Real-Time Detection of Poaching in Wildlife Reserves with Thermal Imaging and CNN Using IoT and ML

Aishwarya Girish Menon
A. Abhishek
Arbaaz Ahmed
Anusha S.N
Kiran M
School of Computer Science and Engineering, REVA University

Abstract—With the advent of European colonization,Poaching and wildlife smuggling has been a vital issue since the 17th century. Illegal hunting of animals threatens the very existence of these organisms and results in the imbalance of the ecosystem and biodiversity of our planet. The objective of the proposed system is to detect poachers using the technologies, Internet of Things (IoT), Unmanned Aerial Vehicles (UAVs), Dashboard cameras, trail cameras and Machine Learning. Detection of poachers is performed using Infrared Cameras that can switch between normal imaging during the day and thermal imaging during the night. This approach reduces a lot of manual labor and improves the precision of detection of poachers with real-time analysis and alerting. Internet of Things is used for data collection and transmission while Convolutional Neural Networks (CNN) trained using TensorFlow, OpenCV and Python aids in the data analysis for classification. An IoT-based system is used for the automated surveillance task and to perform real-time detection on collected live feed. Classification of the live video feed also allows us to keep track of the number of each wildlife species, the number of people and vehicles present in reserve. An Artificial Neural Networks (ANN) model trained with Keras, is used for the detection of gunshots in the wildlife reserves as well. The proposed system is self-maintainable and doesn't require human interaction as it functions on solar energy and uses automated GPS system to traverse through the Wildlife reserves and track real-time location.

***Keywords*—Machine Learning, Convolutional Neural Networks, Thermal Imaging, Internet of Things, Artificial Neural Networks, Gunshot Detection, Poaching Detection**

1. INTRODUCTION

Poaching has been prevailing as a major concern in our society since the advent of European colonization. The need for food, traditionalmedicines and materials, selling of animal parts like Elephant tusks, Tiger fur, Rhino horns as materials or trophies for display, and the recreational hunting of various organisms have tremendously affected the population and survival rate of several animal species. Despite placing strict governmental rules on poaching, humans continue to hunt animals for wildlife smuggling. To protect animals in wildlife reserves from poachers, forest rangers manually conduct patrols on feet or by vehicles throughout the forest reserve. However, it is impossible for the forest rangers to manually traverse and track down poachers at every region quickly with limited manpower.

In some existing projects, the live video feed from a UAV surveillance system is sent to the forest rangers, who then have to manually search for poachers that results in a tedious and erroneous task to monitor the video feeds all day and night.

Moreover, trekking through the forests with wild animals at night, pose threats to the lives of the forest rangers. Recent research and techniques to resolve this issue has been focused on machine learning concepts like computer vision, object detection and on thermal imaging. There is a need for a system that runs real-time detection on the live video and audio feed from the surveillance system that either traverses through the wildlife reserves or tracks every region of the reserve at all times at every region of the reserve to survey for poachers. The system alerts the forest rangers if any anomaly is detected.

Internet of Things refer to network of interconnected devices that exchange information, control, and interact with the surroundings and every other device within that network without human interaction. ArtificialNeural Networks are computational system units or artificial replicas of the biological brain function that consist of processing nodes to apply activation functions on inputs and provide outputs in the same way the brain neurons of a biological brain process information.Convolutional Neural Networks are ANNs that perform image recognition and processing tasks based on the patterns derived from pixel data of images. Thermal imaging refers to the depiction of infrared radiations reflected off objects in the form of an image in the visible light spectrum.

In this paper, Internet of Things is used for data collection and surveillance, while CNN models classify poachers, forest rangers and different animal species from the live video feed and ANN models classify gunshot or other noise data from the live audio feed. If any anomaly is detected in the detection tasks, the forest rangers are immediately alerted via an SMS-based system.

2. LITERATURE SURVEY

In paper [1], the authors proposed a Mavic Pro 2 Drone system that transmits 4K videos at 30 fps over a distance of 7 km. The video streams from the drone are redirected through a controller via a Wi-Fi connection to a field laptop or remote server using 4G technology. A real-time messaging protocol is used to transmit the video feed to an online server withthe object detection model having a Faster-RCNN network.

In paper [2], the authors proposed a Drone system with a Global Navigation Satellite System (GNSS) for navigation and precise geographic coordinate positioning. The system was used for the detection of Elk species in Siberia. Data was collected using a Wide Area Sony RX1R camera and a Thermal Imaging camera Atom500 and processed manually with a Thermal Image object finder software.

In paper [3], the authors developed a Smart thermal infrared camera UAV system called SPOT. The system used a Faster RCNN network for object detection and Azure Cloud GPU for more processing speed and power. The UAV flight path was manually controlled and a FLIR 640 Vue thermal imaging camera was used for data collection.The object detection model used a Faster RCNN Network for detection and classification.

In paper [4], the authors proposed a Support vector machine-based object detection model. The insights from the analyzed data were sent via SMS to the respective person. The accuracy of the project was quite low using SVM.

In paper [5], the authors created Infrared Radiation Dataset with Aerial images of wildlife reserves with animals and poachers using the FLIR Vue Pro 640 camera with Tamarisk 640 for public use. The captured dataset was labelled using a tool called VIOLA and multiple object detection network models such as YOLO, Faster RCNN and Domain Adaptive Faster RCNN were tested on the dataset to find the most efficient network model.

In paper [11], the authors proposed an end-to-end, distributed, IoT based monitoring system for wildlife reserves based on edge cloud systems. Image classification and tagging were performed using TensorFlow and OpenCV. A synthetic training dataset was created using google stock images. The system was used to identify bears, deer, and coyotes. The IoT system acts as a set of eleven camera traps to collect images around different regions of the wildlife reserve. The collected data was transferred to the edge cloud computing system for storage and processing using Wi-Fi technology.

The disadvantages found in the papers above include the manual control of the UAV system for monitoring the reserve with controllers, use of inaccurate object detection and classification methods as seen in paper [4], lack of real-time alerting with the precise location of animals, poachers or rangers and manual input of data collected from the UAV to an object detection model for classification. In paper [5] data collected is used only for night-time monitoring as infrared waves cannot provide accurate results during the daytime.

3. PROPOSED SOLUTION

The proposed system uses IoT for Data collection and transmission, while CNN and ANN models using ML and Deep Learning aid in data analysis and classification. Raspberry Pi module was chosen as the microcontroller as it is equipped with a fast-processing speed, integrated Wi-Fi chip, stable network connectivity and quick response to program instructions in real-time.

A. Software Terminologies

Convolutional Neural Networks (VideoClassification):

CNN models are a subset of machine learning technology used for analyzing, detecting patterns, and classifying image and video datasets. It consists of 3 main layers. The first layer, convolution or maxpooling layer is used to find the features present at a given region of the feature maps generated to reduce the number of parameters to be learned. In the second layer, fully connected dense layers are used to feed the collected information about the image and its features to each of the neurons in the network and vector operations are performed. The final layer, the output layer is used to obtain the classification details of the given image based on its features upon applying activation functions to the input from the dense layer. Fig.1 represents the working of CNN models.

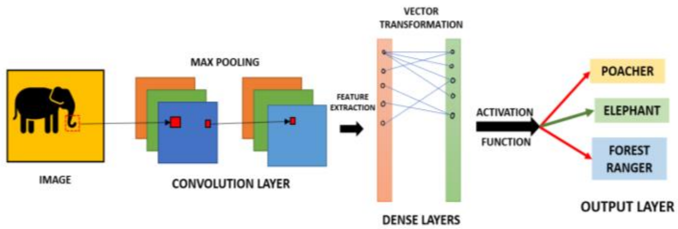

Fig.1. Working of the CNN models.

Artificial Neural Networks (Audio Classification):

Artificial neural networks are a collection of nodes that process the feature inputs fed to them by applying an activation function over them to obtain the final output. It consists of 3 layers, the input layer at which the input audio is fed to the nodes, the hidden layers that process the audio, extract features from its waveform and classify the feature inputs and the output layer at which an activation function is applied to the feature inputs from hidden layers before releasing the final output, that classifies what the audio sound represents. Fig.2 represents the simplified working of the audio classification ANN model.

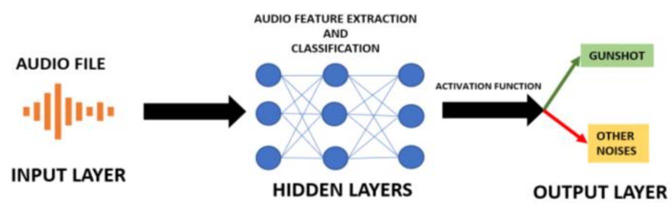

Fig.2. Working of the ANN models.

Internet of Things (Surveillance System):

Internet of Things refer to a network of electronic devices, software, embedded systems, and sensors interconnected between one another for the exchange of data. The interconnection between the trail cameras, drone systems and vehicle dashboard cameras to the central computing system for data exchange is an IoT network.

B. Surveillance System Construction

The collection of data is performed by the autonomous drone systems, trail cameras and vehicle dashboard cameras. These systems are controlled using a Raspberry Pi B+ module acting as the microcontroller or the brain of the system. The drone system chassis is constructed based on a quadcopter design with 4 brushless DC motors and propellors to propel the system into the air. It uses a flight controller for all flight control functionalities and an electronic speed controller to adjust the motor speed.A gyroscope is used to adjust the quad's orientation, acceleration, and velocity.The drone will be powered with a Li-On battery which may be chargedusing Solar panels coupled with Li-On charging modules. A GPS modulewill be used to guide the drone throughout different geolocations in the wildlife reserve, where each pair of geographical coordinates will be pre-programmed on the microcontroller. A magnetometer will help the drone to adjust and identify its current orientation and direction. In order to assist in the detection of Gunshots in the forest reserve, a USB microphone is attached to the USB port of the microcontroller. The collection of thermal or normal video feed of the wildlife reserve in real-time is performed using an Infrared camera unit attached to the drone. The camera activates thermal vision during the night and normal vision during the day.A light intensity sensoris used to detect the intensity of light throughout the day and adjust the camera vision. If the light intensity is low, the thermal imaging vision is activated and when the light intensity is high, the normal imaging vision is activated.The USB microphone present on the drone system transmits live noise feedback from the area where it is patrolling to the central computer. Fig.3 depicts the block diagram of the Drone system construction.

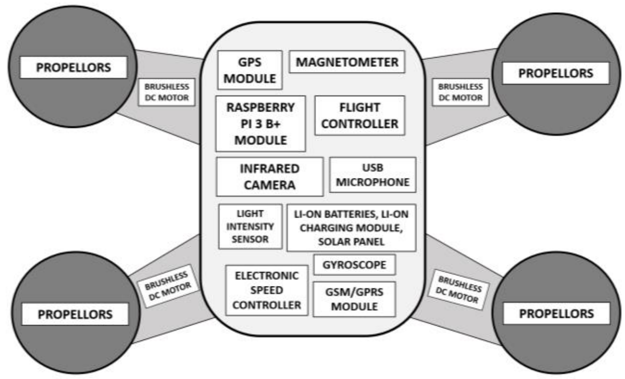

Fig.3. Block diagram of Autonomous GPS Drone system.

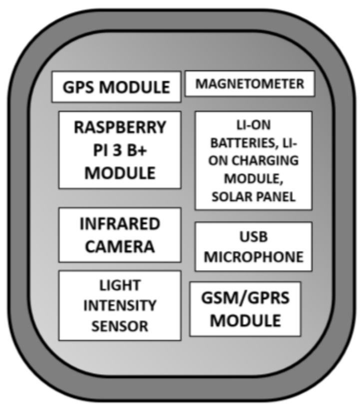

Fig.4. Block diagram of Dashboard and Trail Camera system.

Fig.4 represents the chassis and the components of a trail camera and dashboard camera system. The system consists of a Raspberry Pi B+ module functioning as the microcontroller, a magnetometer to track the system's direction, an infrared camera to stream live video, a USB microphone to record surrounding area noises, a light intensity sensor to detect whether it is day or night, a GSM/GPRS module to send alerts immediately if poachers or gunshots are detected and solar panels with Li-On charging modules and batteries to power the system.

C. Object and Acoustic Noise Detection Systems

The real-time video feed is analyzed and classified using twoCNN object detection models, one for daytime analysis and the other for night-time analysis. The models are pre-trained on a large dataset containing images of poachers, forest rangers and animals in thermal vision and normal vision. The Raspberry Pi B+ module is connected to the laptop using Wireless Connection through a common Wi-Fi network and streamed on-screen using VNC viewer software. Putty.exe is used to establish an SSH connection with the module for network configuration. The training of the CNN and ANN models were performed on Jupyter Notebook, an open-source software integrated development environment (IDE). All collected data such as video feeds, audio feeds, generated alerts and other counting and detection information are stored in the central computing system.

The datasets having normal images and thermal images are first acquired using python and OpenCV scripts. The dataset is then labelled using a tool called *LabelImg*. The labelled images are segregated into train and test folders on a 70:30 ratio. The required file paths are set, and a label map is generated along with TF records. The TensorFlow object detection library is cloned, and a preconfigured model called SSD mobile net v2 is reconfigured and updated to function according to the proposed system. The two CNN models are finally trained, wherein one model trains and tests on the Normal vision dataset for daytime classification and the other model trains and tests on the Thermal vision dataset for night-time classification. The trained models are loaded and run to detect poachers, forest rangers and different animal species in real-time.The activation function used by the CNN model is ReLU6which is a modified version of the rectified linear unit where the activation is limited to a maximum size of 6. An open-source autopilot Map-based system software called ArduPilot is used to pre-program the flight path with its geographic coordinates to the flight controller and microcontroller of the Drone system. The proposed system further verifies whether the object detected by the system as a poacher, or a forest ranger is accurate. This is done by the proposed system by sending the GPS coordinates of the location where the forest ranger or poacher was identified to the central system. The central computer uses python scripts with python libraries like Open Cage Geocoder and Folium to confirm the location of forest rangers in the wildlife reserve by comparing the coordinates from the surveillance system to each of the forest ranger's coordinates. The location of every forest ranger is continuouslytracked by using the IP addresses of their Mobile devices. If the coordinates do not match the coordinates of the forest rangers, then the identified human may be confirmed as a poacher.An acoustic noise detection model created with ANN is used for the detection of gunshots in the wildlife reserves. The USB microphone present on the proposed system transmits live noise feedback from the area where it is patrolling, to the central computer. The central computer converts the received noise data to a wav file using python scripts and feeds the wav file to an ANN model that was trained using Gunshot and other noise waves to classify the detected noise using Keras.Mel-frequency Cepstrum (MFCC) was used for feature extraction during data preprocessing. A python package called Librosa is used to perform audio analysis on the audio datasets. Keras is an open-source software library that aids in the creation of ANN models. The model used ReLU and SoftMax activation functions with Adam optimizers. SoftMax is a function that turns a vector with k real values into a vector with k real values having a total sum of 1. ReLU is a linear function that outputs input directly if it's positive else it outputs zero.

D. Working of the proposed system

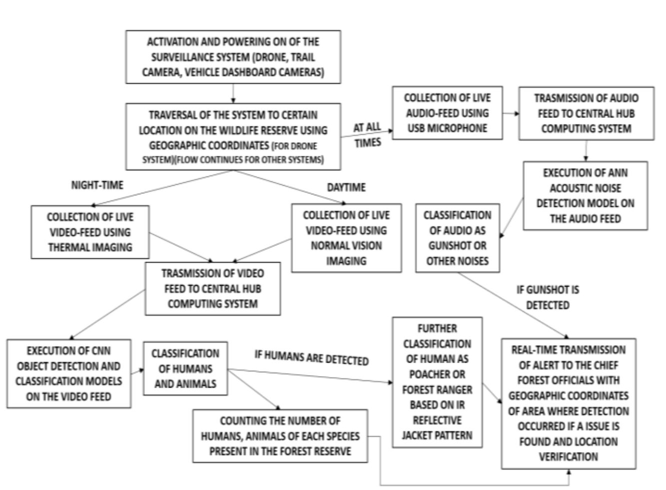

Fig.5. Flow diagram of processes of the proposed system.

Fig.5 represents the complete process of the proposed system as a flow diagram. The system begins operation upon activation and powering it on. For drone systems, the drone traverses to a certain location on the wildlife reserve using the pre-programmed geographic coordinates on the flight controller. The system (drone, dashboard camera or trail camera) collects normal vision live videofeeds during the day and thermal vision live videofeeds during the night. The USB microphone simultaneously collects the live audiofeed. The collected live audio and videofeeds are sent to the central computing system.

CNN models are run over the live-video feeds to detect humans and animals. It also returns the count of the number of humans and the number of animals of different species via SMS to the chief forest rangers. If humans are detected by the CNN model, it further classifies the human as a poacher, ora forest ranger based on a unique pattern reflected by the IR reflective jackets worn by forest rangers on the live videofeeds. At the same time, the ANN model is executed on the live audiofeed for the detection of gunshots. If the CNN and ANN models detect poachers or gunshots at a particular location, it performs the real-time transmission of alerts to the chief forest officials with the geographic coordinates of the location where the anomaly was detected via an SMS system using the GSM/GPRS module. Further verification of the location of poachers and forest rangers is performed by verifying the cellular IP address location coordinates of the forest rangers to ensure that the detected human being is a poacher or not. If the coordinatesdo not match with the coordinates of any of the forest rangers, the detected human is verified as a poacher.

4.Experimental Results

For the purpose of demonstration on how the surveillance system would stream the video data to the central computing system and the how object detection models are run on the live feed, a Raspberry Pi 3 B+ module with NOIR IR camera has been used. The CNN and ANN models are run inside the Raspberry Pi module, on the real-time data collected by the NOIR camera. Fig.6 below is an image of the microcontroller and camera.

Fig.6. Raspberry Pi 3 B+ and NOIR IR camera.

During daytime, the surveillance system collects live video feed using normal vision and transmits the video to the central computer system. A CNN model is then run on the live video feed to classify any object it detects as Forest ranger, poacher or any animal species. The objects classified as humans initially and they are further classified as Forest rangers if the model detects any special patterns reflected by the reflective jackets worn by them, on the live feed. The other objects classified as humans are further classified as Poachers, if it doesn't detect any unique reflective patterns on the human. The model was trained using a dataset having 1036 images consisting of poachers, forest officials and elephants as the animal species in the forest reserve. The images of poachers and rangers used for training are custom collected dataset while the images of elephants used for training are google images. The labels used for labelling and classifying the images were poacher, forest ranger and elephant. The model was trained for 10,000 epochs and the average loss score was 0.171. The average accuracy was 86-98%. The daytime dataset was split at a percentage of 70% for training and 30% for testing.

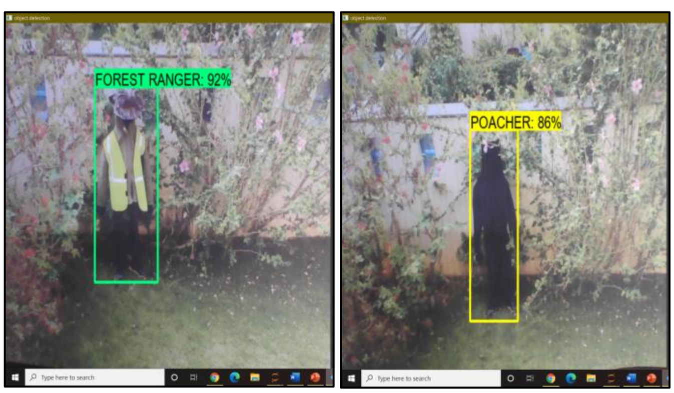

Fig.7. Detection and Classification of forest rangerand poacher usingDaytime CNN model.

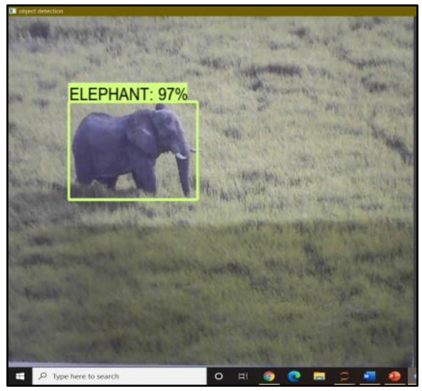

Fig.8. Detection and Classification of Elephants using Daytime CNN model.

Fig.7 and Fig.8 depict the real-time detection and classification of poachers, forest rangers and elephants when the CNN model is run on the live normal vision video-feed from the surveillance system.During Night-time, the surveillance system collects live video feed using thermal vision and transmits the video to the central computer system. A CNN model is run on the live video feed to classify any object it detects as forest ranger, poacher, or any animal species.The objects classified as humans initially and they are further classified as Forest rangers if the model detects any special patterns reflected by the reflective jackets worn by them, on the live feed.The other objects classified as humans are further classified as Poachers, if it doesn't detect any unique reflective patterns on the human as done in daytime model. The model was trained using a dataset having 400 images consisting of poachers, forest officials and elephants as the animal species in the forest reserve. The dataset of poachers, rangers and elephants used for training were obtained from BIRDSAI[7] dataset and google images. The labels used for labelling and classifying the images were poacher, forest ranger and elephant. The model was trained for 10,000 epochs and the average loss score was 0.233. The average accuracy was 90-100%. The night-time dataset was split at a percentage of 70% for training and 30% for testing.Fig.9 and Fig.10 depict the real-time detection and classification of poachers, forest rangers and elephants when the CNN model is run on the live thermal vision video-feed from the surveillance system.

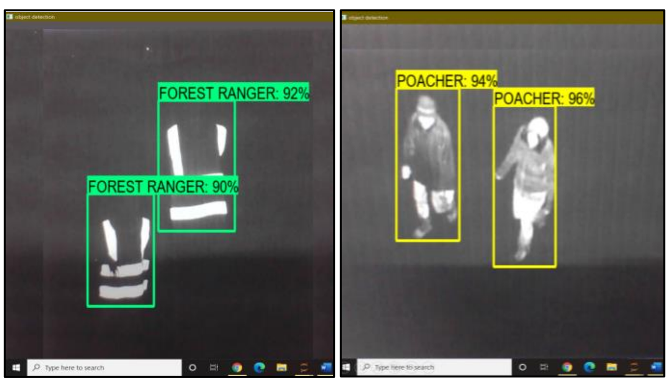

Fig.9. Detection and Classification of forest rangers and poachers usingNight-time CNN model.

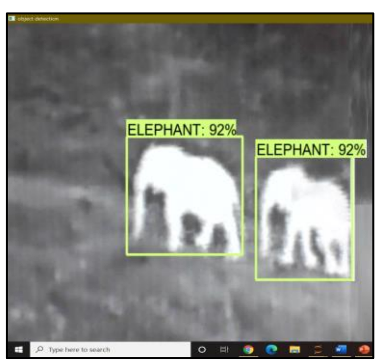

Fig.10. Detection and Classification of Elephants using Night-timeCNN model.

The USB microphone present on the drone system transmits live noise feedback from the area where it is patrolling to the central computer. The central computer converts the received noise data to a wav file using python scripts and feeds the wav file to an ANN model that was trained using Gunshot and other noise waves to classify the detected noise using Keras. The ANN model was trained for 250 epochs and acquired an average accuracy of 80%. The dataset was split at a percentage of 80% for training and 20% for testing. Fig.11 represents the output obtained upon classifying an audiofeed detected as gunshot noise with the label *gun_shot*. The graph plotted using matplotlib depicts the waveform of the detected audio as a time series graph for visualization.

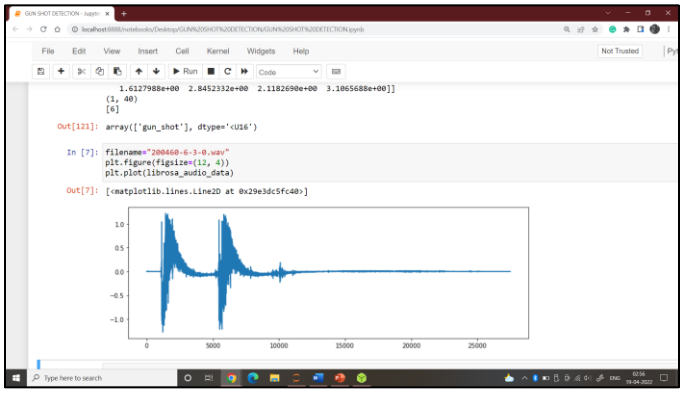

Fig.11. Detection of Gunshots from recorded live audiofeed.

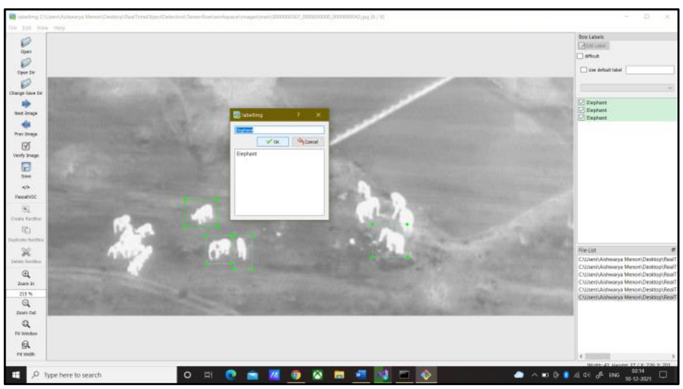

Fig.12. Labelling of dataset images using *LabelImg* tool.

Fig.12 represents the labelling process of the collected dataset using *LabelImg* tool.

5. Conclusion

Poaching has been a prevailing issue in the modern society that affects the biodiversity, ecosystem and the very existence of the flora and fauna of our planet. Thermal imaging and Machine learning have emerged as novel technology capable of resolving poaching. In this paper, the proposed system combines two disciplines of computing: the Internet of Things and Machine Learning. Real-time video streams from the surveillance system are evaluated and classified using a convolutional neural network model. An ANN model is further used for the detection of gunshots in the wildlife reserves. Real-time alerts are transmitted to the chief forest rangers or officials upon the detection of poachers or gunshots in wildlife reserve.The proposed system will provide aid to the Governmental, Commercial and NGO organizations for the protection, preservation, and conservation of endangered animals in the wildlife reserves.The estimated cost of the system is cost-effective compared to commercial systems in the present-day markets. Multiple models of the proposed systems may be deployed using the proposed plan at the cost of one high-end system currently available in the market.The proposed system will also reduce the dangers faced by forest rangers during the night patrols in wildlife reserves with predatory animal species by reducing the manpower required for monitoring each region of the reserve manually.

6. Future Work

In this paper, the Convolutional Neural Networks model for detecting Poachers in wildlife reserves and an Artificial Neural Networks model for detecting gunshots were developed and tested on a small-scale surveillance system. In future, a fully functional IoT-based autonomous GPS enabled surveillance system (using Drones, Trail cameras, Vehicle dashboard cameras) with the ML models [6], [8], [9], [10], [12] may be implemented for commercial and public use. The technology of using laser ID reflectors to uniquely identify forest rangers may also be implemented in future.

7. Acknowledgement

We would like to thank REVA University for all the guidance and support provided to us in the completion of this technical research paper and project.

8. References

[1] C.Chalmers, P.Fergus, S.Wich, and A.C. Montanez, “Conservation AI: Live Stream Analysis for the Detection of Endangered Species Using Convolutional Neural Networks and Drone Technology”, Arxiv, Oct 2019.
[2] V. Atuchin,A. Prosekov, A. Vesnina, andA. Kuznetsov, “Drone-Assisted Aerial Surveys of Large Animals in Siberian Winter Forests”, DOI:10.21203/rs.3.rs-55278/v1, 2020.
[3] E. Bondi, F. Fang, M. Hamilton, D. Kar, D. Dmello, J. Choi, R. Hannaford, A. Iyer, L.N. Joppa, M. Tambe, andR. Nevatia, “SPOT Poachers in Action: Augmenting Conservation Drones with Automatic Detection in Near Real Time”, IAAI-18, 2018.
[4] Basha, S. M., Poluru, R. K., & Ahmed, S. T. (2022, April). A Comprehensive Study on Learning Strategies of Optimization Algorithms and its Applications. In *2022 8th International Conference on Smart Structures and Systems (ICSSS)* (pp. 1-4). IEEE.
[5] K. Edemacu, J.W. Kim, B. Jang, and H.K. Park, “Poacher Detection in African Game Parks and Reserves with IoT: Machine Learning Approach”, ICGHIT-2019, DOI 10.1109/ICGHIT.2019.00011, IEEE, 2019.
[6] A. Puri andE. Bondi, “Space, Time, and Counts: Improved Human vs Animal Detection in Thermal Infrared Drone Videos for Prevention of Wildlife Poaching”, KDD 2021 Fragile Earth Workshop, 2021.
[7] E.Bondi, R.Jain, P.Aggrawal, S. Anand, R. Hannaford, A. Kapoor, J.Piavis, S. Shah, L. Joppa, B.Dilkina and M.Tambe, “BIRDSAI: A Dataset for Detection and Tracking in Aerial Thermal Infrared Videos”, IEEE Xplore, 2020.
[8] C. Chalmers, P. Fergus, C.C. Montanez, S.Longmore, and S. Wich,“Video Analysis for the Detection of Animals Using Convolutional Neural Networks and Consumer-Grade Drones”, Journal of Unmanned Vehicle Systems, 2021.
[9] M. Mulero-Pázmány, R. Stolper, L.D.van Essen, J.J.Negro, andT. Sassen,“Remotely Piloted Aircraft Systems as a Rhinoceros Anti-Poaching Tool in Africa”, PLoS ONE, vol: 9(1), 2014.
[10] D. Brown and D. Schormann, “Poacher Detection and Wildlife Counting System”, SATNAC 2019, 2019.
[11] A. Rosales, N. Golubovic,C. Krintz, and R. Wolski, “Where's The Bear? Automating Wildlife Image Processing Using IoT and Edge Cloud Systems”,DOI :247-258. 10.1145/3054977.3054986., 2017.
[12] I. Zualkernan, S. Dhou, J. Judas, A.R. Sajun, B.R. Gomez andL.A. Hussain, “An IoT System Using Deep Learning to Classify Camera Trap Images on the Edge”, MDPI, Computers, 2022.