"Object Recognition Using CNN Algorithm"

Mr. Dhairya J. Nagda, Mr. Nitin R. Pardhi, Mr. Viraj V. Mohite, Mr. Shubham Sinha, Mr. Sumit Talreja, Mr. Aditya Gourishetty, Mr. Sparsh Sharma

Abstract- This chapter intends to present the main techniques for detecting objects within images. In recent years there have been remarkable advances in areas such as machine learning and pattern recognition, both using convolutional neural networks (CNNs). It is mainly due to the increased parallel processing power provided by graphics processing units (GPUs).

This defense system of ours has two units- one is the control unit (to control mobility) and the other is the motion tracking unit. This robot would be a remote operator would be getting a live video feed from the camera to help him manually control both the above-mentioned units of the rover. The rover is also capable of automatically tracking movement of objects in its range of vision.

I. INTRODUCTION

Deep learning is a technology inspired by the functioning of human brain. In deep learning, networks of artificial neurons analyse large dataset to automatically discover underlying patterns, without human intervention.

In deep learning, a computer learns to classify images, text and sound. The computer is trained with large image datasets and then it changes the pixel value of the picture to an internal representation, where the classifier can detect patterns on the input image. Deep learning for image classification is becomes essential use of machine learning method.

To increase performance the application of neural networks to learning tasks that contains more than one hidden layer. Deep learning is part of a broader family of machine learning methods based on learning data representation, as opposed to hard code machine algorithms. One of the most frequently used deep learning method for image classification is the convolutional neural network (CNN). CNN learns directly from the image data, thus eliminating manual feature extraction.

Common problem in image classification using deep learning is low performance because of over fitting. To increase performance and preventing over fitting large dataset and model used. CNN have fewer connection and hyper parameter that make CNN model easy to train and perform slightly worse than other models.

In this paper, a deep learning convolutional neural network based on keras and tensor flow is deployed using python for object classification. Here, three models have been trained a tested which can be used in applications such as detection of animals, generic items (fan, watch, guitar etc) and automobiles.

II. BASIC TERMINOLOGY

In this paper, a deep-learning algorithm based on convolutional neural network is implemented using python, tflearn and augmentation for object classification. In this paper we can state that a CNN with higher layer performs classification process with much higher accuracy. In this paper, a modified CNN architecture that combines multiple convolution and max pooling layers for higher level feature learning is proposed. This paper states the utilization the Caltech 101 dataset.

1.Caltech 101 Dataset

Caltech 101 contains a total of 9,146 images, split between 101 distinct object categories (faces, watches, ants, pianos, etc.) and a background category. Provided with the images are a set of annotations describing the outlines of each image, along with a MATLAB script for viewing.

2.Convolutional Neural Network (CNN)

Convolutional Neural Network is a special type of feed forward artificial neural network, which inspired by visual cortex. In CNN, the neuron in a layer is only connected to a small region of the layer before it, instead of all the neurons in a fully connected manner, so CNN handle fewer amounts of weights and also a smaller number of neurons.

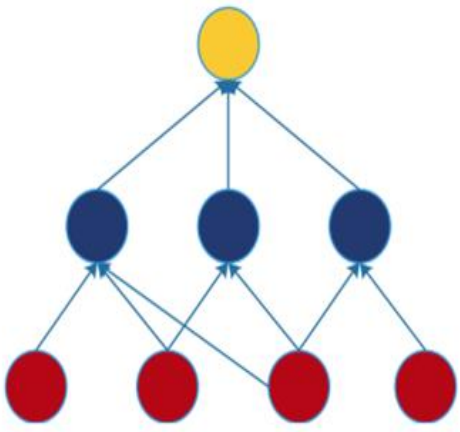

Fig. 2.2 Convolutional Neural Network

3.TensorFlow

Relu Activation Function: Relu F(x) = max (x, 0), is mostly used deep learning activation function, for hidden layers. A rectified linear unit has output „0" if the input is less than „0", and raw output otherwise". Relu is the simplest non-linear.

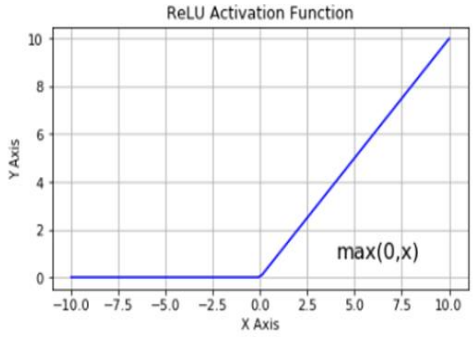

Fig. 2.3 Relu activation function

III. LITERATURE REVIEW

Hasbi Ash Shiddieqy, Farkhad Ihsan Hariadi, Trio Adiono "Implementation of Deep-Learning based Image Classification on Single Board Computer"[1], In this paper, a deep-learning algorithm based on convolutional neuralnetwork is implemented using python and tflearn for image classification, in which two different structures of CNN are used, namely with two and five layers and It conclude that the CNN with higher layer performs classification process with much higher accuracy. Rui Wang, Wei Li, Runnan Qin and JinZhong Wu "Blur Image Classification based on Deep Learning" [2].

In this paper, a convolution neural network (CNN) of Simplified-Fast-Alexnet (SFA) based on the learning features is proposed for handling the classification issue of defocus blur, Gaussian blur, haze blur and motion blur four blur type images. The experiment results demonstrate that the performance of classification accuracy of SFA, which is 96.99% for simulated blur dataset and 92.75% for natural blur dataset, is equivalent to Alexnet and superior to other classification methods.

Sameer Khan and Suet-Peng Yong "A Deep Learning Architecture for Classifying Medical Image of Anatomy Object" [3], In this paper, a modified CNN architecture that combines multiple convolution and pooling layers for higher level feature learning is proposed. In this, medical image anatomy classification has been carried out and it shows that the proposed CNN feature representation outperforms the three baseline architectures for classifying medical image anatomies.

Ye Tao, Ming Zhang, Mark Parsons "Deep Learning in Photovoltaic Penetration Classification" [4], this paper proposed a deep learning-based algorithm to differentiate photovoltaic events from other grid events, and it conclude that a deep convolutional neural network can achieve higher classification accuracy than a fully connected model.

IV. RESEARCH METHODOLOGY

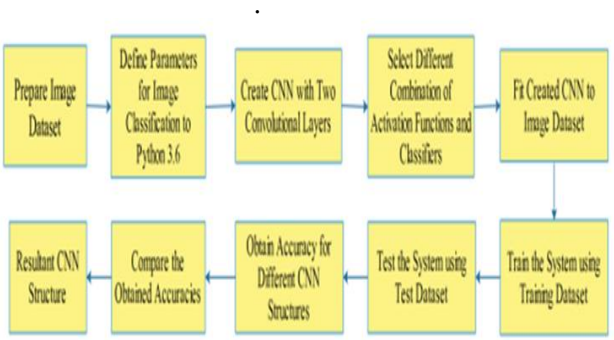

Fig. 4.1 Proposed Methodology

The flow diagram of proposed methodology is shown in fig. 4.1. Each block of proposed flow diagram is clearly labeled and represents processing steps. Using this methodology, we compare four different structure of CNN, with four different combinations of classifiers and activation functions.

Fig. 4.1 Proposed Methodology In first step image dataset is prepared, there are 25 categories of data with around 27000 images, where 22000 images used for training and 6000 images used for testing purpose.The data present within the Caltech 101 dataset is not sufficient to create an accurate model that classifies this wide range of categories.

Data augmentation has been used to increase the number of images through basic operations like zooming, shrinking, shearing, horizontal flip, vertical flip, rotation etc.

In second step, define parameters for image classification to python. In third step create CNN with three convolutional layers, then we select different combination of activation functions and classifiers for comparison purpose. In next steps, we fit the created CNN to image dataset and Train, Test the system with training and test datasets respectively.

Finally, we obtain the accuracy for different CNN structures and compare these accuracies for performance measurement, and then get the resultant CNN structure.

Data Augmentation

Recent advances in deep learning models have been largely attributed to the quantity and diversity of data gathered in recent years.
Data augmentation is a strategy that enables practitioners to significantly increase the diversity of data available for training models, without actually collecting new data.

Data augmentation techniques such as cropping, padding, and horizontal flipping are commonly used to train large neural networks. However, most approaches used in training neural networks only use basic types of augmentation.

While neural network architectures have been investigated in depth, less focus has been put into discovering strong types of data augmentation and data augmentation policies that capture data invariances.

Recently, Google has been able to push the state-of-the-art accuracy on datasets such as CIFAR-10 with Auto Augment, a new automated data augmentation technique.

Auto Augment has shown that prior work using just applying a fixed set of transformations like horizontal flipping or padding and cropping leaves potential performance on the table. Auto Augment introduces 16 geometric and color-based transformations, and formulates an augmentation policy that selects up to two transformations at certain magnitude levels to apply to each batch of data.

These higher performing augmentation policies are learned by training models directly on the data using reinforcement learning.

V. EXPERIMENTAL SETUP

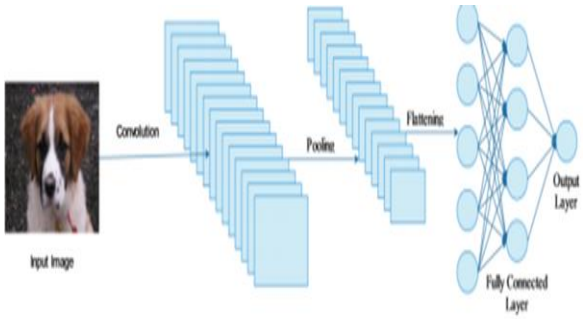

Fig. 5.1 Convolutional neural network model

In this paper, we perform experiments on windows 10 in python 3.6 on CPU system and create the CNN model based on keras and tensorflow libraries.

The CNN model used for experiments is shown in fig 5.1. This model mainly consists of four layers including, convolutional, pooling, flattening and fully connected layers.

Fig. 5.1 Convolutional neural network model for convolutional layer, the size of input image is set to 100*100 pixels with 3 channels (RGB). To extract the features from the image we use filters of size 5*5 pixels. For pooling layer, we use a window of size 5*5 pixels, which used to compress the original image size for further processing.

For performance measurement we use two activation functions namely, Relu (Rectified linear unit), Tanh (Hyperbolic tangent), and two classifiers namely Softmax, Sigmoid. In experiment, we use combination of these activation functions and classifiers, and analyze that which combination gives better classification accuracy for object classification.

CONCLUSION

Deep learning is a learning method for data analysis and predictions, now a days it has also become very popular for image classification problems.

In this paper, a deep learning convolutional neural network based on keras and tensorflow is deployed using python for binary image classification. In this study, we compare four different structures of CNN on CPU system, with different combinations of classifier and activation function.

With experiments, we obtained results for each combination and observed that for binary image classification, Relu activation function and Softmax classifier combination gives better classification accuracy (90.54%) than any other combination of activation function and classifier.

So, we conclude that on CPU system, Relu activation function and Softmax classifier gives better classification accuracy for object recognition.