Big Data Image and Video Analysis Using Deep Learning

Chunchu Srinivasa Rao[1†] , Dr Venkateswara Rao Naramala[2†],
Dr V.Sitharamulu[3], Dr. Jujjuri. Rama Devi[4], Rama Krishna Paladugu [5]
[1]Assistant Professor in CSE (Data Science) department, R.V.R & J.C College of Engineering, Guntur, Andhra Pradesh, India.
{Mail Id: nivasaraman@gmail.com} **Orcid Id:** 0000-0001-6906-3657
[2] Professor in CSE (AL & ML) department, R.V.R & J.C College of Engineering, Guntur, Andhra Pradesh, India.
{Mail Id: vnaramala@gmail.com} **Orcid Id:** 0000-0002-4146-2119
[3] Associate Professor in CSE department, Institute of Aeronautical Engineering (IARE),Dundigal, Hyderabad .Telangana, India.
{Mail Id: vsitaramu.1234@gmail.com } **Orcid Id:** 0000-0002-3419-1074
[4] Sr. Assistant Professor in CSE department, PVP Siddhartha Institute of Technology,Vijayawada, Andhra Pradesh, India.
{Mail Id: k.ramakarthik@gmail.com } **Orcid Id:** 0000-0002-2747-624X
[5]Assistant Professor in CSE department, R.V.R & J.C College of Engineering, Guntur, Andhra Pradesh, India.
{Mail Id: mails4prk@gmail.com } **Orcid Id:** 0000-0002-0480-882X

Abstract
Thousands of closed-circuit television cameras are linked to the internet, according to IDC (International Data Corporation), supplying 65 percent of all streaming data, which is greater than data from transactions, medical records, entertainment, and social media combined. Streaming data is information generated in real time by thousands of data sources, which typically deliver data in small batches. This video and image stream contains useful information about traffic, weather, and the environment, among other things. Machine Learning takes data and turns it into a programme that can complete a task. Deep learning has demonstrated to be far more promising than the current system. Deep learning's Deep Convolution Neural Network (DCNN) is suitable for a wide range of image and video processing workloads.

Keywords:
Machine Learning, DCNN, Streaming, Deep Learning.

1. Introduction

The main source of big data is now video data. Because of the complexities, pace, and volume of video data, public security and other surveillance applications demand rapid, intelligent video processing in real time. [1]
In the field of video surveillance and intelligent camera systems, abnormal event detection is critical. The majority of existing approaches are not object-aware in the literature, meaning that they do not distinguish between distinct objects during processing. [2]
Hundreds of thousands of cameras are connected to the Internet, allowing data to be transmitted in real time (videos or periodic images). The photos contain information that can be used to figure out what's going on in the scene, such as traffic, weather, and the environment. Many challenges face analyzing data from these cameras, including (i) obtaining data from geographically dispersed and diverse cameras, (ii) creating a software environment that allows users to analyze massive amounts of data from the cameras at the same time, and (iii) allocating and managing processing and storage resources. Despite the fact that CCTV is the most commonly utilized method for traffic monitoring, issues develops, necessitating manual traffic flow checks. [4]. For a traceability system using deep learning, A video-based approach for detecting objections is proposed. The surveillance footage is first gathered, the convolutional neural network model is trained off-line after an annotated photo database of the target item, such as people or vehicles, is constructed. The trained model is used to build and implement a system for real-time target detection and recognition. For the traceability application, a deep learning-based detection technique is effective. [5]

With millions of public network cameras throughout the world catching numerous occurrences, The vast volume of a system to retrieve, save, and analyse visual data from cameras is required. The information gleaned from the data will eventually help us comprehend the planet better. In order to meet the needs of the analysis, such a system would need to allocate and manage a large number of resources. [6]
Because video surveillance data is rapidly growing, efficiently storing and querying large surveillance films, as well as querying performance and storage fault tolerance, is a challenge. Thanks to modern cloud computing and big data approaches, intelligent

Manuscript received April 5, 2022
Manuscript revised April 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.4.99

analysis of large-scale video data is now possible. Using an HBase-based method, process, store, and query surveillance footage. We employ a distributed storage architecture, slicing videos into small bits and storing them in HDFS, before extracting video data using Hadoop preprocessing. [7]
Because smart monitoring cameras combined with sophisticated video analytics algorithms can monitor and predict unusual behaviour or incidents, video surveillance has become an important part of modern city security and protection systems. Massive amounts of video data pose significant challenges for analytics, storage, and retrieval in the Big Data era as the surveillance network expands. [9].

As the network era approaches, video surveillance equipment is becoming more integrated with evolving IT systems. In the field of video surveillance, big data technology has advanced at a rapid pace. The goal of big data surveillance video data mining is to provide a solution for intelligent video surveillance in the big data era by integrating big data technology and video surveillance companies flawlessly. [10]

2. Literature Survey

Surveillance systems do not provide target identification, particularly in the given scenario. [5] {people, tractors, bicycles, cars etc}.Existing system can only detect events without bridging the spatial and temporal association of a number of unusual events. [9]. Time series data is well-suited to neural network architectures such as Long-short Term Memory [LSTMs] and Recurrent Neural Networks (RNNs). When it comes to training, hundreds of video clips are available, they are resource intensive and time consuming. CAM2 is a protein that plays a role in (Continuous Analysis of Many Cameras).Analyzing and streaming thousands of cameras is a challenge. CAM2 is the world's first and only open system for analysing real-time picture and video streams. [6] It pulls data from chosen cameras and is in charge of putting our analytic methods into action on streams.It is available at https://cam2.ecn.purdue.edu/ [6][3].
Other alternatives for camera streaming are http://www.webcams.travel[3]
www.wonderground.com/webcams [2].

The problem of expanding the typical stream processing framework for video analysis, particularly for situation awareness, must be addressed. Data model, operators, and vocabulary for representing complicated circumstances, as well as the QoS (Quality of Service) standards and algorithms required to fulfil them. Current data representation (e.g., relation and arrable) and querying capabilities are insufficient to infer long-term R&D difficulties. [11].

3. Problem statement

Massive surveillance video must be retained for months, if not years, resulting in high storage expenses. The number of false alarms is increasing, necessitating manual processing. There are no stream categorised groups in the current system. Deep learning can be used to solve all of these video surveillance camera limitations.

Objectives of the proposed work with justification:

- Perform Stream grouping for surveillance video.
- Loop through the entire video file.
- Use convolution neural network (CNN) functions to analyze.
- Develop model to target detection and event detection.

4. Proposed methodology

Use Convolution Neural Network (CNN) to analyze stream data as frames and use frameworks to perform analysis.
CAFFE platform for Deep learning, Convolution Architecture for Fast Feature Embedding (CAFFE) is a deep learning framework. It is an open source framework and developed by UCBerkely. [5] It supports many different types of deep learning architecture towards image classification, image segmentation, CNN etc. Yahoo has also integrated CAFFE with Apache spark to create CAFFE on spark-a distributed deep learning frame work. Analysis computing resources Analyzing a single image from 70000 cameras every minute equates to 9TB of visual data each day. It's critical to use the cloud for large data. [6] We can use different Amazon EC2 instances, CAM2 and CAFFE platform together for the study of mages and videos.

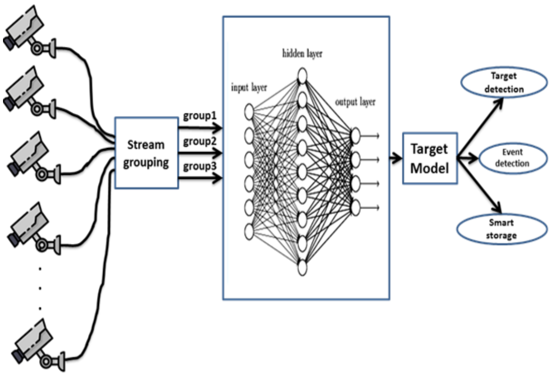

Fig. 1. Surveillance data stored and processed by CNN to generate target model for target detection

5. Experimental data analysis and results

a. Existing Work

For a long time, video analysis and fusion have been practised, and a variety of approaches have been created for this goal. As object recognition, categorization, and tracking have advanced to the point where they can be successfully used in commercial and other systems (for example, Google images, Facebook, image-based commerce, and the VIRAT project and so on), the focus has switched to situation awareness and fusion.
Perhaps pre-relational DBMS data processing and analysis approaches can be likened to the current level of video processing. To write searches and queries, both hierarchical and network DBMSs need the usage of a navigational programming language.
It can be thought of as video frame stream processing (one or more streams) with an acceptable representation that captures significant visual information.

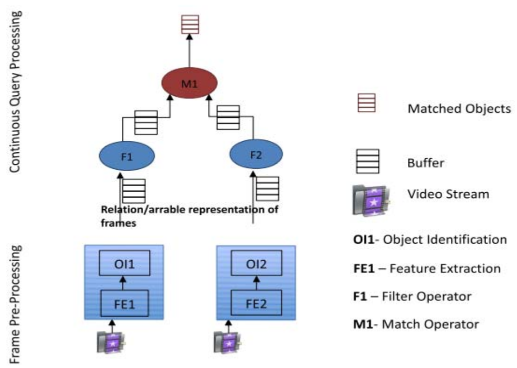

Fig. 2 Video Stream processing architecture using arable representation model

b. Proposed Work

Take any Action video datasets that are routinely used in action recognition studies or Surveillance stored data into consideration. There are no stream categorized groups in the current system.
To perform analysis, we consider surveillance video. By using Convolution Neural Network (CNN) to analyze stream data as frames and use framework to perform analysis.

Fig. 3 Frames generated using framework for a surveillance video [Source from https://viratdata.org/]

Stream grouping can be performed repeated over all the frames generated from the framework and apply CNN to the analyze the data. Use the target model to detect type of action in the video suspected as abnormal.

Perform Stream grouping is the proposed work and to analyze the relevant frames from the huge amount of surveillance videos. Very less and useful video from huge surveillance is stored for further processing.

Table.1 Comparison on surveillance video- action recognition using existing and proposed model

#UCF101 (UCF101 Human Actions dataset)	#Videos analyzed	#Modality	#Total No. of actions	#Actions Identified
Arable representation [11](Existing)	300	RGB	101	72
Two-Stream Inflated 3D ConvNet (I3D) [13] (Existing)	400	RGB	101	97
DTD Three-Stream CNN [14] (Existing)	100	RGB	101	98
Stream grouping using convolution Neural Network (Proposed)	420	RGB	101	100

UCF101 (UCF101 Human Actions dataset) is made up of real-life YouTube videos. There are 13320 videos in all in 101 activity categories. [12]

6. Conclusion

Expected outcome of the proposed system is to extract data from video like Object tracking in video, identify their properties, recognize people actions and events., situation awareness that lead for adequate immediate reaction.

References

[1] Weishan Zhang, Liang Xu, Zhongwei Li, and QinghuaLu,:A Deep Intelligence Framework for Online Video Processing,, IEEE Software (Volume: 33, Issue: 2, Mar.-Apr. 2016)

[2] XianghaoZang, Ge Li, Zhihao Li, Nannan Li, WenminWang:An Object aware Anomaly Detection and Localization in Surveillance Videos, , 2016 IEEE Second International Conference on Multimedia Big Data (BigMM)

[3] Ahmed S. Kaseb, Everett Berry, YoungsolKoh.:A System for Large-Scale Analysis of Distributed Cameras, 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), 3-5 Dec. 2014.

[4] HyeongsoonIm; Bonghee Hong; Seungwoo Jeon; JaegiHong :Bigdata analytics on CCTV images for

collecting traffic information 2016 International Conference on Big Data and Smart Computing (BigComp), 18-20 Jan. 2016.

[5] Bing Tian; Liang Li; Yansheng Qu; Li Yan: Video Object Detection for Tractability with Deep Learning Method, 2017 Fifth International Conference on Advanced Cloud and Big Data (CBD), 13-16 Aug. 2017.

[6] Ahmed S. Kaseb; Anup Mohan; Yung-Hsiang Lu: Cloud Resource Management for Image and Video Analysis of Big Data from Network Cameras, 2015 International Conference on Cloud Computing and Big Data (CCBD), 4-6 Nov. 2015.

[7] Weishan Zhang; Yuanjie Zhang; Liang Xu; Faming Gong: Hbase Based Surveillance Video Processing, Storage and Retrieval, 2016 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI), 20-21 Oct. 2016.

[8] Haitao Zhang; Bin Xu; Jin Yan; Lujie Liu; Huadong Ma: Proactive Data Placement for Surveillance Video Processing in Heterogeneous Cluster, 2016 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), 12-15 Dec. 2016.

[9] Zhenfeng Shao; JiajunCai; Zhongyuan Wang: Smart Monitoring Cameras Driven Intelligent Processing to Big Surveillance Video Data, IEEE Transactions on Big Data (Volume: 4, Issue: 1, March 1 2018).

[10] Honghua Xu; Ming Fang; Li Li; Ye Tian; Yanfeng Li: The value of data mining for surveillance video in big data era, 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), 10-12 March 2017.

[11] S. Chakravarthy, A. Aved, S. Shirvani, M. Annappa1, and E. Blasch2Adapting Stream Processing Framework for Video Analysis, ICCS 2015 International Conference On Computational Science,Volume 51, 2015.

[12] A. R. Z. KhurramSoomro and M. Shah, “Ucf101: A dataset of 101 human action classes from videos in the wild,” 2012, cRCV-TR-12-01.

[13] Joao Carreira, Andrew Zisserman, Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset, CVPR, 22 May 2017 (v1), last revised 12 Feb 2018 (this version, v3)]

[14] Yemin Shi; Yonghong Tian; Yaowei Wang; Tiejun Huang, Sequential Deep Trajectory Descriptor for Action Recognition With Three-Stream CNN, IEEE Transactions on Multimedia, Volume 9 Issue 7, 2017.

Chunchu Srinivasa Rao received the B.Tech degree from J.N.T.U Hyderabad in 2008 and M.Tech. degree from Acharya Nagarjuna Univ. in 2012. Working as a an assistant professor (from 2012) in the Dept. of Computer Science and Engineering, R.V.R & J.C College of Engineering, Guntur, Andhra Pradesh, India. His research interest includes Big Data Analytics, Machine Learning and Data Science. He is a member of ACM Professional and life member of IAENG.

Dr Venkateswara Rao Naramala received the PhD degree from Acharya Nagarjuna University in 2017 and M.Tech. degree from Andhra University in 2004. Presently working as a Professor in the Dept. of Computer Science and Engineering (AI & ML) , R.V.R & J.C College of Engineering, Guntur, Andhra Pradesh, India. His research interest includes Artificial intelligence, Machine Learning and Pattern recognition. He is a life member of ISTE and member of ACM.

Dr V.Sitharamulu Associate Professor, Department of Computer Science and Engineering,Institute of Aeronautical Engineering (IARE) Dundigal, Hyderabad - 500 043,Telangana, India.

Dr.Rama Devi Jujjuri received the B.Tech degree from RVR & JC College, Guntur in 2004, M.Tech. degree from Andhra University, Visakhapatnam in 2006 and and her Ph.D. in 2020 from Gitam University in Visakhapatnam. Working as a Sr. Assistant Professor in the Dept. of Computer Science and Engineering, PVP Siddhartha Institute of Technology, Vijayawada ,Andhra Pradesh, India. Her research interests include Data Mining, Machine learning, Deep learning and Data Science. She is a member of ACM Professional, Faculty coordinator of PVPSIT ACM Student Chapter and life member of ISTE.

Rama Krishna Paladugu received the B.Tech degree from J.N.T.U Hyderabad in 2005 and M.Tech. degree from Acharya Nagarjuna Univ. in 2007. Working as a an Assistant Professor in the Dept. of Computer Science and Engineering, R.V.R & J.C College of Engineering, Guntur, Andhra Pradesh, India. His research interests include Machine learning, Deep learning and Data Science. He is a member of ACM Professional and life member of IAENG.