Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 7 July 2021: 12311–12321

Research Article

Crime Prediction, Analysis and Criminal Tracking

Diksha Arora[a], Harsh Gupta[b], Mohammad Taha Ali[c], Dr. Jaspreet Kaur[d]

[a,c,d,e] Department of Electronics and Communication, Galgotias College of Engineering and Technology, Greater Noida, India
Email: [a]aroradiksha1409@gmail.com, [b]harshgupta6201@gmail.com, [c]mohammadtahaali11@gmail.com, [d]jas.kaur@galgotiacollege.edu

Abstract

With the advancement in technology and rapid growing crime rates in every part of country, there is a huge need/requirement of an automated system to reduce the crime rate. The proposed model predicts the crime by recognizing the behavioral patterns in real-time environment and gives alert about the time and place to the assigned authorities. Also, it detects the type of crime mainly by detecting the weapon used and detects the face of the criminal by face recognition technique and track him. Models which already exist either lack in above mentioned tasks or in accuracy/efficiency but our model performs the mentioned tasks with a good accuracy score thus avoiding any false positives.

Keywords: *OpenCV, Deep Convolutional Neural Network, Recurrent Neural Network, Long-Short Term Memory, Face Embedding, Cascade Classifier, Darknet, Inception v3, Transfer Learning*

Introduction

Image processing is a developing and under research technology that is growing faster day by day in fighting against crime cases, investigation and criminal tracking. Fast and accurate identification of a bizarre activity is very important in safeguarding any society. With the evolving smart cities, the unification of crime detection systems is in severe need to improve security. From the past few years, the number of CCTV cameras installed to surveil private and public spaces has exponentially increased. People are strongly relied on typical video surveillance to achieve the above-mentioned objective. But in large cities, manually supervising thousands of CCTV footages creates a burden for the supervising authorities which lead to hike in error rates. The solution is to create a model which detects the anomalous scenarios and perform various weapon detection techniques with a high accuracy score. Criminal detection and tracking are the crucial part of the model which can extract criminal image and track the criminal if he appears in any of the CCTV footages in a particular area. The extracted image can also be considered as an evidence for trials and prosecution.

In particular, the work would be beneficial for independence living conditions, where people live alone, warning system will greatly improve their security. For this purpose, we have used a combination of Deep learning techniques like DCNN, RNN and LSTM to extract frames from a video set and process it for various suspicious activities recognition.

Deep Convolutional Neural Network (DCNN)-DCNN is a class of deep neural networks mostly applied to analyze visual imagery. For example, image and video recognition and is based on shared-weights architecture.

Recurrent Neural Network (RNN)- RNN contrary to feed forward neural networks, use their local memory to process random sequences of inputs. RNNs are helpful as they are not bounded by the length of an input and can use temporal relation to predict finer meaning.

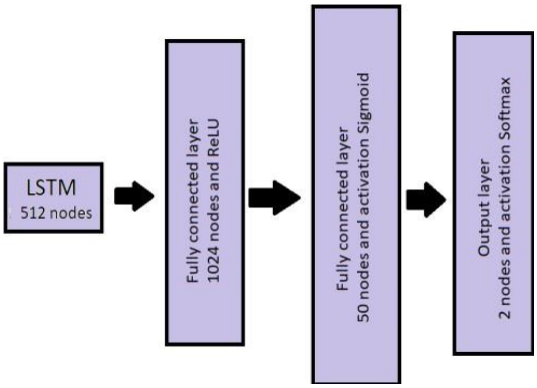

Fig.1 Recurrent Neural Network

Long Short-term Memory (LSTM)- LSTM is a kind of RNN that uses special units as an addition to standard units and includes a "memory cell" that maintains information and data in memory for longer period of time.

It is able to learn appropriate features from an image or video with high accuracy levels varying with the algorithm and number of layers used making it more efficient in terms of memory and less complex as compared to NN. We have implemented a transfer learning technique using the weights of a pre-trained Inception v-3 model. Advantage of transfer learning include avoiding training of CNN and save memory and time. LSTM is used as the first layer in the RNN model. It is a very powerful neural network technique used for image recognition it is a great tool for dealing with sequential data and can be used for text generation.

Assisting forensic based video analysis, a deep- learning based material and tracking algorithm that can detect and identify potential suspects and tools from the frames is proposed in the model.

Object detection is the primary module in the project. In crime scene study, photos and videos are the main aid to analyze the scene. However, due to presence of massive data, the task to detect the desirable object is very tedious. For this work, the model is based on Darknet and YOLO v3, which automatically detects objects in a real-time environment. The main focus was to detect weapons like guns. And for the live criminal tracking part, we have used face embedding algorithm where each face is converted into a vector also called deep metric learning.

Face embedding- It is the process of extracting high quality features from faces and return the numerical vector that represent each detected face in the image in 1024-dimensional space.

Thus, the working model is able to quickly detect crime in a real-time environment through CCTV footages and generate appropriate security alerts. It is also able to detect weapons like guns, and through face detection and recognition, it will help in tracking the criminal.

The rest of the research paper is organized in the following way- The second part of the research paper will tell us about the works similar to this project, the third part will cover the proposed model and methodology, i.e., the design and implementation. The paper will then discuss about the results achieved while deploying the model.

Conclusion, acknowledgment and the references are mentioned at the end.

The purpose of using these techniques isn that unlike the conventional neural networks (NN), CNN is based on the concept of featture learning.

Research Objectıve

The research paper aims to train a model which predicts the crime, detects the weapons used and tracks the criminal face recognition. The paper explains implementation of video classification for crime detection. It uses various Deep Learning tools and techniques to frame a good architecture of video classifier. The project tries to detect crime in a video by classifying videos as: -

1. Crime videos
2. Normal videos

And by using different object detection and face recognition techniques, the weapon detection and criminal tracking is carried out.

The work will further help the security agencies to predict, detect and track the crime and criminals in a particular region.

Related Works

In Chakravarthy's research, they proposed a crime detection model which performs over a large data set which is adapted by various video surveillance systems to act as an intelligent alert system, which would lower the load of work on supervising authorities [1]. They made a project scheme which is based on the DCNN for the crime detection using the Hybrid Deep Learning algorithm.

Saikia's research presents a real-time system based Faster R-CNN (Region-based Convolutional Neural Network), which detects tools which might be found in an indoor environment [2]. The accuracy obtained in this paper is 74.33%.

In Xiao's work, we see work on advanced forensic video study techniques to help in the criminal investigation [3]. They use a compatible contrast limited adaptive histogram equalization to improve the surveillance camera footage quality and abet the video-based forensic analysis. They also make use of a deep an object detection and tracking algorithm based on deep learning which detects and identify crime suspects and objects from footages. The overall accuracy of their project is found to be 87.1%.

In Nakim's research we see a proposed model to detect threatening object to predict whether a crime has happened in a specific image [4]. They make use of Tensor Flow open-source platform, Non-Linearity, Re Lu, Convolutional Neural Layer, fully connected layer and dropout function of CNN to achieve an accurate result. This paper has a testing accuracy of 90.2%.

In Weiyi's research, they build a model using training data set that does data cleaning and transformation [5]. Data visualization helps in the analysis of data set. Their model mainly predicts the crime and gives an accuracy of 78.9%.

In Kormath's paper, they work upon the detection and enhancement of the fuzzy videos taken from the surveillance cameras [6]. The paper talks about crime being detected and enhancement of the features of the image by either contrast ratio enhancement, image defogging, or other techniques.

In Devishree's work, they talk about everything related to detection of crime, the processes involved like dataset formation, training of model, gradient descent, experiment analysis and result analysis [7]. Their model recognizes guns, knives and blood stains.

In this research paper they also tell us about TensorFlow, CNN, Re LU and max pooling related algorithm to get the model trained. This model shows the accuracy result of 90.2%.

Methodology

The project deals with three working models- Crime detection, weapon detection and criminal tracking. The proposed framework of each model is stated in the following sections.

A. Crime Detection

The first network is the convolutional neural network to extract high-level features of the images. A pre- trained model, Inception V3 was used and transfer learning technique was applied. Inception V3 is developed by Google. It is trained on the ImageNet Large Visual Recognition Challenge dataset.

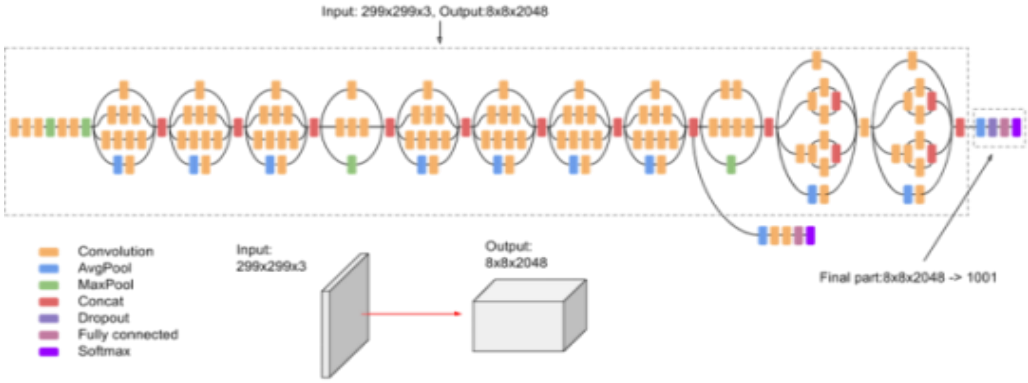

Fig. 2 Inception-v3 architecture

The second network used is the recurrent neural network. In CNN, the inputs & outputs are not dependent on each other but in RNN, the outputs from previous step are fed as inputs to the current step. RNN are mainly used to predict the successive word in a sentence, by remembering the previous words. Hidden state in RNN remembers information about the sequence. The layer has first layer as LSTM layer followed by two hidden layers, the first one

has 1024 nodes with the activation function ReLu and the second one has 50 nodes with activation sigmoid. The output layer has 2 nodes, the two classes- Crime and normal, and has the activation SoftMax.

About Dataset- The dataset this project used is a collection of 40 crime videos and 30 normal videos. We collected crime videos from different sources. They are mostly CCTV footages recorded at the time of the crime.

Data preparation- Each video is converted into frames and 40 frames are manually selected from each video which depicts the crimes happening. The 40 frames of each videos are stored separately.

For transfer learning we only used the output from the last pooling layer of Inception v3. Each frame was

passed into the inception-v3 network and a 2048 vector was extracted. Dataset has 40 crime videos, hence 40*40 i.e., 1600 2048-vectors for crime videos and 40*30 i.e., 1200 2048-vectors for normal videos. 40 frames for each video were taken and grouped to represent them as a sequence for our RNN. The final list has a length of 2800(1600+1200). It was converted into a 3D array of dimension (70,40,2048) where 70 is the total number of videos, 40 is the sequenced frames of each video and 2048 represents the length of vector extracted. It is split into training data and test data in the ratio 0.8:0.2. The model is finally trained on this 3D array.

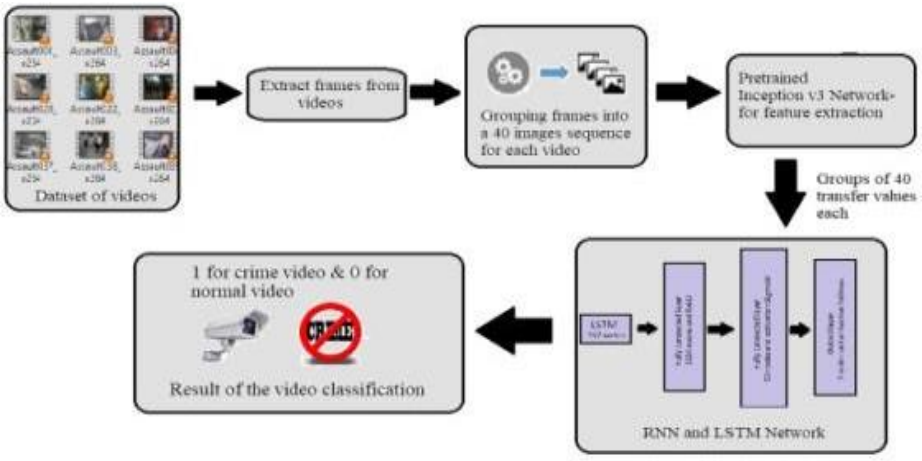

Fig.3 Architecture of Crime classification

B. Weapon Detection

The framework of the real-time gun detection includes the following- The DarkNet package is used to implement YOLO v-3 to train and get the weights on our gun dataset. The dataset contains .jpg files of guns and each image is associated with a label file which has this information in the syntax:

<x_center><y_center>

YOLO (You Only Look Once) is an extremely fast object detection algorithm mainly used for real time detection. It is faster than Faster RCNN and RCNN and is widely being used. It can accurately detect objects in images and videos.

After training, the trained weights are extracted. OpenCV is used for input and video processing . YOLO v3 used the already trained weights to perform the detection tasks on the input video given, and we get the output as a video with rectangle drawn around the gun (if found).

C. Criminal Tracking

The third and the last section of the project is tracking the criminal by face detection method.

The frames are extracted from crime detected video frames and are further analyzed for positives and the frame with the most crime prediction rate is selected for detecting the face of the criminal. The assigned

authorities have the right to add the face of the criminal in their criminal database, if the face found is not already present, followed by the commencement of live detection of the criminal in all the nearby cameras. The assigned authorities can also feed in the face of any other criminal in our model to run face recognition for it in CCTV cameras.

When recognizing faces, there may be slight differences in human faces but there are some features which are common to all and our model focuses on those features. To recognize a face, it is important that we detect/locate a face in an image/video. Face recognition is a method of verifying the identity of an individual using their facial features. There are many algorithms use for face recognition but face recognition by using deep learning is more accurate. The central idea of this section was performing face embedding which converts the face into a vector using deep metric learning technique.

The first task is the face detection in the video frames. Face is detected using face recognition library which is created by Adam Geitgey. The library will provide the coordinates of face so this face is extracted for further analyzing. Features are extracted from the face using face embedding, then the neural network takes the face as an input and gives the most important features of that face which is called embedding in machine learning language.

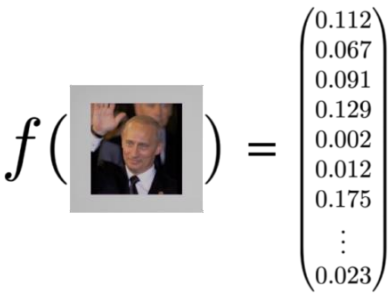

Fig. 4 Face Embedding

The neural network is trained for giving similar vectors for the face that are similar. The logic behind this is though some of the facial features may vary with time or with the image quality but the vectors associated with the faces are in very close vector space. The training would take a lot of time and computation power so a pre-trained network by Davis King on a dataset with approximately 3 million images were used. All the images were transferred in our data to this pre-trained network to get the respective embedding for next step.

The final step is comparing faces which processes the new image which is not in the dataset using the same network which is used above and then compare these embedding with the embedding which are already extracted. Then it is checked if the image is any close or similar to any other embedding.

The model was implemented using OpenCV and Python. The libraries used were dlib and Face recognition.

OpenCV-It is an image and video processing library mainly used for face detection, reading, photo editing, optical character recognition, and a lot more.

dlib-It is a library maintained by Davis King and is used in implementing "deep metric learning" and constructing face embedding used in the actual recognition processes.

face_recognition- It is originally designed by Adam Geitgey, and revolves around dlib's facial recognition process. It super easy to work with and should be installed after installing dlib.

Experımental Result

The dataset was divided into two categories crime videos and normal videos and the model was trained with 0.8 as training set and 0.2 as test set. The trained model was tested on the test set and it gave an accuracy of 92.8% which implies that the video classification was implemented successfully and our model can detect the abnormal behaviors and crime patterns. These abnormal behaviors were detected using the deep convolutional neural network and RNN network using video classification. DCNN extracts the high attributes

from the frames using the LSTM. The frames with crime probability were then used for further analysis like weapon detection.

Fig. 5 shows a crime being detected successfully detected on a frame of a CCTV footage.

Fig. 5

As shown in Fig. 6 and Fig.7, the detection of crime weapon was successfully executed. The model was able to detect the Gun in the video frame.

Fig. 6

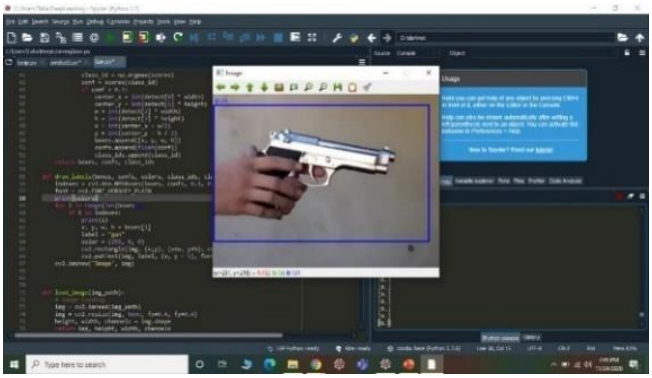

Fig. 7

Analyzing the crime positive frames face detection was successfully implemented. Fig. 8 shows the face detection and recognition using face detection technique.

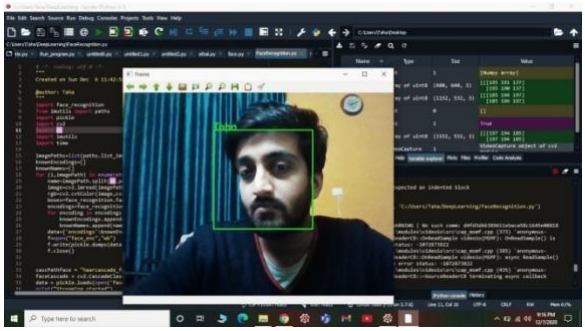

Fig. 8

Summarizing, the objectives of our research attained with good accuracy score and experimental observations were held successfully.

Conclusıon & Future Scope

As we know that India is a rapidly evolving country. Also, the crime rate in India is increasing day by day. It has become difficult for the security agencies to handle these crimes and track the criminals. So, our proposed model helps in identifying, detecting and tracking criminals with the good accuracy.

Our model does the video classification and identifies the crime with the accuracy of 92.8% and detects the weapon and send the positive alert to the assigned authorities.

In further works, we will work upon analyzing the crime scenes and draw the crime graph to find out the crime rate of that particular area. And, we will increase the accuracy of our model to make it more efficient.

References

[1] Sharmila Chakravarthy, Steven Schmitt, Li Yang, "Intelligent Crime Anomaly Detection in Smart Cities using Deep Learning ". Computer Science and Engineering Department University of Tennessee at Chattanooga, 4th IEEE International Conference on Collaboration and Internet Computing, 10 Nov 2018

[2] Scene Evidence Analysis Surajit Saikia, E. Fidalgo, Enrique Alegre and Laura Fernandez-Robles, “Object Detection for Crime Using Deep Learning.” Springer International Publishing AG 2017 S. Battiato et al. (Eds.): ICIAP 2017, Part II, LNCS 10485, pp. 14–24, 2017

[3] JianyuXiao, Shanchang Li (Member, IEEE), and Qinliang Xu, “Video-Based Evidence Analysis and Extraction in Digital Forensic Investigation ”. , 2169- 3536 2019 IEEE. Received April 1, 2019, accepted April 16, 2019, Date of publication April 26, 2019

[4] Mohammad NakibRozinTanvir Khan, Md. Sakibul Hasan, ” Crime Scene Prediction by Detecting Threatening Objects Using Convolutional Neural Network”. Department of Computer Science and Engineering, BRACC University,13th April 2017

[5] Weiyi Shi, "The Application of Image Processing in the Criminal Investigation". Advances in Computer Science Research, Volume 71, 4th International Conference on Machinery, Materials and Information Technology Applications (ICMMITA 2016).

[6] SulthanaAbdussamedKormath, Nafsila, Minhaj, ShanidMalayil,"Criminal Assault Analysis and Security Using Image Processing & Machine Learning". Volume 8, No.2, International Journal of information systems& Computer Sciences. March-April 2019

[7] Devishree D. S, Divakar K. M, Ashini K. A., ArnavSingh Bhardwaj, Mohammad Younis," Crime Scene prediction and analyzing its accuracy with frames using Deep Neural Network". Volume5,0Issue 2, International Journal of Advance Research,Ideas and Innovations in Technology, July 2016