Deep Learning Based Image Classification Using Small VGGNet Architecture

Kishan K Nayak, Prabhudev N, Rohan A R, Supriya S, Nayana R
Reva University Bangalore Karnataka
nayana.r@reva.edu.in

***Abstract*: One of the most intriguing and important areas of computer vision is a bracket. Deep literacy is important in image brackets because it allows for the grouping of analogous images. Certain adaptations to the images must be made before feeding them to the training model, like normalization, image resizing, slate-scaling, and so on. Following image-processing, the coming task is to prize features from the images, draw boxes around the objects in the image, and eventually classify the images grounded on the discovery of objects in the images. In a nutshell, a bracket is a simple task for humans but a delicate task for machines. For this situation, we'll utilize the Small VGGNet (VGG-Visual Geometry Group) Architecture, which is a sort of convolutional brain organization, to characterize pictures utilizing profound learning. The certified model is approved by utilizing three arbitrary pictures from Google. We explored different avenues regarding scaling the picture to various sizes. The profound learning model, which utilizes a Small VGGNet design to zero in on picture arrangement, performed incredibly well, accurately ordering the pictures.Google Photos has already implemented image classification in their app but uploading 1000s of images to the cloud requires a strong internet connection because computation is done on Google's server.**

***Keywords*— Classification, Small VGGNet, deep learning, CNN, computer vision.**

I. INTRODUCTION

Classification is one of the interesting and primary fields in computer vision. In image classification, deep learning plays the important role in grouping similar images. Before feeding the image to the trained model certain changes must be made to the images like normalization, image-resizing, grey-scaling.

After pre-processing the image, the next task is to extract features from the images, draw the boxes around the objects in the image, and finally classify the images based on the detection of objects in the images. In nutshell, classification is an effortless task for humans and a challenging task for machines. Because classification is a component of supervised machine learning, we will train the model using labeled data. In supervised machine learning, both the input and output are known (for example, if we have a cat image(x), we will label it as cat(y), where x denotes the self-determining variable and y denotes the reliant on the variable). The trained model consists of predefined patterns that are compared with an object that resides in images to classify to the suitable categoryConsider a scenario in which we need to categorize 1000 images of animals into a suitable category. If we use any application, GUI, or non-GUI, we will require a large amount of bandwidth to upload the images to the cloud, where computation occurs. Instead, we will provide the trained model with an application with a simple graphical user interface that will classify more than 1000 images at once, create a folder corresponding to the label, and copy the images to the matching folder. Instead, we will provide the trained model with an application with a simple graphical user interface that will classify more than 1000 images at once, create a folder corresponding to the label, and copy the images to the matching folder.

We aim to discover hidden patterns in the dataset using machine learning and deep learning algorithms. which could help us correctly identify data points that our algorithm hasn't seen yet. Classification is the systematic grouping and categorizing of objects based on their characteristics. By using data to train the algorithm, image classification was created to bridge the gap between computer vision and human ideas. Image classification is accomplished by categorizing the image according to the content of the vision.

In this context, we will investigate image classification using deep learning. A few years ago, we used machine learning, a subset of artificial intelligence, to achieve image classification.

The main disadvantage of computer vision is that it can only retrieve a limited set of structures from images while unscrambling and is incapable of extracting distinguishing functionalities from the training data set. A deep learning model is being developed to decompose knowledge with a homogeneous structure in a way that is close to how humans think and act. Deep learning does this by combining many algorithms into a stacked structure known as an artificial neural system (ANN) or a neural network.

Neural networks are used in machine learning to execute operations by processing a training set. Thousands or millions of correlation nodes make up a neural network.

The vast bulk of those nodes is "feed-forward," meaning data flows only in one stream. Each single nodes collect data from different nodes and send it to other nodes for processing. When the model is learning at the initial stage, we will assign arbitrary values to the node weights, when training data flows continuously to the neural network, the node weights are adjusted automatically after each iteration till the output of that neural network is precise.

II. LITERATURE SURVEY

Neural networks are called computational or Mathematical models that are used to find an optimal solution. They're made up of neurons, which are the fundamental computational units in neural networks. Artificial Neural Network is another name for it (ANN). ANNs come in a variety of shapes and sizes. The convolutional neural network, or ANN, is a software program that mimics the workings of human brain neurons and networks. Most neural nets are highly successful at their intended functions, such as segmentation and sorting, and mimic their more highly complex peers.

Feedback ANN - The contribution of feedback ANN, which is the output of individual neurons, is injected back into the network so that the model accuracy increases with each iteration. Since it incorporates information back into itself, the feedback network is suitable for solving optimization problems. A feed-forward network is a basic neural network that has one or more layers of neurons and an input layer, an output layer, and one or more layers of neurons. The network's strength can be seen by going over its input and evaluating its output based on the associated neurons' group actions, and the output is determined.

The image classification process consists of two steps: training the system and testing. The training process essentially takes the distinguishing characteristics of the images (forming a class) and creates a unique description

III. RELATED WORK

The challenge of image bracket from a large dataset is the subject of a recent exploration paper by Jianxin Wu etal. (1). In image bracket, the support vector machine (SVM) classifier has proven to be veritably effective.

According to the paper by Fuliang Wang and Feng Wang etal.

(2), Artificial Neural Network is able of representing AND, OR, and NOT while efficiently handling noisy data. Also, Monica Bianchinietal.

(3) also bandy the artificial neural network bracket technique. Monica Bianchini etal.

(4) go over the artificial neural network bracket technique as well. SerafeimMoustakidisetal.

(5) describe a new fuzzy decision tree in which knot demarcation are enforced using double SVMs. Lizhen Lu etal.

(6) developed the Decision Tree classifier, which divides the input into orders to determine class class.

A Spatial – Contextual Support Vector Machine for Ever Tasted Image Bracket is bandied by Cheng-Hsuan Li etal. (7). Different image bracket styles have their own set of benefits and downsides. In image bracket, some styles combine two or further classifiers. However, it's allowed to be more effective if a classifier can rightly prognosticate. The birth of a pattern or point from the available input datasets requires image bracket.

IV. IMAGE CLASSIFICATION STEPS

Image Classification includes the following steps:

a) Image Acquisition: collect images for post-processing

b) Image Pre-Processing: Image pre-processing techniques include image transformation, noise removal, and atmospherically correction

c) Image Feature Extraction: Extracting the image pattern is the most important feature. Classification: Using appropriate methods that compare the image pattern to images in the database, the images are classified into predefined categories based on extracted feature

1.Reducing the dimensionality of an image to the 1D (Single array) helps to feed the network easily.

2.Dividing the image pixels by 255, we will get values in 0's and 1's, this is done to normalize the image

3.For categorical values, we use a one-hot Encode function to represent categorical values in numerical form

4.Define the model architecture, it can be a sequential model with n-dense layers

5.Create a model and make predictions with it.

i. IMAGE CLASSIFICATION METHODS

The following approaches are used to classify images:

1) Support Vector Machine: In a high-dimensional region, this approach generates a series of hyper-plane that can be used for classification or regression. The hyperplane's excellent separation. SVM employs a non-parametric approach with binary classifiers to handle more input data efficiently. The accuracy is determined by the hyperplane used. The structure of the SVM algorithm is more complicated than that of other methods. Therefore, there is a lack of accountability in the results.

2) Artificial Neural Network (ANN): An artificial neural network is a form of artificial intelligence that mimics some of the functions of the human brain. An ANN is made up of several layers. Neurons are the building blocks of a neural network scheme. Weighted connections link neurons in all layers to neurons in the layers above and below them. The accuracy is determined by the number of inputs and the network structure. The ANN process is a non-parametric approach. This approach classifies the feedback quickly, but the training process takes a long time. It's tough to choose the best architecture.

3) Decision Tree: A decision tree is a graph of choices that resembles a tree. The decisions that must be taken care graphically represented by each branch. It is a non-parametric supervised approach. It splits the input into groups that are all the same. At each intermediate point, this approach allows for the approval or rejection of class marks. After classification, this approach gives a set of instructions that must be followed.

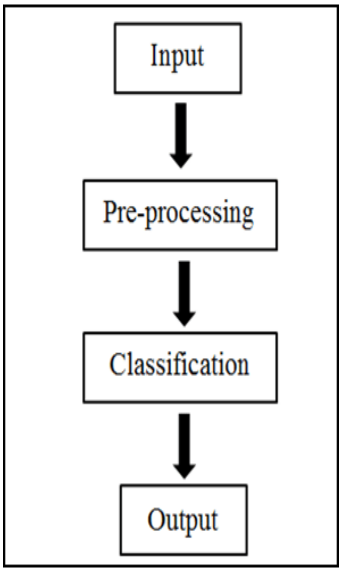
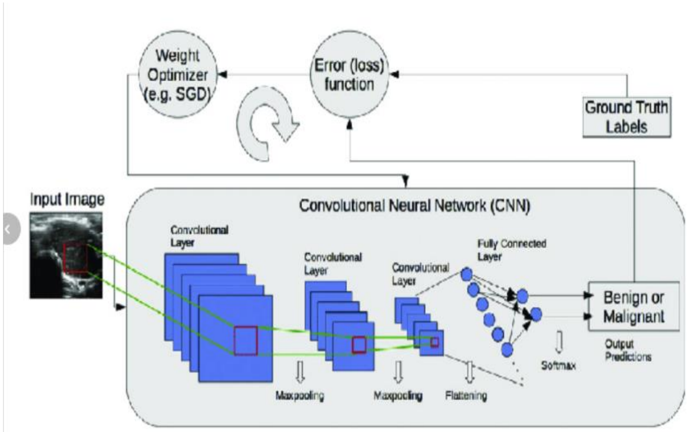


Table 1: Advantages and Disadvantages of image classification methods

Method	Advantages	Disadvantages
1.Support Vector Machine	• Deliver unique solution. • Very efficient than other methods. • Avoid over-fitting	• High algorithm complexity • Run slowly
2. Artificial Neural Network	• Robust to noisy training dataset • Very efficient for large dataset	• High computational cost • Lazy learner
3. Decision Tree	• Require little efforts from users • Easy to interpret and explain	• Splits are very sensitive to training data set • High classification error rate

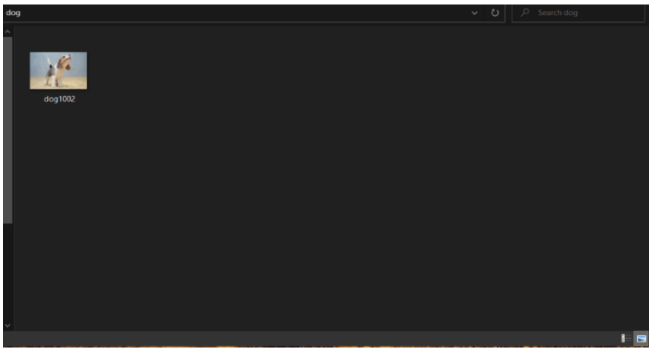
Demonstration
i)Input image:

ii)Internal processing of image: -



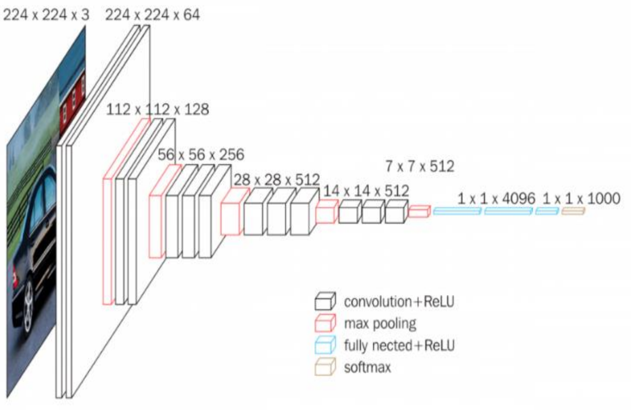
Output: -

- Detection



- Automatic folder creation based on the detected image

- **Adds the detected image to the appropriate folder**



iii) Small VGGNet Architecture

Input: - VGG accepts an RGB image with a resolution of 224x224 pixels. To maintain the target image size constant for the ImageNet contest, the programmers cropped out the central 224x224 patch on each image.

Convolutional Layers: -The receptive area of VGG's convolutional layers is extremely tiny (3x3, the smallest size that captures in all directions). Several 1x1 convolution detectors convert the input into a linear form before moving it through a ReLU unit. To retain image quality after convolution, the convolution stride is set to 1 pixel.

Fully connected layers: -VGG includes three completely connected layers, one in each class, with 4096 channels in the first two and 1000 channels in the third.

Hidden layers: -VGG's hidden layers all use ReLU (a huge innovation from Alex Net that cut training time).In any CNN, training time and memory consumption are important things that are to be considered, so to avoid this problem VGGNet doesn't use Local Response Normalization (LRN) because it causes problems also and doesn't improve accuracy.

ii) CONCLUSION

This paper explains how image classification works. We implemented a method that uses a convolutional neural network to extract and select essential features from any given image and classify the images into appropriate categories. In comparison to other classifiers, the Convolution neural network may have high precision. On both basic CPUs and GPUs, consistency and accuracy are evaluated. As a result, we've decided that Convolution Neural Networks are a safer choice for image classification. This system may be used for biometric identification, for example.

Finally, this study is about image classification using deep learning via the Keras framework. It has three objectives that have been met throughout this research. The conclusions are inextricably tied to the objectives since they can assess whether all the objectives have been met. All the data obtained were extremely amazing, it is feasible to conclude. This study focuses on the deep neural network (DNN), which is particularly useful in picture categorization technologies. Beginning with the construction, training model, and classification of photos into categories, the DNN technique was examined in further depth. Epochs' functions in DNN allowed for precision control while also preventing issues like overfitting. Deep learning implementation utilising the Keras framework also produced positive results, as it is capable of simulating, training, and categorising five different varieties of flowers with up to 90% accuracy. Finally, Python was chosen as the programming language for this study because it is compatible with the TensorFlow framework, allowing for Python-based system creation from start to finish.

iii) REFERENCES

[1] Young Jong Mo, Jongheon Kim, Jong Kim, Aziz Mohaisen, and Woojoo Lee, Performance of Deep Learning Computation with TensorFlow Software Library in GPU-Capable Multi-Core Computing Platforms.

[2] Yann LeCun, Leon Bottou, Joshua Bengio, and Patrick Haffner, "Gradient-Based Learning Applied to Document Recognition", Proc. Of IEEE, November 1998.

[3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks" May 2019.

[4] Ahmed, Syed Thouheed, and Sharmila Sankar. "Investigative protocol design of layer optimized image compression in telemedicine environment." *Procedia Computer Science* 167 (2020): 2617-2622.

[5] Fei Li, Justin Johnson and Serena Yueng, "Talk 9: CNN Architectures" May 2017

[6]AnupamAnand,“Image_classification”,May–2018

[7] T. Shreekumar, N. V. Sunitha, N. Suhasini, K. Suma, K. KArunakara, “Blur and Noise Removal from the Degraded Face Images for Identifying the Faces Using Deep Learning Networks”, International Conference on Artificial Intelligence and Sustainable Engineering 2022, pp 325–341, DOI: 10.1007/978-981-16-8546-0_27