

Image Recognition for Object Detection using Deep Learning

Shivani Tyagi¹, Supriya P Panda²

Department of Computer Science and Engineering
Faculty of Engineering and Technology

Manav Rachna International Institute of Research and Studies Aravalli Hill, Faridabad, Haryana, India

Abstract - Object-finding is closely linked to video analysis and image comprehension, which have attracted the attention of researchers in recent years. Hand-crafted features and a learnable subtle structure are the basis of a typical object identification system. Its use is easily defined by building complex functions that combine high-quality content with object detectors and viewing classifiers with a variety of low-level viewing capabilities. Beside fast improvement of in-depth learning, additional controlling tackles have been proposed towards superior, or profound abilities to address issues of concern to traditional systems. These mockups acts inversely in web planning and training approaches etc.

This research paper discusses the smooth and convenient conversation between people and machines. The aim is to design an approach that can detect an object from image automatically without any efforts. Sometimes it is hard to identify the object when there are multiple objects in the image like mobile, tab, computer, person, cars etc. because of illusions and because of faint images, so there is an aim to detect object in an image using deep learning which is an easy way to handle and need not any kind of efforts from the person.

Provided is a evaluation on deep learning built [i] object detection agendas. Analysis commences with a detailed overview on the past of deep learning and its demonstrative tool, namely Convolutional Neural Network (CNN) [iii].

Keywords: Deep learning, Convolution Neural Network (CNN), Machine learning, Object Detection.

1. Introduction

Object Recognition means to apply some kind operations and formulas on images and to detect its component and values. It detects useful things from images and identifies that giving its name. Object detection is an important part of research area. Object recognition includes following steps:

- Importing
- Analyzing
- Output provided

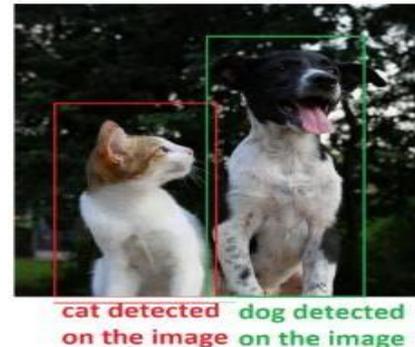


Fig. 1 Detection [i]

Modern technology is transferring all component of real world into digital form in current time. In nature, such component cannot be static but also it could be dynamic such as Object detection or movement, etc. Nowadays, all real-world objects are recognizing and transferring into the digital form with the help of deep learning.

Deep learning recognizes the objects and then matches them from the data entered in it. It detects the objects and produces a suitable output. CNN layers are included in Neural Network category which are useful in object recognition and classification. They are useful for the purpose of object detection, Identification, object, and traffic sign identification and apart from that in power invasion in robots and driving cars.

It includes convolution and pooling which passes through pixels and collects the output. Object recognition means processing of digital images in the form of pixels. Object Detection has many real time applications like urban planning, Agriculture also in some devices which senses the features of objects as Object detection is finding what and where (multiple) objects are in an image.

Machines Recognize images through CNN and pooling which passes through many activation layers. We can Simply say CNN is a scanning which scans images through filters in the form of pixels and then collects information for the group of pixels.

This filters can be of any type which are used for color intensity detection, Edges, curve detection. And this filters are of different sizes. This filters have random values and when this are convolve over the image then filter's random value joins with image's pixels value and give new set of value to identify an image.

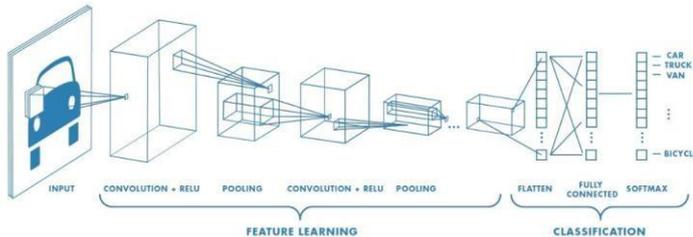


Fig. 2. Flow chart for the process by CCN

1.1 Deep Learning Uses:

- Image Recognition: many companies like Facebook uses deep learning neural networks for scanning of pictures.
- Object Recognition: Deep networks are also used to detect objects within images.

1.2 Applications:

- **Healthcare:** Object detection is widely used in healthcare like in ultrasounds or X-rays. It basically detects the bone cracks, liver lesion etc. and extract many of the medical images. [2]
- **Entertainment:** Deep Learning helps in advertising, designing.
- **Robotics:** Deep learning is also useful for robot sensing application. Deep learning is heavily used these days in building robots to perform human like tasks. Robots are used to understand the environment around it.

I. Artificial Neural Networks:

The human brain comprises a high number of neurons, which are interconnected, and that connectivity helps them to transfer or parallel processing of information in very little time and generate output. It has a great learning capability i.e., to learn from experiences. When machines are given this capability artificially, it is called as Artificial Neural Networks (ANN) and it trains data and try to create better output. It helps in face detection, object detection, identifying gestures. It has a learning algorithm which are actual power of neural networks. Like in google it has a trained data and learning algorithms work's there .Like when some input is given to it , it first convert it in meaningful data and then apply neural networks to it and then it learns and try to improve our data.

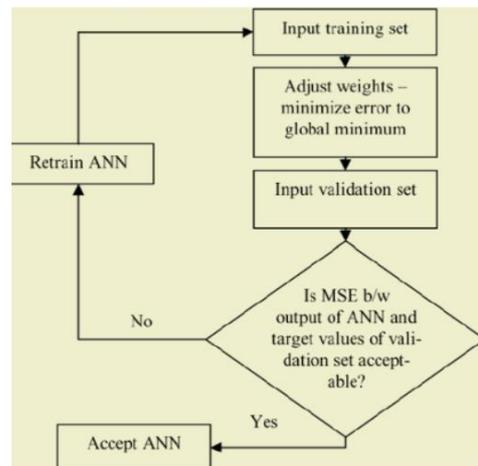


Fig.3. The artificial neural network process flow chart

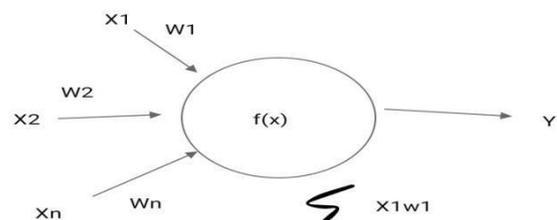


Fig. 4 The Neurons

Any word functions or quadratic functions can be apply but here are certain outputs, and these have some weights now apply linear function as:

$$\Sigma X1 * W1 \dots (1)$$

Here we get certain outputs and then we check the comparability of our values with threshold values to see whether our output is accurate or not. And after seeing output we can say it works same as biological neural networks which simply takes data and produce certain outputs.

II. Convolutional Neural Networks:

In human brain , visual cortex system is trained in a way that can simply detect the object or an image by seeing the features even an image is clear or not but it is so difficult for machines to understand the objects. In 1998, *Yen Lacuna* showed that machines can also detects the features for that CNN was used.

CNN includes convolution and pooling.

These two passes through many activation layers and at last generate final output. It is simply a type of scanning. machine identifies the picture in form of pixels. And it scans the group of pixels with the help of filters and collects information related to it. These filters are of many types with different steps like

- Color filter
- Curve filter
- Edge filer

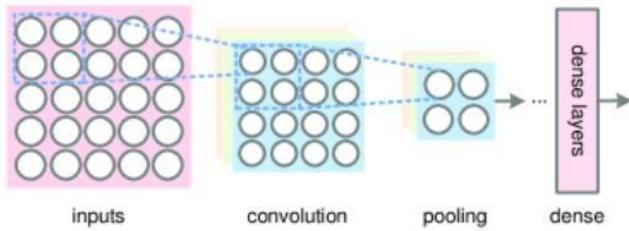


Fig.5. Convolutional Neural Networks Flow/Steps

This can be of different sizes and have random values. And when we evolve this to an image filter's random values added with image's pixels values gives new setup values which helps in identifying the object of the image and similarly all these properties are given to an image to get the final output, then the output generated is fed into fully connected layer which helps in image recognition and finding the object inside that. Some techniques are also used such as normalization, regularization like ReLU or Rectified linear unit functions are used to generate an output.

Pooling: This layer comes after convolution layer. It selects maximum elements from the region covered by filters and in result gives most prominent features of the previous feature map. It reduces the dimensions of an image by reducing the pixels of an image to detect and pull the object from the image. Simply pooling functions gradually reduce the spatial size of the input representation.[8]

Fully connected layer: Then after pooling data goes for flattening (program condenses all the image layers into a single layer image). Flattening converts 2D arrays to 1D array that is it converts it into a vector and then it passes from fully connected layer. ANN has many hidden layers to check nonlinearity and complexity included in feature representation. Fully connected layer uses softmax activation function to generate output (Softmax is an arithmetic expression that converts vector number into vector probability, where probability of each value is proportional to relative scale of each value.) Softmax activation range is from 0-1. It is used in CNN for classification. Basically, used in output layers of network. It allows that the output generated is in the form of classes. Basically, softmax function predicts the probability with respect to classes. Its main aim is to classify high level features of input image that is generated by CNN or Pooling based upon training data. [4].

LITERATURE REVIEW

Manoj Krishna¹, M Neelima², M Harshali³, M Venu Gopala Rao⁴, of Koneru Lakshmaiah Education Foundation (KLEF), at Vijayawada in Andhra Pradesh, India in 2018 tested object detection using deep learning. image

classification is correct for test images. Reliability of deep learning is conveyed. [1].

Md Tohidul Islam, B.M. Nafiz Karim Siddique, Sagidur Rahman, Taskeed Jabid, Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh in 2018 proposed a model in which object detection classifies the food categories using CNN (convolution neural networks). Classified food dataset with 92.86% accuracy having different categories[3].

Xin and Wang Journal in 2019 proposed a model that uses different classifiers to classify an image with some accuracy. From experiment it is observed that the accuracy of other classifiers in training set and test set have lower accuracy than CNN classifier. CNN has highest accuracy than other six commonly used classifiers KNN, SVM [5].

Benuwa, B.B., Zhan, Y.Z., Ghansah, B., Wornyo, D.K. and Banaseka Kataka, F., journal in 2016 proposed a model and did image classification and compare result in three different stages like accuracy, training performance and computational time. and introduces a different way for classification. [6].

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. and Berg, A.C in 2015 proposed a model that in which image classification on more than 1000 images collected from Flickr or ImageNet is applied. Then outperform the best model on images and some of the error may be eliminated and then comparison of computer vision accuracy with human accuracy is done. Each image has one ground truth label. When also applied for each image, it produces a list of categories[7].

Deng, J., Berg, A.C., Li, K. and Fei-Fei, L., in September 2010, exhibit a chain of command aware cost function for recognition and classification and prove that it brings most of the instructive classification outputs. Such demonstrations point to future research directions for image classification of large and critical dataset and benchmarking issues for accessing diff algo[8].

Liu, B., Yu, X., Zhang, P., Yu, A., Fu, Q. and Wei, X. in 2017 believed that well designed feature are still an issue that impact on performance of classifier. So the demonstration was done that the feature extraction method combining with linear SVM (support vector machine) classifier obtains better classification performance[9].

Wiatowski, T., Tschannen, M., Stanic, A., Grohs, P. and Bölskei, H. in 2016, June proposed a new CNN architecture and produced a mathematical framework for analysis. established deformation and translation sensitivity. Apply theory to filters and non-linearities and pooling operators that results in impressive classification[10].

2. Software Requirement

Python

Python is developed by Guido Van Rossum in Netherlands in 1980 and launched in 1991. python is one of the open source, high level, Object Oriented, interrupted programming language having different features than other programming language. According to Programming community index It is one of the top most programming language.

By the use of this language we can make any application quicker and efficiently. Python is used to make desktop, GUI application and web application. It has a different syntax than other language which makes it more easier and consistent for users. It has dynamic type system and automatic management system facility in it. In recent era Python is used by many big companies like Youtube, Instagram, Google etc. The programs written in it are in modular styled to perform different tasks. [11].

Syntax

While writing a python programs there must be no syntax error or must be with a particular structure otherwise the interpreter will generate error messages and not execute the program. Here is a simple example of python program.

```
print("hello world")
```

Deep learning has a great learning capability unlike machine learning it handles huge amount of data. It detects digital images and produces suitable output.

Uses of python:

- 1.Desktop Application
- 2.Web Application
- 3.Database Application
- 4.Networking Application
- 5.Gaming Application
- 6.Data Analysis
- 7.Machine Learning Application
- 8.AI application
- 9.For IOT Application

Python is used for Some Basic applications like calculator, VLC

media etc. Many Web application run on Python like Facebook, Instagram. For database correction and database application we use Python. To communicate Networking device like to make sensors and to make various games application also or in data analysis that is to extract some useful information from a big set of data we use Python.

Keras Library for Deep Learning

Keras is suggested library in Python. Keras is an high level neural network API written in python. or we can say it is a free, open source library available to user. It is more familiar with English language . Sometimes we don't have most of the knowledge about the backend or deep learning so there we can use Keras which provides low cognitive knowledge about deep learning and machine learning.

Mostly researchers and data scientist preferred deep learning framework that is keras. It supports various backend engines and will not put you in one ecosystem. Products can be formed through models easily by Keras. Keras has a GPU and distributed training support.[vi]

Allows for easy and fast prototyping
Supports both convolutional networks and recurrent networks, as well as combinations of the two.
Runs seamlessly on CPU and GPU.

Fig. 6 Keras Library

Convolutional Neural Network (CNN) as shown in Fig.7 is very useful in areas such as image recognition and object classification inside the image. CNN helps in identifying various objects like faces, objects and traffic signs etc.

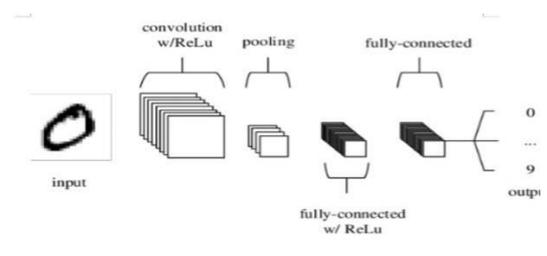


Fig. 7: Convolutional Neural Networks

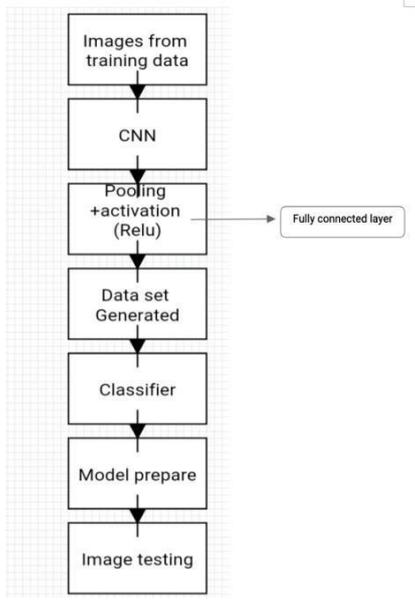


Fig. 8. Data Flow Diagram

2.1 EXPERIMENTAL RESULTS:-

- ✓ In this model, deep learning and CNN is used to classify the data object through an image.
- ✓ The data is trained by multiple data set with the help of tensorflow[iii].
- ✓ After collecting the dataset the pre processing method is used for convert some images into the binary images and obtain and recognize the object.
- ✓ After execution of code the camera will open for real time image recognition for object detection
- ✓ Below image shows the output from the project where image data is getting processed through multiple stages and find and identifying the object present inside the image which is mobile.



Fig.9.a Obtained output

- ✓ Below is the random image taken from the internet and targeted to find weather the project is working fine to get the good accuracy in finding out the object

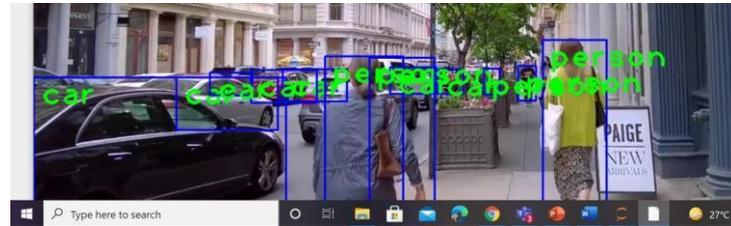


Fig.9.b Obtained output

2.2 FUTURE SCOPE

Image Recognition For Object Detection Using Deep Learning beneficial in production automation, also for agricultural landscape like by tracking satellite imaging object of fields it helps in irrigation monitoring.

Also, will be beneficial in disaster management, or in bio medical science for detection of internal parts of the body or for diagnosis etc. It has some realistic applications, including industrial, scientific and medical applications.

3.CONCLUSION

Because of python’s great learning capability and advantages of executing feature engineering on its own, Image detection using Deep Learning has been a research hotspot in recent years. This paper provides a detail review on DL-based Image detection framework. A model on AI to detect a test accuracy and test loss of an image is conducted using deep learning tool CNN was experimented. CNN is used because of its core advantage, as CNN automatically detects the important features of an image without human supervision unlike its predecessors. An attempt is made to make a model that uses a hand on artificial intelligence to predict the image's accuracy and test lose using powerful tool of This model is able to predict the data set images with accuracy more accuracy. This model is further tuned to minimize the loss function for better output. The large dataset collection of images is taken from multiple r training and testing the data.

REFERENCES

- [1] Krishna, M.M., Neelima, M., Harshali, M. and Rao, M.V.G., 2018. Image classification using deep learning. *International Journal of Engineering & Technology*, 7(2.7), pp.614-617.
- [2] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH. (2014). p. 580–587. doi: 10.1109/CVPR.2014.81
- [3] Hassairi, S., Ejbali, R. and Zaid, M., 2015, November. Supervised image classification using deep convolutional wavelets network. In *2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 265-271). IEEE.
- [4] Islam, M.T., Siddique, B.N.K., Rahman, S. and Jabid, T., 2018, October. Image recognition with deep learning. In *2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)* (Vol. 3, pp. 106-110). IEEE.
- [5] Xin, M. and Wang, Y., 2019. Research on image classification model based on deep convolution neural network. *EURASIP Journal on Image and Video Processing*, 2019(1), p.40.
- [6] Benuwa, B.B., Zhan, Y.Z., Ghansah, B., Wornyo, D.K. and BanasekaKataka, F., 2016. A review of deep machine learning. In *International Journal of Engineering Research in Africa* (Vol. 24, pp. 124-136). Trans Tech Publications Ltd
- [7] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. and Berg, A.C., 2015. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3), pp.211-252.
- [8] Deng, J., Berg, A.C., Li, K. and Fei-Fei, L., 2010, September. What does classifying more than 10,000 image categories tell us?. In *European conference on computer vision* (pp. 71-84). Springer, Berlin, Heidelberg.
- [9] Liu, B., Yu, X., Zhang, P., Yu, A., Fu, Q. and Wei, X., 2017. Supervised deep feature extraction for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(4), pp.1909-1921.
- [10] Wiatowski, T., Tschannen, M., Stanic, A., Grohs, P. and Bölcskei, H., 2016, June. Discrete deep feature extraction: A theory and new architectures. In *International Conference on Machine Learning* (pp. 2149-2158).
- [11] Roychowdhury, S., Diligenti, M. and Gori, M., 2018, January. Image Classification

Bibliography:

- i. https://en.wikipedia.org/wiki/Deep_learning
- ii. Cross Entropy: https://en.m.wikipedia.org/wiki/Cross_entropy
- iii. <https://www.tensorflow.org/learn>
- iv. <https://towardsdatascience.com/object-detection-with-convolutional-neural-networks-c9d729eedc18>
- v. [https://en.m.wikipedia.org/wiki/Rectifier_\(neural_networks\)](https://en.m.wikipedia.org/wiki/Rectifier_(neural_networks))
- vi. Keras Documentation: <https://keras.io/>
- vii. https://en.m.wikipedia.org/wiki/Stochastic_gradient_descent

BIOGRAPHY



Dr. Supriya P. Panda has done M.Sc from IIT, Kanpur, and pursued MS, Ph.D. (BGSU, Ohio, USA), where she worked as a Teaching Fellow. After that, she began her professional career at BITS, Pilani (Rajasthan) in 1990, where she worked for over a decade. Best Teaching Fellow at BGSU, Ohio, USA during MS and Ph.D(1985-90)



Shivani Tyagi has done B.tech from Manav Rachna University, Faridabad in computer science and currently pursuing M.tech in computer science and networking from Manav Rachna International University and working as an Associate Software Engineer at Accenture Solutions.