ABSTRACT

As we all know that wildlife ecosystem is in a critical situation, like various species in ecosystem have increased in the list of endangered species. So, not only preserving the species but also, we need to detect the location of those animals in order to protect and prevent them from poaching and it becomes easier to provide them the certain medication and also priorly alerting the nearby locality in case of any animals around the certain village. So, we as a human beings have a responsibility to save the animal species as well as to maintain the balanced wildlife ecosystem.

The conservation of biodiversity depends on precise accurate time to time information. All this data about the forest ecosystem is obtained through motion activated cameras which are placed in certain areas of the forest which covers almost all animals. These cameras are also known as camera traps because it clicks the image of the animals at every instant periodically. These camera traps or motion activated cameras play an important role in fetching the data of forest biodiversity. The reason for using this motion activated cameras because its cheap, an effective method to fetch the data and non-intrusive. The data that is generated from camera traps are collected through the computers which are poised to dramatically increase the efficiency of the image based on the data collected through survey.

TABLE OF CONTENTS

Chapter 1:

INTRODUCTION

1.1 Need of Python in ML

Python is the most certain language for implementing machine learning because it contains large number of libraries. Considering machine learning topics are mostly on mathematical optimization, probability and statistics, Python consists of libraries that help data scientists perform various operations to process the data easily. Python is the best choice for machine learning because it is easy to understand which helps quickly to validate the data without any errors. Also, Python is a simple and easy language to preprocess the millions of data due to availability of large ecosystems. With the help of large library ecosystem developers can solve complex problems or tasks without performing large coding. Because of the versatility of python testing can be done easily. Those tests can be performed in any operating systems like Windows, Linux etc.

1.2 Introduction to Machine Learning

Machine learning (ML) is the ability to study the computer algorithms which tends to improve automatically through experience. It is a subset of Artificial Intelligence. Building of model based on sample data using Machine learning algorithms is called as “training data”, in order to predict or make decisions without the use of direct programming to perform the tasks and to test the model we take a small sample of data called as “testing data”. There are three types of Machine Learning Techniques: Supervised Technique where the outcome of the certain task is known and the model is trained in a similar way, Unsupervised Technique where the outcome is unknown and this technique involves clustering, Reinforcement Technique where the model is built in such a way that the model learns and gets experience from the environment. The algorithms of Machine learning are widely used in applications like image processing, email filtering, computer vision etc. In our project we use Machine Learning Algorithms to detect animal species in an image and classify them according to their kinds.

1.3 Introduction of Machine Learning Libraries

- OpenCV: OpenCV is an optimized library which mainly focuses on real-time applications. It is a multi-Platforms for C++, Python and Java interfaces and supports operating systems like Linux, MacOS, Windows, iOS, etc. using OpenCV we can develop real-time computer vision applications like face detection and object detection.
- TensorFlow: TensorFlow is a Python library which is an end-to-end open-source platform for machine learning algorithms. It has a vast community resource, huge libraries, comprehensive, flexible ecosystem of tools that lets developers or researchers to easily build models or deploy machine learning powered applications and to push the state-of-the-art in ML.
- NumPy: NumPy belongs to the python library and it stands for Numerical Python. It is a library which consists of multidimensional array objects and a collection of computing operations for processing those arrays. With the use of NumPy we can perform mathematical and logical operations.
- PIL: PIL is the Python Imaging Library which helps in providing the python interpreter with image processing and editing capabilities.

1.4 Project Introduction

To better understand the complexities of the nature and animal ecosystems and to protect them, a detailed knowledge of animal behavior, their location and the total number of them is helpful to protect and track them. The wildlife studies as mentioned in the above statement need tracking as well as location devices to record them. To record them we make use of motion activated cameras which are also known as camera traps since they are cheap and compare to other detection devices the data can fetched easily through this device in an efficient way. This motion activated cameras have revolutionized wildlife ecology study and conservation study over the last few years. These cameras have also become an important essential tool for ecologists which enables them to study about the animals or nature like distribution of species over wildlife ecosystem. Although this motion activated camera collect the images in a large volume and turn those images into useful information. The information from the images are extracted through computer vision which is done by humans manually. Humans review all the images manually and label each image this burden of reviewing each and every image is time consuming as well it is much costlier because there is a chance of missing a valuable knowledge from the image. This time consuming is the major disadvantage of this manual reviewing of data. To eradicate this major problem, we use Deep Learning Techniques to ease the process of reviewing the large-scale data.

Chapter 2:

LITERATURE SURVEY

[1] In the paper titled "**Animal Detection Using Template Matching Algorithm**" published in the year 2020 by the "I K Amri and Murinto". Animal detection is useful to prevent animal- vehicle accidents and using this will increase human and wildlife safety, it will detect the animals entering the road and warn the vehicles travelling in that direction through audio, video signals. After survey (their survey) the most appropriate method is selected for animal detection is by using Template Matching Algorithm. Template matching is a method which identifies small parts of an image and relates that to the template image. Normalized cross correlation is introduced in this paper to perform template matching. In signal processing, cross-correlation is used in this paper to measure the similarities between two waveforms as a component of time-slack applied to one of the waveforms. Most commonly, template matching is used for searching a long duration signal to identify the particular feature. Here they have observed that the template matching method worked well for animal detection. Hence, the false positive rate for animal detection is 13.3 %.The proposed technique has shown 84.18% efficiency for animal detection.

[2] In the paper titled "**Identifying, Counting, and Describing Wild Animals in Camera Trap Images with Deep Learning**" published in the year 2018 by the "Mohammad Sadegh Norouz Zadeh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S. Palmer, Craig Packer, and Jeff Clune PNAS" Here the location and activities of animals in the wild is known prior using deep learning. This paper involves for collecting the movements of wildlife by gathering camera trap image data accurately along with the ability of involuntarily and also a motion sensor is used. Here they extract the data or pre- process the data using Deep Learning Technique. A Deep convolutional neural network is used in such a way that it is trained to count, recognize, and demonstrate the behaviors of different varieties of breed from Serengeti dataset of 3.2-million image Snapshot. The accuracy calculated by Deep Learning Algorithm is the effectiveness in highlighting the significance of using deep neural networks to automate the extraction of data from camera trap images. In this paper they used two pipeline stages to improve their model: first they pre-process the empty data from the dataset and in second pipeline extracting the dataset that contains animals without empty label. They used the multi-task learning technique to train the model and tested the ability of state-of-the-art computer vision methods called DNNs to automatically extract information from images in the SS dataset, the largest existing labelled dataset of wild animals. They first demonstrated that DNNs algorithm can perform compared from the given SS dataset, although performance is worse for rare classes. The system will classify only self-confident images, and also it automates animal were labelled empty by identification

accuracy rate to 99.3% of given data. But even though performance remains likewise 96.6% accurate of being efficient, saving more than 8.4 years of human labelling effort on the given 3.2-million-image dataset.

[3] In the paper titled "**Context R-CNN: Long Term Temporal Context for Per-Camera Object Detection**" published in the year 2020 by the author "Sara Beery, Guanhang Wu, Vivek Rathod, Ronny Votel, Jonathan Huang California Institute of Technology" In this paper they use the static monitoring cameras to fetch the useful information. They have adopted an attention-based approach which allows the Context R-CNN model to index into a long-term memory bank which can be constructed on a per-camera basis and also includes object detection performance boosting the contextual features. They apply Context R-CNN to two methods: one is vehicle detection in traffic cameras and species detection using camera traps. They build Context R-CNN on top of single frame detection models. Additionally, building on intuitions that moving objects exhibit periodic behaviour and tend to appear in similar locations, they hope to inform their research predictions by conditioning on instance level features from contextual frames. Because of this lastest requirement, they choose the Faster R-CNN architecture as the requirement where it is competitive highly as their base detection model and provides suitable choices for extracting new level features. Their method is easily applicable to any two- stage detection framework. An image returns a collection of class agnostic bounding box proposals when it is first passed through a Initial stage region proposal network (RPN) after which running non-max suppression, These box proposals are then passed into the second stage, which extracts instance-level features via the ROI Align operation which then undergo classification and box refinement. In this process, they will contribute a implement model that shows temporal context of a camera, and it also shows that static camera setting is important, beneficial is attention-based temporal context. Context R-CNN is effective and capable of passive-monitoring sampling strategies which provides data streams with low and irregular frame rates.

[4] In the paper titled "**Deep learning improves acoustic biodiversity monitoring and new candidate forest frog species identification (genus *Platymantis*) in the Philippines"** published in the year 2021 by the "Ali Khalighifar, Rafe M. Brown, Johana Goyes Vallejos**"** In this Paper Monitoring devices are used for detection of certain specific-species which show a good result in overcoming the challenges that are associated for discovering the species rates of lagging frog. These devices generate data at paces faster than it can be analysed. They have a TensorFlow Inception v3 method for designing a robust, which is a automated species identification system for frog species of Philippine, using a single-note spectrograms audio. They explored two concepts: (1) performance of our deep learning model in discriminating closely-related frog species based on images representing advertisement call notes, and

(2) the potential of this platform to accelerate new species discovery. The methods used for implementing is CNN method. Convolutional neural networks (CNNs) are used for pattern recognition and image classification tasks the CNNs are also a subset of DNNs. CNNs is the ability to perform eliminating the need for hand-crafted feature extraction along with automated feature extraction. CNN architecture is built on three types of layers: Convolutional layers, pooling layers and fully connected layers. They use a transfer learning technique. Transfer learning is method where classification task A acquires experience from classification task B. One of the most successfully implemented models of transfer learning is Inception v3 —a CNN, implemented in Tensor Flow. This CNN consists of 48 layers, and is trained for more than 1 million images from the ImageNet Inception v3 method is widely recognized for getting greater results compared to other models in challenges involving of classifying images into different varieties of classes. TensorFlow identified species results of 94% accuracy for overall correct identification rate. In this paper they calibrated models for classification challenges using different numbers of training steps, and considered two factors to find optima for training TensorFlow correct identification rate and processing time.

[5] In the paper titled "**Deep Convolutional Neural Network Based Species Recognition for Wild Animal Monitoring**" published in the year 2018 by the author "Guobin Chen, Tony X. Han, Zhihai He Roland Kays, and Tavis Forrester" In this Paper they are discussing about our society Which is less concerned about the well-being of wildlife species. Due to this camera-trap technologies are used for wildlife monitoring where they are available at a reasonable cost, rapidly deployable easy to maintain. The data images that are captured by the camera traps are used first; the method used for sequence to segment out the moving foreground EVOC method is applied to the image. They have implemented two image classification algorithms:(1) Species Recognition Baseline using BOW based Image Classification. The BOW image classification algorithm is simple and quite robust to deformation this follow the famous bag-of-words model to do the classification. Then made use of the histogram to represent the image, and use linear SVM as the classifier. We train a code book using k-means clustering and for training image we have sampled 1000000 features from the images. (2) Deep Convolutional Neural Network (DCNN) based image classification algorithm for classification of wild animals on these very challenging camera-trap imagery data and to extract features to represent the block and designed a DCNN with 3 convolutional layers and 3 max pooling layers. The convolutional layer has a convolutional kernel with a size of 9 × 9, while pooling layer has a kernel with a size of 2 × 2. The DCNN based image classification algorithm can achieve superior performance over most of the state-of-the-art image classification algorithms with on the collected camera-trap dataset, we compared the BOW model with our DCNN algorithm for species recognition. The performance comparison of DCNN is good and has greater Accuracy compared to BOW mode. we can predict that

the recognition algorithm DCNN based species can improve quickly as it can finally achieve the goal of automatic species recognition for camera-trap data and large learning capacity.

[6] In the paper titled "**Wildlife Spotter Project"** published in the year 2017 by the "Australian Organizations and Universities**"** Australia is a vast country and has several varieties of wildlife species. Wildlife Spotter is an science project undertaken by many organizations and universities in Australia, taking help from citizens as volunteers to classify species from millions of images captured by camera traps. These cameras are located in the nationwide: tropical rainforests, dry rangelands, and around the cities, set up to automatically snap colour, high-definition images day and night. Till now more three million images were completed. To classify and process the enormous volume of images they invite them for joining for image analysing, the project invites volunteers to citizens as a scientist to classify the image. The main goal of the project is, through analysing captured images, to assist researchers study Australian wildlife populations, behaviours and habitats to save threatened species and preserve balanced, diverse, and sustainable ecosystems. The Wildlife Spotter project is categorized and divided into six sub-projects, specifying on Australia based on separated natural areas are: New South Wales that has Central mallee lands, South-central Victoria Tasmanian nature reserves, far north Queensland, coastal forests in New South Wales, Northern Territory arid zone. Volunteers participate the project by registering online accounts, logging in the Web-based image classification system and manually labelling the displayed images, one by one. The introduced image to a specific species will be assigned by the user by clicking the appropriate category from a given list of animals. In case of defective problem, by blank image or image problem, then the user labels image as no animal view. For attaining reliable classification accuracy, each image in the dataset is repeatedly processed and checked by a different user to label. For example, most classified images in the Victoria dataset each was done by five citizens. The datasets that are collected from the camera traps are in greater volume and in less quality, which basically takes more processing time and can lead to inconsistent labelling. In this work, they aim at building a practical, fully automatic animal recognition framework for Wildlife Spotter project, freeing scientists from the burden of manual labelling, while dramatically reducing processing time.

[7] In the paper titled "**Orientation Robust Object Detection In Aerial Images Using Deep Convolutional Neural Network"** published in the year 2019 by the author "QQ Liu, JB Li - Procedia Computer Science**".** Object detection in aerial images is challenging due to various factors like object colors, aspect ratios, cluttered backgrounds etc. Here they use Deep Convolutional Neural Network (DCNN) features from combined layers to perform orientation robust aerial object detection. The object detection plays a vital role in image identification to know where to place the particular image. Recent

studies aim to find features that are invariant to specific transformations. Image segmentation localizes object candidates, and then they are classified with an SVM classifier trained on the orientation invariant features. The feature selection Procedure is based on the recent advance of disentangling learning, which shows that it is proper to use separate groups of features to model distinct factors. For orientation robust DCNN feature extraction, they utilize AlexNet architecture. By the method of using selective search, it effectively reduces the region proposals about 60% without performance loss. According to the PASCAL VOC object detection evaluation protocol, a detected bounding box and a ground truth is recognized as matched if their overlap is larger than 50%. The results show that the DCNN features from combined layers are efficient when performing or robust aerial object detection. It also shows that the t-SNE analysis and visualization can be used to find proper DCNN layers.

[8] In the paper titled "**Animal species classification using machine learning techniques"** published in the year 2019 by the author "Fahad Alharbi, Abrar Alharbi, and EijiKamioka**".** Machine learning is widely used in animal detection using camera traps, especially to find animal predators that are of great threat to the humans. The machine learning techniques used in this paper are like Support Vector Machine(SVM) and Multilevel perception(MLP) to classify them as predators or pet based on the database that contain features like ears, eyes. The major problem in this research is image classification and SVM is the best linear methods in image classification and pattern recognition because it has the capability to get good generalization with the limited numbers of samples. First the dataset was created containing images of 10 animals (5 predators, 5 pet). Secondly the region of interest was drawn manually. Then statistical information was collected by methods like Mean, Standard Deviation and perimeter. dataset was divided into 2 sets one was training set which had 150 images and the other was testing dataset that had 50 image information in which 25 of them were predators and the rest were prey. The SVM accuracy rate for both the classes is 78% whereas of MLP is 82%. This proves that the neural network have a better ability to learn and work for non-linear data. Finally, this paper concentrates on the unique characteristics of the predator which help the nearby community to be aware of the presence of the predator as well as help farmers to protect livestock from animal attacks.

[9] In the paper titled "**Deep Residual Learning for Image Recognition"** published in the year May 2020 by the author "Ruhua Wang, Chencho, Senjian An, Jun Li, Ling Li, Hang Hau**".** This paper addresses the problem of degradation using deep residual learning framework. This paper shows Deep residual nets are easy to optimize than 'plain nets' and Deep residual networks can easily obtain accuracy gain from greatly increased depth providing better results compared to old works. There are two kinds of networks plain and residual networks. In Residual network, shortcut connections are used which in turn the network into residual version which can be used when both input and output have the

same dimensions. Their work uses batch normalization immediately after each convolution and before activation. Stochastic Gradient Descent (SGD) which is a simple and effective approach is used for fitting linear classifiers and regressors under loss functions. This method on the ImageNet test set achieves 3.57% error. The paper also go through various other techniques like Identify and projection shortcuts, Deeper Bottleneck Architecture, state-of-art methods object detection on PASCAL and MS COCO. This work proves deep residual learning is efficient and more reliable for image recognition and helps to find the content in the image as it uses residual networks.

[10] In the paper titled "**Edge Machine Learning for animal detection, classification and Tracking"** published in the year 2020 by the author "Amanda Tydén and Sara Olsson**".** Prevention and maintaining the existing bio diversity is the most crucial part as we hear every time that a particular animal species are added into the endangered species list. The use of machine learning and camera traps will have great advantage as the system require minimum human works and low latency and also generate alarms in case of any danger. Transfer training is used to utilize pre-trained data and thereby reducing training time and amount of training data. Model SSD mobileNet V2 is employed which has the ability to detect more images per second makes it more accurate in object tracking. The concept of deep machine learning like CNN (convolutional Neural Network) is used for better object detection as well as tracking. The input given to the CNN model is tensor, which is a generalization of matrix in higher dimensions. Activation function is the output value which is calculated by each neuron by applying a non-linear function. This function represents a neural output as function of its input. Sigmoid and Softmax are common classification functions used for loss computation. Other methods like Localization loss, Regularization loss are used to increase the ability of the network to locate the object in image. The training data typically contains classes of labels that contain how to handle all other objects that may appear, this is referred to as open set recognition. In tracking, objects can be assigned with unique ID and extra information can be retrieved such the time duration and location and a tracking model generally consist of detector and a tracker. Finally, this paper provides the research based on the efficiency of the deep machine learning techniques and the models they have used as well the datasets they have employed. It also helps to understand different frameworks like edge machine learning which leads to maximum output of the image captured.

[11] In the paper titled "**Quantum-soft QUBO Suppression for Accurate Object Detection"** published in the year 2020 by the author "Junde Li and Swaroop Ghosh**".** From past decades non-maximum suppression has been adopted for object detection. In this paper they first remove redundant detections into Quadratic Unconstrained Binary Optimization (QUBO) framework that consists of detection score from each bounding box and overlap ratio between pair of bounding boxes. Redundant

object detections are removed by Quadratic Unconstrained Binary Optimization (QUBO) framework. Tabu Search like classical algorithms reduces time complexity from exponential to polynomial at the cost of reduced accuracy and shows the accuracy rate of 87.87%. Quantum-soft QUBO Suppression (QSQS) algorithm for fast and accurate detection by utilizing quantum computing advantages is used to solve QUBO problem. The results proves The results show the proposed QSQS method improves mAP from 74.20% to 75.11% for PASCAL VOC 2007 through two level enhancements.

[12] In the paper titled "**Fauna Image Classification using Convolutional Neural Network**" published in the year 2020 by the author "Kavish Sanghvi and Ishani Saha". Convolutional Neural Network is the best choice for image recognition, processing and classification. This paper proposes a fauna image classifier using convolutional neural network, which classify images of different species and animals captured in forest environments to obtain desired accuracy, and help ecologists and researchers in neural network, artificial intelligence. This paper uses Dn-CNN denoiser which uses the signal sent by receiver for filtering. Dn-CNN denoiser can handle blind Gaussian noise with unknown noise levels too. To obtain the best results for object identification and training of the convolutional neural network, it is necessary to provide input image with enhanced features as training sample. The aim of the training algorithm is to train a neural network such that the error is reduced between the network output and the desired output. Their model has successfully trained with the accuracy rate of 91.84%, and classified images with the accuracy rate of 99.77. Finally, the proposed fauna image classification using convolutional neural network can be used extensively for fauna image classification.

[13] In the paper titled "**Validating the practicality of utilising an image classifier developed using TensorFlow framework in collecting corrugation data from gravel roads"** published in the year 2021 by the author "Osama Abu Daoud, Omar Mahmoud Albatayneh and Khaled Ksaibati**".** This paper employs the technologies like deep learning and image classifiers by collecting the data from gravel roads. The image classifier used in this research is created by TensorFlow framework. The image classifier will classify the severity of corrugation on gravel roads into five levels. The data will be evaluated in two sections: visual inspection and the other one is developed image classifier. The main purpose is to provide a Smartphone detector for gravel roads corrugation. This classifier will enhance the data collection process and provide the decision-makers and local agencies with a cost-effective data collection tool. A confusion matrix determines the achieved accuracy by using gravel roads. Confusion matrics are used to visualize important predictive analysis like recall, accuracy etc. They give direct comparisons of values like true positives, False positives, False negatives and true negatives. The confusion matrix showed the accuracy level of 83% in the developed image classifier of practical

field.

[14] In the paper titled "**Analyzing and Improving the Image Quality of StyleGAN"** published in the year 2020 by the author "Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila**".** This paper utilizes StyleGAN architecture which yields better results in generative image modeling. Having the ability of improving image quality, this path length regularizer provides an advantage that the generator becomes much easier to invert. Quantitative analysis of the quality of the image is a difficult process. The difference in the density of two distributions in the high dimensional feature space of an InceptionV3 classifier is measured using Frechet inception distance (FID). Precision and Recall (P&R) delivers an additional visibility by quantifying the percentage of generated images that are similar to training data and the percentage of training data that can be generated, respectively. Progressive growing has been very useful in stabilizing high-resolution image synthesis. Finally this configuration has achieved 40% faster than original StyleGAN. This is because of weight demodulation, lazy regularization, and code optimizations.

[15] In the paper titled "**Deep Learning Object Detection Methods for Ecological Camera Trap Data"** published in the year 2018 by the author "Stefan Schneider, Graham W. Taylor, Stefan C. Kremer**".** In this Paper They are discussing about the ecologists in the Population uses camera traps as a basic approach to monitor animal population sizes and to manage the ecosystems around the world. Due to the more Usage of camera Traps the popularity of the camera trap methodology grew rapidly thereafter as they respond to motion, which generally corresponds with an animal entering the frame. They have used deep learning techniques to identify, detect, quantify, and determine animal species of wildlife within camera trap images by object detection. This Method is basically used for species classification, as the images are often messy, with animals being partly obstructed, positioned at varying distances from the camera, cropped out of the image, or extremely close to the camera. They have Implemented their capabilities by training and comparing two deep learning object detection classifiers:(1) Faster Region-Convolutional Neural Network It's the first approach where an image is taken and by using an image segregation algorithm and it is segregated into a series of different sized boxes, and each region segregated is passed through a CNN. Faster R-CNN introduces a Region Proposal Network (RPN) to the framework, that enables nearly cost-free region proposals. (2) Next approach for object detection is YOLO, it divides an image in the form of grid, with each grid cell acting as the origin for predefined 'anchors' relevant to the classification size of interest. When training and using YOLO, output classifications are returned for every anchor in a single iteration the dataset that is used for classification is Reconyx Camera Trap dataset along with the self-labelled Gold Standard Snapshot Serengeti data set. In the Recent Technologies and advancements in the field of computer vision and deep learning

have given rise to reliable methods of object detection. Faster R-CNN shows promise for accurate and greater efficiency analysis for camera trap data whereas YOLO is often less accurate due to the static nature of the anchor boxes Faster R-CNN overperformed the YOLO v2.0 classifier on the two given data sets with a good accuracy rate of 93.0% and 76.7%.

[16] In the paper titled "**Deep CNNs for large scale species classification"** published in the year 2019 by the author "Raj Prateek Kosaraju, Deng et al**".** The world contains several thousands of species of animals. Large Scale image classification is a quite challenging problem within the field of computer vision. As the real world contains billions of different varieties, understanding the performance of popular techniques and models is vital in order to apply them to real world tasks with the recent advances and Technologies by using deep learning, and with the availability of large datasets, we can attempt to solve this problem in a feasible way. We will adopt the best performance of a CNN architectures along with the methods like transfer learning and data pruning for evaluating classification performance and discovering the best work. They have implemented appropriate techniques and most efficient CNN based deep learning architectures on the dataset from iNaturalist 2019 Challenge to perform large scale species classification. The dataset used above can be suitably categorized into Amphibians, Reptiles, Birds, Insects, Fungi and plants with a total of many different species. Methods that are using the dataset pruning and transfer learning Techniques are shown to outperform models trained without using either of the two techniques. The dataset was undergone for Preprocessing to remove unwanted image and then resizing of images was done such that they are of the suitable size for the CCN architecture we use. We split the train/validation images set into 90% train and 10% validation Transfer Learning has been a popular technique used to improve the performance of classifiers over image classification tasks. There have been many popular CNN architectures that have been known to perform well on image classification tasks. Some of them are AlexNet was the first widely popular CNN architecture that beat traditional image classification techniques. Densenet was built on top of networks proposed earlier to address the vanishing gradients problem. ResNext is an extension of the original ResNet deep residual network architecture. The standard residual block will replace with another one which uses split-transform Merge strategy method that is used in inception modules. ResNext performs the best for large scale classification compared to other architectures. ResNext has the least validation error and greater efficiency among other architectures.

[17] In the paper titled "**Image Classification using Convolutional Neural Network**" published in the year 2020 by the author "Kavish Sanghvi, Adwait Aralkar , Saurabh Sanghvi and Ishani Saha". Image classification is one of the common and basic tasks in computer vision, and it has drawn attention in recent times. Monitoring of wild animals efficiently and suitably in their natural habitats is very

much essential to inform management and conservation decisions regarding the migration patterns, habitat protection, wildlife species, and grouping species of same animals together and also rehabilitation. Today, with the growth of technologies and with necessity applications of Artificial Intelligence, with different fields Deep Learning, Machine Learning and Neural Networks, and its subsets have Increased their momentum. The neural network developers are "training" the network in such a way to be "intelligent" and "independent" and has become a data centric model. The training needs appropriate tools and software such as the classifiers, which analyze, extract useful features and feed huge amounts of data. The above features are then used next time it is fed data to observe a pattern and train the network. Convolutional neural network is a special algorithm of artificial neural networks and deep learning, which can take an input image, assign importance to various aspects in the image and be able to differentiate one from another and aimed for effective image recognition, processing and classification. The classifier used in this paper is a fauna image classifier which uses convolutional neural network, that will be used to classify images of different animals and species which are captured in dense forest environments to achieve the good accuracy, and the researchers and population ecologists in artificial intelligence, zoological domains and neural network, study further to improve environmental, habitat and extinction patterns. A convolutional neural network (CNN) is developed and trained for efficiently classifying the images with greater accuracy results. The proposed method for classification of fauna images using convolutional neural network gives an accuracy of 91.84%, and classified images with 99.77% accuracy. For the purpose of training the model technologies like Leaky ReLU, Tensorflow, VGG16, etc have been used. It uses Leaky ReLU for implementation of convolutional neural network with fauna image classification. The efficiency convolutional neural network architectures and different activation functions was compared, and we noticed that ReLU activation function and VGG16 model to be most accurate and appropriate for classification of image.

[18] In the paper titled "**A Systematic Review About Use of TensorFlow for Image Classification and Word Embedding in the Brazilian Context**" published in the year 2020 by the author "Thereza Patrícia Pereira Padilha , Lucas Estanislau Alves de Lucena". Due to the recent Advancement of Technologies, there is huge demand for Artificial Intelligence (AI). Some Machine Learning applications include image recognition, natural language processing, speech recognition, search engines, , handwriting recognition and robot locomotion we can find several frameworks for Machine Learning applications, such as TensorFlow which uses Google's framework , Azure that uses Microsoft framework and AWS TensorFlow. The popularity among developers who are intending to apply machine learning to overcome challenges, such as to provide personalized and real-time feedbacks and preparing teachers to use intelligent tools in order to anticipate interventions. Thus, this paper intention is to present results of a systematic review that Brazilian papers use Google's AI framework to identify

and analyze, known as TensorFlow. There are many varieties of applications that TensorFlow can be used for, such as word embedding, image classification, recommendation systems, chatbot building and stock prediction. However, the systematic review many focused only on applications such as image recognition and word embedding because they are the two which was most used by TensorFlow. These two types of applications have transformed basic everyday actions such as writing a faster message on WhatsApp, unlocking a mobile screen through facial recognition or Email by proposing a set of words according to the context presented. Image classification applications serve to unlock a cell phone screen, to find a suspect, or even to identify an unknown object. Basically, the image classification process uses algorithms that search, compare, and try to find relationships among a given image and also other stored images in a database. Word embedding applications are other tasks that can be implemented by TensorFlow when you are writing a message, this type of application suggests some words according to the meaning presented, reducing the user's time and effort. The systematic review model proposed was adapted, which consisted of 3 phases: (1) Input – definition of one or more research questions, search engine, search string, and inclusion and exclusion criteria; (2) Processing- search for papers according to criteria and it defined search location and (3) Output – list of filtered papers with your relevant information. The systematic review of papers is made about word embedding and image classification applications with the TensorFlow framework written in Brazilian context. After a systematic search, 12 papers were researched in this area, 2 for word embedding applications and 10 being for image classification.

[19] In the paper titled "**Deep Learning methods for Animal Recognition and Tracking to Detect Intrusions**" published in the year 2020 by the authors "Ashwini V Sayagavi, Sudarshan T S B, and Prashanth C Ravoor". In the last few years, there has been a constant rise in number of reported human-animal conflicts. The major reason for this condition is due to increased rate of forest reduction. They stray close to humans mainly for food. Hence it is necessary to have a automated system to detect any animal sighting in the nearby locality. Recent technologies like artificial intelligence and deep neural network have made easier to tackle these problems. This project proposes to use YOLO object detection model for the detection of animal or humans. This model is trained to identify 6 different entities; in which 1 is human and the other 5 entities are elephant, zebra, giraffe, lion and cheetah. Once the object is detected the animal is tracked using CSRT to determine its intentions, and based on the obtained information, alerts are sent to the concerned authorities. The device used in the project is Raspberry Pi devices equipped with cameras. Deep Convolutional Neural Networks (DCNN) are known to be accurate, and outperform all other existing methods in the task of image classification. Object detection systems predict regions of interest within images and classify entities within those regions. YOLO is a DCNN object detection model which has good performance both in terms of accuracy as well as speed

of inference. The CSRT tracker is used to track animals effectively, since it is both fast and accurate. The input to the tracker is the centroid of the detected bounding box. This tracker is robust and reduces the need for continuous object detection, which is costly and compute intensive. In the end, this method achieved an accuracy of 98.8% and 99.8% to detect animals and humans respectively.

Chapter 3:

PROBLEM STATEMENT

Checking of wildlife species in their environment is important to study their movements. The detailed and accurate information about wildlife behavior and their location across the geographical areas helps us to transform our ability to manage, study and conserve the species as well as the biodiversity ecosystems. Here we investigate the data collected by camera traps and recognize the type of species and its location using Deep Learning Technique.

Chapter 4:

OBJECTIVE

Advances in Artificial Intelligence and Image Processing are changing/challenging the way people interact with digital images and video.

Animals involve or attack into human settlements primarily for the need of food which in turn may cause conflicts that results into injury to humans, animals, or both. A fully automated monitoring system that helps in detecting animal movements and notifies that is through the location unit to the concerned authorities can help to reduce causalities. Computer Vision is one of the easiest and best choice of technology that can help in solving these kinds of associated problems. The system that is referred here is in context of a network of cameras running image processing software.

4.1 Detects animals in the image:

Our Project detects animals in image. Though there are many challenges in capturing the image all the challenges are eradicated using the Machine Learning Technique i.e., using Deep Learning Methods like CNN.

4.2 Classifies Animals based on their species:

After detecting the animals, the image processing is done where the animals are segregated into their kind of species and the result of which species they belong is displayed.

4.3 Manual work is reduced by providing automated reliable detection system:

The manual monitoring would consume more time, tiring and may not accurate too. Therefore, it is the suitable automated model for easy monitoring and detecting.

4.4 Challenges in the dataset is removed:

In this project, we mainly focus on classification of animal species, which is a challenging work for humans. Images that are taken from motion activated cameras are not always perfect, and many images may contain animals that are too far away, sometimes close to the camera, or animal may visible partially. In addition, different weather conditions like lightning, rainy, shadows, or foggy can make the identification process difficult.

Some of the challenges are shown below:

4.4.1 Illumination:

Images are illuminated poorly and it becomes difficult to recognize the animal in the image the example is shown here.

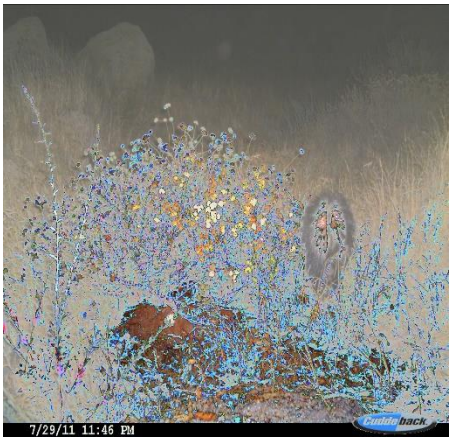
4.4.2 Motion Blur:

The shutter speed of the motion activated camera is not fast enough to remove the motion blur, so animals in the image are sometimes blurry. The example contains a blurred coyote.

4.4.3 Weather Conditions:

Poor weather like rain, lightning or fog in the background disturbs the clicking of image by the camera

4.4.4 Camera Malfunction:



Sometimes there can be a camera malfunction which may cause strange discolorations.

4.4.5 Close to the Camera:

It is hard to capture animals when it is very close to the camera, causing a forced perspective.

Chapter 5

SYSTEM REQUIREMENTS SPECIFICATION

5.1 Software Requirements

- Operating System: Windows 10
- Programming Language: Python 3.9 (latest version)
- Libraries: TensorFlow, OpenCV, PIP, NumPy, Pandas
- IDEs: Jupyter Notebook

5.2 Hardware Requirements

- Processor: 1GHz or faster
- Memory: 4GB or above

Chapter 6

METHODOLOGY

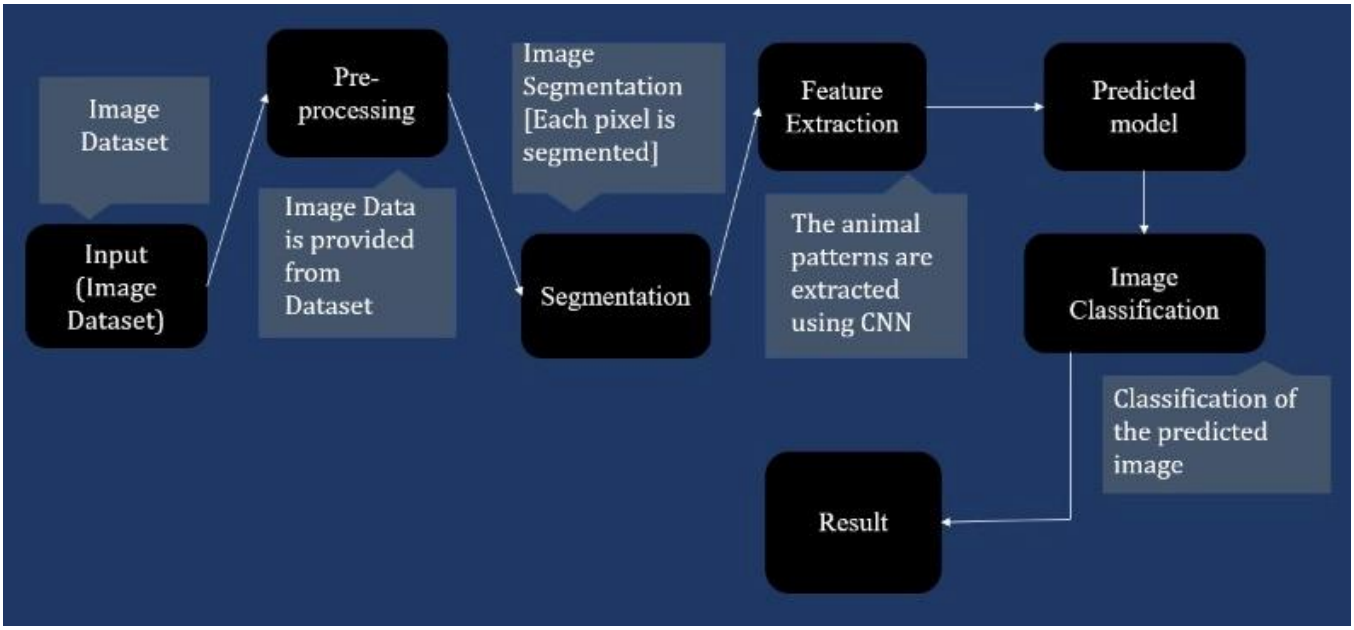


Fig: Flow Diagram

- The dataset collected is the input of our project. This dataset undergoes pre-processing technique where the unwanted data or the empty data are removed from the dataset
- After pre-processing of data, the images in the dataset are passed to the segmentation process where the background as well as foreground is formatted using the clustering techniques. The most used clustering technique for image processing is K means clustering method.
- The segments of the image are passed to the Deep learning method which is a supervised technique where the outcome of the task is known and the model is trained in the similar way. For image processing the suitable method is Deep CNN Algorithm which checks each pixel of the image and predicts the model this model is the output of the CNN algorithm.
- The predicted model of undergoes image classification method where several images are converted into pre-determined classes it's like predicting and classifying the probability that an image belongs to each class.
- The model also undergoes object detection technique where the classified images are used to locate the instances of predefined object classes within images.
- Finally, according to the user input in the front end the model processes the task in the backend the displays the output of the animal. The example of the outcome of our project is shown below:

6.8 Expected Outcome

Fig: Image 1

Fig: Image 2

The images 1 and 2 shows the possible outcome from the camera trap.

The picture taken from camera trap is processed using Deep learning Techniques and the animals can be classified based on their classes. Here the camera senses every motion that happen in its surrounding.

In Image 1 our model has tendency to detect all the animals present in the image which is trapped by motion sensored camera

In Image 2 though there in a background challenge our algorithm identifies the animal in it.

Chapter 7:

CONCLUSION

We have proposed an approach of detecting the animals in the image of the camera traps which helps to avoid them from poaching, prevent them from extinction and also through the images of the camera traps we can prevent all the illegal activities done in the forest areas. This approach helps to conserve and maintain balanced ecosystem. In this phase of the project, we reviewed the literature on the automatic detection and classification of Animal Species. This literature helps us understand the challenges that we face in the image dataset to identify the animals in a particular image. According to the literature experimental results we came to know that convolutional neural network helps in Animal Identification and helps to eradicate the challenges that we have mentioned in the dataset thet comes in the dataset and also the accuracy rate of this neural network is far better than other neural networks. This helps the ecologists to study about wildlife ecosystem and the forest officer to keep track on the activities that go throughout the forest areas.

7.1 Future Enhancements:

In terms of Future Enhancements, we can see that after collecting the data samples we can combine both detection and classification model to increase the accuracy but in future we need to build a model in such a way that the computational efficiency is also met along with the best accuracy of the model and also decrease the storage overhead at the time of training and inference to maintain the better performance gains.

Chapter 8:

REFERENCES

- Working of Deep Learning https://searchenterpriseai.techtarget.com/definition/deep-learning-deep-neural-network
- Applied Deep Learning Concepts: Part 4 – Convolutional Neural Network https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2#a86a
- Animal Detection Using Template Matching Algorithm---I K Amri and Murinto (2020)
- Deep Residual Learning for Image Recognition---Ruhua Wang, Chencho, Senjian An, Jun Li, Ling Li, Hang Hau (26th May 2020)
- Identifying, Counting, and Describing Wild Animals in Camera Trap Images with Deep Learning---Mohammed Sadegh Norouzzadeh, Anh Nguyen, Margaret Kosmala, Ali Swanson, Meredith Palmer, Craig Packer, Jeff Clune (15th November 2018)
- Wildlife Spotter---ABC's citizen science project (1st August 2016)
- Context R-CNN: Long Term Temporal Context for Per-Camera Object Detection--- Sara Beery, Guanhang Wu, Vivek Rathod, Ronny Votel, Jonathan Huang California Institute of Technology (2020)
- Deep learning improves acoustic biodiversity monitoring and new candidate forest frog species identification in the Philippines---Ali Khalighifar, Rafe M. Brown, Johana Goyes Vallejos (17 January 2021)
- Orientation Robust Object Detection in Aerial Images Using Deep Convolutional Neural Network-- QQ Liu, JB Li - Procedia Computer Science, 2019
- Deep Convolutional Neural Network Based Species Recognition For Wild Animal Monitoring---Guobin Chen, Tony X. Han, Zhihai He Roland Kays, and Tavis Forrester (2018)
- Animal species classification using machine learning techniques--Fahad Alharbi*, Abrar Alharbi, and Eiji Kamioka (2019)
- Edge Machine Learning for animal detection, classification and Tracking-- Amanda Tydén and Sara Olsson (2020)
- Quantum-soft QUBO Suppression for Accurate Object Detection--Junde Li and Swaroop Ghosh(2020)

- Fauna Image Classification using Convolutional Neural Network--Kavish Sanghvi and Ishani Saha(2020)
- Validating the practicality of utilising an image classifier developed using TensorFlow framework in collecting corrugation data from gravel roads--Osama Abu Daoud, Omar Mahmoud Albatayneh and Khaled Ksaibati(2021)
- Analyzing and Improving the Image Quality of StyleGAN--Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila(2020)
- Deep Learning Object Detection Methods for Ecological Camera Trap Data -- Stefan Schneider, Graham W. Taylor, Stefan C. Kremer(2018)
- Deep CNNs for large scale species classification -- Raj Prateek Kosaraju, Deng et al(2019)
- Image Classification using Convolutional Neural Network -- Adwait Aralkar , Saurabh Sanghvi(2020)
- A Systematic Review About Use of TensorFlow for Image Classification and Word Embedding in the Brazilian Context -- Thereza Patrícia Pereira Padilha , Lucas Estanislau Alves de Lucena(2020)
- Caltech camera traps. http://lila.science/datasets/caltech-camera-traps. [2019]