IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-8, Issue-1, Jan 2020 E-ISSN: 2347-2693

Yolo Deep Learning Model Based Algorithm for Object Detection

Amandeep Kaur[1*], Deepinder Kaur[2]

SUS Tangori, India

DOI: https://doi.org/10.26438/ijcse/v8i1.174178 | Available online at: www.ijcseonline.org

Accepted: 12/Jan/2020, Published: 31/Jan/2020

***Abstract*-** With the growth of deep learning and digital image processing, it is require knowing about the facts of deep learning. Deep learning is major factor of object detection. In existing research R-CNN algorithm used for detect the objects from an image while in this proposed research method we are using YOLO (you only look once) algorithm to detect the different objects in a single image. This is less time consuming because we only recognize a image once and we detect the whole objects in an image.

***Keywords*-** Digital image processing, image recognition, Accuracy, time complexity, histogram, K-MEANS, YOLO

I. INTRODUCTION

Object detection is use to detect the instances from an image. We only detect some instances from a class of image. Class of image means similar type of instances they we detect easily lie people, car etc. that is visible by direct eyes [1]. Many possible locations are not visible directly it only detect by the use of computer objects. Researcher did work from many years on object detection from an image. The first efficient detector algorithm introduced by Viola-Jones in 2001. They detect the face on web camera in real time [2]. That was the first stunning demonstration in computer vision.

1.1 Category Objects Detection:

Object detection method divided into two categories:

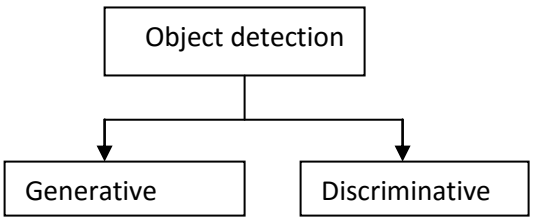

Fig 1.1 categories of object detection

Generative Object detection: It is probability-based model in which some conditions are applying on the pose that similar to background. In this model, decision based on posterior probability that how much pose are identical.[1]

Discriminative Object detection: This model builds a classifier that is finding the discrimination between the sub-images or images which containing the objects or which are not. This classifier is using to minimize the mistakes of over fitting.

For these models, computational tools used which scan the whole image and search the possible poses from scanned image.

1.2 Methods of Object Detection:

For Object detection, generally two methods used:

1. Machine Learning
2. Deep Learning

These are the two main approaches use for object detection.

1. **Machine Learning:** In Machine Learning approach
   - First, define what type of features having an image [2]. These feature approaches are:
     i. detection framework based on Haar Features
     ii. Scale invariant Feature Transform(SIFT)
     iii. Histogram of oriented gradients(HOG) features
   - Apply the SVM (Support Vector Machine) Technique

2. **Deep Learning:** For end-to-end Object Detection, Deep Learning method is use without defining the features [2]. This learning based on the Convolutional Neural Network (CNN) Deep Learning Approaches are:
   i. Region Proposals
   ii. Single Shot MultiBox Detector
   iii. You only look once

II. REVIEWOF LITERATURE

Various authors in the past have done studies related to object detection, machine-learning methods object detection Algorithm, various Deep Learning Algorithm. Few studies reviewed by me concerning my research proposal are as follows:

[1] Geethapriya. S et.al(2019) researched on object detection by using single neural network to detect the boundary of the image, but this researcher only detect a

single object from one picture not detect the all pictures this is the main drawback of this research.[17]

[2] Wenging Chu et.al(2018) focused on object detection with the combination of R-CNN with CRF. It detects those objects, which are inside the image, but is also work on regions. By using CRF on R-CNN for improve the accuracy. In addition, make object detection regions more visible and divided into categories [7].

[3] Spyros Gidaris et.al(2016) proposed a novel object localization technique by use LocNet architecture. LocNet provides the information of boundary of the object inside the search region.
It improved on mAP for IoU thresholds on PASCAL VOC2007. Thus, it provides the independent box proposal method for set of sliding window by the use of LocNet[8].

[4] Gemma Roig et.al (2015) researched on multi class object detection in multi camera field by the use of CRF (Conditional Random field). In this research researcher used CRF with energy minimization model to handle the optimization problem when detection of multiple objects in multiple cameras.

[5] Navneet Dalal et.al (2005) found the machine learning algorithm HOG (Histogram of Oriented Gradient) for human detection. This researcher worked on large range of pose variation on human images. It form a new database, which is improve the spatial binning, high normalization etc. features in human detection [10].

[6] Ren S.He et.al (2015) worked on region proposal algorithm to find the object location. It worked on region proposal network (RPN) that shares convolutional features of the image which predict object bound at each position. In this researcher used the Fast R-CNN for object detection in which end-to-end regions provided by the RPN[11].

[7] Wei Liu et.al(2016) worked on object detection by the use of deep learning algorithm SSD(single shot multibox detector). SSD is use for eliminating proposal generation, resample the stages for detection and give the result in full computational single network. In this researcher compare the accuracy the dataset which computed by R-CNN with SSD [12].

[8] Joseph Redmon found a new deep learning approach YOLO for object detection. YOLO process the image in 45 frames per second. It performs all method including DPM and RCNN for wide margin between the regions of the image. All processing done in a single pipeline no need for different regions in an image. It improves the time complexity for the object detection.

[9] Albert Soto i Serrano examined that with the evolution of AI car companies are researching new features of cars. In this, they used deep learning algorithm for new changes in automotive sector. For improving the feasibility for driving assistance or autonomous driving, they used YOLO with neural network for the object detection. It improved the feasibility of object detection from a running car [14]

[10] Z. Shen et.al(2014) found an edited version of CNN for maximum entropy for a particular vector. The particular vector not connected to image category it differs from the image features. This advance version used when limited training data is available for improve the generalization. It improved the mean average precision with SVM [15].

[11] R. B. Girshicks(2015) examined Fast Region Based Convolutional network method (Fast R-CNN) for improve the testing speed and training and improve the detection accuracy of the object. It trained the VGG16 network for improvement in speed ,less time complexity and also 9times faster than the previous algorithm R-CNN[16].

[12] Sandeep Kumar et.al[2017] used EASYNET model in digital image processing for detection and reorganization object in images. It detected those images, which captured by single shot while it is clear or blur. It only detects the front object and subtracts the background image. Thus, it only detects the front objects.

[13] Şaban Öztürk et al[2018] studied extraction algorithm in object detection how to extract the useful results. Object detection detect the all objects which are not useful this technique is use only those which object are useful by using GLCM, LBP, LBGLCM, GLRLM and SFTA technique and image parts are classified by SVM, KNN, and Boosted tree. This technique is very time consuming and very costly.

[14] Shi Na et al.[2016]found that clustering analysis method is one of the main analytical methods in data mining, the method of clustering algorithm will influence the clustering results directly. This paper discusses the standard k-means clustering algorithm and analyzes the shortcomings of standard k-means algorithm, such as the k-means clustering algorithm has to calculate the distance between each data object and not all cluster centers in iteration make the efficiency of clustering is high. This paper proposes an improved k-means algorithm in order to solve this question, requiring a simple data structure to store some information in iteration, which used in the next iteration. The improved method avoids computing the distance of each data object to the cluster centers repeat, saving the running time. Experimental results show that the improved method can effectively improve the speed of clustering and accuracy, reducing the computational complexity of the k-means.

[15] Marcin Wo ´zniak , Dawid Połap[2018] focused on only 2D object detection via clustering algorithm which only detect and recognize the 2D object with CNN Algorithm. It

only detect the clustered features but sometimes it show difficulty to show clear objects in an image.

[16] Khushboo Khurana[2013] found the technique for detect the multiple objects from an image either the object at the front or in the background. This researcher used the Scale-Invariant Feature Transform (SIFT) technique for detecting the multiple objects from a single object. But this technique is not much useful in real life it did not detect the clear objects and drawback of this technique it did not detect the clear objects.

[17] Ross Girshick [2015] proposed Fast R-CNN and SPPnet technique for object detection and recognize the objects. It improves the quality of detection and improves the threads in previous results. But this technique is too costly for detection because it used two technique.

[18] Şaban Öztürk et al[2018] studied extraction algorithm in object detection how to extract the useful results. Object detection detect the all objects which are not useful this technique is use only those which object are useful by using GLCM, LBP, LBGLCM, GLRLM and SFTA technique and image parts are classified by SVM, KNN, and Boosted tree. This technique is very time consuming and very costly.

III. PROPOSED WORK

In this, we discussed the implementation of new algorithm for detection and recognition of K-MEANS and Faster R-CNN algorithm.

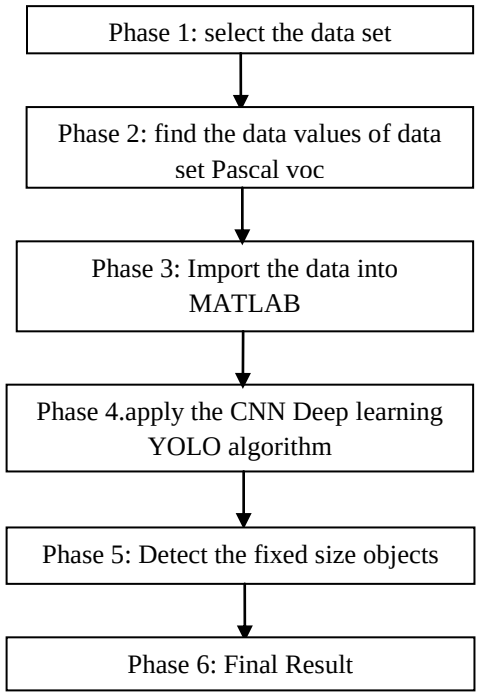

Figure 3.1 Proposed work

Figure 3.1 steps of current research

IV. CONVOLUTIONAL NEURAL NETWORK (CNN)

4.1 Convolution:

- CNN's main is convolution. It is a mathematical term, which is use to merge two functions and make third set.
- In above CNN definition three functions depth. Height and width make a single vector in which all things or we can say all features defined.
- Convolution is mainly use in image transformation from 2D matrix image to 3D matrix.ss

Convolutional Neural Network is Deep Learning based method. Convolutional neural network is different from the neural network [4]. In Neural Network input transform through hidden layer while CNN input depends upon the three dimensions:

1. Height of the object
2. Width of the object
3. Depth of the object

All of these upper dimensions organized in a single Box.

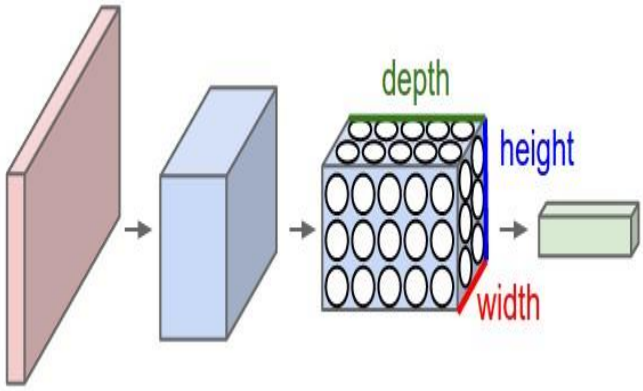

Figure 2: "The red input layer holds the image, so its width and height would be the dimensions of the image, and the depth would be 3 (Red, Green, Blue channels)".

Source: http://cs231n.github.io/convolutional-networks/
When everything is complete all dimensions reduced into a single vector depth dimensions.

V. DEEP LEARNING ALGORITHM

Deep Learning Algorithm having advance set of models as compare to the Machine Learning in which we define first the features which is also known as unsupervised learning but we can say deep learning is a supervised learning in which every features is explained and neural network give the proper build model.
These are following approaches, which used to object detection from an image in CNN:

1. **AlexNet :** This is the first deep learning algorithm approach which introduce in 1980's when CPU starts doing work on neural network then AlexNet is use to

speed up the CPU[6]. This approach is old but it still use for all neural tasks either it is a computer vision or speech recognition.

2. **GoogleNet:** This model formed by Google. It contains 22 layers for end model[6]. This model called Inception Model.

3. **Region Proposals:** The following are Region Proposal approaches:
   i. **R-CNN:** Ross Girshick proposed a method which work on huge region data. This Model work on an image, which divided into 2000 regions.
   ii. **Fast R-CNN:** It is the enhancement of the previous model by same research. In this, we did not need of 2000 regions. It directly works on convolution operation sson image and feature map generated.
   iii. **Faster R-CNN:** This model is same as upper region based model only difference in this it takes less time and give high speed for image detection, face recognition etc.

4. **SSD (single shot multibox detector):** SSD is use eliminating the need of region. It also recovers the accuracy and improves the low-resolution image.

5. **YOLO:** YOLO stands for YOU ONLY LOOK ONCE. All the previous algorithms only detect the object within the image region. This algorithm is much more different from the region-based algorithms. In this, a single convolution network show in bounding boxes and finding the probability of these boxes.

VI. RESULTS AND DISCUSSION

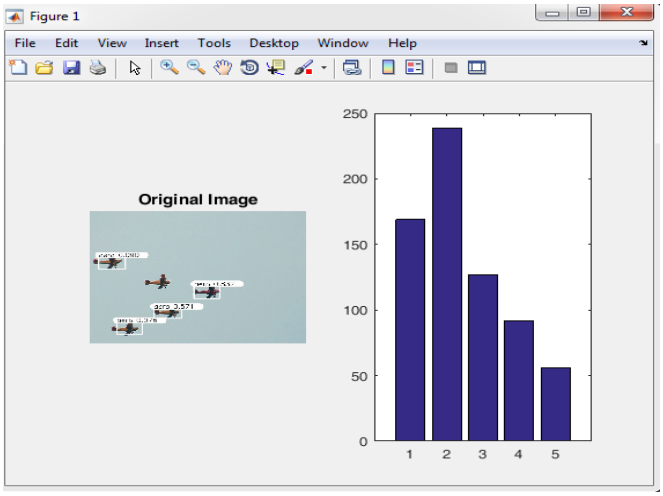

Figure 6.1 proposed image histogram

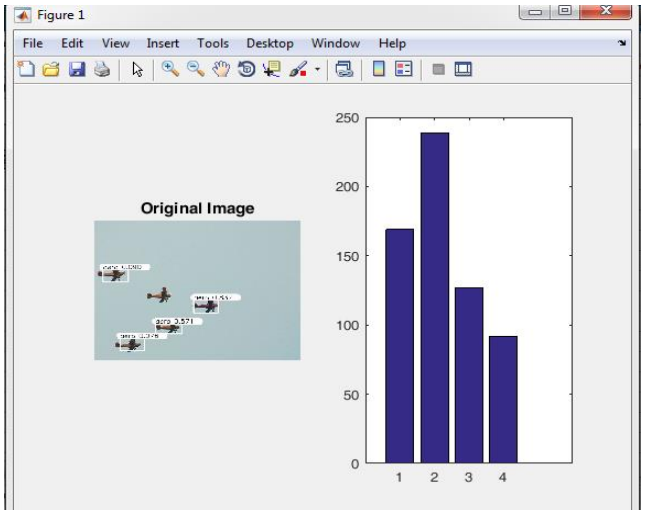

Figure 6.2 exiting image histogram

In the following figure time complexity of different objects detected with previous algo R-CNN and proposed method of YOLO algorithm

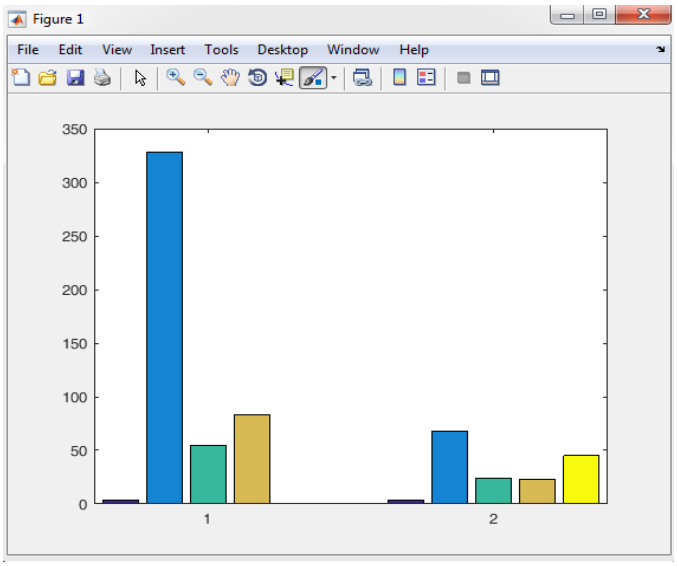

Figure 6.3 compare time complexity object detection

VII. FUTURE SCOPE AND CONCLUSION

In proposed research we are try to find the all objects from a single detection by using YOLO algorithm. This algorithm is useful because of image detection can handle easily and all objects are detected once in this YOLO algorithm. In this technique we improve the time complexity and accuracy of an image in which we are detect the objects.

REFERENCES

[1] Yali Amit and Pedro Felzenszwalb, University of Chicago
[2] Ross, Girshick (2014). "Rich feature hierarchies for accurate object detection and semantic segmentation" (PDF). Proceedings

of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE. doi:10.1109/CVPR.2014.81
[3] "ETHZ - Computer Vision Lab: Publications". Vision.ee.ethz.ch. Retrieved 2013-10-09.
[4] Jussi Jokela "PERSON COUNTER USING REAL-TIME OBJECT DETECTION AND A SMALL NEURAL NETWORK" TURKU UNIVERSITY OF APPLIED SCIENCES
[5] http://cs231n.github.io/convolutional-networks/
[6] ImageNet Classification with Deep Convolutional Neural Networks by Alex Krizhevsky University of Toronto kriz@cs.utoronto.ca Ilya Sutskever University of Toronto ilya@cs.utoronto.ca
[7] Wenqing Chu, Deng Cai " Deep feature based contextual model for object detection" Neurocomputing 275, 1035–1042, **2018**
[8] S. Gidaris , N. Komodakis , Locnet: improving localization accuracy for object de- tection, in: Proceedings of the IEEE Conference on Computer Vision and Pat- tern Recognition, pp. 789–798, **2016.**
[9] Gemma Roig_1;2 Xavier Boix_1;2 Horesh Ben Shitrit2 Pascal Fua2 "Conditional Random Fields for Multi-Camera Object Detection"1ETHZ, Zurich (Switzerland) 2EPFL, Lausanne (Switzerland)
[10] Navneet Dalal, Bill Triggs. Histograms of Oriented Gradients for Human Detection. Cordelia Schmid and Stefano Soatto and Carlo Tomasi. International Conference on Computer Vision.
[11] Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: NIPS, **2015.**
[12] Wei Liu1, Dragomir Anguelov2, Dumitru Erhan "SSD: Single Shot MultiBox Detector" UNC Chapel Hill 2Zoox Inc. 3Google Inc. 4University of Michigan, Ann-Arbor 1wliu@cs.unc.edu, 2drago@zoox.com.
[13] Joseph Redmon University of Washington pjreddie@cs.washington.edu Santosh Divvala Allen Institute for Artificial Intelligence" You Only Look Once: Unified, Real-Time Object Detection"
[14] Albert Soto i Serrano" YOLO Object Detector for Onboard Driving Images" , ESCOLA D'ENGINYERIA (EE), UNIVERSITAT AUTONOMA DE BARCELONA (UAB)
[15] Z. Shen and X. Xue. Do more dropouts in pool5 feature maps for better object detection. arXiv preprint arXiv:1409.6911, **2014.**
[16] R. B. Girshick. Fast R-CNN. CoRR, abs/1504.08083, **2015.**
[17] Geethapriya. S, N. Duraimurugan, S.P. Chokkalingam, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, **Volume-8**, **Issue-3S**, **February 2019.**

