IMAGE CLASSIFICATION BASED ON DEEP LEARNING USING PYTHON AND KERAS

JETTY MANU PRANEETH[1], P.P.M PRASAD[2]
[1]*M.Tech. student, Bapatla Engineering College, Bapatla, A.P., India*
[2]*Department of ECE, Bapatla Engineering College, Bapatla, A.P., India,*
[1]manujetty25@gmail.com,[2]manohar.padyala@gmail.com
[1]Associate Professor, Department of E.C.E, Bapatla Engineering College,
manohar.padyala@gmail.com

ABSTRACT: Deep learning algorithms can be considered as one of the subcategories of the machine learning algorithms. The deep learning algorithm mainly aims at obtaining different abstract representations of the raw data at multiple levels in multiple layers. In recent times, numerous deep learning algorithms have been proposed to solve traditional problems that occur in artificial intelligence. Image classification is a fascinating topic in deep learning. Specifically, the image classification comes under the category of computer vision. This paper gives an overview of various deep learning approaches and their recent developments and then gives an overview of how to classify an image dataset using a convolution neural network, an open-source neural network library (KERAS), along with python on a CIFAR-10 dataset. First, we will explore the CIFAR-10 dataset, and then we will train our neural network with the help of python and KERAS. Finally, in this paper, we summarize the recent trends and challenges we generally face in the design and training of deep neural networks

Keyword: Image classification System (ICS), deep learning, computer vision, developments, applications, trends, challenges.

1. INTRODUCTION:

Deep learning is an important subfield of machine learning based on Artificial Neural Networks. Deep learning Algorithms can be defined as the class of machine learning algorithms that actively extracts the features of a very high-level data in multiple layers from the given raw input data. Deep learning is very successful in real-time applications and can be applied in various fields and applications like computer vision, speech classification, speech recognition, image classification, natural language processing, etc. The word Deep learning comes from the fact that we use multiple layers of network in our algorithms. In Deep Learning we teach the algorithm to obtain and transform the raw input data at each level into a slightly more abstract and composite representation. Image classification can be classified as one of the applications of Deep Learning. In this paper, we deal with the classification of different images using a single dataset. The Image classification application is the process where we take the raw input which may be a matrix of pixels, and in the first layer we abstract the pixels and encode edges; the second layer we try to encode and later compose of the way the edges are arranged; in the third layer we try to encode various distinct features like mouth and eyes, and the fourth layer may recognize that the image contains a face. Generally, the deep learning process is employed in the case of image classification because by using deep learning we can train the computer to learn which features to optimally place in which level on its own. Though this is a highly advanced process the methods involving Deep Learning do not completely remove the need for personal intervention, for example, depending on the numbers of layers and the size of the layers the abstraction obtained can be of varying degrees. In this paper, we employ a new take on image classification method by using a CIFAR-10 dataset, which contains a total of 60000 images, in which 50000 for training and 10000 for testing the deep learning algorithms and from there we design a Convolutional Neural Network with the help of python and KERAS and deploying deep learning algorithms to obtain the result with high accuracy.

2. RECENT DEVELOPMENTS:

In the recent years, Deep learning has seen a tremendous growth many new methods pertaining to deep learning have been introduced and have been extensively employed in various fields and many real-time applications like computer vision and as a consequence, a large number of related approaches have emerged. Based on the basic concepts that they are derived from, the methods of Deep Learning can be divided into four categories they are: Convolutional Neural Networks (CNNs), Restricted Boltzmann Machines (RBMs), Auto encoder and Sparse Coding.

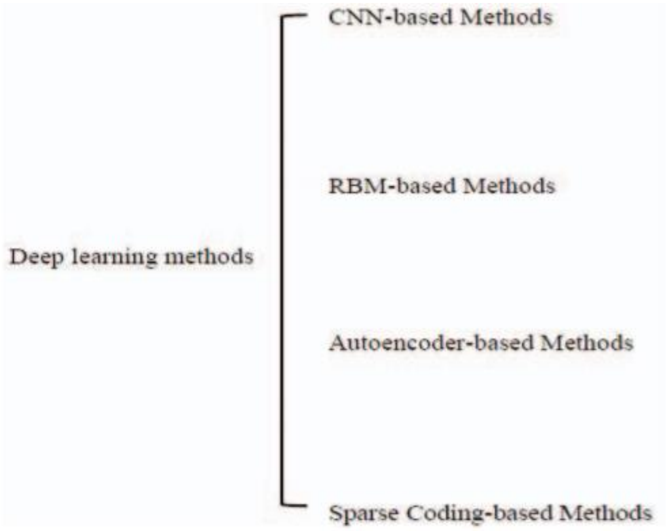

Fig: 1. Categorization of the deep learning methods.

The above figure represents about the four different methods of deep learning.

2.1 Convolutional Neural Networks (CNNs):

In deep learning the **convolutional neural networks** (**CNN**, or **ConvNet**) is defined as a class of deep neural networks. They are most commonly used in the applications of visual image analysis. The Convolution Neural Networks are also known as **shift invariant** or **space invariant artificial neural networks** (**SIANN**), they are categorised based on their shared-weights architect and their characteristics. The Convolution Neural Networks have applications in various fields such as image and video recognition, image classification, medical image analysis, natural language processing,[5] and financial time series and many more. In general, a CNN consists of three main layers, which are called 1.convolutional layers, 2.pooling layers, and 3.fully connected layers. Each layer plays different roles during the training and testing.

In general the Convolution Neural Networks (CNN) architecture for the classification of image is shown as layer-by-layer. There are two stages for training they are forward and backward stage. The main goal of the forward stage is to make sure that the input image is presented with the current parameters (weights and bias) in each layer. And in the Second stage the backward stage computes the gradients of each parameter with chain rules. All the parameters are updated based on the gradients, and are prepared for the next computation on the forward stage. After sufficient iterations and obtaining the necessary data the training part will be done.

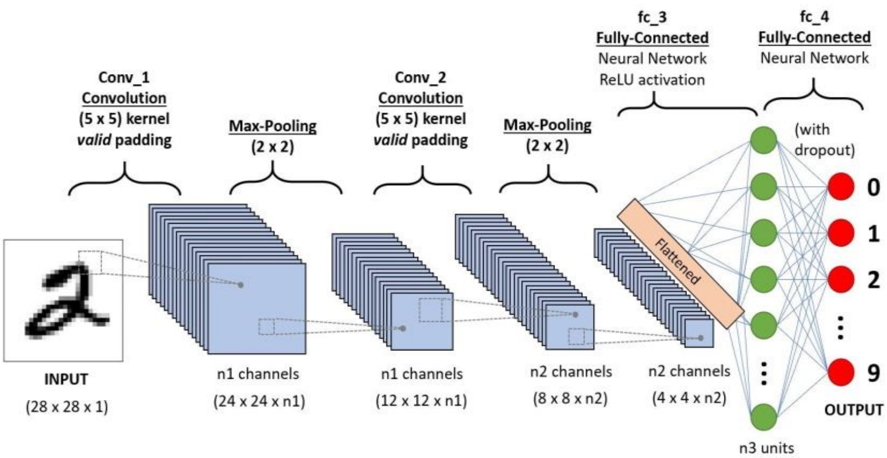

Fig: 2. General Architecture of Convolutional Neural Network

2.2Pooling of Layers in CNN:

The pooling layer follows a convolutional layer in a Convolutional Neural Network. In general, the Pooling layers are used to reduce the dimensions of feature maps and network parameters. The pooling layers are similar to convolutional layers in a way that they are also translation-invariant this is because their computations take neighbouring pixels into account. The most commonly used pooling techniques are 1. Average pooling and 2. Max pooling, these two methods are the most commonly used strategies. Fig. 3 gives an example of a max-pooling process. For example, if we take an 8x8 feature image, the output can be reduced to 4x4 dimensions, and with the max-pooling operator, we can obtain the size 2x2. Between the two pooling methods (Average and Max) the max-pooling method gives the best results. The pooling methods are used to reduce the computational cost by reducing the number of parameters to learn, it also provides the basic translation invariance to the internal representation.

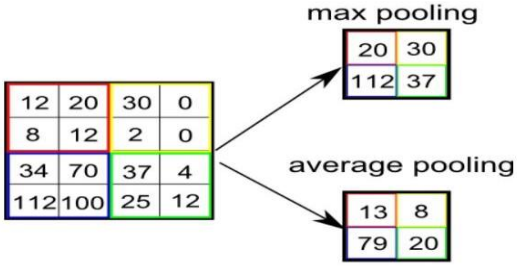

Fig: 3. Pictorial representation of Max and Average pooling methods

3. Image Classification:

Image classification can be defined as the task of extracting abstract information from a multiband raster image. The resulting abstract image form can be used to create thematic maps that can be used in image classification. The main problem of the image classification is to categorize all the pixels of a digital image into one of the defined classes. Image classification is most critically used in the case of digital image analysis. Image classification is an application that can be done in both supervised and unsupervised classification. In supervised classification, we select samples for each target class, we train our network on these target class samples and then classify new samples,in unsupervised classification, we group the sample images into clusters of images having similar properties and then we classify each cluster into our intended classes. In our paper use a CIFAR-10 dataset as the input data. The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images. The dataset is divided into five training batches and one test batch, each with 10000 images.

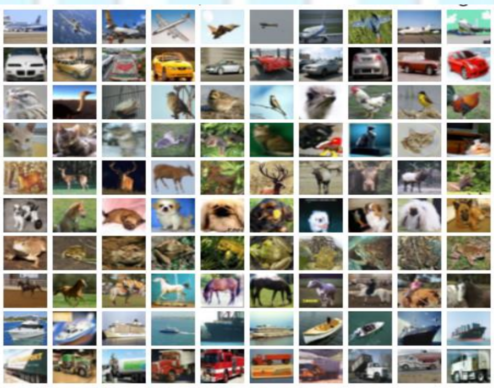

Fig: 4. A CIFAR-10 dataset

To obtain the CIFAR-10 dataset for our paper there isn't a need to download the dataset separately, as it can be imported directly from one of the KERAS databases. The requirements for the development and execution of an image classification project are KERAS and Tensorflow software's installation. KERAS can be classified as an open-source neural network library that is written in Python. It supports Tensor Flow, Microsoft Cognitive Toolkit, etc. It is designed as a user-friendly module that can be used for fast experimentation with deep neural networks. KERAS contains numerous implementations of commonly used neural-network building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to make working with image and text data easier to simplify the coding necessary for writing deep neural network code. The steps for the implementation of Image Classification using KERAS are, after installing KERAS first we have to load all the required datasets from the internal modules. After that, we have to plot some images to visualize the dataset. Import and create a Convolutional Neural Network and convert the pixel values and normalize the dataset. Also, create a sequential model, add the layers, and start to train the data. Calculate the accuracy, save the model, and start to make predictions based on the model.

3.1 APPLICATIONS AND RESULTS:

Deep learning has been very successful and is widely adopted in various fields and applications like computer vision and its subfields like, image classification, object detection, image retrieval, and semantic segmentation, face and human pose recognition, which are very important for understanding an image. Using KERAS for image classification in deep learning we can obtain images with high accuracy even with very abstract data.

Fig: 5. Expected result on the GUI.

4. TRENDS AND CHALLENGES:

Despite the great progress achieved in understanding the theory of deep learning, a significant amount of topics like understanding in evolving and optimizing the CNN architectures are still yet to be explored and should be improved. Another major challenge is that of training with limited data that is larger models tend to demonstrate more potential. However, because of the shortage of availability of the data that is required for training may determine the size and learning ability of the models. Also, it is expensive to obtain fully labelled data for training. More powerful models such as deep learning related algorithms have moved forward the-state-of-the-art results of various computer vision tasks by a large margin and it becomes more challenging to make more progress.

5. CONCLUSION:

This paper presents a comprehensive review of deep learning and gives us a categorization scheme to analyze the existing deep learning method. The deep learning algorithms are divided into four categories. This paper also gives us details about various applications of deep learning and in particular, describes the Image Classification application and its implementation using Convolutional Neural Networks and KERAS software. This paper describes the challenges and the trends in designing and training deep neural networks, along with several directions that may be further explored in the future.

REFERENCES:

[1] Xin Jia, "*image recognition method based on deep learning*", 2017 29th Chinese Control and Decision Conference (CCDC).

[2] A. Bordes, X. Glorot, J. Weston, *"Joint learning of words and meaning representations for open-text semantic parsing"*, at AISTATS, 2012.
[3] D.C. Ciresan, U. Meier, J. Schmidhuber "*Transfer learning for Latin and Chinese characters with deep neural networks*", at IJCNN, 2012.
[4] T. Mikolov, I. Sutskever, Chen K, "*Distributed representations of words and phrases and their compositionality*", at NIPS, 2013.
[5] N. Srivastava, G. Hinton, A. Krizhevsky, "*Dropout: A simple way to prevent neural networks from over fitting*", The Journal of Machine Learning Research: 1929-1958,2014.