JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# Using Convolutional Neural Networks for Image Classification

**Prof. N.V. Gawali, Kanishk Pagare, Shubham Daundkar,, Dipak Wagh**

Computer Engineering Department,

Pune District Education Association's College of Engineering, Pune.

**Abstract –** *In the ever so rapidly growing tech world with lots of advancements happening each day, we need image classification for data generation, automation and AI. Image classification is a method of classification based on the statistics in this image. It is a subset of Deep Learning. This paper focuses on classification on images using CNN. CNN is a convolutional neural network which is a part of Machine Learning and Artificial Intelligence. CIFAR-10 dataset has been used for the classification.*

***Key Words***: Machine Learning, Artificial Intelligence, CNN, CIFAR-10 dataset, Image classification.

## 1. INTRODUCTION

Today, with the increasing instability, requirement and applications of artificial intelligence, fields like machine learning, and its subsets, deep learning and neural networks have gained huge momentum. The intent of the classification process is to classify all pixels in a digital image into one of several classes. Usually, multi-spectral data is used to execute the classification process, and the spectral pattern present within the data for each pixel is used as the numerical basis for classification. The objective of image classification is to identify and portray the structures occurring in an image, as a unique grey colour. Image classification is possibly the most important part of digital image analysis. Classification between objects is a complex task and therefore image classification has been an important task within the field of computer vision. Classification of an image is done into classes. There are theoretically n number of classes in which a given image can be classified. Manually checking and classifying images could be a mind-numbing task especially when they are large in number and therefore it will be very useful if we could automate this entire process using computer vision. The developments in the field of self-driving vehicles also serve as a great example of the use of image classification in the real-world. The applications include automated image organization, random photos and videos, visual search for better product discovery, large visual databases, image and face recognition on social media, and many more; which is why, we need image classifiers to achieve maximum possible accuracy.

**Working of CNN:**

Convolutional Neural Network (CNN) are designed to recognize visual patterns directly from pixel images with minimal pre-processing. It is a special architecture of artificial neural networks. CNN is involves two very simple elements, namely convolutional layers and pooling layers. The elements of CNN, such as convolutional and pooling layers, are relatively straightforward to understand. The challenging part of using CNN in practice is how to design model architectures that use these simple elements in the best way. CNN is vastly popular because of its architecture and it does not require feature extraction. The system learns to do feature extraction and the core concept is, it uses complexity of image and filters to generate constant features which are passed on to the next layer. The most commonly used architectures of CNN are LeNet, AlexNet, ZFNet, GoogLeNet, VGGNet, and ResNet

**Categories of CNN layers:**

1. **Input Layer:** A raw image is provided as an input containing width and height 32 each and deepness of 3.
2. **Convolution Layer:** a dot product is calculated between the filters and patch of an image. For an instance, we have 8 filters for a layer then output will 32*32*8
3. **Activation Function Layer:** An activation function such as sigmoid, RELU is applied by this layer.
4. **Pooling Layer:** This layer is intermittently interleaved in the covnets therefore reducing the size of volume thus reducing memory and preventing it from over-fitting.
5. **Fully-Connected Layer:** by taking output of the previous layer as input, it yields 1D array whose size is equal to the number of classes.

The basic CNN architecture is as follows:

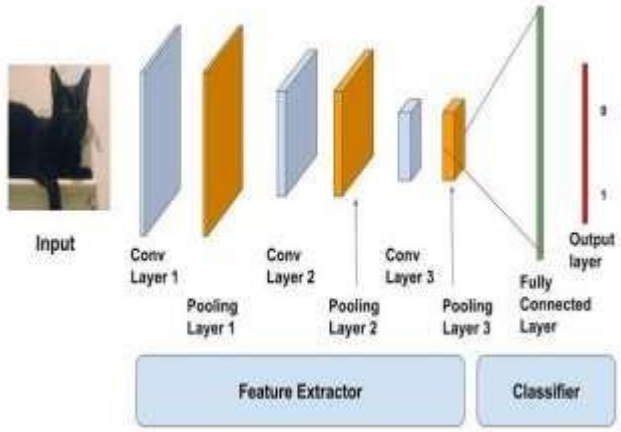


**Fig: CNN Architecture**

**Working of Image Classification:**

1. **Image Pre-processing**: In this process, the image data is filtered by supressing unwanted distortions in the image so that the classifiers can benefit from this improved data to work on. Steps for image pre-processing includes Reading image, Resizing image, and Data Augmentation (Gray scaling of image, Reflection, Gaussian Blurring, Histogram, Equalization, Rotation, and Translation).

2. **Detection of an object**: Detection refers to the localization of an object which means the breakdown of the image and classifying the position of the object of interest.

3. **Feature extraction and training**: This is a key step where statistical methods are used to identify the patterns in the image; features that might be unique to a particular class and will later on, help the model to differentiate between different classes. This process where the model learns the features and patterns from the dataset is called model training.

4. **Classification of the object**: This step classifies identified objects in the image data into predefined classes by using a suitable classification technique that compares the image patterns with the target patterns.

**CNN Models:**

The Alex Net, Google Net and ResNet50 are the most used CNN for classification of objects from the images.

The construction of those networks' compares in terms of internal layers and techniques used. GoogleNet accomplish various sizes of convolutions which concatenate the filters for the following layer. In estimate to Alex Net, it makes use of output of the previous layer due to the fact the input to the layer. ResNet is a short name for Residual Network. Instead of seeking to learn some features, we try to learn some residual. Residual is a subtraction of function found from enter of that layer. ResNet does this with the useful supply of right away connecting enter of nth layer to a few (n+x)th layer. By using this technique, training of units to CNN has become easier which has result in development within the accuracy rate.

**Comparative Study of CNN Models:**

500 images are separated into 400 teaching and learning images and 80 testing images for every category creating a dataset of 40000 several images. These 80 groups are oragnized into 15 super categories. The selected classes for learning and testing are Buildings, Trains, Desert, River, City, Streets and Nature

| CIFAR-10 | | Alex Net | Google Net | ResNet50 |
|---|---|---|---|---|
| Image classes | Buildings | 0.0% | 70.1% | 50.2% |
| | Trains | 22.4% | 73.9% | 56.2% |
| | Desert | 84.9% | 62.8% | 36.8% |
| | River | 00.0% | 73.12% | 33.5% |
| | Streets | 23.6% | 79.2% | 34.5% |
| | Trees | 31.5% | 94% | 92% |
| | Cars | 32% | 78% | 34% |
| | Mountains | 87% | 73% | 63% |

Fig 3: Performance of CNN on CFIR-10 test datasets

| CIFAR-10 | | Alex Net | Google Net | ResNet50 |
|---|---|---|---|---|
| Image classes | Dry Fruits | 6.0% | 70.1% | 49.82% |
| | Toffees | 32.4% | 74.9% | 53.2% |
| | Benches | 94.9% | 68.8 | 36.8% |
| | Desks | 10.0% | 73.12% | 34.5% |
| | Laptops | 28.6% | 79.2% | 37.5% |
| | Smartphones | 27.3% | 28% | 36.5% |
| | Projectors | 35% | 25% | 35% |
| | Desktops | 47.56% | 89% | 78% |
| | Televisions | 41.5% | 95.78% | 93.4% |

Fig 4: Performance of CNN on CFIR-10 test datasets

**Results:**

The prediction accuracy rate is done by testing the various images using different CNN network. Following table summarize the accurate rate of the images.

| | CLASSES | ALEX NET | GOOGLE NET | RESNET50 |
|---|---|---|---|---|
| PREDICTION ACCURACY RATE | Buildings | 45 | 35 | 55 |
| | Trains | 16 | 17 | 23 |
| | Desert | 0 | 4 | 7 |
| | River | 2 | 7 | 18 |
| | Streets | 1 | 17 | 10 |

| | | | | |
|---|---|---|---|---|
| | Mountains | 3 | 35 | 7 |
| | Trees | 4 | 3 | 6 |
| | Smartphones | 0 | 0 | 0 |
| | desktops | 6 | 18 | 8 |

Fig 5: Accuracy rate of various images done by CNN Models

The above table clearly summaries that Google Net and ResNet50 has better precision in classifying the objects in the image compared to Alex Net.

**Conclusion:**

The work evaluated the prediction accuracy of 3 different convolutional neural networks (CNN) on most prominent education and check datasets considerably CIFAR10. The fundamental resolution was to seek out the accuracy of different networks on same datasets and analyze the stability of prediction by each of those CNN. . We have a trend to target our study on fifteen categories of every dataset solely.

**References:**

**[1] G. Padmini[1], Pacchipulusu Kiran[1],Sri Gadu Srinivasa Rao[2] (2022) "Deep Convolutional Network for Image Recognition Using Gradio"**
**[2] Siddhant Dani, Prof. P. S. Hanwate, Hrishikesh Panse, Kshitij Chaudhari, Shruti Kotwal (2021) "Survey on the use of CNN and Deep Learning in Image Classification"**
**[3] Sakshi Parikh, Kartik Parekh, Hrithik Chauhan, Dhruv Desai (2020) "Image Classification using CNN"**

**[4] Mr. Utpal Shrivastava, Dr. Vikas Thada (2019) "Optimized Multi Class Classification of Images Using Deep Learning"**
**[5] *Ridhima Dhawan (2018)* "Classification of Image using Convolutional Neural Network"**
**[6] Kavish Sanghvi (2020) "Image Classification Techniques" https://medium.com**

**[7] www.github.com**