e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:06/June-2021 Impact Factor- 5.354 www.irjmets.com

OBJECT DETECTION BY USING CONVOLUTIONAL NEURAL NETWORKS

Mr. Kiran Kumar.M*[1], Soumya.P*[2], Thanuja.T*[3], Sandhyarani.B*[4], Tejaswini.S.Y*[5]

*[1]Assistant Professor, Department Of Computer Science And Engineering, Madanapalle Institute Of Technology And Science, Madanapalle, India.

*[2,3,4,5]B. Tech IV Year, Department Of Computer Science And Engineering, Madanapalle Institute Of Technology And Science, Madanapalle, India.

ABSTRACT

Object recognition has become a significant assignment for different purposes in our everyday lives. AI methods have been utilized for this assignment from prior however they are utilized for the arrangement of picture based species to extricate the list of capabilities. This undertaking of choosing the list of capabilities assists with choosing the ideal article discovery. To conquer the article order issue, this application proposes an exchange learning-based profound learning technique. The distinctive convolutional neural networks (CNN) are concentrated in this work. Here for the improvement in the outcome, the greater part casting a ballot plot is utilized. The general work is completed on the CUB 200-2011 dataset. The outcomes got have shown mind boggling improvement in the exactness of the proposed work when contrasted with the distinctive CNN models. Here the Japanese funnies (manga) are utilized for the assessment.

Keywords: Deep Learning, Object Detection, Training, Convolutional Neural Network.

I. INTRODUCTION

The convolutional neural network has been widely used for object detection work. The pre-processing tasks needed to perform very much lower as compared to other classification algorithms. Object detection is used in various fields, especially for face recognition. Here in the proposed work, various types of birds are classified. CNN is used for the classification task. In the fields like image recognition and classification, the neural network employed here is the convolutional neural network. It is a category of the neural network, which worked efficiently in the field of image recognition or object detection. CNN has worked well in identifying a face, objects, and traffic signboards. The idea behind its good work is that it is an extension of deep learning algorithms. In the convolutional neural network, there is no need of providing high costs and large resources because the standard algorithm can work. By applying the relevant filters, the spatial and temporal dependencies can be captures by the convolutional network. The architecture of the convolutional neural network is having the connectivity patterns as that of the neurons in the human brain and was inspired by the visual cortex.

II. LITERATURE SURVEY

1. Roy, S. K., Krishna, G., Dubey, S. R., & Chaudhuri, B.

Hyperspectral picture (HSI) order is generally utilized for the investigation of distantly detected pictures. Hyperspectral symbolism incorporates changing groups of pictures. Convolutional Neural Network (CNN) is perhaps the most regularly utilized profound learning-based technique for visual information handling. The utilization of CNN for HSI arrangement is likewise noticeable in ongoing works. These methodologies are for the most part dependent on 2D CNN. Though, the HSI order execution is exceptionally subject to both spatial and unearthly data. Not many techniques have used the 3D CNN because of expanded computational intricacy.

2. Wang, C., & Peng, Z. Design and Implementation of an Object Detection System Using Faster R-CNN. International Conference on Robots & Intelligent System (ICRIS).

Late advancement in object recognition is incredibly determined by the achievement of district proposition approaches and locale-based convolutional neural organizations (R-CNNs). In this paper, we planned and carried out an article identification framework utilizing a quicker CNN strategy that offers full-picture convolutional highlights with a recognized organization, in order to empower almost without cost locale recommendations. Improvement of this framework depends on the past work on Faster R-CNN. Results show that with this strategy, we could accomplish high precision while identifying objects

3. Parintorn Pooyoi; Punyanuch Borwarnginn; Jason H. Haga; Worapan Kusakunnira.

Pictures from CCTV cameras can be utilized for dissecting fiasco circumstances in a specific region. Snowfall is one of the climate conditions that could cause cataclysmic events in Japan. It is feasible for a machine to recognize snow and imprint these regions in that picture. There are existing convolutional neural organization-based (CNN-based) systems that can accomplish high exactness in an article characterization task. Be that as it may, these structures can't characterize or stamp the influenced region then, at that point show the outcome. To resolve this issue, this paper proposes a strategy to foster a model utilizing CNN structures, with the exchange learning procedure.

III. PROBLEM STATEMENT

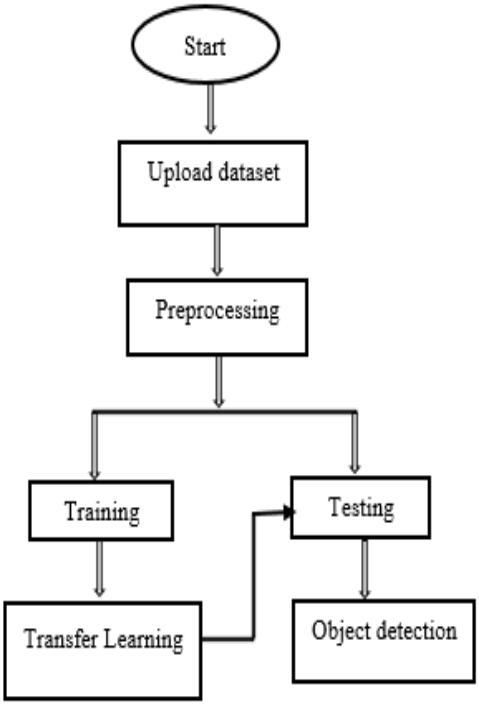
In this digital era, there is a tremendous growth in the area of artificial intelligence and machine learning. The aim of doing such experiments is to build machines that mimic the detection capability of humans. This growth has provided various options to the researchers. Also, the introduction of machine learning started using deep learning approaches to provide a big field for research. The fields like artificial intelligence, speech recognition, face recognition, object detection are the various areas of the applications of deep learning. The convolutional neural network has been widely used for object detection work. The pre-processing tasks needed to perform very much lower as compared to other classification algorithms. Object detection is used in various fields, especially for face recognition. Here in the proposed work, various types of birds are classified. CNN is used for the classification task.

IV. PROCESS FLOW

The process flow Contains Certain Steps

- Upload Dataset
- Preprocessing
- Training
- Testing
- Object Recognizing

4.1 ARCHITECTURE:



V. PREDICTING THE RESULTS WITH THE HELP OF MODELS

The name of an image is predicted after fitting the appropriate data in the required models for the prediction. The models that are used to detect the object are Google Net, ResNet50, VGG16, VGG19, AlexNet. ResNet50 has given better results among the five models. The algorithms that are used to detect the object are Convolutional Neural Network, Transfer Learning.

5.1 Convolutional Neural Network Algorithm:

In deep learning, a convolutional neural network is a class of deep neural networks, most commonly applied to analyze a visual image. The name itself indicates that the neural network calculates a mathematical operation which is called a convolution neural network. CNN is a different type of neural network which uses convolution instead of normal matrix multiplication in more than one layer. A convolutional neural network consists of three layers which namely called the input layer, output layer, hidden layer.

In a CNN, the hidden layers combine the layers which perform the mathematical operation. Normally it includes a layer that performs the matrix multiplication or other dot product. A convolutional layer must have the following attributes:

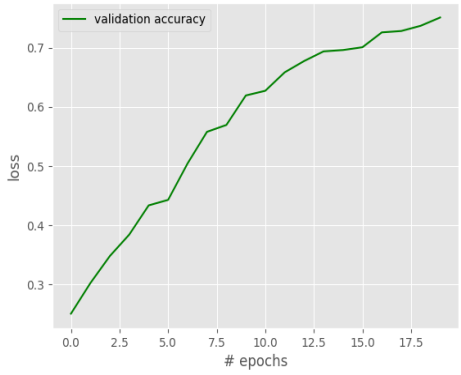
- Convolution kernels consist of width and height which are called hyper-parameters.
- The input channels and output channels are called hyper-parameters.
- The input channels of the convolution filter must equal the number of channels of the input feature map.
- The hyperparameters of the convolution operation are padding size and stride.

CNN for image classification works as it takes input in the form of an image and gives the output as a category of that image. The CNN convolves already learned features along with the input data, and it uses a two-dimensional convolution layer. That is ideal for the processing of two-dimensional pictures. In image classification, the CNN has input, output, and hidden layer. The hidden layer is having some layers like the convolution layer, Relu Layer, pooling layer, and the fully connected layer. The CNN works based on the feature extraction within itself, and it does not require manual extraction of the feature. This feature of CNN makes it suitable for image classification.

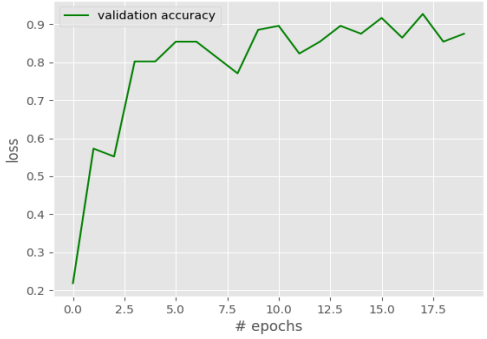
5.2 Transfer Learning:

This is the learning technique where the pre-trained convolutional neural network is retrained in some other dataset by using the weights of that particular network. This method of training makes the CNN work fast, where the random initialization is done. Some of the pre-trained CNN models are VGG16, VGG19, Alex Net, Google Net, Res Net this all are commonly used for transfer learning.

Alex Net: This model consists of 25 layers that are distributed as five are of convolution; three are of pooling and three for fully connected layers and seven for the rectified linear units.

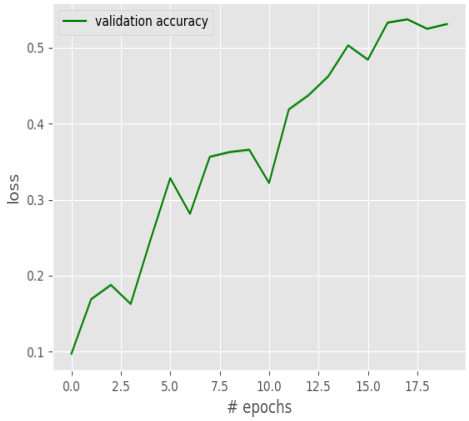


VGG-16 and VGG-19: In this, the VGG-16model having 41 layers and the matrix used by all the convolution layer is 3*3.The number of layers used by the VGG-19 model is different from the VGG-16 model.

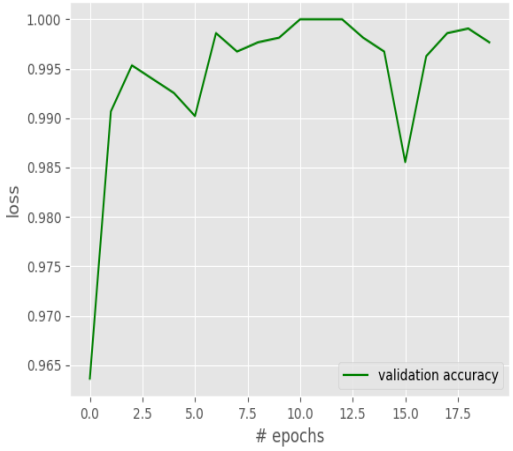


e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:06/June-2021 Impact Factor- 5.354 www.irjmets.com

Google Net: As this module has the inception modules, so it becomes a complex architecture to handle. In this, the addition of a large number of filters is done and sequentially stacking of layers is done due to which the astronomical calculation and memory cost is there.



ResNet-50: In this, some transition layers were ignored by giving preference to lower layers. These 177 layers are in total, which is having direct connection that increases its performance.



VI. RESULTS

Data Set:

- The COCO Dataset has 121,408 images
- The COCO Dataset has 883,331 object annotations
- The COCO Dataset has 80 classes

Here, the dataset is applied to the proposed CNN model. This decreases the learning rate. The hardware used for the experiment consists of an Intel(R), CPU N3540 @2.16GHz, 64bit OS. The rescaling of images is done as the pre-trained CNN works on a fixed-size image. The accuracy parameter is used to evaluate the proposed approach. The tabular form is used to present the numerical value obtained from the experiment.

S.NO	MODELS	ACCURACY
1	Google Net	81.5%
2	ResNet50	99.6%
3	VGG16	87%
4	VGG19	87%
5	AlexNet	81.5%

The algorithms that are used Google Net, ResNet50, VGG16, VGG19, Alex Net. Among them, ResNet50 has given the best result for the parameters that are taken into consideration. The Google Net has given the accuracy of 81. 5%, the ResNet50 model has also given the accuracy of 99.6%, the VGG16 model and VGG19 model have given t he same accuracy 87% and Alex Net with the accuracy of 81.5%.

VII. CONCLUSION

So the conclusion is the convolutional neural network with transfer learning and the majority voting scheme works well in comparison with the actual work. The proposed work gives fine granularity to work done earlier. The results have shown an extreme improvement in the field of object detection. Here, the COCO(Common Objects In Context) dataset is used to detect different types of objects. The 5 different pre-trained models of CNN were used, and their results are added with the majority voting scheme, and the results obtained a high accuracy of 97.45%. Also, it is known that there is always a chance for improvement. So in our work, there may present some limitations that can be analyzed further. In the future, some other classification methods will be applied. Also, the presented model can be tested on different datasets for evaluation.

VIII. REFERENCES

[1] Umit BUDAK, Abdulkadir ŞENGÜR, Aslı Başak DABAK and Musa ÇIBUK, Transfer Learning-Based Object Detection and Effect of Majority Voting on Classification Performance, 2019 International Artificial Intelligence and Data Processing Symposium (IDAP)

[2] Parintorn Pooyoi; Punyanuch Borwarnginn; Jason H. Haga; Worapan Kusakunniran…Snow Scene Segmentation Using CNN-Based Approach with Transfer Learning, 2019 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON).

[3] Wendi Cai; Jiadie Li; Zhongzhao Xie; Tao Zhao; Kang LU, Street Object Detection Based on Faster R-CNN…2018 37th Chinese Control Conference(CCC)… 10.23919/ChiCC.2018.8482613

[4] Hideaki Yanagisawa; Takuro Yamashita; Hiroshi Watanabe….A study on object detection method from manga images using CNN…. 2018 International Workshop on Advanced Image Technology (IWAIT)

[5] Yiming Zhang; Xiangyun Xiao; Xubo Yang…..Real-Time Object Detection for 360-Degree Panoramic Image Using CNN… 2017 International Conference on Virtual Reality and Visualization (ICVRV)

[6] B N Krishna Sai; T. Sasikala…Object Detection and Count of Objects in Image using Tensor Flow Object Detection API… 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)

[7] Jinsu Lee; Sang-Kwang Lee; Seong-Il Yang…An Ensemble Method of CNN Models for Object Detection, 2018 International Conference on Information and Communication Technology Convergence (ICTC)

[8] D. A. Bashar, "Survey on Evolving Deep Learning Neural Network Architectures," J. Artif. Intell. Capsul. Networks, vol. 2019, no. 2, pp. 73–82, 2019, DOI: 10.36548/jaicn.2019.2.003.