IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

DEEP LEARNING BASED VIDEO STREAM ANALYTICS WITH CLOUD UPDATES

[1]Murali C, [2] Mr. N Praveen ME,
[1]$\overline{M.E}\,(CSE)$, [2]Assistant Professor,
[1]Department of Computer Science and Engineering,
[1]RVS Educational Trust's Group of Institutions, Dindigul, India

Abstract: The aim of this work is to design and investigates the problem of video stream analytics by proposing recommended solutions. In today's world, The Internet of Things (IoT), which connects gadgets like sensors and video cameras, is becoming more and more popular as people use the internet recurrently. As a result, massive volumes of streaming data are being generated quickly. A quick and effective analysis of this data is necessary for applications that need low latency response, like augmented reality, video surveillance, and autonomous cars. Current methods store and process this data using machine learning-based analytics by utilizing cloud infrastructure. Restrictions on network capacity and latency between the data source and cloud make it difficult for this centralized strategy to offer real-time analysis of large-scale streaming data. A system for large-scale, high-performance data analytics, called Real Edge Stream (RES), is the work that is being suggested.

The proposed approach investigates the problem of video stream analytics by proposing (i) Filtration and (ii) Identification phases. The filtration phase reduces the amount of data by filtering low-value stream objects using configurable rules. The identification phase uses deep learning inference to perform analytics on the streams of interest. The project consists of stages which are mapped onto available in-transit and cloud resources using a placement algorithm to satisfy the Quality of Service (QoS) constraints identified by a user.

I. INTRODUCTION

Video cameras are the most versatile form of IoT devices. The number of video cameras have grown at an unprecedented rate to increase public safety and security around the globe. France and UK have 1 million and 6 million CCTV cameras installed respectively. They are currently being monitored by human personnel who incessantly stare at a grid view screen to monitor actions or events of interest. Video surveillance has traditionally been performed by operators watching one or more video feeds, the limitations of which are well known. About 90 percent of incidents are missed after 20 minutes as the concentration of the CCTV operators drops.

II. NEED OF THE STUDY.

1.1. OBJECT DETECTION- AN OVERVIEW

Object detection is a phenomenon in computer vision that involves the detection of various objects in digital images or videos. Some of the objects detected include people, cars, chairs, stones, buildings, and animals. The major concept of YOLO is to build a CNN network to predict a (7, 7, 30) tensor. It uses a CNN network to reduce the spatial dimension to 7×7 with 1024 output channels at each location. YOLO performs a linear regression using two fully connected layers to make 7×7×2 boundary box predictions. To make a final prediction, we keep those with high box confidence scores (greater than 0.25) as our final predictions.

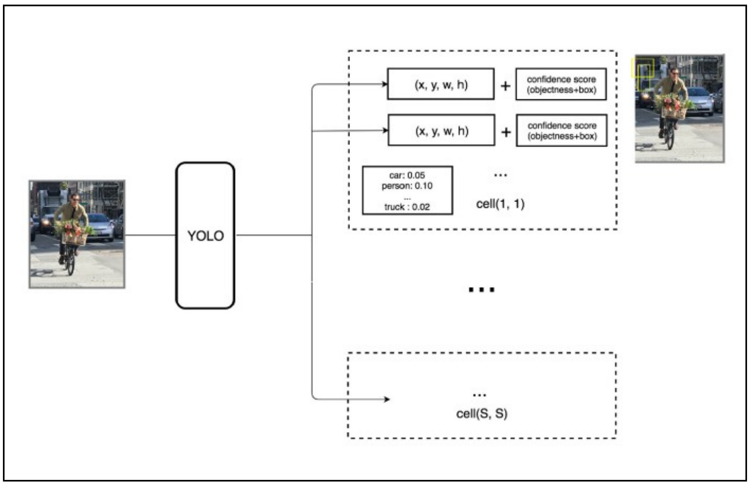


Figure 1.1 Overview- YOLO detection

Compared to other region proposal classification networks (fast RCNN) which perform detection on various region proposals and thus end up performing prediction multiple times for various regions in an image, Yolo architecture is more like FCNN (fully convolutional neural network) and passes the image (nxn) once through the FCNN and output is (mxm) prediction. This the architecture is splitting the input image in mxm grid and for each grid generation 2 bounding boxes and class probabilities for those bounding boxes. Our network uses features from the entire image to predict each bounding box. It also predicts all bounding boxes across all classes for an image simultaneously. This means our network reasons globally about the full image and all the objects in the image. The YOLO design enables end-to-end training and real time speeds while maintaining high average precision.

1.2. PROBLEM STATEMENT

Video analytics relieves the operator by automating the surveillance process and can proactively detect, recognize and respond to events coming from video streams. However, performance and accuracy are the two major concerns in video analytics. The accuracy of the analytics process has been improving by leaps and bounds by novel approaches in deep learning models over recent years. Analytics using deep learning demand powerful compute and storage resources which are offered by many existing cloud platforms. Therefore, most existing video stream analytics systems employ cloud infrastructure to perform operations such as object classification and recognition on the incoming data streams. However, performance is a major concern for cloud-based stream analytics systems due to network latency and limited bandwidth.

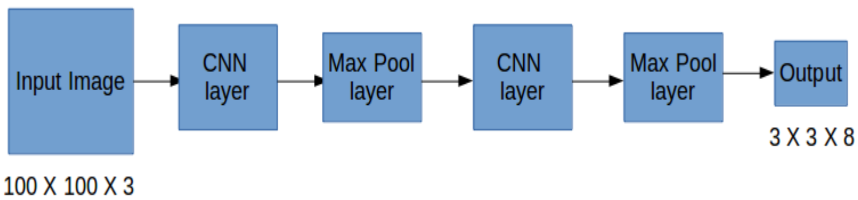
III. *MODULES:*

There are 4 Modules:

- Object Detection
- Find Object location
- Find Object movement
- Update to cloud

1.3.1 OBJECT DETECTION

In object detection, module, the input of video stream is considered. The input video stream is identified for objects using R-CNN techniques and it primarily use regions to localize the objects within the image. The network does not look at the entire image, only at the parts of the images which have a higher chance of containing an object. YOLO first takes an input image from video stream. The framework then divides the input image into grids. Image classification and localization are applied on each grid. YOLO then predicts the bounding boxes and their corresponding class probabilities for object.



During the testing phase, we pass an image to the model and run forward propagation until we get an output y.

1.3.2 FIND OBJECT LOCATION

Object localization from the input video stream can be identified. Image classification from object detection, which goes through a ConvNet that results in a vector of features fed to a SoftMax to classify the object. Neural network has a few more output units that encompass a bounding box. In particular, we add four more numbers, which identify the x and y coordinates of the upper left corner and the height and width of the box (bx, by, bh, bw).

1.3.3 Find object movement

The Python's built in deque datatype to efficiently store the past N points the object has been detected and tracked at. Libraries imutils also used by collection of OpenCV and Python convenience functions. By checking the X, Y co-ordinates values and tracking them, the object movement can be detected.

1.3.4 Update to Cloud

Similarly, the machine learning stages for object detection were executed on two cloudlets followed by an object filtration stage. Finally, object recognition is performed on the cloud platform for the filtered frames. This configuration uses all the in-transit nodes from the data source to the destination and routes the data to the cloud platform.

IV. CONCLUSION

An architectural approach for supporting scalable real-time video stream processing using edge and in-transit computing. Current approaches to stream analytics are based on the use of centralized cloud platforms. With the increase in data volumes and velocity, a centralized analysis approach of this kind becomes infeasible for high-performance applications due to limited uplink bandwidth, variable latency and congestion between the data sources and the cloud platform. This approach to video analysis consists of a filtration phase followed by an identification phase. The filtration phase allows objects of low-value to be filtered (discarded) by the edge and in-transit nodes using configurable rules from a user. The identification phase performs deep learning inference on the objects of interest.

V. References

[1] Abdullah, A. Anjum, M. F. Tariq, Y. Baltaci, and N. Antonopoulos, "Traffic monitoring using video analytics in clouds," in Proc. IEEE/ACM 7th Int. Conf. Utility Cloud Comput., 2014, pp. 39–48.

[2] Chen, T.-S. Chen, T.-W. Huang, L.-C. Yin, S.-Y. Wang, and T.-C. Chiueh, "Intelligent urban video surveillance system for automatic vehicle detection and tracking in clouds," in Proc. IEEE 27th Int. Conf. Adv. Inf. Netw. Appl., 2013, pp. 814–821.

[3] Kafka, "A high-throughput, distributed messaging system" 2014 (Online) URL : kafka.apache.org

[4] Pereira, M. Azambuja, K. Breitman, and M. Endler, "An architecture for distributed high performance video processing in the cloud," in Proc., 2010, pp. 482–489.

[5] Ryu, D. Lee, M. Jang, C. Kim, and E. Seo, "Extensible video processing framework in apache hadoop," in Proc. IEEE 5th Int. Conf. Cloud Comput. Technol. Sci., 2013, vol. 2, pp. 305–310.

[6] Shan, F. Porikli, T. Xiang, and S. Gong, Video Analytics for Business Intelligence. Berlin, Germany: Springer, 2012.