ISSN 2320 - 2602

Volume 12 No. 6, June 2023

International Journal of Advances in Computer Science and Technology

Available Online at http://www.warse.org/IJACST/static/pdf/file/ijacst011262023.pdf

https://doi.org/10.30534/ijacst/2023/011262023

Night Vision Thermal sensor based Animal Movement Observation using CNN and YOLO v3

Y. Rakesh Kumar[1], Donthula Bhavana Yadav[2], K. Jagruthi[3], K. Laxmi[4], Priyanka[5]

[1] Department of Electronic and Communication Engineering, G. Narayanamma Institute of Technology and Science (for Women), Shaikpet, Hyderabad, Telangana, India-500104, rakeshyacharam@gnits.ac.in

[2] Department of Electronic and Communication Engineering, G. Narayanamma Institute of Technology and Science (for Women), Shaikpet, Hyderabad, Telangana, India-500104, donthula.bhavana@gmail.com

[3] Department of Electronic and Communication Engineering, G. Narayanamma Institute of Technology and Science (for Women), Shaikpet, Hyderabad, Telangana, India-500104, jagruthi193@gmail.com

[4] Department of Electronic and Communication Engineering, G. Narayanamma Institute of Technology and Science (for Women), Shaikpet, Hyderabad, Telangana, India-500104, laxmireddy1599@gmail.com

[5] Department of Electronic and Communication Engineering, G. Narayanamma Institute of Technology and Science (for Women), Shaikpet, Hyderabad, Telangana, India-500104, priyankachauhangnits418@gmail.com

Received Date : April 25, 2023 Accepted Date : May 18, 2023 Published Date : June 07, 2023

ABSTRACT

Developing an automated system for wildlife detection and recognition using thermal cameras can be challenging nowadays, which have many applications in wildlife conservation and management. This includes many challenges like detecting false positives and negatives and also environmental factors like temperature and humidity. Developed a system for the classification of animals using one of the developed algorithms, the computer-aided CNN algorithm and Python Flask web application that loads a pre-trained convolutional neural network (CNN) model for image classification (cat or dog) and allows the user to upload an image for classification. The wildlife conservationists need a system, which uses a thermal images or videos and detects the animal presence in video. Temperature difference is used to distinguish the animals and can develop the system which uses images to detect the animals in a given video. Using this option, the conservationists will identify the animal using the developed system which is built with yolov3 (You Look Only Once, version 3) algorithm. Finally, if a pet animal is surrounded by a wild animal, or if both pet and wild animals are present in the video, an alarm indication is given in order to protect domestic animals from the wild animals and also safeguard the agriculture fields.

Key words: CNN, Python Flask web application, Thermal cameras, Temperature difference, Wildlife detection, YOLO v3

1. INTRODUCTION

Thermal imaging technology for wildlife detection has many advantages such as being able to detect animals in low light conditions and through foliage. Developing an automated system for wildlife detection and recognition using thermal cameras can be challenging but promising. Machine learning algorithms can be used to detect and classify animals based on their thermal signature. But, the system should be designed to handle variations in thermal signature due to environmental factors, such as temperature and humidity and also include challenges involved in distinguishing between different animals and avoiding false positives. The automated system can have numerous applications in wildlife conservation and management, such as detecting poaching activity and monitoring endangered species [3]. Optimizing the animal detection process using thermal imaging technology, such as using image processing techniques will improve the contrast of thermal images and adjusting the temperature range to match the animal's body temperature [2].

Firstly, the classification of thermal images is based on CNN algorithm and Flask web application which allows uploading the thermal image for classification and also mentions the probability of identification of thermal image. Later, used YOLO v3 algorithm to detect the animal movement in a thermal video and alarm indication is given for specific animals according to user. The YOLOv3 algorithm is a state-of-the-art object detection algorithm that has been adapted to work with thermal images. Proposed method uses a dataset of thermal images which achieves high detection accuracy.

2. LITERATURE REVIEW

CNN-based approach has high accuracy i.e.., 91.3% in classifying animals in thermal imagery, using small dataset of thermal images due to the limited availability of such data, which is a challenge in this field of research [8].

A pre-trained CNN model (ResNet50) and fine-tuning it for the animal classification task with a collected dataset of thermal images of animals, includes cows, deer, dogs, and cats, and used it to train and test the model, gave high accuracy in the experiments, with an overall accuracy of 96.7% on the test set [7].

A vast number of studies have been carried out in the field of thermal imaging and YOLO v3. The YOLOv3 algorithm is a powerful tool for real-time person detection and tracking in thermal images. The algorithm is able to accurately detect and track individuals in various scenarios, including low-light and crowded environments.

The performance of the algorithm is affected by various factors, such as the size of the training dataset and the choice of hyper parameters. Further optimization of these factors can potentially improve the accuracy of the algorithm [1]. Thermal cameras were utilized to track the movements of huge ungulates in an Indian tropical jungle. They record the nighttime behavior of the creatures by capturing their patterns of movement. Thermal video processing is a powerful tool for security applications, as it can detect objects and people based on their heat signatures, even in low light or obscured conditions. Its application in security improved detection rates and reduced false alarms. There are various techniques for processing thermal video data, including background subtraction, image segmentation, and object recognition. Each technique has its own strengths and weaknesses, and the choice of technique depends on the specific security application [4].

The YOLOv3-based system outperforms other detection methods, achieving high accuracy and fast processing times. the effectiveness of the proposed system through experimental evaluations, which show that the system can achieve high precision and recall rates for vehicle detection in thermal imagery. There are limitations to the system using YOLO v3, such as the need for more diverse training data to improve performance on different environmental conditions [5]. Thermal imaging is also used to track the behavior of arboreal creatures that are active at night in an Australian forest. They discovered that thermal cameras offered a non-invasive technique to observe these creatures in their natural habitat while being able to pick up on minute behavioral variations, such as changes in resting posture.

YOLOv3 performs significantly better than YOLOv2 and YOLOv3-tiny in terms of accuracy and detection speed. To further demonstrate the effectiveness of the system using YOLOv3, there are some experiments conducted on a real-world dataset. The experimental results show that the proposed system achieves high accuracy and efficiency in detecting and counting even in challenging scenarios such as and crowded scenes. YOLOv3-based detection and counting system outperforms both models, which are like Faster R-CNN and SSD, two widely used object detection models, in terms of accuracy and detection speed [6].

3. PROPOSED DESIGN

3.1 PROPOSED DESIGN FOR THERMAL IMAGE CLASSIFICATION USING CNN

Classification of thermal images using CNN involves various steps. Figure 1 shows the steps for thermal image classification with the probability of identification using CNN and Python Flask web application.

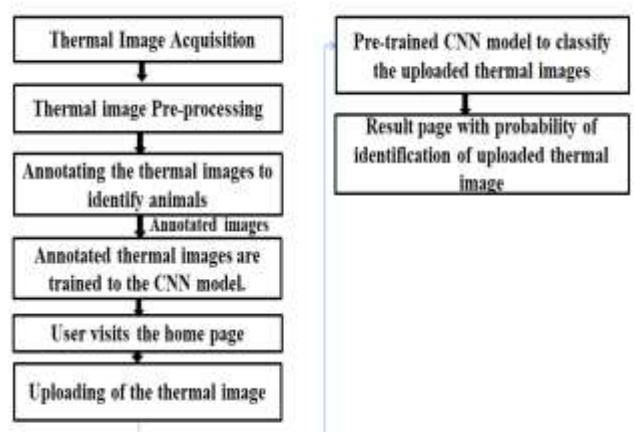

Figure 1: Block schematic for thermal image classification with the probability of identification using CNN and Python Flask web application

1. Thermal image acquisition: Thermal camera is used to capture thermal images of animals. Thermal cameras detect infrared radiation emitted by animals, which is then converted into temperature values and displayed as an image.

2. Thermal Pre-processing: Pre-processing can involve several steps, which is done before training to CNN model, such as resizing, normalization, and filtering. Resizing is done to ensure that all images have the same dimensions. Normalization can be used to adjust the temperature range of the images. Filtering can be used to remove noise or enhance certain features.

3. Annotating the thermal images to identify animals: Thermal images are annotated with labels to indicate the presence of objects or regions of interest. This process involves manually marking areas of the image that correspond to specific objects or features. The annotations can be stored as separate files.

4. Annotated thermal images are trained to CNN model: Once thermal images have been pre-processed and annotated, they can be used to train a CNN model. The annotated images are used as input to the model, and the model learns to associate the input images with the corresponding labels. During training, the model adjusts its weights and biases to minimize the difference between the predicted labels and the ground truth labels. After training, the model can be used to predict the labels of new, unseen thermal images.

5. User visits the home page: Later user visits the home page. When a user navigates to the root URL of the web application, it renders the "home.html" template.

6. Uploading of thermal image: Upload the thermal image which is to be classified in a home.html page.

7. Pre-trained CNN model to classify the uploaded thermal images: After extracting the uploaded image from the request and saves it to a path on the server specified by the "UPLOAD_FOLDER" variable. Then, the "classify()" function is called with the path of the uploaded image and a pre-trained CNN model called "cnn_model". The "classify()" function returns a label and a probability associated with the input image.

8. Result page with probability of identification of uploaded thermal image: The results are then rounded to two decimal places and passed to the "classify.html" template along with the filename of the uploaded image. The "classify.html" template is then rendered, and the results are displayed on the page.

3.2 PROPOSED DESIGN FOR ANIMAL MOVEMENT OBSERVATION USING YOLO v3

To observe the animal movement in a thermal video using YOLO v3 involves various steps. Figure 2 shows the steps for animal movement observation using YOLO v3 and thermal sensor.

1. Image Acquisition: Data set consisting of images of various animals collected from Kaggle website.

2. Image Pre-processing: The captured images are pre-processed to enhance the image quality and remove any noise. The pre-processing techniques can include image filtering, thresholding, and image enhancement as explained below.

- Image Filtering: It is a technique used to smooth out or sharpen images by applying a mathematical function or filter function which can be designed to reduce noise or to enhance specific features in the image, to each pixel in the image. Some common types of image filters include Gaussian filter, Median filter, and Bilateral filter.

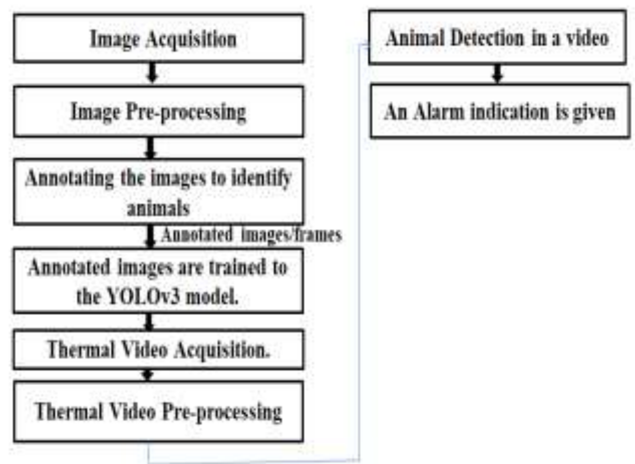

Figure 2: Block schematic for animal movement observation using YOLO v3 and Thermal sensor

- Thresholding: It is a technique used to convert gray scale images to binary images by separating pixels into two groups based on a threshold value. Pixels with values above the threshold are set to one, while pixels with values below the threshold are set to zero. This is often used to segment images and extract specific objects or features from an image.

- Image Enhancement: It is a set of techniques used to improve the visual appearance of images by increasing the contrast, brightness, or sharpness of the image. Some common techniques include histogram equalization, contrast stretching, and sharpening. These techniques can be used to improve the clarity and visibility of images, and to bring out specific features in the image.

3. Annotating the images to identify animals: Then collected images are annotated, where annotation is the process of labeling or marking specific objects, regions, or features of an image, video, or text data to make it understandable for

machine learning models. Annotation has been done using VIA (VGG Image Annotator).

4. Annotated images are trained to YOLO v3 model: Later, the pre-processed annotated images are then fed to the YOLOv3 algorithm, which is a deep learning algorithm that can detect and localize objects in real-time.

5. Thermal Video acquisition: The first step in video preprocessing is the acquisition of video data from a thermal camera or other sensor or any other website. The quality and resolution of the video data can have a significant impact on the subsequent analysis.

6. Thermal Video Pre-processing:

Thermal Video stabilization: Video stabilization is the process of removing unwanted camera shake or jitter from video data. This can be achieved using software or hardware techniques, such as optical stabilization. Data can be affected by noise, which can reduce the accuracy of subsequent analysis. Techniques such as spatial filtering, temporal filtering, and wavelet de-noising can be used to reduce noise in video data. Preprocessing of the thermal video data can be done in several ways before using it for detection and tracking using YOLOv3. These steps include:

- Contrast stretching: Thermal video data can have low contrast, which can make it difficult to detect people. This technique is applied to enhance the contrast of the video data.
- Background subtraction: This technique is used to detect moving objects in the thermal video data. This helps to identify people who are moving in the video.
- Spatial filtering: This technique will be applied to remove noise from the thermal video data. This helps to improve the accuracy of person detection.
- Segmentation: Segmenting the thermal video data to isolate people from the background. This helps to improve the accuracy of detection and tracking.

7. Animal Detection in a video: The YOLOv3 algorithm uses a convolutional neural network (CNN) to detect the specified animal in a image and classify them into different categories.

8. An Alarm indication is given: Once it identifies the specified animal in a video, an alarm is given.

4. RESULTS AND DISCUSSION

We have two request methods “POST” and “GET”. “POST” method is used to send data to a server to create or update a resource. “GET” method is used to request data from a specified resource. Here, “GET” method used to get template home.html to request data stored in it. “POST” method used to send data to update home.html. The data contains label and its probability in percentage. The developer enabled server in debug mode which allows going through each line of code, evaluating the apps variables, methods in program .py file and how well code is working.

Defined a Flask web application that allows users to upload an image and classify it as either a cat or a dog. The code is composed of several functions and uses a pre-trained CNN model to perform image classification. Then imported the necessary packages, including TensorFlow, Flask, and OS. It also defines a path to the directory where the uploaded images will be stored and a path to the directory where the pre-trained model is stored. The preprocess_image () function takes an image as input and resizes it to a square image of 192x192 pixels. It then normalizes the pixel values to be in the range [0,1]. The load_and_preprocess_image() function reads an image from a file path and preprocesses it using the preprocess_image() function. The classify () function takes a pre-trained CNN model and an image file path as input. It loads and preprocesses the image, reshapes it to have a batch dimension of 1, and passes it through the model to obtain a probability score for each class (cat and dog). It then returns the predicted label (cat or dog) and the probability score for the predicted class.The home() function defines the home page of the web application.

The upload_file() function is called when the user uploads an image. If the request method is GET, it returns the home page. If the request method is POST, it saves the uploaded image to the directory specified earlier and calls the classify () function to obtain the predicted label and probability score. It then returns a template with the image file name, predicted label, and probability score.The send_file() function is used to send the uploaded image back to the client.Finally, the if __name__ == "__main__": statement runs the Flask application.

The main function apply_yolo_object_detection takes an image as input and returns the same image with bounding boxes drawn around the detected objects and their class labels. It does this by first creating a blob from the input image using cv2.dnn.blobFromImage function and then setting this blob as the input to the pre-trained neural network. The output from the neural network is then processed to extract the bounding box coordinates, class indexes, and class scores of the detected objects. Non-maximum suppression (NMS) is applied to filter out overlapping boxes and only the most probable bounding boxes are kept. Finally, the bounding boxes are drawn on the input image and their class labels are added using the draw_object_bounding_box function.

The draw_object_bounding_box function takes an image, the index of the object class, the bounding box coordinates, and the index of the bounding box as input and returns the same image with a bounding box drawn around the object and its class label added to it. It also plays a alarm sound if the detected object is any specified animal by user. The draw_object_count function adds a label to the image indicating the number of objects of the desired class found in the image. The start_video_object_detection function captures a video and applies object detection on each frame of the video. It displays the processed frames on the screen with a reduced window size. It helps in various fields such as in forest areas, domestic animal protection areas etc.

The accuracy of a model is a crucial metric for any machine learning task. CNN models are well-known for their high accuracy in image classification tasks, and with proper training and tuning, can achieve high accuracy levels. YOLOv3, on the other hand, is primarily designed for object detection tasks and has demonstrated high accuracy in detecting objects in complex scenes. However, the accuracy of object detection in thermal videos using YOLOv3 can be influenced by the quality of the input video, the size of the objects being detected, and the complexity of the background.

Another important metric to consider is the speed of the model. CNN models can be relatively slow, especially if the input images are high resolution or if the model has many layers. On the other hand, YOLOv3 is designed to be a faster model for object detection. It uses a single neural network to process the entire image and can detect objects in real-time.

Figure 3: Identification of cat in a video

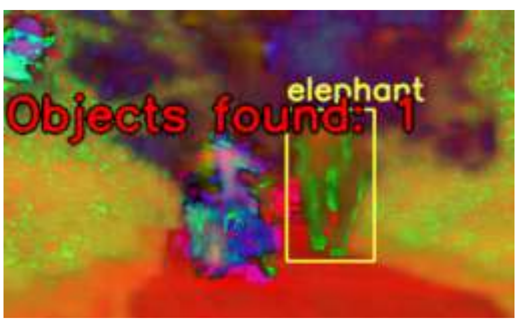

Figure 4: Identification of elephant in a video by an alarm

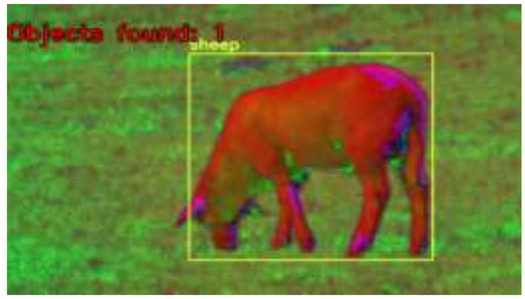

Figure 5: Identification of sheep in a video

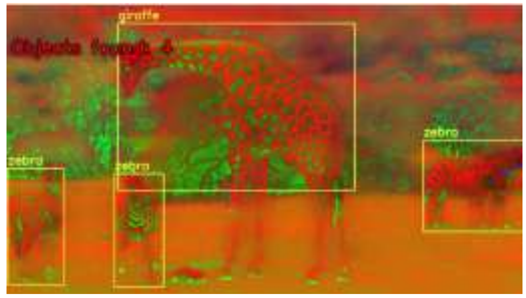

Figure 6: Identification of giraffe and zebra in a video

Model size is also an important consideration, especially for web applications where the model needs to be loaded quickly. CNN models can be large, especially if they are pre-trained models or have many layers. YOLOv3 is relatively small in size, which makes it easier to deploy in a web application.

While CNNs can perform image classification tasks well, they are not specifically designed for object detection tasks. YOLOv3, on the other hand, is designed for object detection tasks and can detect objects in real-time. This makes YOLOv3 a better choice for the detection of animal movement in thermal video. Figure 3 shows the detection of a cat and also labeled as one object is found in a thermal video. Figure 4 shows the detection of an elephant in a video by an alarm as the user want to detect wild animal. Figure 5 shows the identification of a sheep in a video. Figure 6 shows the identification of zebra and giraffe in a video.

Based on the above metrics, it appears that using YOLOv3 for object detection in thermal video and CNN for image classification is a good approach for a Flask web application. However, the choice of model ultimately depends on the specific requirements of the application and the type of data being processed.

We have taken the following training parameters to train the thermal images to YOLO v3:

1. Batch: This is the number of images used in each iteration of training. Here, batch=64, which means that the model will process 64 images in each iteration of training.

2. Subdivisions: This is the number of mini-batches that will be processed before the gradients are updated. In this case,

subdivisions=1, which means that the gradients will be updated after each batch of 64 images.

3. Width: This is the width of the input image that will be used during training. Here, width=416.

4. Height: This is the height of the input image that will be used during training. Here, height=416.

5. Channels: This is the number of channels in the input image. In this case, channels=3, which means that the input images are RGB images.

6. Momentum: This is the momentum used in the optimization algorithm. Here, momentum=0.9.

7. Decay: This is the weight decay used in the optimization algorithm. In this case, decay=0.0005.

8. Angle: This is the maximum angle of rotation used for data augmentation during training. In this case, angle=0, which means that the images are not rotated during training.

9. Saturation: This is the maximum saturation factor used for data augmentation during training. In this case, saturation=1.5, which means that the saturation of the images can be increased by up to 50%.

10. Exposure: This is the maximum exposure factor used for data augmentation during training. In this case, exposure=1.5, which means that the exposure of the images can be increased by up to 50%.

11. Hue: This is the maximum hue shift used for data augmentation during training. In this case, hue=0.1, which means that the hue of the images can be shifted by up to 0.1.

12. Learning rate: This is the learning rate used in the optimization algorithm. Here, learning_rate=0.00261.

13. Burn in: This is the number of iterations during which the learning rate will be increased linearly from zero to the specified learning rate. In this case, burn_in=1000.

14. Maximum number of batches: This is the maximum number of batches that will be processed during training. In this case, max_batches=2000200.

15. Policy: This is the learning rate policy used during training. In this case, policy=steps, which means that the learning rate will be decreased according to a step function.

16. Steps: These are the iteration steps at which the learning rate will be decreased. In this case, steps=1600000, 1800000, which means that the learning rate will be decreased at iteration 1600000 and 1800000.

17. Scales: These are the factors by which the learning rate will be decreased at each step. In this case, scales=0.1, 0.1, which means that the learning rate will be decreased by a factor of 0.1 at each step.

5. CONCLUSION

The project successfully demonstrates the two approaches which serve different purposes, the YOLOv3 and thermal sensor approach is focused on detecting animal movement, while the CNN model approach is focused on classifying images. The YOLOv3 and thermal sensor approach is more effective in low light conditions and can detect animal movements that traditional cameras would miss. The CNN model approach is more suitable for tasks such as image classification and object recognition, where the focus is on identifying specific objects within an image. Both approaches have their strengths and weaknesses, and the choice of approach will depend on the specific requirements of the task at hand. The two approaches can be complementary, and in some cases, it may be beneficial to combine them to achieve more comprehensive results.

Thermal sensors and YOLO v3 can be used to detect animal movement near border areas, helping border security forces to identify potential threats and take appropriate action and also can be used in search and rescue operations to locate missing or injured animals, as well as human beings, in areas with low visibility or difficult terrain.

6. FUTURE SCOPE

The future scope of this technology is vast, as advancements in AI and machine learning will further enhance the accuracy and efficiency of animal movement detection. It can be integrated with other technologies like drones and GPS to provide real-time tracking and monitoring of animal movements. With further research and development, this technology can have a significant impact on various fields, including wildlife conservation, agriculture, and security.

REFERENCES

[1] J. Ma, J. Fan and L. Tian, "An Object Detection Method for Thermal Images Based on YOLOv3 Algorithm," Journal of Physics: Conference Series, vol. 1549, issue 1, 2020.
[2] Yu Oishi, Hiroyuki Oguma, Ayako Tamura, Ryosuke Nakamura and Tsuneo Matsunaga, "Animal Detection Using Thermal Images and Its Required Observation Conditions," MDPI, 2018. [Online].Available: https://www.mdpi.com/2072-4292/10/7/1050
[3] P. Christianen, Kim Arild Steen, Rasmus Nyholm Jorgensen and Henrik Karstoft "Automated Detection and Recognition of Wildlife Using Thermal Cameras," MDPI, 2014. [Online]. Available: https://www.mdpi.com/1424-8220/14/8/13778

[4] M. C. Burl and A. R. Liggins, "Thermal video processing for security applications," Proceedings of SPIE - The International Society for Optical Engineering, vol. 7349, 2009.

[5] Y. Gao, J. Zhang and J. Liu, "YOLOv3-based Automatic Vehicle Detection in Thermal Imagery," Infrared Physics & Technology, vol. 110, p. 103473, 2020.
[6] S. Sahana and S. K. Sahoo, "An Efficient YOLOv3 Based Vehicle Detection and Counting System for Intelligent Transportation Systems," International Journal of Intelligent Transportation Systems Research, vol. 19, issue 3, 2021.
[7] M. J. Genge, A. R. Mahajan, G. R. Naik, S. A. Patil, V. M. Wadhai, "Animal classification using thermal imagery and deep learning," Journal of Applied Remote Sensing, vol. 12, issue 3, 2018.
[8] J. C. R. Menezes, S. Mitra, S. K. Sahoo, "Animal classification in thermal imagery using a convolutional neural network," Journal of Electronic Imaging, vol. 29, no. 1, p. 013026, 2020.