ISSN No. 0976-5697

Volume 11, Special Issue I, May 2020

International Journal of Advanced Research in Computer Science

CONFERENCE PAPER

Available Online at www.ijarcs.info

STARS - DL

Prithviraj Hiremath
Department of C&IT
Reva University
Bangalore, India
prithvirajh98@gmail.com

K Sai Tarun
Department of C&IT
Reva University
Bangalore, India
tarunmahesh3@gmail.com

Sowmya Sundari L K
Department of C&IT
Reva University
Bangalore, India
sowmyasundarilk@reva.edu.in

Abstract – The purpose of this project is to detect, track and notify the intruding objects. The system will have a clear view of the surveillance area with a high pixel camera or a surveillance drone. The camera will be fixed in a position that the surveillance area is covered. It will identify interruption and perceive the object by contrasting its features with the features of the objects that are stored in the database.

If a feature matches the intruding object,then it will get tracked and will notify the authorities. Thus, the image will be captured, proceed and the authorities can take actions necessary. The above system will work as a detection, recognition and alertthe intrusion which will be of great use for Military applications. Our system ensures low execution time.

Keywords: detect intrusion, recognition, notify the authorities

I. INTRODUCTION

JSTARS gives the layout of ground situation information through communication via the cloud with the air force, ground posts and army mobile ground stations and centers of military analysis far from the point of conflict. JSTARS can determine the direction, speed and patterns of military activity of ground vehicles and helicopters.
STARS – DL is a surveillance system that will check or survey the intended area if there is any object which is not supposed to be present. The system will notify the authorities to check the anomaly in the surveillance area and take necessary action.
We are using yolo as our base algorithm and python to input the surveillance videos. As our project will be using high resolute cameras or drones to get real-time surveillance as input, we are going to use the cloud to analyses the data.
We are using mobilenet to upload the dataset of the gun images for training the system to identify the guns in the surveillance area.

II. LITERATURESURVEY

In September 1996, JSTARS was first affirmed for f creation for 14 airplanes, the remainder of which was conveyed in August 2002. Three further airplanes were conveyed between February 2003 and March 2005. The 116th Air Control Wing works the JSTARS airplane at Robins Air Force Base in Georgia. The 116th is a 'mixed wing' with both aviation based armed forces and air national gatekeeper workforce.

Improved Object Detection for Air Force Joint Surveillance Target Attack Radar System descriptors, outskirt identification and example coordinating or utilizing progressively complex strategies, for example, course classifiers with boosting. The adequacy of these strategies made them basic in some spots. In any case, they have a few impediments. As these frameworks depend on metal recognition, they can't distinguish non-metallic firearms. They are costly to be utilized in numerous spots as they require to be joined with X-beam scanners and Conveyor belts. They are not exact because they respond to every metallic item.

According to [2] the grids of histograms of oriented gradient descriptors outperform existing feature sets of human detection it is also called as 'HOG'.

[1] and [3] papers the algorithm they used can alert the human operator when a firearm or knives are visible in the image using CCTV cameras footage.

The [4] reference paper conveys that the inferring of 3D human pose from marker-free images is a challenging task. So, they used many algorithms to get the 3D poses and have used silhouettes as visual inputs.

The problem of minimizing false positives alerts and solve it by building the key training dataset guided by the results of a deep Convolutional Neural Networks (CNN) classifier and assessing the best classification model was proposed by [5] IEEE paper.

There is a paper on a similar to this project "A Computer Vision-based Framework for Visual Gun Detection using Harris Interest Point Detector" [6] which is for gun detection that was our base paper as we are going to detect other objects too.

III. OBJECTIVES

- Identifies the objects which are threats.
- To Provide proof of Concept for building, deploying and scaling the STARS
- To Provide Algorithm implemented using Convolutional Neural Networks (CNN) and related models.

IV. BLOCK DIAGRAM

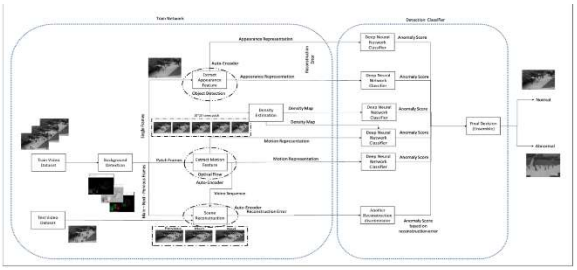

Figure 1 block diagram of deep learning process for the system

The above diagram Figure1 explains the deep learning process of the system. The images are given continues to the system to learn the normal layout of the surveillance area which is stored in the system. The system keeps on learning the images to identify every single movement in the area

When there is an abnormal or intrusion in the surveillance area then it will automatically send an alert to the authorities

V. METHODOLOGY

The proposed paper is based on deep learning techniques called Tensor Flow for detecting objects in video. The main components are considered for this method. The first component is the extraction and learning of the feature and the second component is the detection of objects andthe step by step process is as given below.

A.Pre-Processing.

The initial step before beginning to concentrate and learning highlights is to gauge and expel the foundation. The foundation is unique for various situations as there are different techniques for its expulsion. For example, the foundation may incorporate void spaces or road fringes. Right now, foundation estimation depends on most event of recurrence between video outline patches [6]. The foundation estimation ventures from the outset, a histogram is produced for each edge of the video which depends on pixels and their area in the picture. At that point the histogram of the edges in each fix is contrasted and one another, and the greatest qualities per fix are distinguished as the foundation and are accordingly grayed. Evacuating the foundation will diminish the expense of the figuring and the preparation time. This progression is considered as a piece of train organize.

With the addition to the estimation, a training network has four fundamental segments. The profound system for separating the appearance includes utilizes a stacked demising auto-encoder with a 6 encode layer and a similar structure of the interpreting layer. Each edge is convolving to connect with 10X10 window size, and it incorporates walk and cushioning. All casings standardize in parallel mode. This system has 6 layers and 6 same structures in the interpret layer which is more profound than the current strategies. The distinguished items are called appearance portrayal. This yield is utilized in distinguishing stage and is used as a contribution to thickness estimation part to build the exactness of estimation

C. Detection Component

In the detection component, learned features that are generated in the training network are given to a classifier. Features are given as individual and combined features to these networks. Reconstruction error and visible features are given in to the network as a combined feature since the visual feature or object detection with a reconstruction error can be a strong feature for the detection of anomalies in the system. The error for the corresponding framework will make the detection more accurate which is at 96%.

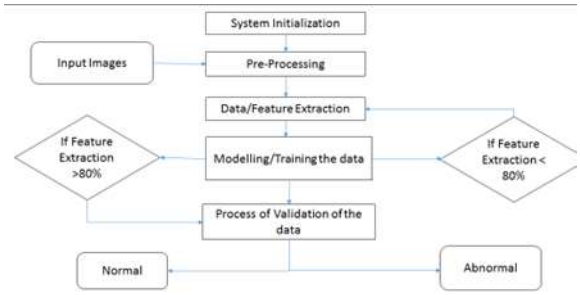

Figure 2 flow chart for the input and output of the system

The above flow chart Figure 2 shows the how the algorithm worksInput is given to the pre-processing which in turn is connected to the data extraction if the feature extraction is less than 80% accurate then it will be sent back to the modeling the data then it processes the data and gives the output as normal or abnormal data which in turn will notify the authorities.

The input of the images or videos will be given to the algorithm by a python code.

VI. ALGORITHM

The Real-time face detection system is built on Convolution Neural Network. YOLO - You Only Look Once the algorithm is paired with the GPU of the system for faster image processing and computations using CUDA -Computer Unified Device Architecture.

The dataset involves the feeding of several aspects of the face using the FACE TRAINING script built using Python and TensorFlow MTCNN. Weights of the dataset are downloaded from the repository and it executed.

We are also using Mobilenet as our base algorithm for gun detection to get more accurate

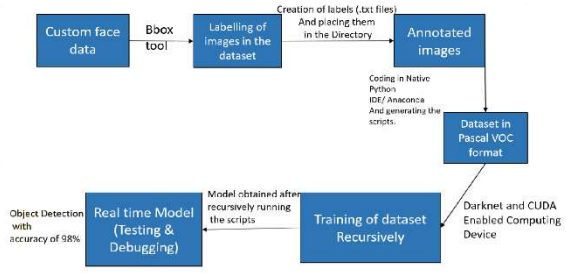

Figure 3 block diagram for yolo algorithm

The block diagram Figure 3 above shows how the YOLO algorithm works step by step.

Step1: the face data is given as the input to the Bbox tool.

Step2: the images are labeled after grouping them in the dataset.

Step3: after storing them in the directory the images are annotated.

Step4: dataset is analyzed in pascal VOC format and is sent to CUDA enabled computing devices.

Step5: the training of dataset (recursively) until the models are obtained.

Then the real-time model testing and debugging which give a 98% accuracy rate.

VII. Modules Identified

STARS - DL project has 5 different modules.

They are as follows:

1. data collection
2. data processing
3. modeling/training
4. testing
5. validation

VIII. RESULTS

We found out that our system works well in all situations except illumination change. It works with 91.66% true high-resolution case of images of guns with backgrounds.

The system also works in all different cases due to the accuracy of the system.

The system is not able to perform in all color-based change because the color-based segmentation algorithm we used is not able to segment the imageaccurately due to which when we are extracting gun color from the input image we are getting only some part of the gun which affects the performance of the system this is resulting in the delay of the outcome.

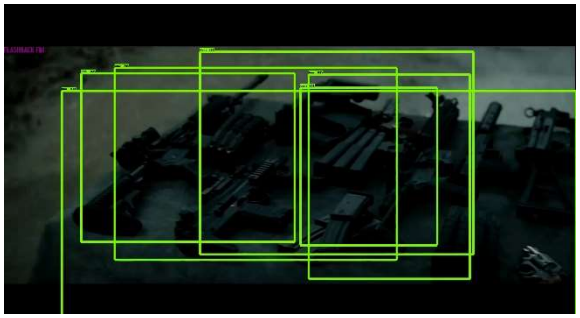

Figure 4 output of a video input from a movie clip

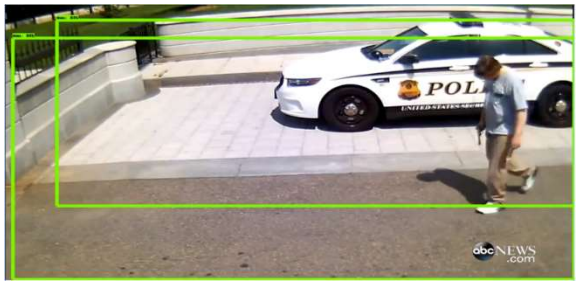

Figure 5output of street camera footage at the white house

As the above figure (4)(5) show the results of the tested video inputs given to the system

IX. CONCLUSIONS

A deeplearning (DL) based Object detectionin avideo surveillance camera is presented. One preferred position of this strategy is the utilization of profound learning systems in all train and recognition parts. The two primary segments of this technique are assessed dependent on certain measurements and with the UCSD dataset which is the most celebrated Object identification dataset. Another advantage of this strategy is the seclusion of the training arrange stage. Along these lines, it can use as a pre-training organize incomparable works.

For additional improvement, it is conceivable to include a segment that can add depictions to every discovery classifier or the last one; or it is conceivable to include a part in the identification stage which can restrict the Object precisely.

X. REFERENCES

[1] Grega, Michal, Matiolanski, Andrzej, Guzik, Piotr and LeszczukMikolaj. Automated Detection of Firearms and Knives in a CCTV Image. Multidisciplinary Digital Publishing Institute - Sensors. Published 1 January 2016.

[2] Dalal, Navneet and Trigg, Bill. Histograms of Oriented Gradients for Human Detection. Computer Vision and Pattern Recognition. Published 20 June 2005.

[3] Grega, M., Lach, S. Sieradzki, R. Automated recognition of firearms in surveillance video. In Proceedings of the 2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), San Diego, CA, USA, 25–28 February 2013; pp. 45–50.

[4] Chen, C.; Zhuang, Y.; Xiao, J. Silhouette representation and matching for 3D pose discrimination A comparative study. Image Vis. Computer. 2010, 28, 654–667.

[5] Automatic Handgun Detection Alarm in Videos Using Deep Learning Roberto Olmos 1, SihamTabik 1, and Francisco Herrera1,2.

[6] Rohit Kumar Tiwari and Gyanendra K Verma. A computer vision-based framework for visual gun detection using the Harris interest point detector. Procedia Computer Science, 54:703–712, 2015.