Survey Paper on Deep Learning Based Distributed Smart Surveillance Architecture using Edge and Cloud Computing

Mr.Tipu Sultan[1], Prof.Shrikant Nagure[2]

[1]T.E.Computer Engineering, RMD School of Engineering,Warje,Pune,Maharashtra,India
[2]Asst.Prof.ofComputer Engineering, RMD School of Engineering,Warje,Pune,Maharashtra,India

---***---

***Abstract* - The** *smart surveillance aspect is one of the most important services provided in a smart city. This smart surveillance applications are equipped with camera sensors to AutoDetect and identify potential regions through automated object detection methods. This new method is less expensive and in this project it is used as a standalone platform for image processing. The area will be processed through a deep learning-based image Classification by Using Raspberry pi in order to ensure high efficiency and accuracy. It cannot only help to detect object rapidly and accurately, but also can reduce big data storage needed to be stored in smart surveillance systems. Traditional surveillance system records video footage to a storage device continuously.*

KeyWords: *Raspberry,deep learning,surveillancee* ,performance,neural system,computerized**.**

1.INTRODUCTION

In the developed and in members of the developing world, surveillance societies have started to emerge. Monitoring cultures are societies that function, in part, because of the extensive collection, recording, storage, analysis, and application of knowledge on individuals and groups in those societies as they go about their lives. Retail loyalty programs, website cookies, national identity schemes, routine health screening, and no-fly lists all qualify as surveillance. Each features, at different conditions, the conventional number of data concerning people with the specific purpose of administering, regulating, training or changing what they do in the future. Deep learning has achieved considerable thought in recent years and has given remarkable performance than previous handcrafted backgrounds in transportation type memory. for fifinegrained VTR, and introduced a convolutional neural system (CNN) method to identify vehicles from various angles. Another VTR system, in which the classification was conducted based on the mixture of social and global features derived from the CNN model, was introduced by Fang et al. . Sochor et al used a 3D bounding box to remove the vehicle pictures from the video streams, and then fed these picture to a CNN model in order to enhance the recognition performance. a faster region CNN model was used to identify vehicles, developed by another CNN to recognize the vehicle types. Although the above-mentioned works have achieved a high standard accuracy even for fine-grained vehicle type identification, these works heavily rely on big-scale interpreted surveillance vehicle type data. This paper proposes to review the recognition problem from the way of real-world importance where tagging sufficient data in surveillance pictures is extremely time-consuming and exhausting. Visual data provided by the surveillance cameras in the hospital and other healthcare environments can be used for search support and management. Liu et al proposed a computer image based system for following the position of hospital patients in the bed. An essential application of surveillance videos in the hospital is recognizing when patients have dropped to the ground. Automatic alert and quick detection and active response to a dropped patient can have significant results for the health of the patient, particularly for elderly and children. Fall detection based on hospital surveillance videos is studied by various researchers. Shahzad et al. introduced a fall detection algorithm based on various kernel learning. Abobakr et al. studied the application of random arrangement forests for error detection.

Why Smart Surveillance?

In our occupied life we don't have enough time to monitor and to keep a watch on everything. From every family most of the features are going, or even in malls and hospital to monitor each and every term is not possible. It's 21st century and we want to create smartly to perform our life better, easier and protected, so rather of sitting at once place for greater why not take the security in our pocket.

2. LITERATURE SURVEY

Boult et al., 1999, stated that video surveillance involves following an area for important events.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

Perimeter security frequently needs watching areas that support trespassers just cover and shield[8]. By definition, such "interesting" areas have restricted visibility. Furthermore, points of interest generally attempt to hide themselves within the cover sometimes adding camouflage to further decrease their clarity. Such marks are only visible "while in motion". The merged result of limited visibility distance and destination visibility severely decreases the value of any panning-based approach. As a result, these conditions call for a wide field of view, and are a natural application for omnidirectional VSAM (video surveillance and monitoring)[4].
They described an omnidirectional tracking system. They considered domain application restrictions and background on the paracamera. They used frame-rate Lehigh Omnidirectional Tracking System (LOTS) and explained some of its unique characteristics. In particular, the order combined production depends on adaptive multi-background modeling and a quasi-connected-components system. Those key elements are explained in some detail, while other parts are reviewed[7].

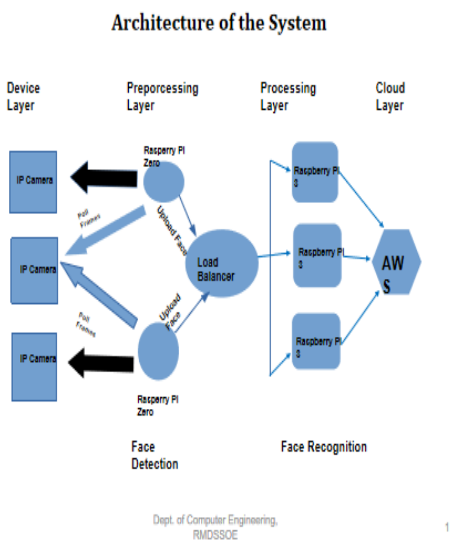
Weiming et al., 2004, declared that observed surveillance in changing scenes, particularly for humans and media, is currently one of the most current research topics in computer image. It has a extensive spectrum of possible uses, including access control in particular areas, human description at a range, crowd flux statistics and barrier division, discovery of unusual behaviors, and interactive surveillance using multiple cameras.
In general, the processing framework of optical surveillance in changing scenes includes the next stages: modeling of conditions, detection of motion, kind of moving objects, tracking, recognition and variety of behaviors, human association, and fusion of data from multiple cameras[12].

They examine current improvements and global policies of all these stages. Lastly, they investigated feasible study areas like occlusion handling, a mixture of two and three-dimensional tracking, a mixture of motion study and biometrics, anomaly discovery and behavior prediction, content-based
8 retrieval of surveillance videos, action recognition and natural language information, union of knowledge from various sensors, and distant surveillance.

Valera and Velastin, 2005, viewed the current situation-of-the-art in the construction of computerized visible surveillance systems so as to present researchers in the area with a review of progress achieved to date and to know areas where more research is required. The capacity to identify objects and humans, to explain their actions and communications from data received by sensors is required for automated visual surveillance[5].

The growing need for intelligent visual surveillance in commercial, law implementation and service demands makes automated visual surveillance systems one of the principal current application areas in computer vision. The importance of the study is on analysis of the creation of smart distributed automated surveillance systems. The survey decides with a review of possible expected places.
Mohamed et al., 2006, declared that optical surveillance systems have earned a lot of attention in the last several years. They performed a visual surveillance system that is based on the integration of change exposure and visual tracking to produce more reliable execution. Motion detection is performed using an algorithm that connects temporary variance with background modeling methods. The tracking algorithm mixes motion and condition information into an surface model and uses a particle filter structure for tracking the target in subsequent frames. The systems was examined on a wide ground-truthed data set including hundreds of color and FLIR image chains. A performance evaluation for the operation was performed and the ordinary evaluation decisions are presented.

1.2 ARCHITECTURE PROPOSED

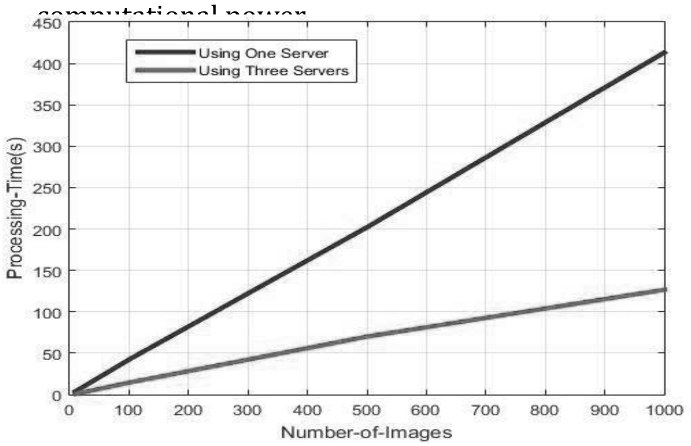


A. Device layer :- Device Layer (DL) contains an array of IP cameras. Web server in IP cameras receives for
incoming applications and serves video footage to various clients concurrently.System overview. RTP protocol is applied for this communication and OpenCV has native support to stream into RTP

B. Preprocessing layer :- Preprocessing layer (PPL) performs the related duties as Edge Layer in captures faces from raw footage from DL and feeds it to upper layer, PL.Main duty of Preprocessing layer is to decrease bandwidth by preparing sensor data before sending it to Processing layer. This also spreads computing power more regularly in the cluster. Processing Layer should only process what it receives from this layer rather than Device layer.

C. Processing layer :- Processing Layer (PL) accepts requests from Preprocessing layer. Requests contain detected faces and other metadata such as time of the capture, filename and camera location. Processing layer applies face recognition to received face against known people database. PL requires more computa- tional power than PPL because face recognition is done in this layer. If a match is found, frame is ignored and result of the recognition is saved to the Cloud database

D. Cloud layer :- Cloud Layer (CL) is used for two reasons: persist recog- nition results and perform another
face recognition. In order to persist recognition results and unknown faces, we chose a Cloud based database and object storage.Face recognition can be performed by custom Cloud deployments or using vendor's computer vision API offerings. Another way is to deploy a Tensorflow instance to the Cloud and perform further processing which may require more computational power



Recognition Processing Time (one server vs. three servers)

Images number	Processing time (seconds)
10	3.73
100	42.29
500	202
1000	413.11

Images number	Processing time (seconds)
10	0.81
100	14.41
500	96.97
1000	126.6

CONCLUSIONS

In the conclusion we conclude that every person wants to be in a better and secure world, this survey paper has covered most of the algorithms and the advantages and disadvantages of every paper that has declared so far and the work that has made over this design. To provide better security and safety new designs are completed and also that are cost-effective and to make it better further research is going on.in future we can make everything wireless in this project and also better video quality which will be easy to operate and in a way the security data can be carried and observed anywhere.

ACKNOWLEDGEMENT

We would like to thank our guide Prof.Shrikant Nagure for supporting and guiding us in our survey.

REFERENCES

[1] A. Geitgey. (2018, Sep) ageitgey/face recognition: The world's simplest facial recognition api for Python and the command line.

[2] A. A. Wazwaz, A. O. Herbawi, M. J. Teeti, and S. Y. Hmeed, "Raspberry pi and computers-based face detection and recognition system," in 2018 4th International Conference on Computer and Technology Applications (ICCTA).

[3] D. King. (2017, Jul) dlib c++ library: High quality face recognition with deep metric learning.

[4] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, “RTP: A Transport Protocol for Real-Time Applications,” Internet Requests forComments, RFC Editor, RFC 3550, July 2003.

[5] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, “Labeled facesin the wild: A database forstudying face recognition in unconstrained environments,” in Workshop on faces in’Real-Life’Images: detection,alignment, and recognition, 2008.

[6] P. Menezes, J. C. Barreto, and J. Dias, “Face tracking based on haar-like features and eigenfaces,” IFAC Proceedings Volumes, vol. 37, no. 8, pp. 304–309, 2004.

[7] Conference: Visual Surveillance, 1999. Second IEEE Workshop on, (VS'99).

[8] q IEE, 2005 IEE Proceedings online no. 20041147 doi: 10.1049/ip-vis:20041147.

[9] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, “RTP: A Transport Protocol for Real-Time Applications,” Internet Requests for Comments, RFC Editor, RFC 3550, July 2003. [Online]. Available: http://www.rfc-editor.org/rfc/rfc3550.txt

[10] M. Michael, J. E. Moreira, D. Shiloach, and R. W. Wisniewski, “Scale- up x scale-out: A case study using nutch/lucene,” in 2007 IEEE International Parallel and Distributed Processing Symposium, March 2007, pp. 1–8.