Grenze International Journal of Engineering and Technology, Jan Issue

Real-Time Snake Detection with Alert Systems using Deep Learning

Jyothika K Raju[1], Maitri V J[2], Dr Hemavathy R[3], Dr Ramakanth Kumar P[4] and Dr T Shankar[5]
[1-4]Computer Science & Engineering, R.V.College of Engineering, Bengaluru, India
Email:jyothikakraju.cs20@rvce.edu.in, maitrivj.cs21@rvce.edu.in, hemavathyr@rvce.edu.in, ramakanthkp@rvce.edu.in
[5]Mentor, CCTV Research Center, R.V.College of Engineering, Bengaluru, India
Email: tshankar.68@gmail.com

***Abstract*— Snake interactions can pose serious dangers to public safety and biodiversity preservation in both human-populated areas and animal environments. The creation of a Snake Detection From Video Footage and Alert System utilizing AI and ML technologies has attracted interest as a solution to this problem. The system attempts to reliably identify snakes and issue timely alerts, reducing human-snake conflicts and improving coexistence. It does this by utilizing real-time video processing, object detection techniques like YOLO, and intelligent decision-making. Current studies and projects in this area have shown that it is possible to identify snakes in video footage using machine learning techniques. Many researchers have used deep learning techniques, such as YOLO, to identify snakes in real time and with accuracy. The ability of contemporary systems to adapt to various situations, nevertheless, is still a study area that needs further attention. Real-time processing requirements may also encounter difficulties in contexts with limited resources. This gap will be filled by the proposed project's adaptive Real-Time Snake Detection with Alert Systems Using Deep Learning. Real-time snake recognition and alarm creation from a model's video feed are among the goals. With the help of these goals, the project hopes to provide a thorough and useful solution that will aid in the protection of both humans and wildlife in areas where snakes are a common occurrence.**

***Index Terms*— YOLOv8, object detection, snake detection, alert system, safety, CVAT**

I. Introduction

Due to the wide range of snake species, their aptitude for camouflage, and the complexity of natural habitats, snake detection from video footage is a difficult undertaking. However, recent developments in machine learning have made it possible to create cutting-edge snake identification algorithms that are highly accurate.
Convolutional neural networks (CNNs) are one of the methods used most frequently for snake detection. For image identification applications, CNNs are a class of deep learning algorithms that works well. They may be taught the characteristics that tell snakes apart from other objects using vast databases of tagged snake photographs. A CNN can be used to identify snakes on live video footage once it has been taught. Every frame of the video will be scanned by CNN, which will look for any areas with snake-like characteristics. The presence of a snake in these areas is then confirmed through further analysis. As new machine learning algorithms are created, snake detection technology continues to advance. The development of highly precise snake detection systems, however, has been made possible by the advancements made in recent years and can be utilized in a number of applications, including wildlife monitoring, pest control, and safety for citizens.

Grenze ID: 01.GIJET.10.1.510_2

There are many motivations for snake detection and sending alerts which include:
Public safety: Snakes can be harmful, especially to young children and pets. People can be warned of snakes in their vicinity so they can take precautions to stay away from them using snake detection devices.
Wildlife monitoring: Monitoring snake populations and their movements is possible with the use of snake detection equipment. This knowledge can be utilized to safeguard snake populations and better understand the ecology and behavior of snakes.
Pest control: Snake detection systems can be used for pest control to identify and monitor snakes that are a problem, including venomous snakes or snakes that harm crops. Effective pest management measures can be targeted with the help of this information.
Scientific research: Data regarding snake behavior and ecology can be gathered using snake detection systems. Researchers could apply this knowledge to generate better conservation plans and a better understanding of snakes.

II. LITERATURE SURVEY

The architecture outlined in the paper, "Reptile Recognition based on Convolutional Neural Network" by Condro Kartiko et al. attempts to recognize the reptile without using backdrop removal procedures. The first layer will be supplied with a 64x64 input image, which will then be processed via the layers and mapped to one of the 14 reptile groups. The findings of the study improved our comprehension of the CNN technique's process for identifying and classifying reptiles, and future work will involve increasing the amount of input data required by the model to operate more properly[1].
The paper “Image Enhancement and Object Recognition for Night Vision Surveillance” demonstrates that, when compared to overnight, the performance of a surveillance system using a standard camera during the day is noticeably superior. The underlying issue with nighttime surveillance is that things shot by traditional cameras have minimal contrast with their surroundings due to a lack of ambient visible light. As a result, the image is taken in low-light circumstances with an infrared camera and then enhanced using various enhancing algorithms based on the spatial domain to provide an image with greater contrast. To conduct the classification, a fully connected layer of neurons is used after the convolutional neural network [2].
The study “A Survey on Snake Species Identification using Image Processing Technique” demonstrates that, despite not being required, snakes play an important role in the ecology. Deforestation is destroying the snakes' habitat. In some regions of the world, snake identification is critical for treatment planning, even though it is not always possible. To solve this issue and put a stop to it, a proposed method that can consistently identify snake species is introduced [3].
The paper “A Review on Object Detection and Tracking in Video Surveillance” presents that video surveillance research is now ongoing in the subject of security. An intelligent video surveillance system is expected to locate and track the object. This paper investigates effective video surveillance system tactics. The functions of a video surveillance system are separated into three categories: object detection, classification, and tracking. Objects in photos can be found utilizing techniques such as Spatio-Temporal Filtering, Optical Flow, and Background Subtraction. The techniques are examined in terms of accuracy rate and temporal complexity [4].
The paper “A comparative study on image-based snake identification using machine learning” reveals a variety of cutting-edge machine learning techniques, ranging from holistic to neural network algorithms, have been evaluated for the first time in this study to determine how accurate they are. In this study, a dimension reduction strategy principle component analysis (PCA) and linear discriminant analysis (LDA)] is integrated with the holistic approaches [k-nearest neighbors (kNN), support vector machine (SVM), and logistic regression (LR)] as well as the feature extractor. The classifier integrated with PCA does not yield more than 50% accuracy in holistic techniques (kNN, SVM, LR). Using LDA to extract the important features, on the other hand, significantly improves the classifier's performance. MobileNetV2 is given in this paper as a powerful deep convolutional neural network solution for categorizing snake photos that can be deployed to mobile devices. This discovery paves the way for the creation of mobile applications for detecting snake photographs [5].
According to the study "A Paper on Night Vision Technology," the various "Night Vision" approaches are considered inventions because they enable us to see in complete darkness and adjust our vision in weak light. This invention is a synthesis of several distinct approaches, each with its own set of pros and downsides. Illumination, thermal imaging, and low-light imaging are three of the most well-known methodologies. A variety of night vision devices (NVDs) generate images in light levels fading into utter darkness, as well as describe several applications for night vision technology to handle diverse challenges caused by low light settings[6].

III. AI AND INNOVATION

The snake detection and alert system can benefit greatly from the addition of AI and ML components. This enhances the solution's robustness, effectiveness, and domain-specific adaptability. First off, by employing machine learning, the system can properly identify different snake species by learning from a wide dataset of video footage. In areas where different species of snake coexist, adaptation is essential. Second, AI-driven real-time video processing makes sure that the system can evaluate incoming frames quickly, sending timely alerts to users, and so improving human safety in snake-prone areas. Third, the use of computer vision-based object detection techniques enables the system to properly identify and localize snakes inside video frames, decreasing false positives and lowering the danger of alerting fatigue. Furthermore, the ML-based decision-making elements support validation of detections, ensuring high accuracy and minimizing pointless alarms. A key component of snake detection from video footage and alarm systems is computer vision. For the system to identify snakes and issue alerts, it must be able to "see" and decipher visual data from the video frames. Snakes are located and identified within video frames using computer vision techniques, notably object detection algorithms. Popular options for this task include object detection algorithms like YOLO (You Only Look Once) or SSD (Single Shot Multibox Detector). These methods are ideal for the system's real-time video processing needs since they can locate snakes with bounding boxes in real-time or almost real-time. Since the technology is real-time, it helps with proactive preventive steps like alerting emergency personnel or medical institutions in the case of snake encounters. Future updates and improvements are also possible thanks to the AI model's adaptability, which guarantees that the system will continue to develop and function well over time.

IV. METHODOLOGY

The methodology starts with data collection, which entails gathering and annotating video footage of several snake species. The dataset's diversity is increased using preprocessing techniques like frame extraction and data augmentation. The object detection engine is the YOLOv8 method, which was developed using annotated data to identify snakes in video frames. The methodology used in this study can summarized as:

1. *Setting up the dataset*: Create a dataset with annotated images of snakes. Draw bounding boxes around the important objects in the photos. In total, around 650 images were chosen.
2. *Preprocessing*: Resize the photos to a scale that is suitable for the YOLO model as part of the preprocessing step. Set the pixel values to a suitable range, such as 0 to 1. Create the dataset, make the training and testing subsets.
3. *Model selection*: Choose a suitable YOLO model variant (such as YOLOv3, YOLOv4) based on the requirements for accuracy and speed. The YOLOv8 model has been selected.
4. *Model Training*: Use the chosen architecture to initialize the YOLO model, and load pre-trained weights if necessary. Apply techniques like stochastic gradient descent (SGD) or the Adam optimizer to the labeled dataset to train the model. During training, the model learns to distinguish between background elements and snakes by improving the loss function, which measures the discrepancy between expected and actual bounding boxes.
5. *Hyperparameter Tuning*: Change the YOLO algorithm's hyperparameters, such as learning rate, batch size, number of training iterations, and anchor box sizes, to improve the model's performance. Iterative experimentation and validation may be required on the test dataset.100 epochs have been used.
6. *Model Evaluation*: Using the test dataset, the trained model's performance is assessed for accuracy, recall, and mean average precision (mAP).
7. *Real-Time Processing*: Implement a pipeline to process video frames in real-time while applying the snake classification and object identification models to each frame.
8. *Alert Generation*: Design a system to generate alerts when snakes are found and notify the appropriate people.

V. INSIGHT INTO YOLOV8 ALGORITHM

A deep learning technique called YOLOv8 uses convolutional neural networks (CNNs) to find snakes in photos. They can be trained on a dataset of labeled photos to learn the characteristics that set different things apart. An algorithmic technique for locating areas of interest in an image is called region suggestion networks. To ascertain if an object is present, these zones are subsequently subjected to additional analysis. It employs a multi-scale training methodology. This indicates that different-sized photos were used to train the algorithm. As a result, the algorithm has the ability to recognize objects at various scales, which enhances its accuracy. To increase the algorithm's precision, it uses data augmentation. A technique called data augmentation includes generating new photos from existing ones in order to artificially increase the dataset's size. This lessens the likelihood that the algorithm will overfit to the training set of data.

The YOLOv8 model makes use of a CNN with 21 layers divided into 5 stages. Large object detection occurs in the first step, while small object detection occurs in the final stage. A collection of predetermined boundary boxes called anchor boxes is also used by the model to aid with item identification. The YOLOv8 algorithm initially creates a grid of cells out of the image. The system then forecasts the likelihood of a snake appearing in each cell. Additionally, if a snake is present, the system estimates its bounding box. To produce a final detection of the snake, the predictions from each cell are then integrated. To remove duplicate or unnecessary detections, the algorithm employs a method known as non-maximum suppression (NMS). The program has been proven to recognize snakes in photos pretty effectively. On a dataset of labeled snake photos, the system was able to attain an accuracy of more than 95%.

The YOLOv8 architecture is composed of the following modules:

- *Backbone*: The backbone is a collection of convolutional layers that extract relevant characteristics from the input image. YOLOv8's backbone is based on the ResNet50 architecture however it has been altered to include a few extra layers.
- *SPPF*: The Spatial Pyramid Pooling (SPP) module is used to enhance the network's receptive field and improve object identification accuracy at different scales.
- *C2f*: The C2f module was introduced in YOLOv8 as a new module. It combines high-level data from the backbone with contextual information to increase detection accuracy.
- *Detection heads*: The detection heads are in charge of estimating the bounding boxes and class probabilities for objects in an image. YOLOv8 has five detection heads, each of which detects objects at a different scale.
- *Loss function*: The loss function is used to train the YOLOv8 model. The loss function is a hybrid of the Intersection over Union (IoU) loss and the categorization loss.

VI. High Level Design

A system architecture is the model that defines the structure, behavior, and more views of a system

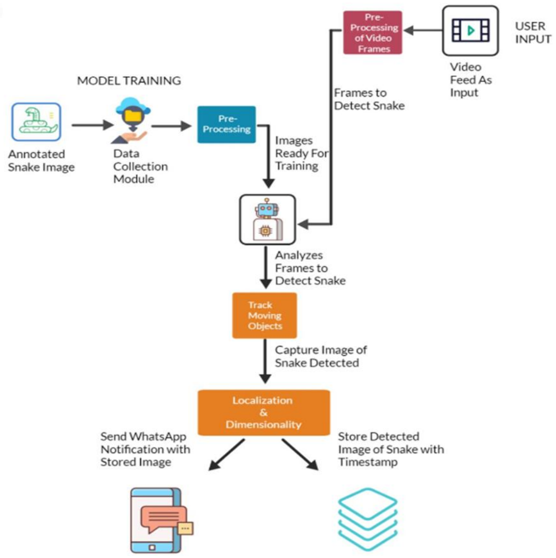

Figure 1. Architecture Diagram

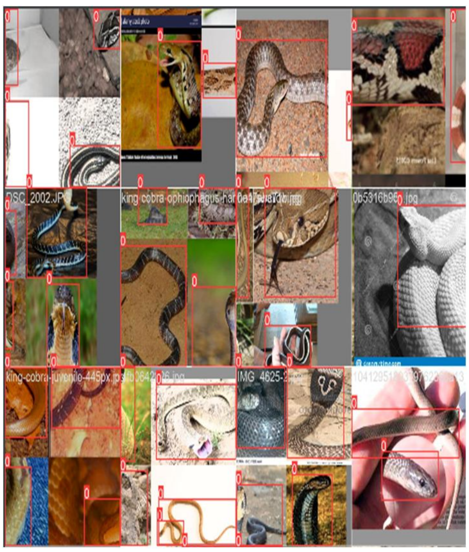

Figure 2. Snapshot of the annotated dataset

Fig. 1 displays the system architecture for enhanced object detection in snakes, which includes numerous critical components that work together to provide the needed functionality.

Overview of the system architecture includes:

- *Image/Video Input*: The system obtains image or video input including snakes for item detection.
- *CVAT Annotation*: The input data will be annotated utilizing CVAT, which allows annotators to indicate regions of interest (ROIs) for snakes in videos or image frames.
- *YOLOv8 Model Training*: The annotated data is utilized to train the YOLOv8 deep learning model. Focusing on the annotated ROIs, YOLOv8 is taught to recognize snakes.

- *Real-time Object Detection*: After the model has been trained, it is deployed to conduct real-time object detection on new input data. The system detects capsules and pills by applying the learned YOLOv8 model to the input pictures or video frames.
- *Snake Detection Alert*: If a snake is discovered, the system sends out an alert. Python code is used to send a notification to the appropriate persons, such as a WhatsApp message.

The experimental dataset contains 650 photos of snakes of various species. To increase their usefulness, these images were heavily labeled with the Computer Vision Annotation Tool (CVAT). The annotations focused on identifying the snakes that were there. The photographs were tagged and then converted to the YOLO format for additional in-depth analysis and application. Fig. 2 depicts the annotated dataset, where each item has been outlined with bounding boxes. These boxes act as an indication as to whether the object falls under the "Snake" class.

VIII. RESULTS

The results of the Real-Time Snake Detection with Alert Systems Using Deep Learning are showcased below:

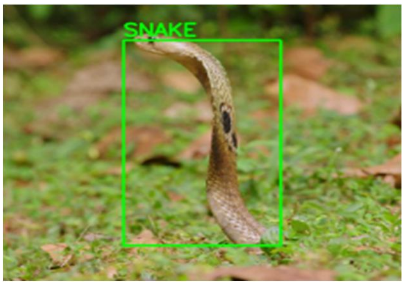

Fig 3. Snake Detected from the video footage

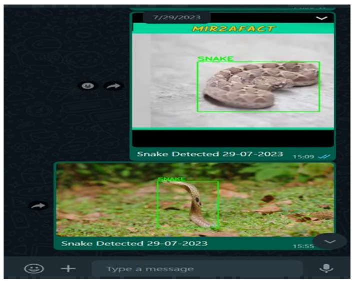

Fig 4. Whatsapp alert

Fig 3 indicates the frame in which the snake was detected, displaying it to the viewer and drawing a bounding box across the area of the image where the snake was discovered. This project has live monitoring enabled. Fig 4 demonstrates the way to utilize the Pywhatkit Python library. It has aided in the development of WhatsApp alerts. Pywhatkit includes a number of functions, such as the ability to transmit messages, photos, and URLs via Whatsapp. It depicts the alert that was sent to the appropriate in-charge officer. The outcome can be evaluated using performance and evaluation metrics.

A. Performance Metrics

A Confusion Matrix summarizing classification outcomes, training metrics including box, classification, and deformable convolutional losses, as well as equivalents for validation, are among the performance metrics for YOLO-based object identification. While Precision (B) focuses on bounding box accuracy, the mean Average Precision (mAP50) examines overall detection performance with a 0.5 IoU threshold. The confusion matrix is set up as a table that lists the various kinds of results: The accuracy with which the snake was accurately identified from the video footage is shown in the image below. The maximum accuracy is represented by the darker color. As can be seen from the figure, the training model's accuracy is high since the box on the x-axis corresponds to a snake, and the snake's y-axis has a dark color, ensuring that the prediction is usually accurate.

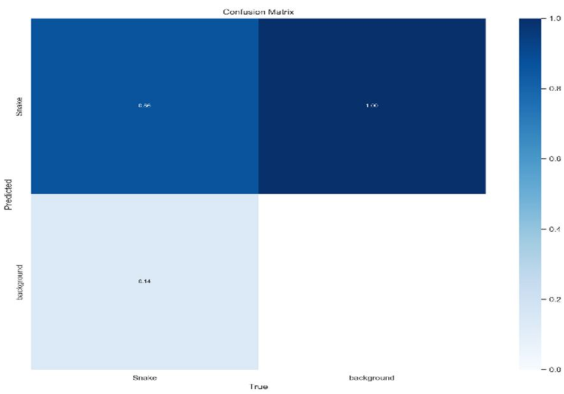

Fig 5: Confusion matrix for snake detection

Fig 5 illustrates the confusion matrix for the snake detection. Each row corresponds to a true class or ground truth. Each column corresponds to a predicted class by the model. The number of instances that belong to that specific combination of true class and predicted class is represented by the intersection of each row and column. The True Positives (TP), False Negatives (FN), False Positives (FP), and True Negatives (TN) results of the model's predictions are quantified in the confusion matrix. True Positives (TP), the number of instances which are correctly predicted as "snakes." False Negatives (FN), the number of instances incorrectly predicted as "snakes". False Positives (FP), the number of instances incorrectly predicted as "not detected" when there are actually "snakes" in the footage. True Negatives (TN), the number of instances correctly predicted as "not detected" when the video does not contain snakes. The values in the confusion matrix can be used to determine the model's accuracy, precision, recall, F1-score, and other performance metrics

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \qquad (2)$$

The accuracy of a classification model is measured using Equation (2) by the ratio of correctly predicted instances (including true positives and true negatives) to the total number of predictions.

$$Precision = TP / (TP + FP) \qquad (3)$$

Precision is the proportion of cases out of all instances that the model has identified as positive that are accurately classified as positive (true positives), according to Equation (3).
Evaluation Metrics: The F1 Confidence Curve and PR Curve are used as evaluation measures.

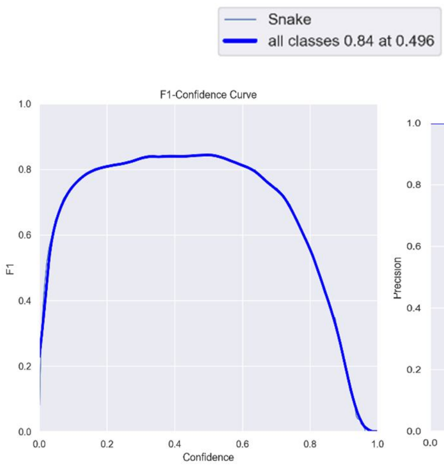

Fig 6: F1 Confidence Curve

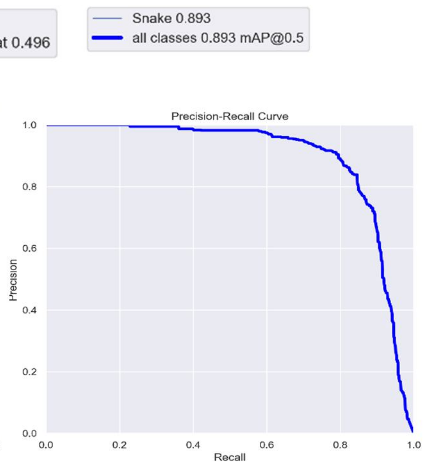

Fig 7: Precision - Recall Curve

The F1 Confidence Curve is a graphical depiction that shows how the F1 score evolves with various confidence thresholds and is used to evaluate the effectiveness of classification models.The F1 Confidence Curve for the YOLO-trained model is shown in Fig 6. The solid line represents all classes, whereas the slender blue line shows a prediction of snakes. This graph illustrates how the model's F1 score changes for different classes and levels of confidence. The anticipated numbers in the figure are more accurate since the thin line overlaps with the thick line.The F1-score, a single scalar statistic, strikes a balance between recall and precision. It gauges how well the model performs when the ratio of positive to negative samples is unbalanced. It is the precision and recall harmonic mean.

$$F1 = 2 * (Precision * Recall) / (Precision + Recall) \qquad (3)$$

A visual depiction used to evaluate the effectiveness of a classification model is called the Precision-Recall (PR) Curve. The PR curve makes it simpler to comprehend how recall and precision are traded off when the classifier's decision threshold is altered. Fig 7 displays the Precision-Recall (PR) Curve for the effectiveness of the YOLO-trained model in detecting and recognizing snakes. This graph illustrates how the model's precision and recall vary across different confidence levels, providing information about the model's ability to properly identify snakes. The ratio of true positive predictions to all anticipated positives (true positives plus false positives) is known as precision. The ratio of correctly predicted positive outcomes to all actual positive outcomes (true positives plus false negatives) is known as recall.

VIII. CONCLUSION

The Real-Time Snake Detection with Alert Systems Using Deep Learning is an innovative and advanced solution designed to enhance human safety. Leveraging cutting-edge AI and ML technologies, the system efficiently analyzes real-time video streams to detect snakes accurately. The YOLOv8 algorithm forms the core of the system, enabling rapid object detection and localization of snakes within video frames. Through extensive training on diverse datasets, the model achieves high accuracy in identifying different snake species, making it applicable in various environments. The system's real-time video processing capabilities enable timely alerts to be triggered, notifying users of potential snake encounters and prompting them to take precautionary measures. With future enhancements aimed at multi-species detection, behavior analysis, and adaptive learning, this project lays the foundation for a comprehensive and intelligent snake detection system, contributing significantly to human safety and wildlife conservation efforts.

REFERENCES

[1] Kartiko, C., Prasetiadi, A., & Usada, E. (2020). Reptile Recognition based Convolutional Neural Network. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(3S), ISSN: 2278-3075, January 2020.

[2] Bhandari, A., Kafle, A., Dhakal, P., & Joshi, P. R. (2018). Image Enhancement and Object Recognition for Night Vision Surveillance.

[3] Joshi, P., Sarpale, D., Sapkal, R., & Rajput, A. (2018). A Survey on Snake Species Identification using Image Processing Technique. International Journal of Computer Applications, 181(29), ISSN: 0975-8887, November 2018.

[4] Angadi, S., & Nandyal, S. (2020). A Review on Object Detection and Tracking in Video Surveillance. International Journal of Advanced Research in Engineering and Technology, 11(9), 1033-1042.

[5] Mahdi Rajabizadeh and Mansoor Rezghi. A comparative study on image-based snake identification using machine learning, Published on 27th September 2021

[6] Tamilchelvan, P., & Greater Noida. (2019). Night vision technology: A review. Jetir.org.

[7] Review on Infrared Imaging Technology by Fujjin Hou, Yan Zhang, Yong Zhou, Mei Zhang, Bin Lv, and Jianqing Wu Published on 6th September 2022

[8] Amit, Yali, Pedro Felzenszwalb, and Ross Girshick. "Object detection." Computer Vision: A Reference Guide (2020): 1-9.

[9] Palanivel, N., Keerthana, V., Monisha, N., & Sandhiya, S. (2018). OBJECT DETECTION AND RECOGNITION IN DARK USING YOLO. Journal of Data Acquisition and Processing.

[10] Agrawal, P., & Khunteta, A. (2017). A Literature Survey on Applications of Image Processing for Video Surveillance. International Journal of Innovative Science and Research Technology, 2(1), January.