IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⋇ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11415

Image Classification Using Machine Learning

Dipak R. Chavhan[1], Tejas A. Borole[2], Lubdha P. Bonde[3] Pratiksha A. Bharule [4]

U.G Student, Department of Computer Engineering, Shram Sadhana Bombay Trust's, College of Engineering and Technology, Jalgaon, India[1-4]

Abstract: The use of machine learning and specifically neural networks is a growing trend in software development, and has grown immensely in the last couple of years in the light of an increasing need to handle big data and large information flows. Machine learning has a broad area of application, such as human-computer interaction, predicting stock prices, real time translation, and self -driving vehicles. Large companies such as Microsoft and Google have already implemented machine learning in some of their commercial products such as their search engines, and their intelligent personal assistants Cortana and Google Assistant. The recognition and classification of the diversity of materials that exist in the environmentaround us are a key visual competence that computer vision systems focus on in recent years. Understanding the identification of materials in distinct images involves a deep process that has made usage of the recent progress in neural networks which has brought the potential to train architectures to extract features for this challenging task. This project uses state of-the-art Convolutional Neural Network (CNN) techniques and Support Vector Machine (SVM) classifiers in order to classify materials and analyze the results. Building on various widely used material databases collected, a selection of CNN architectures is evaluated to understand which is the best approach to extract features in order to achieve outstandingresults for the task. The results gathered over four material datasets and nine CNNs outline that the best overall performance of a CNN using a linear SVM can achieve up to 92.5 % mean average precision, while applying a new relevant direction in computer vision, transfer learning. By limiting the amount of information extracted from the layer before the last fully connected layer, transfer learning aims at analyzing the contribution of shading information and reflectance to identify whichmain characteristics decide the material category the image belongs to. The results of the comparison emphasize the fact that the accuracy and performance of the system improves, especially in the datasets which consist of a large number of images.

Keywords: Image Classification, Convolutional Neural Network, recognition, TensorFlow, CIFAR-10, Support Vector Machine

I. INTRODUCTION

This work aims at the application of Convolutional Neural Network or CNN for image classification. Lillsand and Kiefer defined image processing as involving manipulation of digital images with the use of computer. It is a broad subject and involves processes that are mathematically complex. Image processing involves some basic operations namely image restoration/rectification, image enhancement, image classification, images fusion etc. Image classification forms an important part of image processing. The objective of image classification is the automatic allocation of image to thematic classes . Two types of classification are supervised classification and unsupervised classification. The process of image classification involves two steps, training of the system followed by testing. The training process means, to take the characteristic properties of the images (form a class) and form a unique description for a particular class. The process is donefor all classes depending on the type of classification problem; binary classification or multi-class classification. The testingstep means to categorize the test images under various classes for which system was trained. This assigning of class is donebased on the partitioning between classes based on the training features. Since 2006, deep structured learning, or more commonly called deep learning or hierarchical learning, has emerged as a new area of machine learning research. Several definitions are available for Deep Learning; coating one of the many definitions from Deep Learning is defined as: A class of machine learning techniques that exploit many layers of nonlinear information processing for supervised or unsupervisedfeature extraction and transformation and for pattern analysis and classification.

II. IMAGE CLASSIFICATION

Computational models of neural networks have been around for a long time, with McCulloch and Pitts proposing the initialmodel. Neural networks are built up of layers, each of which is coupled to the others to form the network. The strength of the connections between each pair of neurons in a feed-forward neural network (FFNN) can be thought of in terms of neuralactivation. The neurons in FFNN are connected in a directed method with defined start and stop points, i.e, the input layer and the output layers. The hidden layers are the layers that lie between these two. The goal of learning is to reduce the error between the output received from the input by adjusting the weights that goes into the input layer. The weights are adjustedby process of back propagation (in which the partial derivative of the error with respect to last

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 7.39 ⁝ Vol. 11, Issue 4, April 2022
DOI: 10.17148/IJARCCE.2022.11415

layer of weights is calculated).The process of weight adjustment is repeated in a recursive manner until weight layer connected to input layer is updated.

III. RELATED WORK

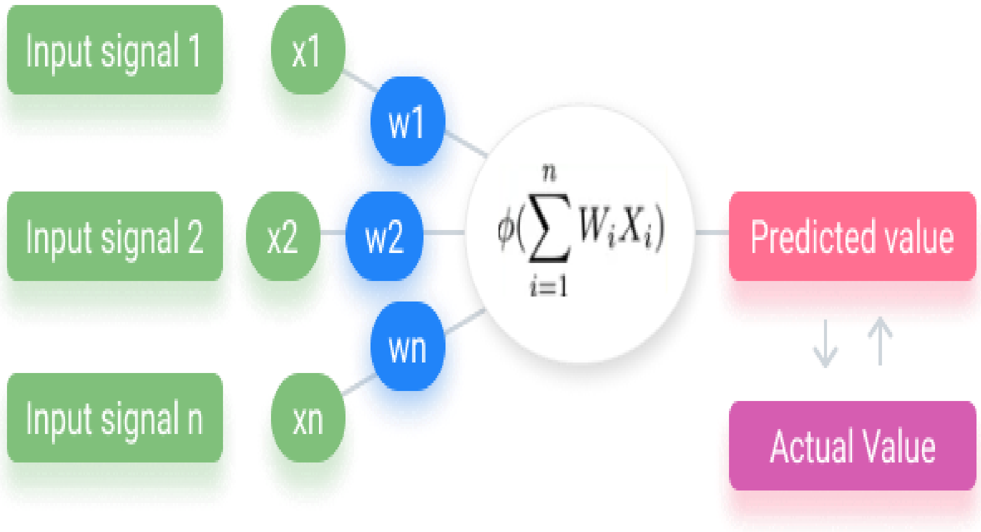
An image in the form of pixels is analysed by a computer. It accomplishes this by treating the image as an array of matrices, the size of which is determined by the image resolution. Simply speaking, picture classification is the study of statistical data utilising algorithms from a computer's perspective. Image classification is accomplished in digital image processing by automatically grouping pixels into defined groups, or "classes.” The algorithms divide the image into a series of key attributes, reducing the workload for the final classifier. These qualities help the classifier figure out what the image is aboutand which class it belongs to. The most significant stage in categorising an image is the characteristic extraction procedure,which is followed by the remainder of the steps. Image classification, particularly supervised classification, is also reliant hugely on the data fed to the algorithm. A well-optimized classification dataset works great in comparison to a bad dataset with data imbalance based on class and poor quality of images and annotations.

IV. METHODOLOGY

CONVOLUTIONAL NEURAL NETWORK

Selection process parameters plays an important role in software development as it helps to choose the best suitable model. A CNN is a framework developed using machine learning concepts. CNNs are able to learn and train from data on their own without the need for human intervention. In fact, there is only some pre-processing needed when using CNNs. They develop and adapt their own image filters, which have to be carefully coded for most algorithms and models. CNN frameworks have a set of layers that perform particular functions to enable the CNN to perform these functions.

CNN Architecture and Layer: The basic unit of a CNN framework is known as a neuron**.** The concept of a neuron is basedon human neurons. These are statistical functions that compute the weighted average of inputs and then apply an activationfunction to the output. Layers are a cluster of neurons, with each layer having a particular function.



Steps followed in the Algorithm –

- STEP 1= load the data set
- STEP 2= Have a glance at the image
- STEP 3= Convert the pixel values of the dataset to float type and then normalize the dataset.
- STEP 4= Import the required layers and modules to create our convolution neural net architecture.
- STEP 5= Split the data into train and test sets
- STEP 6= Train the algorithm
- STEP 7= Evaluating algorithm and getting predicted model.

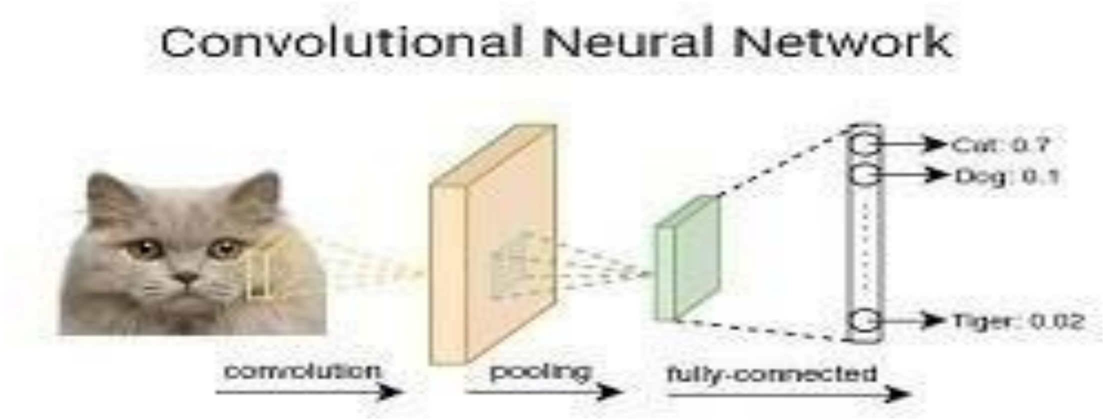
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⋇ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11415

CNN layers can be of four main types: Convolution Layer, ReLu Layer, Pooling Layer, and Fully-Connected Layer.



- **Convolution Layer:** A convolution is the basic process of applying a filter to an input to produce an activation. The convolution layer has a series of trainable filters with a limited receptive range but which can be applied to all of the data presented. In Convolution neural networks, the fundamental building elements are convolution layers.

- **ReLu Layer:** A convolution is the simple process of applying a filter to an input to produce an activation. The convolution layer has a series of trainable filters with a limited receptive range but which can be applied to all of the data presented. In convolutional neural networks, the fundamental building elements are convolution layers.

- **Pooling Layer:** The result of all neurons in the layer is collected in this layer and processes this data. The primary task of a pooling layer is to lower the number of factors being considered and give streamlined output.

- **Fully Connected Layer:** This layer is the final output layer for CNN models that flattens the input received from layers before it and gives the result.

Any project's **Cost Estimation** step is critical. It forecasts if the project investment is sufficient or whether there will be a capital shortage. It depicts the total cost of the project's development. Cost estimation should be completed prior to beginning development to avoid wasting time and resources and project failure. The formula is used to calculate thecost of any software project,

- $C = (aL)^b$
- C = cost of project,
- a = 1.4(constant),
- b = 0.93(constant),
- L = size of code = 2100 lines(avg)

V. RESULT AND DISCUSSION

Since the vast amount of image data we obtain from cameras and sensors is unstructured, we depend on advanced techniques such as machine learning algorithms to analyze the images efficiently. Image classification is probably the most important part of digital image analysis. It uses AI-based deep learning models to analyze images with results that for specific tasks already surpass human-level accuracy. Accuracy of the application is low because of large dataset. Image classification applications are used in many areas, such as medical imaging, object identification in satellite images, traffic control systems, brake light detection, machine vision, and more.

 ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11415

Normalising the dataset by dividing the pixel values by 255.

```
X_train[0]/255
```

```
array([[[0.23137255, 0.24313725, 0.24705882],
        [0.16862745, 0.18039216, 0.17647059],
        [0.19607843, 0.18823529, 0.16862745],
        ...,
        [0.61960784, 0.51764706, 0.42352941],
        [0.59607843, 0.49019608, 0.4       ],
        [0.58039216, 0.48627451, 0.40392157]],

       [[0.0627451 , 0.07843137, 0.07843137],
        [0.        , 0.        , 0.        ],
        [0.07058824, 0.03137255, 0.        ],
        ...,
        [0.48235294, 0.34509804, 0.21568627],
        [0.46666667, 0.3254902 , 0.19607843],
        [0.47843137, 0.34117647, 0.22352941]],

       [[0.09803922, 0.09411765, 0.08235294],
        [0.0627451 , 0.02745098, 0.        ],
        [0.19215686, 0.10588235, 0.03137255],
        ...,
        [0.4627451 , 0.32941176, 0.19607843],
        [0.47058824, 0.32941176, 0.19607843],
        [0.42745098, 0.28627451, 0.16470588]],

       ...,
```

fig shows Normalising the given data by dividing the pixesl value

Defining the model

Creating CNN model using Sequential() function with filter value 32 which will detect 32 different features in image. Filter size (kernel_size) is 3x3. Using activation function as relu and imput shape is 32x32x3. By using MaxPooling, selecting the maximum element from the feature map covered by the filter. Adding another convolution and MaxPooling layer for trial and error. We can add multiple layers depending upon our requirement. The flatten() function coverts multi-dimensional input into single dimension so you can model your input layer.

```
cnn = models.Sequential([
        #cnn
        layers.Conv2D(filters=32, kernel_size=(3,3), activation='relu', input_shape=(32,32,3)),
        layers.MaxPooling2D((2,2)),

        layers.Conv2D(filters=64, kernel_size=(3,3), activation='relu'),
        layers.MaxPooling2D((2,2)),

        #dense
        layers.Flatten(),
        layers.Dense(64, activation = 'relu'),
        layers.Dense(10, activation = 'softmax')
    ])
```

Compiling the model

Compiling the model by using adam optimizer. To compute crossentropy loss between the labels and predictions through 'sparse categorical croosentropy'and using metrics accuracy for evaluating the model.

Fig shows defining the model.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11415

Compiling the model

Compiling the model by using adam optimizer. To compute crossentropy loss between the labels and predictions through 'sparse categorical croosentropy'and using metrics accuracy for evaluating the model.

```
[ ] cnn.compile(optimizer = 'adam',
              loss = 'sparse_categorical_crossentropy',
              metrics = ['accuracy'])
```

Fig shows compilation of the model.

Epochs for indicating the number of passes for the entire training dataset.

```
cnn.fit(X_train, y_train, epochs = 10)
```

```
Epoch 1/10
1563/1563 [==============================] - 65s 41ms/step - loss: 1.7859 - accuracy: 0.4157
Epoch 2/10
1563/1563 [==============================] - 57s 36ms/step - loss: 1.2523 - accuracy: 0.5584
Epoch 3/10
1563/1563 [==============================] - 57s 37ms/step - loss: 1.0924 - accuracy: 0.6208
Epoch 4/10
1563/1563 [==============================] - 58s 37ms/step - loss: 0.9849 - accuracy: 0.6590
Epoch 5/10
1563/1563 [==============================] - 59s 37ms/step - loss: 0.9139 - accuracy: 0.6844
Epoch 6/10
1563/1563 [==============================] - 60s 38ms/step - loss: 0.8525 - accuracy: 0.7052
Epoch 7/10
1563/1563 [==============================] - 59s 38ms/step - loss: 0.7973 - accuracy: 0.7243
Epoch 8/10
1563/1563 [==============================] - 59s 38ms/step - loss: 0.7534 - accuracy: 0.7411
Epoch 9/10
1563/1563 [==============================] - 59s 38ms/step - loss: 0.7101 - accuracy: 0.7554
Epoch 10/10
1563/1563 [==============================] - 59s 38ms/step - loss: 0.6796 - accuracy: 0.7643
<keras.callbacks.History at 0x7fc2bf893910>
```

Fig shows epochs for the entire training dataset

VI. CONCLUSION

In conclusion, this research is about image classification by using machine learning via framework TensorFlow. It has three (3) objectives that have achieved throughout this project. The objectives are linked directly with conclusions because it can determine whether all objectives are successfully achieved or not. It can be concluded that all results that have been obtained, showed quite impressive outcomes. The convolutional neural network (CNN) becomes the main agenda for this research, especially in image classification technology. CNN technique was studied in more details starting from assembling, training model and to classify images into categories. The roles of epochs in CNN was able to control accuracy and also prevent any problems such as overfitting. Implementation of image classification by using framework TensorFlow also gave good results as it is able to simulate, train and classified with up to 80 % percent of accuracy. The next step is toanalyze the solution and study ways to improve the system. Some improvements could be carrying by collecting more quality data, trying more convolutional neural network architectures, or redesigning the vision system.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 7.39 ⋇ Vol. 11, Issue 4, April 2022
DOI: 10.17148/IJARCCE.2022.11415

REFERENCES

[1] Gregor,K., Danihelka, I., Graves, A., Rezende, D. J., Wierstra, D. (2015). DRAW : A Recurrent Neural NetworkFor Image Generation. https://doi.org/10.1038/nature14236

[2] P.Korytkowski, M., Rutkowski, L., Scherer, R.(2016). Fast image classification by boosting fuzzy classifiers. Information Sciences, 327,175-182.https://doi.org/10.1016/j.ins.2015.08.030

[3] S Chauhan, k., Ram, S. (2018) . International Journal of Advance Engineering and Research Image Classification with Deep Learning and Comparison between Different Convolutional Neural Network Structures using Tensorflow and Keras, 533-538.

[4] Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional neural networks. arXiv preprint arXiv:1311.2901 (2013).

