IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CCTV SURVEILLANCE USINGDEEP LEARNING

[1]Natuva Sruthi, [2]Navyashree M, [3]Divya S, [4]Spoorthi S Nadiger,[5]Asha MS

[1,2,3,4] Student, Department of CSE, Dayananda Sagar Academy of Technology and Management, Bangalore, India.

[5] Professor, Department of CSE, Dayananda Sagar Academy of Technology and Management, Bangalore, India.

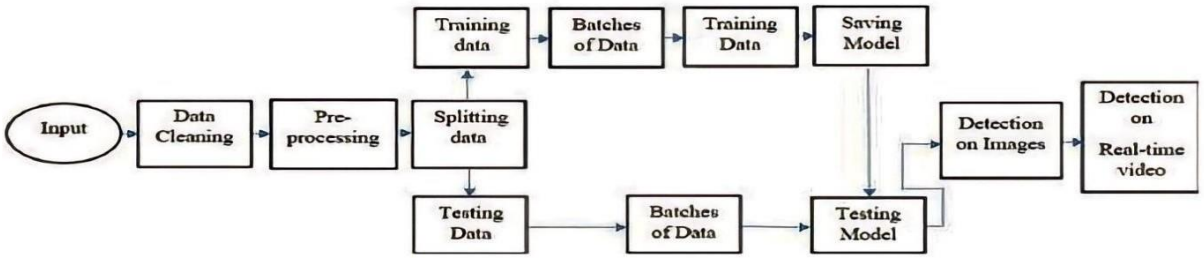
Abstract: Safety and security is a major concern in this day and age. The number of surveillance cameras, often known as closed circuit television (CCTV), has increased significantly globally during the previous several decades.. People use security mechanisms to protect their assets, whether at home or in a business.we focuses on automated object recognition and presents a deep learning approach for autonomously processing security camera pictures. This project's objective was to evaluate and test the most sophisticated algorithms in order to investigate a potential approach for integrating object identification techniques into commercial video surveillance systems. In this post, we suggest offering an intelligent CCTV security system with picture and object detection. For surveillance and live streaming reasons, several USB cameras are mounted at various locations..Previously used in CCTV video analysis, Individuals, pets, and automobiles may all be recognised using object detection algorithms. These algorithms may be further developed to find people carrying weapons in public or prohibited areas, such as guns or sharp objects. Mufflers, helmets,knife and jackets are detected by sending alert messages.Various cameras' live feeds can be seen on a computer or smartphone. We assess the effectiveness of object detection algorithms in terms of speed and accuracy.

Index Terms **-real-time monitoring, facial recognition using deep learning, and security, Faster RCNN, SSD,CCTV,face recognition,Object Detection.**

I. INTRODUCTION

Today, it might be challenging to prevent a third party from accessing extremely confidential information. Data security is crucial because more people are now accessing data from highly secure locations. Due to Pan-Tilt-Zoom operations, poor camera quality, and a variety of working environments, real-world CCTV footage frequently presents more difficulties in object tracking.Moving backgrounds, motion blur, and drastic scale changes are the most pertinent challenges. CCTV cameras have been installed everywhere for security reasons.

Python is the language needed for this project. it has a more complete library, simple syntax, and a solid grasp of the fundamentals, and it can be implemented in any open-source language with CV support.CCTV surveillance is created with the aid of Python's object recognition capabilities and OpenCV MediaPipe. The Python libraries used for image processing the most are OpenCV and MediaPipe. These libraries are used to monitor an object's motion in front of a webcam. In front of the webcam,object detection is possible using this tracking feature.This is both intriguing and difficult because it uses real-time webcam data that is generated by following the motion of unusual activities and object detection.



An important computer vision challenge is the recognition of instances of visual objects of a specific class (such as humans, animals, or automobiles) in digital photographs. Convolutional neural networks, which offer cutting-edge performance in object identification, are being employed more frequently in an effort to create a tracking technique that is more effective. To increase their detection rate in difficult video surveillance scenes, this work investigates the use of heterogeneous training data and data augmentation. In addition,the spatial transformation parameters of the proposed objects are used for modeling, prediction, and autotuning of the detector to improve performance.Real-world CCTV videos and publicly available datasets were used to test the proposed methods. Security is used to identify theft, detect violence, prevent unauthorized entry,and track criminal activity in a given area. As a result, security plays an important role in any unusual activity, so security measures should be applied in a highly secure area.

In the past, detecting thefts and other activities was detected using CCTV footage, but this is a time-consuming and labor intensive process. CCTV is used with deep learning concepts to make it easier to use to overcome this current approach. Deep learning has significantly influenced how the world has adapted to artificial intelligence during the last few years. Some of the well- known object identification methods include the region-based convolutional neural network (RCNN), faster RCNN, Single-Shot Detector (SSD), and You Only Look Once (YOLO). While YOLO works best when speed takes precedence over precision Faster RCNN and SSD have the highest accuracy of these. Faster-RCNN also shares calculations between all proposals instead of performing separate calculations for each. This algorithm detects objects efficiently without sacrificing performance. Unstructured data can be learned using deep learning concepts and can handle large data sets. Deep learning will use datasets to train and generate a finite number of datasets. Large amounts of data can be trained using this concept and the output datasets are also accurate. Another desirable statistic is the accuracy rate. CCTV is used to recognize facial features in the dataset using deep learning. In this way, all CCTV datasets will be identified and also able to distinguish between authorized and unauthorized persons.When an odd incident is identified, the video begins to record. The disadvantages of current technology compared to surveillance systems will be enumerated in this study.

II. LITERATURE SURVEY

[1] Classification of Objects in Video Records using Neural Network Framework Abhiraj Biswas, Arka Prava Jana, Mohana, Sai Tejas S. This paper principally emphasises detection and categorization of several class objects and a single class.The recognition of walkers, anomalous behaviour, and objects are some of these uses. here,the process for object identification and classification strives to categorise objects based on their characteristics and features. Single shot detection frameworks have emerged mostly from sliding-window-based object identification approaches. The Convolutional Neural Network (CNN), which lowered error rates dramatically but at the cost of speed and calculation time, is used in a variety of applications, including facial recognition.To find the objects, the Region based Convolutional Network (R-CNN) employs the Selective Search technique.The Fast RCNN and Faster R-CNN progeny of R-CNN improved the slowness of CNN and RCNN by making the training process end-to-end. Neural networks like CNN and R-CNN are excellent in multiclass object recognition and categorization.For bounding box object detection the Faster RCNN is advised, while the FCN is used to produce the masks that surround each mask..In essence, it is a method for bounding box object recognition that establishes many areas or Regions of Interest (RoI).

[2] A Review of Artificial Intelligence Methods for Data Science and Data Analytics: Applications and Research Challenges By Krishna C V, Rohit H R, Mohana primarily focuses on automating all human tasks that currently call for human intelligence. In our project, images breaks down the data into smaller chunks to help users understand its trends and behaviours. making use of technology The C++, Python, and other libraries are supported by the open source Deeplearning framework. The distributed frame work library is supported by GPU.The majority of deep learning models are supported.We used TensorFlow to perform numerical calculations. Data flow graphs are used to represent computations when they can. It runs faster with Python writing. Here, it is simple to fool image classifiers by using undesirable images, moving images, or any other disturbed images. When it comes to generating accurate results, deep learning is regarded to offer greater benefits than traditional machine learning approaches.

[3] Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications Harsh Jain, Aditya Vikram, Mohana, Ankit Kashyap, Ayush Jain In our paper,By identifying erratic, unexpected, unpredictable, rare events or things that are not thought to be regularly repeating events or regular items in a pattern or items contained in a dataset, the weapon detection approach differs from earlier patterns. In order to locate samples of various kinds of objects, an object detection uses feature extraction and learning approaches or models. Implementation concepts emphasise precise gun detection and classification.The method of object recognition and tracking is developed, trained, and given as a thorough dataset into the algorithm for object detection. A suitable detection algorithm (SSD or fast RCNN) is utilised for the detection of firearms depending on the application.To address the detection issue, the approach employs a number of machine learning models, such as Single Shot Detection and Region Convolutional Neural Network (RCNN) (SSD). The remaining 80% of photos are utilised for training, while just 20% are used for testing. To enable group processing, Tensorflow converts the pictures into the tf record format.

[4]A Survey of Modern Deep Learning based Object Detection Models developed by Syed Sahil Abbas Zaidi, Mohammad Samar Ansari, Asra Aslam, Nadia Kanwal, Mamoona Asghar, and Brian Lee .This study suggests a novel way for locating and identifying every instance of an object, such as mufflers, people, firearms, knives, and helmets, based on data from the actual world.It largely concentrated on image detection. Processing pictures and doing tasks like categorization, There have been significant issues with computer vision, including segmentation, motion estimates, scene interpretation, etc. Its recent exponential rise has been accompanied by the rapid creation of new tools and processes. It is frequently viewed as a supervised learning issue. Here, the detector may create a bounding box around the pictures that were recognised and identify each occurrence of the item classes.

[5] Performance Analysis of Object Detection and Tracking Algorithms for Traffic Surveillance Applications using Neural Networks Naman Jain, Shreesha Yerragolla, Tanuja Guha, Mohana.This paper describes a Utilizing the ideas of convolution layers, a single object identification approach has been developed. The input layer, at least one hidden layer, and an output layer are a few of the layers that make up a neural network. You Only Look Once has been used to detect several items (YOLOv3). Using a single deep convolution neural network, the input is split into a grid of cells, where each cell predicts a boundary box and categorises the input item. It is necessary to identify the object first, track it from frame to frame, and then assess the tracking data. The algorithm's initial effort is to detect any potential objects in the supplied picture. After that, it makes an attempt to identify which object the recognised object most closely resembles before categorising it. The programme then returns the object it has discovered along with how confidently it believes it fits into a certain category.

[6] Background-modelling techniques for foreground detection and Tracking using Gaussian Mixture Model Meghana R K, Yojan Chitkara, Apoorva S, Mohana Background Modelling and Foreground detection .This study used a clever method to create the foreground is detected by subtracting the background region from a model of the backdrop created by extrapolating data from frames and comparing it to each subsequent frame. Using the results of the GMM background subtraction, the feature extraction algorithm classifies the objects and separates them based on that classification. It currently addresses problems like shadowing, occlusions, and lighting changes. One of the most crucial aspects of artificial intelligence is image/video processing. Here, cameras are used to collect photos and video, which are then edited to create the desired effects.Deep processing is required to extract ever-increasing volumes of information because a camera's sensors cannot record all of the information in a scene. Video data is taken from a dataset. depending on the selected background model, does the background subtraction. Use the chosen feature to carry out further object detection. The data is analysed using deep learning techniques,

[7] Application of Firebase in Android App Development-Developed by Chunnu Khawas&Pritam Shah.Firebase is a relatively new technology for handling large amount of unstructured data. It is very fast as compared to RDBMS. This essay aids in learning how to incorporate Firebase into an Android application in accordance with developer needs. Additionally, by eliminating the need for PHP as a third party language to interface with the database, this makes android apps quicker and more effective.It can link between platforms and not just be utilised for Android.The system's backend functions as a database for storing data. Firebase offers features like a backend and real-time database. The application developer is given access to an API that enables the synchronisation of application data between clients and storage on Firebase's cloud. For the Firebase apps, it enables quick and safe file transmission independent of network quality. It is supported by Google Cloud Storage, a service that offers affordable object storage. It may be used by the developer to store user-generated material such as pictures, music, and video.In this part, we've covered topics including authentication, databases, and storage. Additionally, by eliminating the need for PHP as a third party language to interface with the database, this makes android apps quicker and more effective. It can link between platforms and not just be utilised for Android.

[8] Smart Surveillance System using Deep Learning Developed by Dayana R, Suganya M .It should also be highlighted that state-of-the-art accuracy solutions are generally constrained by their reliance on powerful GPUs and huge datasets, indicating that there is still sufficient reason to concentrate analysis attention on numerous old handwritten feature representations. The detection and identification of face characteristics of people using deep learning techniques is covered in this study. The fundamentals of deep learning that are included in this article begin with object detection, action detection, and identification. This initiative is extremely beneficial to businesses and highly sensitive places. The authorities can only be admitted to the appropriate centres with high accuracy recognition.

[9] Visual detection of knives in security applications using Active Appearance Models Developed by Andrzej Glowacz, Marcin Kmieć & Andrzej Dziech. The primary objective of this work is the automated identification of knives in photographs. Due to its accessibility and potential for use as weapons,carrying a knife in public is frequently or categorically forbidden. Security personnel prioritise identifying knives because of this. For instance,using CCTV to monitor the general population is a real-world implementation of the idea of software-based knife detection. An alert is sounded in the event that a knife is found, allowing the human operator to concentrate on the location in issue right away and confirm or deny the discovery.

[10] Anomaly Detection in Videos for Video Surveillance Applications usingNeural Networks Developed by Ruben J Franklin, Mohana, Vidyashree Dabbagol. Anomaly detection is a method for swiftly distinguishing unusual patterns from other patterns, sometimes referred to as outliers. Surveillance footage reveals a number of plausible oddities. The three levels that make up the complete process of anomaly detection in video surveillance are activity detection, image processing, and video labelers. In light of real-time conditions, anomaly detection in videos for video surveillance applications generates correct findings. In this investigation, we were able to detect anomalies in both photos and videos with an accuracy of 98.5%..The usage of real-time video monitoring is beneficial. These needs have led to widespread installation of cameras and the development of video surveillance systems that can grasp their surroundings and identify unusual activities..The major aspect recognises an unusual occurrence when it happens and automatically warns the operator or users. The use of video surveillance greatly improves the management of private life and public space in terms of safety and security. Unsupervised, semi-supervised, and supervised anomaly detection are the three different categories of anomaly detection approaches.Cameras are employed in the procedure for detecting anomalies to gather information about various occurrences that indicate the behaviour of an abnormality in a monitored environment. Anomaly objects are found in both real-time video and images from the COCO dataset. CNN is used to extract frames from a movie and detect objects by extracting characteristics from the necessary region of interest. Normally, the patches that are taken into account,such as motion and appearance, are used to assess and compare the properties. Feature representation, a crucial stage in anomaly detection, is the stage's output. Finding separate occurrences, data points, and observations that differ from the dataset's typical behavioural patterns is the process of anomaly identification. Broad models for anomaly detection based on deep learning yield the fastest and best results.

III. RESEARCH METHODOLOGY

Detecting instances of semantic objects of a certain class (such as people, buildings, or automobiles) in digital photos and videos is the task of object detection, a branch of computer vision and image processing.Every object class has distinctive characteristics that aid in categorization; for instance, all circles are spherical. These unique traits are used for object class recognition.
First, we need to build a model to detect the objects in image or video using TensorFlow.
Using this model as the base we can build a custom object detector for the CCTV.
The plan and procedure used to carry out the study are outlined in the methodology section. This comprises the study's universe, sample, data and data sources, variables, and analytical methodology.

The details are as follows;
Stages to train a model:

1. Data Collection.
2. Annotation.
3. Train-Test Data.
4. Generate TF Records files.
5. Train.
6. Frozen model generation.
7. Inference.

1) Data Collection

- The first step in this procedure is to gather photographs using the custom object (gun, knife).
- gathering pictures with diverse scenarios (eg: captured with different angle based on your use case).
- Clean up the data (remove blur image, unnecessary images).

Figure : Data Collection.

2) Annotation or labeling

We used LabelImg to label the images. It is a tool for graphically annotating images. It uses Qt for its graphical user interface and is developed in Python. The label software will generate an xml file with object information from the image, such as the object's location, name, and coordinates (Xmin, Ymin, Xmax, and Ymax). Your xml file can be saved in either the Yolo format or the Pascal VOC format used by ImageNet.

- Labelling:

i) Click to set the path to the generated.xml file and the input picture directory.
ii) On the image, draw a rectangle-shaped box.
iii) Click next to view the following image after selecting "Save with Pascal voc formate.

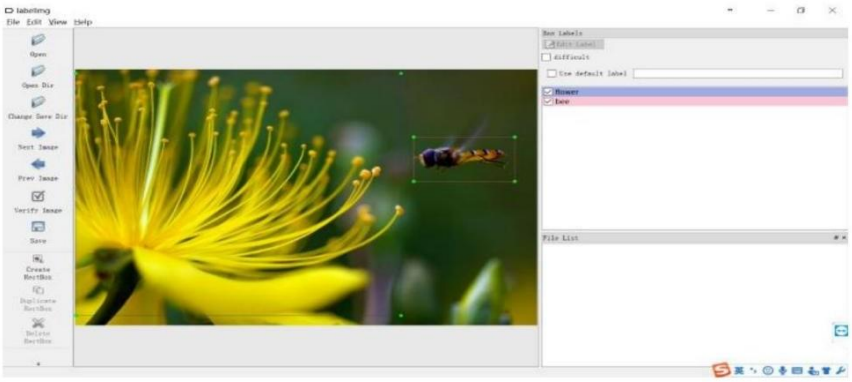


Figure:Annotation using Label Img.

3) Train and test set

Make k-folds of the whole data set for the test and training sets. The remaining 90% (or 80%) of the entire data and their annotation should go in the /train folder, and the remaining 10% (or 20%) should go in the /test folder.

- Generate CSV files:

There are two directories: test and train. Convert each directory's xml file to the appropriate format. files that can be converted

into TFRecord files are csv files.

4) Generate TFRecord

Generate TFRecord files from .csv.

```
# TO-DO replace this with label map
def class_text_to_int(row_label):
    if row_label == 'person':
        return 1
    if row_label == 'level1_helmet':
        return 2
    if row_label == 'level2_knife':
        return 3
    if row_label == 'level3_gun':
        return 4
    if row_label == 'level1_muffler':
        return 5
    else:
        None
```

Figure :TFRecord for the objects

5) Training the model:

The object detection model should be trained to recognise our unique object. Images, matching tf Records, and training and testing data are required for this.It uses the Faster RCNN algorithm. A feature extraction network, which is commonly a pretrained CNN, makes up a Faster R-CNN object detection network. Two trainable subnetworks are then added after that. A region proposal network (RPN) is utilised in the first to create object suggestions, and a prediction algorithm is employed in the second to determine the actual class of the item. Therefore, the RPN that is put after the final convolutional layer is the key differentiator for Faster R-CNN.

The below steps are typically followed in a Faster RCNN approach:

- Give the Convolutional Network a picture as input, and it will return the feature map for that image.
- These feature maps have a region proposal network added to them. This gives you the object suggestions and their confidence score.
- These ideas are subjected to a Region of Interest Pooling (RoI) pooling layer in order to equalise their sizes.
- In order to categorise and output the bounding boxes for objects, the suggestions are finally given to a fully connected layer that has a SoftMax layer and a linear regression layer at its top.

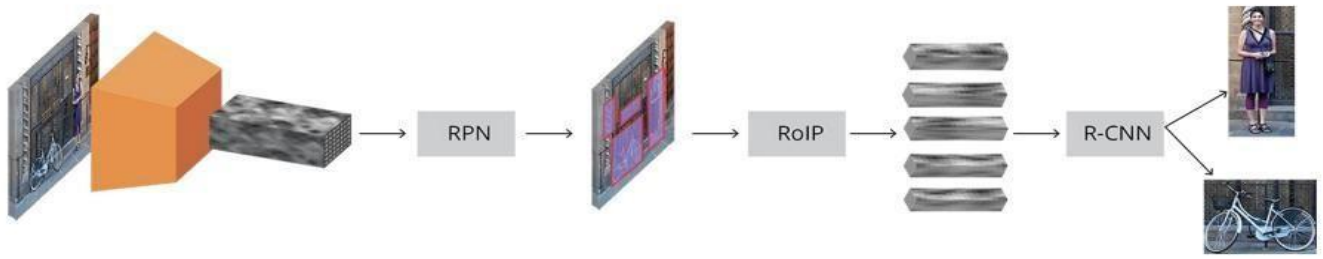


Figure: Faster RCNN.

6) Frozen model generation:

We will put our trained model to the test. We must export the inference graph in order to achieve it. In the context of neural networks, freezing a layer refers to managing how the weights are updated. When a layer is frozen, the weights can no longer be changed.

7) Inference: To produce an output, such as a single numerical score, inference is the act of feeding live data points into a machine learning algorithm (or "ML model").

IV. CONCLUSION

Due to an increase in crime at packed events and unsettling lonely regions, security is always a top issue in all fields. Computer vision is widely used in abnormal detection and monitoring to solve various issues. Due to the increasing need to defend human safety, security, and property, video surveillance systems that can identify and decipher scene and anomaly occurrences are essential for intelligence monitoring.Deep learning is becoming a common practise, and whenever new technology is used, security

and privacy concerns undoubtedly arise. There has been a lot of study done recently on the training and interface modules for deep learning and deep neural networks that preserve security and privacy. Therefore, security and privacy become highly essential problems that cannot be ignored, just as with other technologies.This system aims to build an effective, reliable, secure, fast and an efficient CCTV Surveillance system using deep learning techniques., replacing the manual, unreliable and traditional system. As we have also created dedicated app to get notification and real-time streaming. It converts traditional cameras into smart cameras. This system can be implemented for better results regarding the security.

REFERENCES:

[1] Abhiraj Biswas et al., "Classification of Objects in Video Records using Neural Network Framework", International conference on Smart Systems and Inventive Technology, 2018.

[2] H R Rohit et al., "A Review of Artificial Intelligence Methods for Data Science and Data Analytics: Applications and Research Challenges", 2018 2nd International Conference on ISMAC (IOT in Social Mobile Analytics and Cloud), 2018.

[3] Harsh Jain, Aditya Vikram, Mohana, Ankit Kashyap, Ayush Jain "Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications", the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020) .

[4] Syed Sahil Abbas Zaidi, Mohammad Samar Ansari, Asra Aslam, Nadia Kanwal, Mamoona Asghar, Brian Lee "A Survey of Modern Deep Learning based Object Detection Models",2021.

[5]N. Jain et al., "Performance Analysis of Object Detection and Tracking Algorithms for Traffic Surveillance Applications using Neural Networks", 2019 Third International conference on I-SMAC (IoT in Social Mobile Analytics and Cloud), 2019.

[6] Yojan Chitkara et al., "Background Modelling techniques for foreground detection and Tracking using Gaussian Mixture model", International Conference on Computing Methodologies and Communication, 2019.

[7] Chunnu Khawas, "Application of Firebase in Android App Development "International Journal of Computer Applications 2018.

[8] Dayana R, Suganya M "Smart Surveillance System using Deep Learning"International conference on Smart Systems and Inventive Technology ,2021

[9]A. Glowacz et al., "Visual Detection of Knives in Security Applications using Active Appearance Model", Multimedia Tools Applications, 2015.

[10]Ruben J Franklin et al., "Anomaly Detection in Videos for Video Surveillance Applications Using Neural Networks", International Conference on Inventive Systems and Control, 2020.