where they can disrupt both natural and human systems.

3.6 Image classification

In this study Resnet50 layers, which is one of CNN architectures was implemented. The flow of CNN are Convolution Layer (extract feature with filtering), Strides (shifting pixels over the input matrix), Padding, Rectified Linear Unit (Relu) (introduce non-linearity to the network), Padding Layer (reduce number of parameters), Fully Connected Layer (flatten matrix into vector and feed it to a fully connected neural network layer.)

There are few architecture of CNN (some of the most common is ResNet, VGGNet), in this study ResNet50 architecture was implemented.

Image Classification is a method to extract information of image classes digitally. Shortly, it used to classify images. Image Classification can be supervised (you give label to each images) and also unsupervised (the training model learned to classify the images based on the patterns).Common classification procedures can be broken down into two broad subdivisions based on the method used: supervised classification and unsupervised classification. In a supervised classification, the analyst identifies in the imagery homogeneous representative samples of the different surface cover types (information classes) of interest. These samples are referred to as training areas. The selection of appropriate training areas is based on the analyst's familiarity with the geographical area and their knowledge of the actual surface cover types present in the image. Thus, the analyst is "supervising" the categorization of a set of specific classes. The numerical information in all spectral bands for the pixels comprising these areas are used to "train" the computer to recognize spectrally similar areas for each class. The computer uses a special program or algorithm (of which there are several variations), to determine the numerical "signatures" for each training **class.**

3.6.1 ResNet50 Architecture of CNN

ResNet or Residual Network uses the residual learning instead of trying to learn some features. Residual can be simply understood as subtraction of feature learned from input of that layer. There are some variants of ResNet other than ResNet50. The core idea of ResNet is introducing a shortcut connection that skips one or more layers.

ResNet-50 is a convolutional neural network that is 50 layers deep. Load a pretrained version of the network trained on more than a million images from the ImageNet database. The pretrained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals.

ResNet50 is a variant of ResNet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. It has 3.8 x 10^9 Floating points operations. It is a widely used ResNet model and we have explored ResNet50 architecture in depth.

Keras is a deep learning API that is popular due to the simplicity of building models using it as shown in figure 3.1. Keras comes with several pre-trained models, including ResNet50 that can be used for experiments.

Therefore, building a residual network in Keras for computer vision tasks like image classification is relatively simple.

IV. RESULTS AND DISCUSSION

4.1 General

In this study, the dataset used is downloaded from the source of net and some parts of it was collected manually as well. The dataset consists of total 1595 images of five different animals, i.e. sheep, goat, horse, elephant, donkey. This dataset contains two folders test set and trainset. Out of which 80% of images for training and 20% of images for testing were used. The training set is the material through which the computer learns how to process information. Training data set is used for learning and to fit the parameters of the classifier. Test set is a set of unseen data used only to assess the performance of a fully specified classifier. The study includes 1265 images for training and 330 images for testing

4.2 Results and Discussion

Figure 4.1 The horse object is detected

The figure 4.1.shows the horse animal detection in the grass field video. That video is download from YouTube. This output shows that more than one animal can be detected using CNN ResNet50 architecture. The mother horse got 98.15% accuracy and child horse got 97.22% accuracy.

Figure 4.2 The sheep obect is detected

The above figure 4.2 shows the sheep animal detection in the grass field video. That video is download from shutterstock. This output shows that more than one animal can be detected using CNN ResNet50 architecture. The sheep got 98.17% accuracy.

Figure 4.3 The horse object is detected

The figure 4.3 shows the back pose of the horse animal detection in the farmland. That video is download from shutterstock website. This output shows that horse animal can be detected using CNN ResNet50 architecture. The back pose of the horse got 98.62 % accuracy.

II. Acknowledgment

References

[1] X. Wu, D. Sahoo, and S. C. H. Hoi. Recent advances in deep learning for object detection, 2019.

[2] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi. A survey of the recent architectures of deep convolutional neural networks. Artificial Intelligence Review, Apr 2020.

[3] I. Goodfellow, Y. Bengio, and A. Courville. Deep Learning. MIT Press, 2016. http://www.deeplearningbook.org.

[4] A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. Neural Information Processing Systems. 25. doi: 10.1145/3065386, January 2012.

[5] R. Girshick. Fast r-CNN. In Proceeding 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, 2015, pp. 1440-1448, doi: 10.1109/ICCV.2015.169.

[6] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. International Conference for Learning Representations, San Diego, 2015, 2014.

[7] L. Torrey and J. Shavlik. Transfer learning. In E. Soria, J. Martin, R. Magdalena, M. Martinez A. Serrano, editor, Handbook of Research on Machine Learning Applications. IGI Global. (chapter11). doi: 10.4018/978-1-60566- 766-9.ch011, 2015.

[8] S. Schneider, G. W. Taylor, and S. C. Kremer. Deep learning object detection methods for ecological camera trap data. pp. 321-328. doi: 10.1109/CRV.2018.00052, 2018.

[9] G.S. Cheema and S. Anand. Automatic detection and recognition of individuals in patterned species. In: Altun Y. et al. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2017. Lecture Notes in Computer Science, vol 10536. Springer, Cham, pages 27–38, 2017.

[10] M.S. Norouzzadeh, N. Joshi N. Jojic D. Morris, S. Beery, and J. Clune. A deep active learning system for species identification and counting in camera trap 45 46 Bibliography images. Proceedings of the National Academy of Sciences Jun 2018, 115 (25) E5716-E5725; doi: 10.1073/pnas.1719367115, 2019.

[11] B. Settles. Active learning literature survey. University of Wisconsin, Madison, 2009.

[12] M. Wilber et al. Animal recognition in the Mojave Desert: Vision tools for field biologists. In: 2013 IEEE Workshop on Applications of Computer Vision (WACV), pp. 206-213, pages 206–213, 2013.

[13] N.M. Badrinath, N. Bharat, S. Madhumathi, and B.S. Kariyappa. Pattern recognition using video surveillance for wildlife applications. International Journal of Electronics and Communication Engineering Research and Development (IJECERD), Volume 4, Number 2, April-June (2014), 2014.
[14] C. Geng, S.-J. Huang, and S. Chen. Recent advances in open set recognition: A survey. Page 1–1, 2020. In Proceeding of IEEE Transactions on Pattern Analysis and Machine Intelligence.
[15] M. Hassen and P.K. Chan. Learning a neural-network-based representation for open set recognition, 2018. In book: Proceedings of the 2020 SIAM International Conference on Data Mining, pp.154-162.