

Object Detection through CNN with Deep Learning

Gogireddy Venkata Ashok
Reddy
Pursuing Final Year
Computer Science and
Engineering
KLEF Deemed University,
Vaddeswaram, Andhara Pradesh

Nerella Ganesh Naga Sai
Pursuing Final Year
Computer Science and
Engineering
KLEF Deemed University,
Vaddeswaram, Andhara Pradesh

Potturi Teja
Pursuing Final Year
Computer Science and
Engineering
KLEF Deemed University,
Vaddeswaram, Andhara Pradesh

Senthil Kumar A. M.
Associate Professor,
KLEF deemed to the University
Vaddeswaram, Andhara Pradesh

ABSTRACT

Object detection from images and videos is the main point in the applications of artificial intelligence and computer vision like self-driving cars, robotics etc. In this paper, we have proposed a way to detect the objects in images and videos by a new pre-training strategy through convolutional neural network with deep learning. We are using the ReLU, pooling and fully connected layer methods to increase the accuracy in detecting the objects and the number of detecting objects has increased. We have used coco database in which it has different types of object names with its threshold which are highly used for detecting the objects. We have used 3 different ways of input for detecting the objects which are images, videos and live camera. The algorithm used is regression. We have used YOLO v3 which uses the single neural network and divides the image into regions and predicts the objects.

Keywords

Deep Learning

1. INTRODUCTION

The process of identifying an object from the input. The required format is terminated into three types they are image input, video input and web cam. Object identification is the procedural search of the objects by the frame rate of the pixels by the shape of the picture [1]. There are certain weights which are to be calculated for the identification of different types of objects.

Coco names is the database used in this process [2]. This database consists of a greater number of objects other than any other database supposed in the identification of the objects. Coco names containing the weights which were used in the detection of objects.

A Conventional Neural Networks using Deep learning is the well utilized and best appropriate method in identification of the objects.

The basic methodologies used in CNN for the further process of the detection of the objects.

1. ReLU
2. Pooling
3. Fully Connected Layer

Therefore, In the process of detecting the objects, the objects are classified into two types. Three types they are image input, video input and web cam [3]. Long distance objects and short distance objects, Long distance objects are very small in pixel size where there are some difficulties in this type of objects [1]. While Short distance objects are easy to identify which has the large size of pixels. The modules of object detection are

1. Input
2. Image detailing
3. Extraction of pixel
4. Max polling
5. Fully connected layer
6. Recognized object detected output

The YOLO version3 is major algorithm performed for the operation.

2. OBJECT DETECTION

Object detection through the image has the main input is used for the classification of the objects. The image is stored in the system. Another type of input is recorded video type [4]. While other input method is web cam, Live access of the surroundings can be treated as the input. The object detection has used in Conventional neural network, in this algorithm input under goes three phases as discussed in previous.

The object detection has to maintain the several functionalities in the time of classification of objects [5]. Image has to observe the terms like lighting, scale, pixel range, weather, camera angle. Thus, the various functions have to be taken under consideration [2]. More over the class of the objects has the priority in detection of the objects.

The image and objects are termed into two class object type.

Single class object: The objects in the image are clear and single layered objects.

Multi class object: Image contains the objects in different layers and difficult to retain the object.

When coming to the challenges we have to face two major terms in the process of detection [2].

A. Image challenges:

Challenges through the image is major thing to be consider. If there are any mistakes in the image the whole process goes wrong [6]. So, the challenges through image are priority in process. Lighting, scale, pixel range, weather, camera angle, mode, detailing and color of the image are major challenges from the input. The technical challenges are noise, blur effects and editing.

B. Processing challenges:

Maintenance of weights of the particular objects with uniqueness without combining one weights of object to other weights of the object[9]. Acquiring the high accuracy rate from the objects which are detected.

To overcome the processing challenges in the process threshold weight range accuracy and NMS threshold are used.

3. LITERATURE SURVEY

To undergo this process, some literature survey has taken place by the some of the articles published by the scholars. By that research some of the functionalities has be taken and some are modified according to process mobilities, atmosphere and functions [4]. The some of the concepts which are most used and most accurate one. That are to be given to support the process.

1. R-CNN:

R-CNN has the separate structure rather than the Conventional neural networks. That structure is upon the region extraction, CNN feature computation and classify of regions [8]. By the usage of R-CNN process the object identification method is just 58.5%. The object through comparing with 5000 sample images present in the system storage.

The methods involved in the R-CNN are as follows:

a) Region proposal Generation:

At selected areas about 2000 region partition happens for each image [1]. These is meant for finding the spaces in the object's identification.

b) CNN Based Deep Feature Extraction:

In this part of process, the image undergoes of the high pixel range dimensions mobility with the 4096-dimensional feature image representation for the final region recognition.

c) Classification and Localization:

This part of the process under go with the high level of localization patriated as background image and foreground image.

2. YOLO:

YOLO algorithm is one of the object identification methods are object detection through the regression method.

This YOLO algorithm came up with the extraction of the Conventional Neural Networks, while it has processing details of the max polling of the two fully connected layers [1]. That has the colour matrix shape detailing and the bound box level weights. Which are compared with the confidence weights. The accuracy rate of the YOLO algorithm is 66.4%. Where the Faster R-CNN is 73.2%. When it comes to the probability range 0.732.

4. FEATURE EXTRACTION

In Object detection, the feature extraction is the main algorithm of the process is the basic input image to objects detected output image[10]. Deep learning Conventional neural networks is the algorithm used for the image classification.

Deep learning Conventional neural networks gives the best classification other than the embedding visualization and data augmentation [7]. In the CNN the main term of the image classification is based on the layers present in the image. Later the image pixel and sizes are taken under consideration [6]. These are the main terms in the image classification. So, the CNN is most used algorithm in image classification.

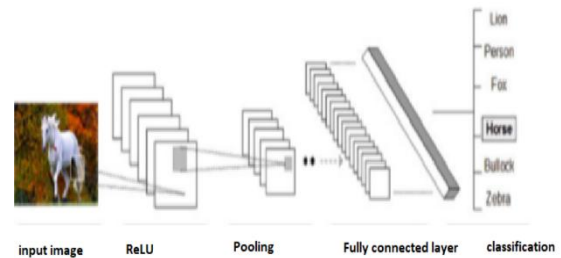


Fig 3.1 Image processing layers.

A. Convention Neural Networks

CNN are classification terms for various objects. In image classification the CNN algorithm has the visual processing images of the according to the layer format [7]. The connections of the layers are divided into the pixel range of the image.

B. ReLU:

Rectified linear unit is the mathematical equation function in the detailing of the layers which were divided in CNN layers.

The functional maintenance of the origin of the material in the region of the smoothing of the objects in the form of both non initiative and initiative methods.

$$f(x) = \ln(1 + e^x)$$

$$f'(x) = \frac{d(\ln(1 + e^x))}{dx} = \frac{e^x}{1 + e^x} = \frac{1}{1 + e^{-x}}$$

The output of the activation function of the ReLU is based upon the max of 0 and x value [5]. Function derived has below.

$$\max(0, x)$$

C. Pooling layers:

The pooling layers major step in the classification of the terms is to find out the major aspects of the in-object detection. The pooling is also a mathematical functional authority [10]. Matrix form of the image pixel are found from the polling layers. In this matrix equation, the width and height are major terms.

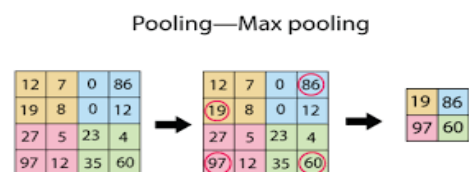




Fig 3.2 Max Pooling Layers.

the range in probability level. While Data augmentation give the accuracy level of 0.75[2]. Coco names has the highest database for the object detection.

Table 1.1: Comparison Results.

Method used Faster R-CNN	Method used YOLO V3
	
Long distance object identification is less compared to YOLO V3	Long distance object detection better than the Faster R-CNN
Accuracy rate 0.75	Accuracy rate 0.97
Database training of the objects comparatively less	Database training of the objects are comparatively more

The above table Faster R-CNN results are found from various research sources.

7. REFERENCES

[1] The Object Detection Based on Deep Learning Cong Tang, Yunsong Feng, Xing Yang, Chao Zheng, Yuanpu Zhou.

[2] Assessment of Object Detection Using Deep Convolutional Neural Networks. Ajeet Ram Pathak, Manjusha Pandey, Siddharth Rautaray and Karishma Pawar.

[3] Multi-target detection in cctv footage for tracking applications using deep learning techniques. A.Dimou, P.Medentzidou, AlvarezGarc, P.Daras, Senior Member IEEE.

[4] Object Detection With Deep Learning: A Review. Zhong-Qiu Zhao , Member, IEEE, Peng Zheng, Shou-Tao Xu, and Xindong Wu , Fellow, IEEE.

[5] P. F. Felzenszwalb et al., “Object detection with discriminatively trained part-based models,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, pp. 1627–1645, Sep. 2010.

[6] Anthony C Davies and Sergio A Velastin, “Progressin computational intelligence to support cctv surveillance systems,” International Journal of Computing.

[7] Alex Krizhevsky, Ilya Sutskever, and Geoff Hinton. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. (2012).

[8] I. Goodfellow, Y. Bengio, and A. Courville.: Deep Learning. MIT Press. (2016).

[9] Liu, A. Ranga, et al, “DSSD: De convolutional Single Shot Detector,”ar Xiv preprint arXiv:1701.06659, 2017.

[10] R. Salakhutdinov, G.E. Hinton.: Deep boltzmann machines. In: AISTATS, (2009)