



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 1, January 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijrset@gmail.com

www.ijrset.com

AI Based Automatic Robbery/Theft Detection using Smart Surveillance in Banks

Archana Yadav¹, S.Sai Swaroop², T.Nagendra Babu³

Mahatma Gandhi Institute of Technology (Department of E.C.E), Gandipet, Kokapet, Hyderabad, India¹

Mahatma Gandhi Institute of Technology (Department of E.C.E), Gandipet, Kokapet, Hyderabad, India²

Mahatma Gandhi Institute of Technology (Department of E.C.E), Gandipet, Kokapet, Hyderabad, India³

ABSTRACT: We propose a CCTV based theft detection along with tracking of thieves. We use image processing to detect theft and motion of thieves in CCTV footage, without the use of sensors. This system concentrates on object detection and motion detection. The security personnel can be notified about the suspicious individual committing burglary using Real-time analysis of the movement of any human from CCTV footage and thus gives a chance to avert the same. For this detection we are using 2 models namely Yolo V3 for weapon detection and LRCN approach for motion detection.

KEYWORDS: Image processing , LRCN approach , Yolo V3, Theft detection , Burglary, CCTV .

I. INTRODUCTION

Closed circuit television systems (CCTV) are winding up increasingly famous and are deployed in numerous banks. As the human factors limit the number of camera views of a single operator, this makes for an immeasurable load for the CCTV operators. According to the "CCTV Operational Needed things Manual 2009", the job of the CCTV operator is to watch and detect, control, recognize, watch/ notice/ celebrate/ obey, and identify situations and people that are possibly harmful to other people and their property. A solution to these problems is to apply image-processing and algorithms to the CCTV footage, which will replace the human operatives and alert the security officials if a terrible situation is found.

II. DESIGN METHODOLOGY

2.1 System Design

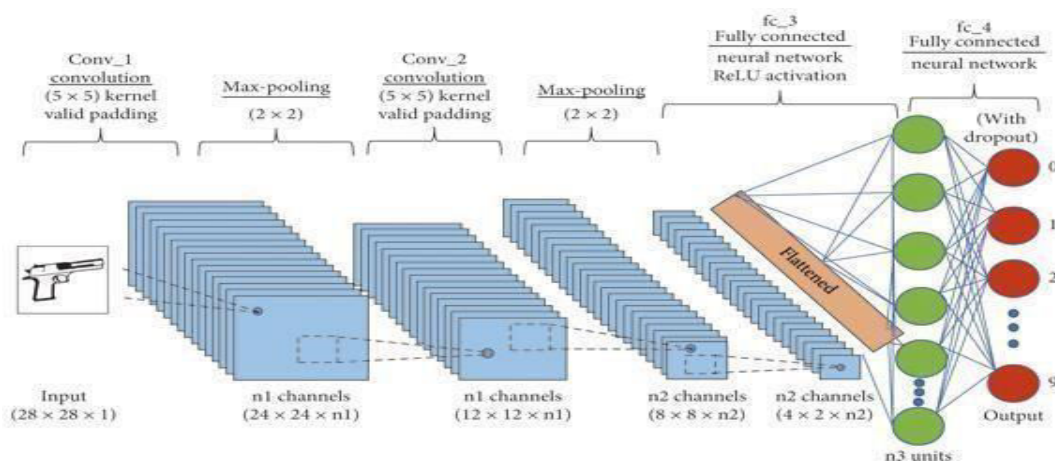
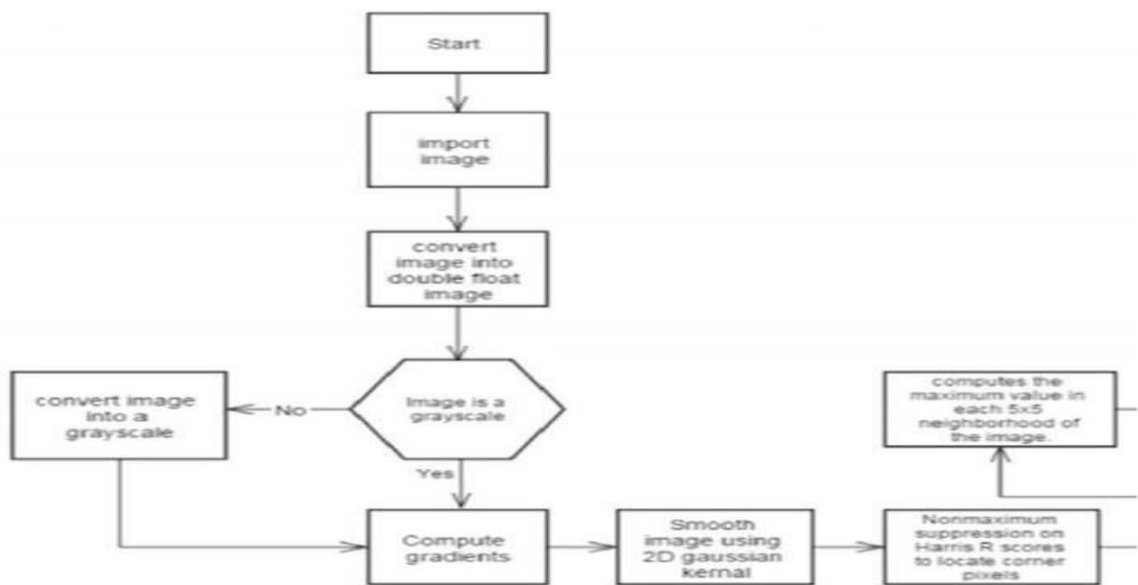


Figure 2.1: Feedforward convolutional neural network (CNN).

YOLO is basically a pretrained object detector. It is a CNN model. A CNN is a deep learning algorithm which can take in a raw input image and assign learnable weights and biases to various aspects/objects in the image. A convolutional layer in CNN model is responsible of extracting the high-level features such as edges, from the input image. This works by applying $k \times k$ filter known as kernel repeatedly over raw image. This further results in activation maps or feature maps. These feature maps are the presence of detected features from the given input. Thus, the pre-processing required is much lower as compared to other classification algorithms, whereas in standard approach, filters are hand-engineered and in CNN these are learned through a number of iterations and training. Figure 2.1 indicates a basic CNN architecture as classification model for 10 different weapons. Subsequently, the next layer is Max-Pooling or Subsampling layer, which is responsible for reducing the spatial size of the convolved features. This is to decrease the computational power required to process the data through dimensionality reduction. This is to decrease the computational power required to process the data through dimensionality reduction. ReLU is a rectified linear unit activation function, which is related to the feature of non-saturating activation. It eliminates undesirable values from an activation map effectively by setting them to nil. Finally, the last layers are fully connected layers transforming the data into a 1-dimensional array. To create a particular long feature vector, the flattened output is fed to a feedforward neural network and backpropagation is applied to every iteration of training. These layers are liable to learn nonlinear combinations of the high-level features as represented by the output of the convolutional layer.

Figure 2.2: Yolov4 Object Detection Algorithm



Long-term recurrent convolutional network (LRCN) main idea is to use a combination of CNNs to learn visual features from video frames and LSTMs to transform a sequence of image embeddings into a class label, sentence, probabilities, or whatever you need. Thus, raw visual input is processed with a CNN, whose outputs are fed into a stack of recurrent sequence models. It is a combination of CNN and RNN, end-to-end trainable and suitable for large-scale visual understanding tasks such as video description, activity recognition and image captioning. The recurrent convolutional models are deeper in that they learn compositional representations in space and time, when the previous models assumed a fixed visual representation or perform simple temporal averaging for sequential processing. And it can be trained or learn temporal dynamics and convolutional perceptual representations as it is directly connected to convolutional network. Long-term Recurrent Convolutional Networks (LRCNs), a class of architectures leveraging the strengths of rapid progress in CNNs for visual recognition problems, and the growing desire to apply such models to time-varying inputs and outputs.

III. IMPLEMENTATION AND RESULTS

3.1 IMPLEMENTATION

In this section, we will implement two models namely YOLO V3 for weapon detection and LRCN model for motion detection. Below is the step-by-step procedure to implement these 2 models on the dataset and get the desired output.

- 3.1.1 Import libraries
- 3.1.2 Visualize the data with its labels
- 3.1.3 Pre-process the Dataset
- 3.1.4 Split the data into Train and Test set
- 3.1.5 Implementing the YOLO V3 model
- 3.1.6 Implementing the LRCN model
 - 3.1.6.1 Construct the model
 - 3.1.6.2 Compile and train the model

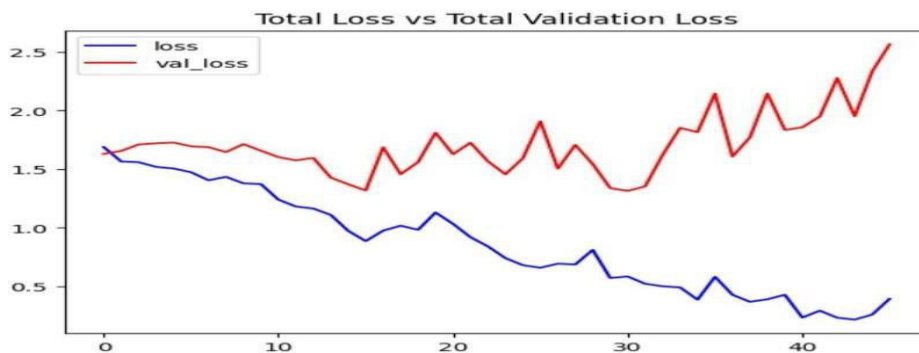


Fig . 3.1.1(a) – Total Loss vs Total Validation Loss

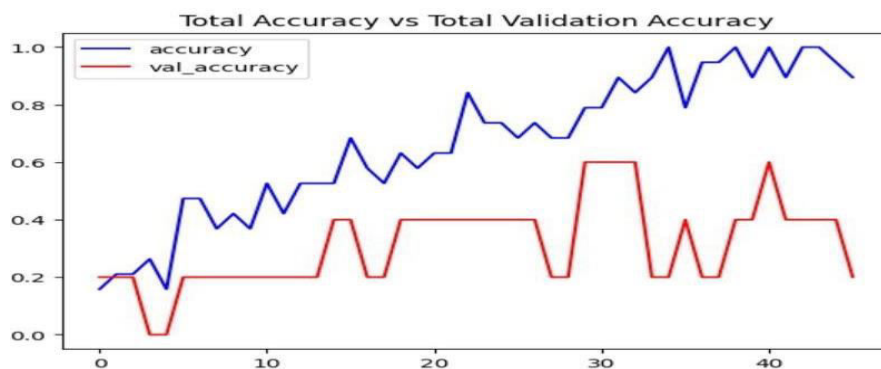


Fig 3.1.1(b) – Total Accuracy vs Total Validation Accuracy

IV. RESULTS

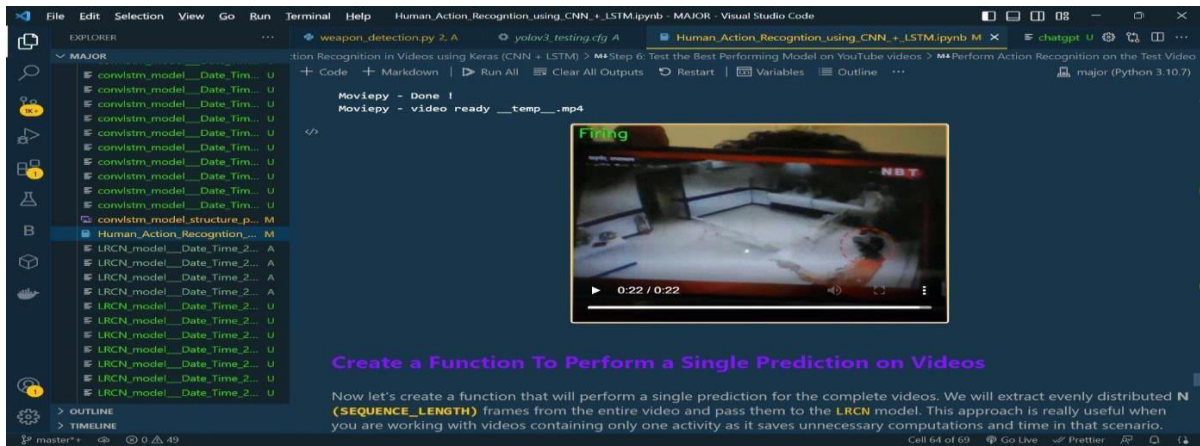


Figure 4.1.1: Output after prediction with label on top left.

The output will be displayed in the editor and will be saved in the temporary memory and the output video contains the label on the top left corner which represents the activity that is predicted by the model.



Figure 4.10: Detecting the activity as Robbery.

V. CONCLUSION AND FUTURE SCOPE

5.1.1 CONCLUSION

In this project, we performed a review of techniques or models used for detecting objects and abnormal motions in banks through smart surveillance. The real time motion detection system with the ability to automatically recognize the video images and the images with moving objects. The existing surveillance and security system can be overcome by a real time detection system with motion. The ability is to automatically analyze the video images and the images with moving or colliding objects once the object is detected alerts are automatically noted or send to the user. A digital surveillance system where the CCTV camera and target for small scale user is developed. The latest technologies are to find the thefts activities and destruction used by video surveillance system. By using this, it is possible to capture and monitor by the user at each and every moment and each and every second of the area coverage.

5.1.2 FUTURE SCOPE

In the future, the scope for Theft/Robbery detection in banks is likely to expand and evolve due to advancements in technology. There is great scope for this Theft detection system if it uses IoT in it to setup alarm rings for the security to alert at the time of suspicious activity. As day by day the technologies are increasing, the manpower work will be reduced simultaneously. So, at that time this system will be used by many banks, ATMs and others for security purposes. Users using this system need not worry about supervising the cameras all the time instead the system will inform the user about the activities happening and will also suggest that the user should take some action or not.



The YOLO V3 model detects weapons in the input video. If any weapon is detected, then the label on the output video will be displayed as Firing.

REFERENCES

- [1]. Mandar Shriram Munagekar, "Smart Surveillance system for theft detection using image processing", 2018.
- [2]. Roberto Olmos, Siham Tabik, and Francisco Herrera, "Automatic Handgun Detection Alarm in Videos Using Deep Learning", 2017.
- [3]. Tasriva Sikandar and Kamarul Hawari Ghazali, "A Review on Human Motion Detection Techniques for ATM-CCTV Surveillance System", 2016.
- [4]. Michał Grega, Andrzej Mاتیolański, Piotr Guzik and Mikołaj Leszczuk, "Automated Detection of Firearms and Knives in a CCTV Image", 2015.
- [5]. Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", 2015.
- [6]. S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition", in IEEE Trans. Pattern Anal. Mach. Intell., 2013.
- [7]. Kushwaha, A., Mishra, A., Kamble, K., & Janbhare, R. (2018). Theft detection using machine learning. In Int Conf. Innov. Adv. Technol. Eng (Vol. 8, pp. 67–71)



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details