Object Recognition Using Deep Learning

G.V.Vinod[1] | A.S.V.S.Anand Gopal[2] | K.Sneha Latha[3] |N.Harish[4]|B.Sai Kumar[5]

[1,2,3,4,5] Department of Electronics and Communication Engineering, Godavari Institute of Engineering and Technology (A), Rajahmundry, Andhra Pradesh, India

Abstract: There are fascinating problems with computer vision,such as object recognition which is the combination of image classification and object detection. Our aim is to identify an object using deep learning.

In this project, we will be using a deep neural network named AlexNet to identify the object in any image. AlexNet is one of the most commonly used techniques for image classification. It can classify images into 1000 different categories, including keyboards, pencils, and other work related equipment, as well as various breeds of pigs,cats,parrots and other living creatures. The object recognition is especially well-suited to identification applications such as face recognition, self-driving cars, identifying illegal note in ATMs etc.

KEYWORDS:object recognition, machine learning, deep learning, existing system, proposed system.

DOI of the Article: https://doi.org/10.46501/GIETEC15

Available online at: https://ijmtst.com/icetee2021.html

As per UGC guidelines an electronic bar code is provided to seure your paper

To Cite this Article: G.V.Vinod; A.S.V.S.Anand Gopal; K.Sneha Latha; N.Harish and B.Sai Kumar[5]. Face Object Recognition Using Deep Learning. *International Journal for Modern Trends in Science and Technology* 2021, *7*, pp. 82-86. https://doi.org/10.46501/GIETEC15

Article Info. Received: 18 May 2021; Accepted: 25 June 2021; Published: 30 June 2021

Copyright © 2021 International Journal for Modern Trends in Science and Technology, ISSN : 2455-3778 http://www.ijmtst.com

INTRODUCTION

Object recognition is one of the most compelling abilities that humans easily possess since childhood. When humans look at a photograph or watch a video, we can readily spot people, objects and visual details. Also humans can easily identify a set of objects that have never been seen before. For example, kids are able to generalize the concept of table or chair after seeing just a few examples. Inspite of this, it is a very difficult task to develop vision systems that match the cognitive capabilities of human beings, or systems that are able to identify the particular identity of an object that has been observed. Generally, artificial intelligence (AI) is a way of making a computer, a computer-controlled robot or a software think intelligently, in a similar manner as an intelligent human. AI is accomplished by studying how human brains think, how human brains learn, decide, and work while trying to solve a particular problem, and then taking the results of this study as a reference for developing intelligent software and systems. The study of human intelligence and its reproduction has brought about several areas which include, but is not limited to neural networks, evolutionary computations, computer vision, robotics, speech processing, natural language processing, planning, machine learning and fuzzy logic.

STRUCTURE OF PAPER

The paper is organized as follows:

Contribution of Paper, briefly introduction of object recognition, machine learning is given in Section 1.Related work is seeing in section 2.and corresponding deep learning is given in section 3. Explanation of existing system of convolution neural network section 4proposed alexnetapproach RNN (methodology) is seeing in section 5. Experimental results can be seen in section 6.Future scope &conclusion,referencescan be seen in section 7.

OBJECTIVES

The main aim of our project is to identify an object using deep learning.In this project, we will be using a deep neural network named “AlexNet”, that has already been trained on a million imagesto identify the object in any image. Hence, it can directly be used for testing and most commonly used for image classification.The object recognition is especially well-suited to identification applications such as face recognition, self-driving cars, identifying illegal note in ATMs etc.

RELATED WORK

The field of AI research was founded at a conference at Dartmouth College in 1956. The attendees, including John McCarthy, Marvin Minsky, Allen Newell, Arthur Samuel and Herbert Simon, became the leaders of that AI research. They and their students wrote programs that were stunning to most people. Computers were winning at checkers, solving word problems in the algebra, proving logical theorems and speaking English. By the middle of the 1960s, research in the U.S. was heavily funded by the Department of Defense and the laboratories had been established around the world. AI's founders were optimistic about the future, Herbert Simon predicted, "machines will be capable, within twenty years, of doing any work a man can do." Marvin Minsky agreed, writing, "Within a generation, the problem of creating 'artificial intelligence' will substantially be solved". They failed to recognize the difficulty of some of the remaining tasks. Progress slowed and in 1974, in the response to criticism of Sir James Lighthill and ongoing pressure from the US Congress to fund more productive projects,both the U.S. and British governments cut off exploratory research in the AI field. In the early 1980s, AI research was revived by the commercial success of expert systems a form of AI program that simulated the knowledge and analytical skills of human experts. By 1985 the market for AI had reached over a billion dollars. At the same time, Japan's fifth generation computer project inspired the U.S and British governments to restore funding for academic research. However, beginning with the collapse of the Lisp Machine market in 1987, AI once again fell into disrepute, and a second, longer-lasting hiatus began. In the late 1990s and early 21st century, AI began to be used for logistics, data mining, medical diagnosis and other areas. The success was due to increasing computational power, greater emphasis on solving specific problems, new ties between AI and other fields and a commitment by researchers to mathematical methods and scientific standards. Deep Blue became the first computer chess-playing system to beat a reigning world chess champion, Garry Kasparov on 11 May 1997.

Copyright © 2021 International Journal for Modern Trends in Science and Technology, ISSN : 2455-3778 http://www.ijmtst.com

According to Bloomberg's Jack Clark, 2015 was a landmark year for artificial intelligence, with the number of software projects that use AI within Google increasing from a "random usage" in 2012 to more than 2,700 projects. Clark also presents factual data indicating that error rates in image processing tasks have fallen significantly since 2011. He attributes this to an increase in affordable neural networks, due to a rise in cloud computing infrastructure and to an increase in research tools and datasets.

DEEP LEARNING

Deep learning is a branch of machine learning which is completely based on artificial neural networks, as neural network is going to mimic the human brain so deep learning is also a kind of mimic of human brain. In deep learning, there is no need to programming everything explicitlys. The concept of deep learning is not new to us. It has been introduced in 1965 for the first time. It has more demand in the market nowadys as we did not have that much processing power and a lot of data in the early days. As in the last 20 years, the processing power increases exponentially, deep learning and machine learning came into the picture. A formal definition of deep learning is- Deep Learning is a machine learning technique that establishes artificial neural networks to mimic the structure and function of the human brains. The question is here is how do we recreate these neurons in a computer. So,we create an artificial neural net where we have nodes or neurons. We have some neurons for input value and some for output value and in between there may be lots of neurons interconnected in the hidden layer. There are some architures in deep learning they are deep neural network, deep belief network(DBN),recurrent neural network.

EXISTING SYSTEM

Convolutional Neural Networks(CNN) define an exceptionally powerful class of models. CNN-based models accomplishing the state-of-the-art results in classification, localisation, and semantic segmentation, amongst others.

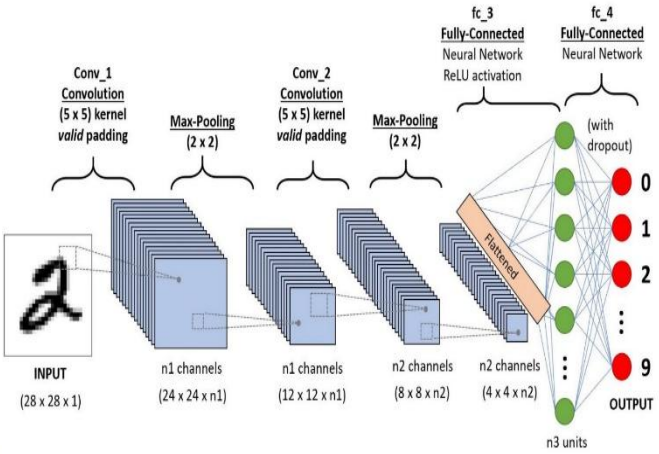

Fig 5.1: CNN Architecture

There are some drawbacks using convolution neural network(CNN):

CNN do not encode the position and orientation of object :

CNNs completely lose all their internal data about the pose and the orientation of the object and they route all the data to the same neurons that may not be able to deal with this type of information. A CNN makes forecastings by looking at an image and then checking to see if certain components are present in that image or not. If they are, then it classifies that particular image accordingly.

Lack of ability to be spatially invariant to the input data :

Artificial neurons output a single scalar. Additionally, CNNs use convolutional layers that, for each kernel, recreate that same kernel's weights across the entire input volume and then output a 2D matrix, where each number is the output of that kernel's convolution with a portion of the input volume. So we can look at that 2D matrix as output of reproduced feature detector. Then all kernel's 2D matrices are arranged on top of each other to produce the output of a convolutional layer. When in the input image we shift the object that we want to detect by a little bit, networks activities i.e. outputs of neurons will not change because of max pooling and the network will still detect the object.

METHODOLOGY

A number of studies have confirmed that intermediary layers in a deep network can capture features that provide a good resolution between object

Copyright © 2021 International Journal for Modern Trends in Science and Technology, ISSN : 2455-3778

inrepresentation and object independence. In this work, several low-level layers of the AlexNet skilled on the ImageNet data set were selected, and each of these layers were examined as a black-box feature extractor. The full structure of the AlexNet is shown in the top half of the Figure. The network was trained using fixed size RGB images from ImageNet.The network structure consists of 8 layers, where the first 5 layers (conv1, conv2, conv3, conv4, conv5) are convolutional and the remaining 3 layers i.e. fc6, fc7, fc8 were fully-connected. The last fully-connected layer (fc8) has the form of a Softmax classifier to categorize an input image into one of the classes used in training. The proposed new object recognition structure, with part of AlexNet embedded as the feature extractor, is illustrated in the bottom half of the figure. This new structure uses the same input format as the AlexNet and consists of several low-level layers of trained AlexNet. These layers thus act collectively as the feature extractor. An RNN unit containing multiple RNN structures, is then added to further process the extracted features, before feeding them into the Softmax classifier that performs recognition on the target data set.

Process Description

The following diagram makes it easier to understand how we proceed:

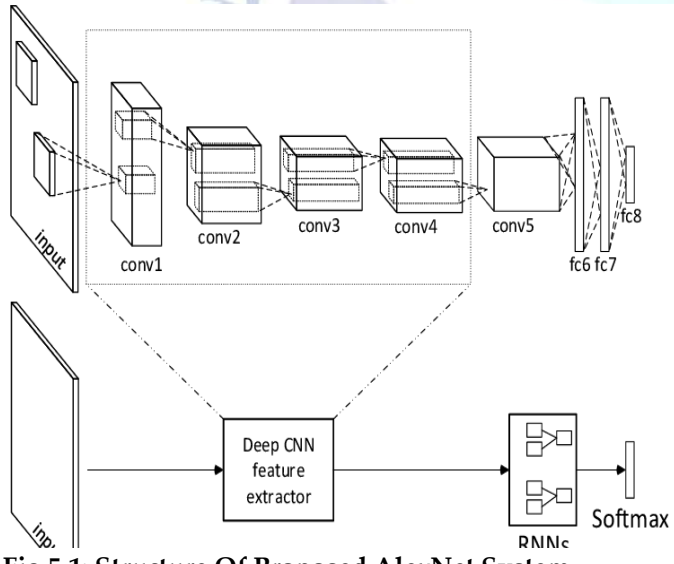

Fig 5.1: Structure Of Proposed AlexNet System

As a part of the new structure, two alternative versions of the AlexNet were used to extract the deep image features: the original version of the AlexNet- 2012, and the AlexNet- 2014. The AlexNet-2014 was more densely connected but had a smaller number of CNN weights in the intermediate layers compared to the AlexNet-2012 version. There were three key reasons behind the use of AlexNet-2014 over the original AlexNet-2012 in this project:

1) It provides slightly higher performance than AlexNet-2012 on multiple datasets;
2) It is computationally cheaper; and
3) It's last three convolutional layers have the same size, which allows for size-independent transferability and comparison of different attributes between layers.

EXPERIMENTAL RESULTS

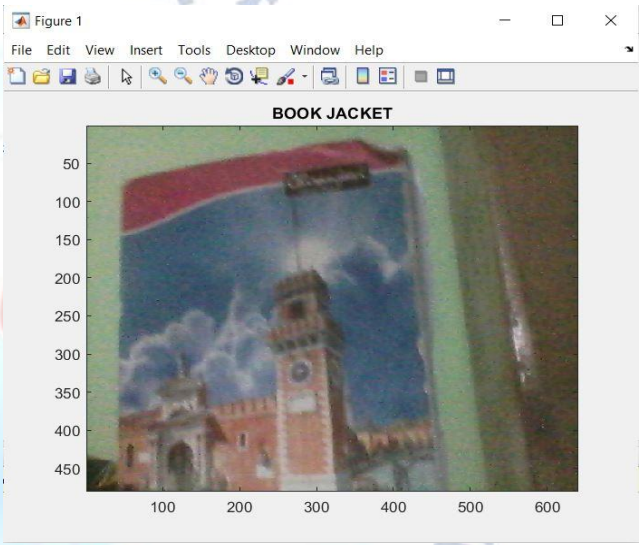

Fig 6.1:Matlab output of book cover/jacket

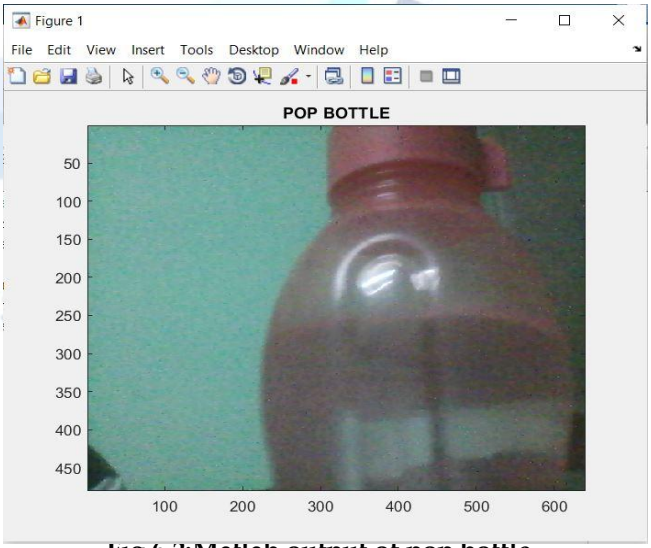

Fig 6.2:Matlab output of pop bottle

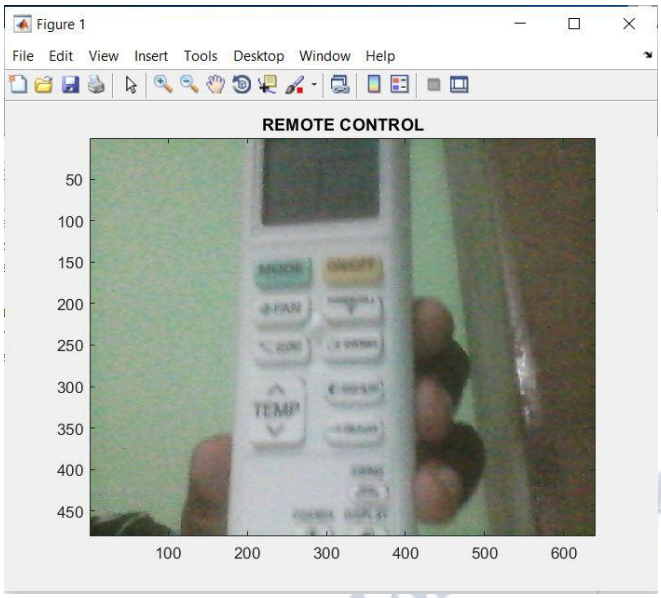

Fig 6.3Matlab output of AC remote control

Different types of objects were tested by capturing those objects with the help of system's web camera in the MATLAB editor and produced results were given above. By observing those results, we can say that AlexNet can identify and recognize various types of objects. Hence through this project, we have done object recognition using deep learning with the help of MATLAB software.

FUTURE SCOPE AND CONCLUSION

This paper presented a new method of feature extraction from a deep Convolutional Neural Network (DCNN) trained on a large dataset named AlexNet and combined with the Recursive Neural Network structure (RNN). The trials conducted on the Washington RGBD image data set have proved that the new proposed method has the benefits of structural simplicity combined with the ability to provide state of the art performance at a low computational cost compared to the AlexNet. The proposed approach requires no training during the feature extraction stage, and can be performed very efficiently. The output features are compact and highly discriminative, and can be used with a simple multi-class classifier in object recognition settings.

REFERENCES

[1] M. Yang, "Object Recognition," Encycl. Database Syst., pp. 1936– 1939, 2009.

[2] "Artificial Intelligence Overview." .

[3] A. Memo, "Review of machine _ deep learning in an artistic context – Machine Intelligence Report – Medium." 2016.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[5] D. C. Cireşan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible , High Performance Convolutional Neural Networks for Image Classification," 2011.

[6] G. Hinton et al., "Deep Neural Networks for Acoustic Modeling in Speech Recognition," pp. 1–27, 2012.

[7] R. Collobert, J. Weston, and M. Karlen, "Natural Language Processing (almost) from Scratch," vol. 1, pp. 1–34, 2000.

[8] C. Couprie, L. Najman, and Y. Lecun, "for Scene Labeling," pp. 1– 15, 2012.

[9] W. M. Guzella, Thiago S., Caminhas, "A review of machine learning approaches to Spam filtering _ WalmirCaminhas - Academia." 2009.

[10] A. Karpathy, "Deep Visual-Semantic Alignments for Generating Image Descriptions," 2015.

Copyright © 2021 International Journal for Modern Trends in Science and Technology, ISSN : 2455-3778 http://www.ijmtst.com