International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 08 | Aug - 2021 ISSN: 2582-3930

Survey paper on Object Identification and Classification Using Deep Neural Network

[1]Rutuja Chimurkar, [2]Pratiksha Katkar, [3]Sanskruti Raut, [4]Vishakha Datar, [5]Prof. Manisha Pise

[1234] Student, [5] Lecturer@

Department of Computer Science & Technology

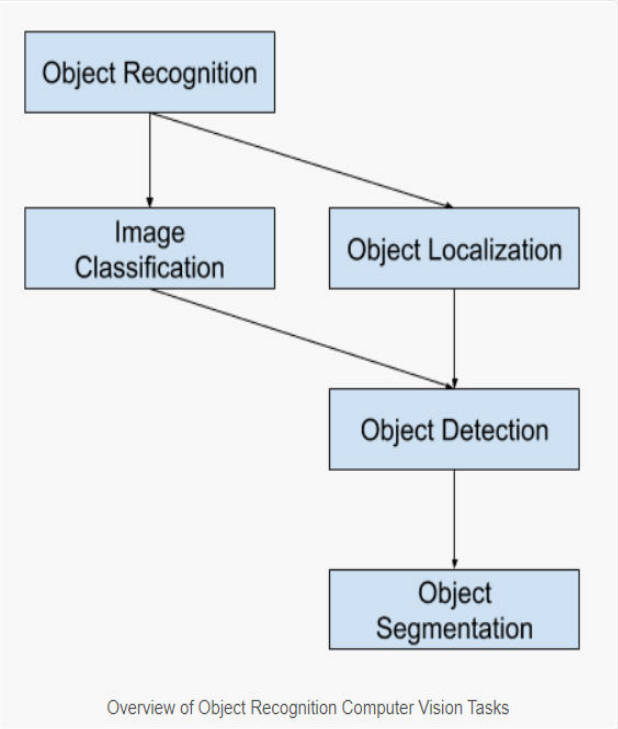
Rajiv Gandhi College of Engineering, Research and Technology, Chandrapur, Maharashtra, INDIA.

Abstract:

Object detection is a process to detect a object and recognize it with its name. Object detection means object identification i.e., to identify the object's class. And object classification means to classify it int its class and label it. Object identification is done by computer vision. Previously it was done with the help of R-CNN deep neural network algorithm, but we have used Fast-R-CNN deep neural network algorithm. We have used datasets to train our model. Our task is to develop an algorithm to classify images of 90 most commonly found objects into different classes by its labels(names). We have also used open CV (open source computer vision).It is the leading open source library for computer vision, image processing and machine learning.

Introduction:

Object detection is detecting a specific object from an image of multiple and complex lines and shapes. Every object class has its own special features that helps in classifying the class. Object class detection use this feature. Example: when looking for circles, we see objects that are at a particular distance from a point i.e., center. Similarly, for face identification we search for eyes, nose, ear and features like skin color and distance between eyes can be found. Here in this project, we will have three tasks. They are as follows: 1. *Image classification* involves predicting the class of one object in an image. 2. *Object localization* refers to identifying the location of one or more objects in an image and drawing abounding box around their extent. 3. *Object detection* combines these two tasks and localizes and classifies one or more objects in an image. One further extension to this breakdown of computer vision tasks is *object segmentation*, where instances of recognized objects are indicated by highlighting the specific pixels of the object instead of a coarse bounding box.



Overview of Object Recognition Computer Vision Tasks

Overview of our project:

Our proposed "Object Identification and Classification Using Deep Neural Network is used in face detection, object tracking, image retrieval, automated parking systems [12]. The number of the applications is increasing in number. The main use of object detection is image classification or more precisely image retrieval. The above system can also used for security and surveillance, advanced driver assistance systems, also known as ADAS, and many others.

Aim & Objectives:

The aim of object detection is to detect all instances of objects from a known class, such as people, cars or faces in an image. Generally, only a small number of instances of the object are present in the image, but there is a very large number of possible locations and scales at which they can occur and that need to somehow be explored. Each detection of the image is reported with some form of pose information. This is as simple as the location of the object, a location and scale, or the extent of the object defined in terms of a bounding box. In some other situations, the pose information is more detailed and contains the parameters of a linear or non-linear transformation. For example, for face detection in a face detector may compute the locations of the eyes, nose and mouth, in addition to the bounding box of the face. The objective of this project is to develop an object recognition system to recognize the 2D and 3D objects in the image.

Algorithms:

Deep Neural network: Deep neural network represents the type of machine learning when the system uses many layers of nodes to derive high-level functions from input information. It means transforming the_data into a more creative and abstract component. There are different types of neural networks and the differences between them lies in their work principles, the scheme of actions, and the application areas. Convolutional neural networks (CNN) are mostly used for image recognition, and rarely for audio recognition. It is mostly applied to images because there is no need to check all the pixels one by one. CNN checks an image by blocks, starting from the left upper corner and moving further pixel by pixel up to a successful completion. Then the result of every verification is passed through a convolutional layer, where data elements have connections while others don't. Based on this data, the system can produce the result of the verifications and can conclude what is in the picture.

Existing system:

ResNet:

To train the network model in a more effective manner, we herein adopt the same strategy as that used for DSSD (the performance of the residual network is better than that of the VGG network). The goal is to improve accuracy. However, the first implemented for the modification was the replacement of the VGG network which is used in the original SSD with ResNet. We will also add a series of convolution feature layers at the end of the underlying network. These feature layers will gradually be reduced in size that allowed prediction of the detection results on multiple scales. When the input size is given as 300 and 320, although the ResNet–101 layer is deeper than the VGG–16 layer, it is experimentally known that it replaces the SSD's underlying convolution network with a residual network, and it does not improve its accuracy but rather decreases it.

R-CNN:

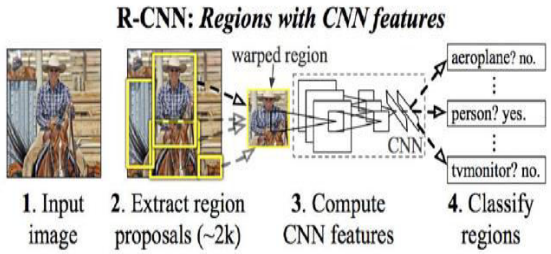
To circumvent the problem of selecting a huge number of regions, Ross Girshick et al. proposed a method where we use the selective search for extract just 2000 regions from the image and he called them region proposals. Therefore, instead of trying to classify the huge number of regions, you can just work with 2000 regions. These 2000 region proposals are generated by using the selective search algorithm which is written below.

Selective Search:

1. Generate the initial sub-segmentation, we generate many candidates' regions

2. Use the greedy algorithm to recursively combine similar regions into larger ones

3. Use generated regions to produce the final candidate region proposals.



Problems with R-CNN:

It still takes a huge amount of time to train the network as you would have to classify 2000 region proposals per image.

It cannot be implemented real time as it takes around 47 seconds for each test image.

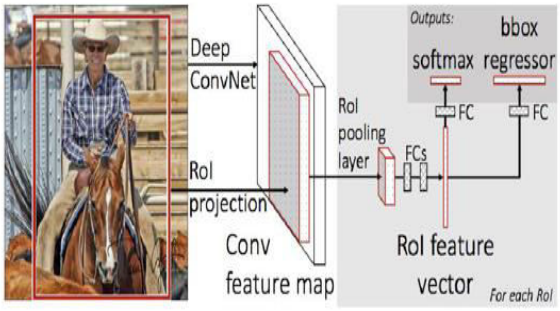
The selective search algorithm is a fixed algorithm. Therefore, no learning is happening at that stage. This could lead to the generation of bad candidate region proposals.

Proposed System:

Fast R-CNN:

The same author of the previous paper(R-CNN) solved some of the drawbacks of R-CNN to build a faster object detection algorithm and it was called Fast R-CNN. The approach is similar to the R-CNN algorithm. But, instead of feeding the region proposals to the CNN, we feed the input image to the CNN to generate a convolutional feature map. From the convolutional feature map, we can identify the region of the proposals and warp them into the squares and by using an RoI pooling layer we reshape them into the fixed size so that it can be fed into a fully connected layer. From the RoI feature vector, we can use a SoftMax layer to predict the class of the proposed region and also the offset values for the bounding

box. The reason "Fast R-CNN" is faster than R-CNN is because you don't have to feed 2000 region proposals to the convolutional neural network every time. Instead, the convolution operation is always done only once per image and a feature map is generated from it.



Training time:

Testing time:

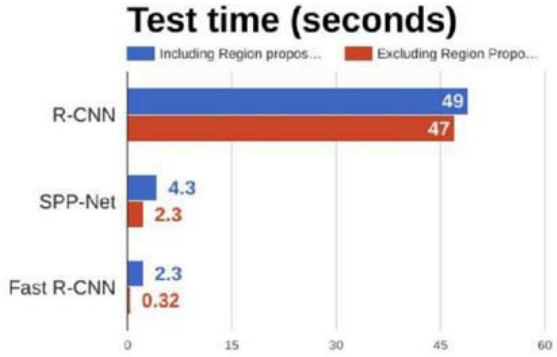


Image AI provides many more features useful for customization and production capable deployments for object detection tasks. Some of the features supported are:- Adjusting Minimum Probability: By default, objects detected with a probability percentage of less than 50 will not be shown or reported. You can increase this value for high certainty cases or reduce the value for cases where all possible objects are needed to be detected.

- Custom Objects Detection: Using a provided Custom Object class, you can tell the detection class to report detections on one or a few number of unique objects.

- Detection Speeds: You can reduce the time it takes to detect an image by setting the speed of detection speed to "fast", "faster" and "fastest".

- Input Types: You can specify and parse in file path to an image, Numpy array or file stream of an image as the input image

- Output Types: You can specify that the detect Objects from Image function should return the image in the form of a file or Numpy array.

Conclusion:

By using the above proposed system for our project i.e., "Object Identification and Classification Using Deep Neural network" , we are able to detect the object more precisely and

accurately and also identify the objects individually with exact location of an object in the picture.

Acknowledgement:

We express our earnest gratitude to our guide Professor Miss Manisha Pise Ma'am, Department of Computer Technology Engineering, for her never-ending encouragement and assistance, constant support, and guidance. We are grateful for her cooperation and her valuable suggestions for this research paper.

We also express our thanks to Dr. Nitin Janwe, Head of the Department of Computer Technology Engineering, for his never-ending encouragement and assistance given to us. We sincerely oblige his guidance and support for this research paper.

References:

- www.wikipedia.org.in
- www.towardsdatascience.com
- www.machinelearningmastery.com
- www.upgrad.com
- www.KDnuggets.com
- Agarwal, S., Awan, A., and Roth, D. (2004). Learning to detect objects in images via a sparse, part-based representation. IEEE Trans. Pattern Anal. Mach. Intell. 26,1475–1490.
- Survey of Object Detection using Deep Neural Networks Mrs. Swetha M S1, Ms. Veena M Shellikeri2 , Mr. Muneshwara M S3, Dr. Thungamani M4- International Journal of Advanced Research in Computer and Communication EngineeringVol. 7, Issue 11, November 2018
- Alexe, B., Deselaers, T., and Ferrari, V. (2010). "What is an object?," in ComputerVision and Pattern Recognition (CVPR), 2010 IEEE Conference on (San Francisco,CA: IEEE)
- Aloimonos, J., Weiss, I., and Bandyopadhyay, A. (1988). Active vision.