THINK INDIA JOURNAL

ISSN:0971-1260
Vol-22- Issue-12-November-2019

Traffic Analysis Using Yolo Algorithm

[1]Shaina Gupta, [2]Manik rakhra

[1]Department of Computer Science and Technology, Lovely Professional University, Phagwara, India.
[2]Department of Computer Science and Technology, Lovely Professional University, Phagwara, India.
[1]Email:-shaina.244968@lpu.co.in
[2]Email:-Manik.23538@lpu.co.in

ABSTRACT- In the present era one of the biggest problems faced by world is traffic management, as there is a growth in population every year, as a result the number of vehicles manufactured are also increasing adding to the existing traffic, With vehicles hitting the roads increasing day by day and the roads merely being with same dimensions the vehicle density on the road has increased .This results in increase in the number of traffic jams being caused. The average time spend by an individual in traffic has also increased. Many lives can be saved by proper traffic analysis and management as many people die because of delayed arrival of ambulance due to the traffic .There can be many ways that this problem can be solved, namely by increasing the dimensions of road and limiting the number of vehicles per individual etc. Which require a lot of investment and changes in the law. Our goal is to reduce this problem to a certain extent by efficient traffic analysis which helps in better management which was not possible in past because of lack of advancements in technology.

INDEX TERMS- Traffic analysis, deep learning, YOLO

I. INTRODUCTION

Traffic Analysis is deep learning based system that helps regulate traffic more efficiently than the system that is in effect in the present world it also helps in gathering various other types of information like number of vehicles, Type of vehicles, vehicle Density and peak of time of the traffic and even identify whether any animals are causing any blocks etc.. These other components when computed can yield even more accurate results which can help prevent long traffic jams. Eventually saving lot of time for people and decrease in air pollution. Our goal is to achieve a better approach in traffic analysis and there by changing the way the traffic is being managed, Our system also tries to reduce the amount of manpower. We take the data from already existing hardware like cctv footage from the cameras already existing and process this data in real time using machine learning to produce efficient result. This can be done with the help of real time object detection methods.

Object recognition is a key innovation behind cutting edge driver help frameworks (ADAS) that empower vehicles to recognize driving paths or perform person on foot discovery to improve street security. Item discovery is additionally helpful in applications, for example, video observation or picture recovery frameworks. In object discovery undertakings we are keen on discovering all item in the picture and drawing purported jumping boxes around them. There are likewise a few circumstances where we need to discover accurate limits of our articles in the process called occasion division.

II. RELATED WORK

Joseph Redmon *et. al.* presented YOLO which is a unified model for object recognition. This model is easy to develop and can be prepared legitimately on full pictures. Not at all like classification system based methodologies, YOLO has been trained on loss function which directly corresponds to identification execution and the whole model is prepared together. YOLO additionally sums up well to new spaces making it perfect for applications that depend on quick, hearty item location. Junyan Lu *et. al.* in their paper introduced a vehicle discovery strategy dependent on YOLO profound learning calculation for ethereal picture is displayed. This strategy coordinates an airborne picture dataset appropriate for YOLO preparing by handling three open datasets. The preparation model has great test outcomes particularly for little articles, turning objects, just as minimized and thick items, and meets the ongoing prerequisites. Next, we will incorporate progressively open elevated picture datasets to expand the number and assorted variety of preparing tests, simultaneously, enhance the YOLO calculation to additionally improve the identification precision.

Peiming Re *et. al.* presented YOLO-PC, a YOLO based continuous individuals tallying approach utilizing limit

ISSN:0971-1260
Vol-22- Issue-12-November-2019

selection. YOLO-PC beats YOLO as it re-trains YOLO arrange, which empowers it to identify more boxes and accomplish higher normal confidence esteem. The limit choice in YOLOPC makes the checking more focused on and its outcome exact and quick. Taking everything into account, this strategy is viable and it is additionally ready to perceive superfluous individuals and overlook them in the checking procedure. YOLO-PC has a wide scope of uses as it can help the improvement of numerous parts of the keen urban communities. On the present premise, we anticipate making more enhancements to this technique by including, for instance, anomalous conduct discovery and kids tallying.

III. CNN ARCHITECTURE

CNNs, which are based on neural networks, constitutes of neurons with the capability to allocate weights depending upon various biases through learning. Each neuron obtains inputs from several sources and generates a weighted sum over the obtained inputs. These obtained weighted sums are further passed to an essential activation function which generates an output . The entire neural network has a loss function. All the inherent qualities pertaining to a neural network can also be found in a CNN. An artificial neural network consists a set of processing units which have been interconnected among each other [Bis95, p. 80-81]. Given a set of input values w0, w1,.., wD, where w0 represents an external input and x1,.., xD are inputs obtained from the processing units which are present within the neural network. The output of processing unit is computed as y = f(z). Where, the activation function f is passed with the weighted sums from inputs passed to neurons z. The output of the function is obtained by utilizing a propagation rule which is applied through the network to map the inputs from various sources to the essential input z. This model of single processing unit defines a single neuron where instead of a propagation rule, the computation of z is done with the help of an adder which results in the weighted sum of all inputs. The visualization of various neural networks can be done through a directed graph3 called network graph.

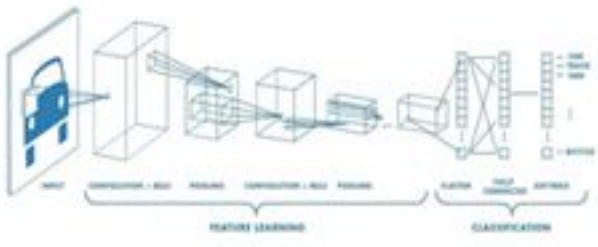

Figure 1:CNN Architecture

Each unit is distinguished through a node which has been labeled according to the output it generates. These units are interconnected among themselves through directed edges. For a single individual processing unit, it has been represented in figure 1 in which the external output has been used for illustration where it would have naturally been omitted. For the sake of convenience the input and output values are to be distinguished. An input unit computes an output y :=x through utilizing a function. These outputs units may arbitrarily accept input values of varying numbers. Here, the network represent a function y(x) whose dimensions are regulated by number input and output units, which essentially constitute the input of the network is accepted through the input units and the output of the network through the output of the network as in figure 4 .. Artificial NeuralNetworks (ANNs) are a design of computational processing systems which take their fundamental architectural ideology through biological nervous of the ecology such as the human brain. ANNs essentially constitute of a high number of computational nodes which are connected among themselves and are termed as neurons for general use. These neurons are entwined in a split among multiple things approach to learn from the input with the goal to optimize the final output. The fundamental structure of an Artificial Neural Network can be controlled to generate a desired model. An input which will be loaded generally containing a multidimensional vector will be passed to the input layers where it will be distributed among the hidden layers. The decisions made by hidden layers will depend upon the output from a previous layer. These layers calculate or compute the change about how the change in stochastic nature within itself changes the condition generally in favor to improves the final output, and this process of making decisions depending on previous data is termed as learning. The process of stacking multiple hidden layers upon themselves is commonly referred to as deep learning. The fundamental architecture of ANN's constitutes a simple three layered feed forward neural network (FNN) which is further subdivided into an info layer, a shrouded layer and a yield layer. This structure is the reason of a couple of fundamental ANN plans, accomplice to Feed-forward Neural Networks (FNN), Restricted Boltzmann Machines (RBMs) and Recurrent Neural Networks (RNNs) are among a couple regardless of the way that there are a couple of something else. The two fundament displaying methods in picture based preparing or recognition assignments are administered and unaided learning. Supervised learning is the process of generating a model through using data which has been pre-labelled as inputs, where these labels act as targets for the model. Each set of training examples will contain input values which are likely vectors which will also have a corresponding single or multiple output values depending upon the nature of the data. The philosophy behind such a preparation procedure to produce models is to decrease the characterization mistake

ISSN:0971-1260
Vol-22- Issue-12-November-2019

created through exact estimation of the preparation model present in the preparation. Despite what might be expected preparing set of solo learning information does exclude any names. Achievement is dictated by the model's capacity to streamline the cost capacity whether to increment or abatement itIn any case, its essential be note that by far most of the image based model affirmation assignments ordinarily depend upon classification through regulated learning. Convolutional Neural Networks (CNNs) are exceptionally like customary ANN as far as their working, since they are worked of neurons which self-streamline themselves through learning. Be that as it may, each neuron will in any case play out the key errand getting an info and play out an undertaking related with it, of which the most broad activity is a calculation of result of scalar finished by a nonlinear characterized work which is the premise of incalculable ANNs. From the beginning of the network which constitutes of a simple input of raw image vectors to the final output of the class score, the entirety of the network will still possess and output a score function of a single perception which is most generally associated as the weight, however it should be noted that the perception may change depending upon the nature of the network and configuration. The loss functions which are associated with their respective classes are present in the last layer ofthe network. It should be noted that the general functions pertaining to neural networks are all applicable. The primary huge unmistakable differentiation among CNNs and standard ANNs is that CNN's are on a very basic level related for the usage of example acknowledgment inside pictures, basically to the picture preparing field. This permits us the capacity to encode picture specific highlights into the neural design, making the system progressively appropriate for picture centered assignments. However, further reduction in parameters is required for the proper set up of a model. A Primary limitation for using tradition artificial neural network is their complexity in processing image data. MNIST database of manually written digits which is utilized for testing most picture based AI models are reasonable for most types of ANN, because of its generally little picture dimensionality of only 28×28. While utilizing this information a solitary neuron in the primary shrouded layer will contain 784 loads (28×28×1, it ought to be noticed that the MNIST information of pictures is standardized to simply high contrast esteems), which is sensible for most types of ANN. However, If we consider a more substantial colorized image input constituting of images with size 64×64, this results in a dramatic increase in the weight associated to a single neuron which is about 12, 288. It should be noted that in order to deal with data which is scaled to such proportion, the network necessary to process said data will need be increases respectively which will be a lot larger than the one used to classify the MNIST digits. Hence, the drawbacks of the traditional ANN's can be understood .

IV. YOLO ALGORITHM

YOLO which means "you just look once" is a picture location calculation which changes object discovery into a solitary relapse issue, which makes jumping boxes and class probabilities from the picture pixels in a single run of the calculation. The calculation accomplishes this by separating the information picture into a S ×S lattice, in which if the focal point of an item falls into a network cell, the said cell is designated the activity for the location of the article present in the cell.

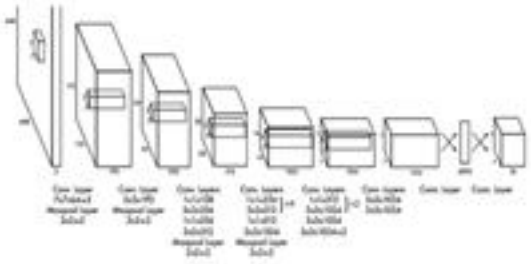

Figure 2:One Layer Of YOLO execution engine

Every boundary box has 5 notations associated with it represented as: (x, y, w, h) along with a box confidence score. The confidence score is a reflection of the likelihood an object being present in the corresponding boundary box and how accurate is the mentioned boundary box in figure 2.The boundary box is normalized by changing the box's width (w) and height (h) in accordance with the image's width and height. Here x and y are offsets to the corresponding cell. Hence, x, y, w and h are all in the range of 0 and 1. Every cell has 20 conditional classes probabilities associated with it, which is the probability that the object present in the box belongs to the said class (one probability per category for each cell). So, YOLO's prediction has a shape of (S,S,B×5+C)=(7,7,2×5+20) =(7,7,30). The class confidence score associated with each prediction box is computed as: class confidence score = box confidence score * conditional class probability. The class confidence score is used to measure both confidence score for classification and localization (location of the object).

V. PROPOSED FRAMEWORK

ISSN:0971-1260
Vol-22- Issue-12-November-2019

The proposed calculation for said venture is called YOLO which means 'you just look once' , a reframed way to deal with object location. Past work on object location used repurposed classifier for discovery of items. Be that as it may, in this methodology we develop object recognition as a relapse issue to spatially isolate bouncing boxes and related class probabilities. One single neural system is utilized to foresee jumping boxes alongside related class probabilities in a single assessment of the calculation. The total of the location pipeline due to being a solitary system can be productively advanced start to finish dependent on identification execution.

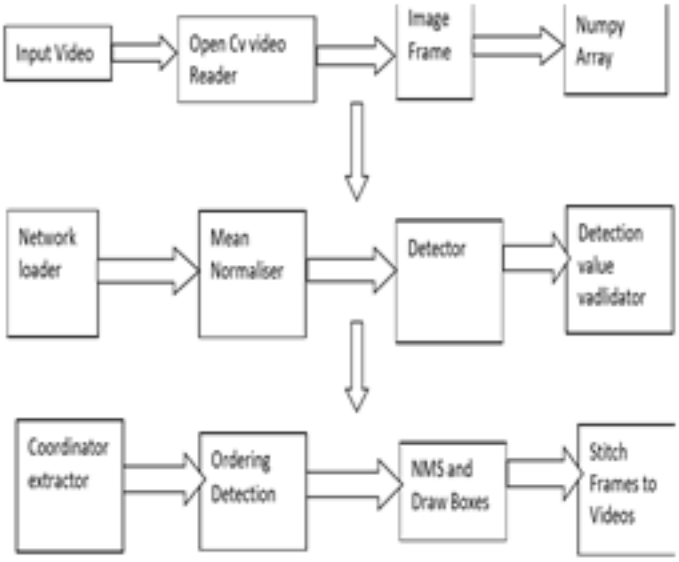

Figure 3:Flowchart of proposed work

Because of the innate relapse nature, the brought together design is very quick. The base YOLO model has a capacity to forms pictures continuously at pace of 45 fps and does as such. A reclassified littler variant of the system named Fast YOLO, forms pictures at the pace of 155 fps while as yet getting and holding twofold the mAP of other continuous locators. Contrasted with the very good quality recognition frameworks working in the discovery field, YOLO makes more confinement mistakes in any case, it is far less inclined to anticipate bogus identifications where the cells are unfilled, and no picture exists. YOLO learns through general portrayals of articles. It beats the more typical and spread-out discovery techniques, including DPM and RCNN, by an all-encompassing edge during the speculation of common pictures to work of art. The YOLO is model executed as a convolutional neural system. The assessment of the thusly produced model is performed on the PASCAL VOC discovery dataset. The underlying layers of the convolutional layered system perform extraction of highlights from the picture. While, the completely associated layers anticipate the yield probabilities and directions. The engineering of the produced YOLO model is essentially enlivened from the Google Net model which is additionally utilized for picture characterization. The created convolutional neural system model comprises of 24 convolutional layers alongside 2 completely associated layers. We utilize a 1 × 1 decrease layers pursued by 3 × 3 convolutional layers in contrast with the beginning modules used by the GoogLeNet. The full system is spoken to in Figure 3. A quicker form of YOLO model is planned and prepared to push the limits of quick item discovery to the greatest conceivable degree. So as to get said speed and make the system quick, a Fast YOLO model uses a NN establishing of a diminished number of convolutional layers (9 rather than 24) and diminished number of channels in those layers similarly. With the special case to the size of the system, all parameters related with preparing and testing are the comparative in all cases among YOLO and Fast YOLO. The last yield of the system is an establishing a 7 × 7 × 30 foreseeing tensor. The ImageNet 1000 class dataset has been utilized to pretrain the convolutional layers. To pretrain the system we use the starting 20 convolutional layers which have been spoken to in the Figure 3 pursued by using a normal pooling layer and completion with a completely associated layer. The system has been permitted to prepare for a total seven day stretch of time and it accomplished a solitary harvest top-5 exactness of 88% on the ImageNet 2012 approval set, which is generally similar to the GoogLeNet models in the Caffe's model zoo. It is then changed over so as to perform recognition. Renetal. shows that joining the convolutional layers alongside associated layers to pre prepared systems can improve execution barely well. Appropriately, we consolidate four convolutional layers with two completely associated layers which have haphazardly instated loads. Detectionrequires fine-grained visual information of images. Subsequently, the addition to the information goals of the system requires a change from 224 × 224 to 448× 448. The subsequent layer can anticipate the class probabilities of both alongside directions of the jumping box. The jumping box's stature and width is standardized by the picture width and tallness which changes their range to fall somewhere in the range of 0 and 1. The jumping box's x and y facilitates are parameterized so they are counterbalances of a specific lattice cell area which makes them likewise to be limited somewhere in the range of 0 and 1.

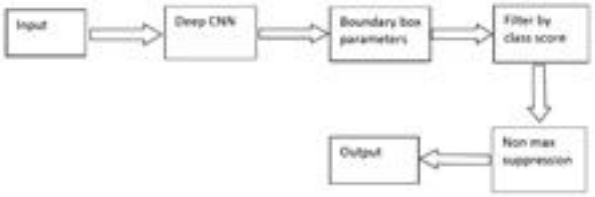

Figure 4:Flow of Prediction Function

ISSN:0971-1260
Vol-22- Issue-12-November-2019

VI. CONCLUSION

This paper we introduced a real time object detection algorithm that can be used for detecting vehicles on road in order to yield even more accurate results which can help prevent long traffic jams. Eventually saving lot of time for people and decrease in air pollution. .We can also count the number of vehicles passing through a signal for a particular period of time and can predict what are peaks timings and what type of vehicles are likely to arrive and how can we regulate it. Our system can also detect any real time possible obstacles like if the road is being blocked by animals like dog, cows etc.

REFERENCES

[1] M. B. Blaschko and C. H. Lampert. Learning to localize objects with structured output regression. In Computer Vision–ECCV 2008, pages 2–15. Springer, 2008. [2] L. Bourdev and J. Malik. Poselets: Body part detectors trained using 3d human pose annotations. In International Conference on Computer Vision (ICCV), 2009.

[3] H. Cai, Q. Wu, T. Corradi, and P. Hall. The cross-depiction problem: Computer vision algorithms for recog-nising objects in artwork and in photographs. arXiv preprint arXiv:1505.00110, 2015.

[4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recogni-tion, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, pages 886–893. IEEE, 2005.

[5] T. Dean, M. Ruzon, M. Segal, J. Shlens, S. Vijaya-narasimhan, J. Yagnik, et al. Fast, accurate detection of 100,000 object classes on a single machine. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Confer-ence on, pages 1814–1821. IEEE, 2013.

[6] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. arXiv preprint arXiv:1310.1531, 2013.

[7] J. Dong, Q. Chen, S. Yan, and A. Yuille. Towards unified object detection and semantic segmentation. In Computer Vision–ECCV 2014, pages 299–314. Springer, 2014.

[8] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Confer-ence on, pages 2155–2162. IEEE, 2014.

[9] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual ob-ject classes challenge: A retrospective. International Journal of Computer Vision, 111(1):98–136, Jan. 2015.

[10] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9):1627–1645, 2010.

[11] S. Gidaris and N. Komodakis. Object detection via a multi-region & semantic segmentation-aware CNN model. CoRR, abs/1505.01749, 2015.

[12] S. Ginosar, D. Haas, T. Brown, and J. Malik. Detecting people in cubist art. In Computer Vision-ECCV 2014 Workshops, pages 101–116. Springer, 2014.

[13] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 580–587. IEEE, 2014.

[14] R. B. Girshick. Fast R-CNN. CoRR, abs/1504.08083, 2015.

[15] S. Gould, T. Gao, and D. Koller. Region-based segmenta-tion and object detection. In Advances in neural information processing systems, pages 655–663, 2009.