

A review:deep learning technique for image classification

Vishali Aggarwal^{1*} and Gagandeep²

M. Tech Research Scholar, Department of Computer Science, Punjabi University, Patiala¹

Associate Professor, Department of Computer Science, Punjabi University, Patiala²

©2018ACCENTS

Abstract

The fundamental reason for the work exhibited in this paper, is to break down the idea of a deep learning algorithm to be specific, convolutional neural networks (CNN) in image classification. A study of deep learning, its strategies, comparison of frameworks, and algorithms is presented. The significance of (sufficient) training has been considered. The advancement has shown imperative execution in various vision assignments, for instance, image identification, question area and sementic division. In particular, late advances of deep learning procedures pass on asking execution to fine-grained image classification which intends to perceive subordinate-level classifications.

Keywords

Convolutional neural networks, Residual learning, Batch normalization, Deep learning.

1.Introduction

Deep Learning, as a branch of Machine Learning, uses algorithms to process information and emulate the thinking procedure, or to create abstractions. deep learning (DL) utilizes layers of algorithms to process information, comprehend human speech, and perceive objects. Data is gone through each layer, with the yield of the previous layer giving contribution to the following layer as appeared in fig1.1. The primary layer in a system is known as the input layer, while the latter is called an output layer. Every one of the layers between the two is called as hidden layers. Each layer is a straightforward, uniform calculation containing one sort of activation function [1].

Feature extraction is another part of Deep Learning. Feature extraction utilizes a calculation to naturally develop significant "features" of the information for reasons for training, learning, and understanding.

Deep Learning methods are one promising road of research into the computerized extraction of complex information portrayals (highlights) at high levels of abstraction. Such calculations build up a layered, hierarchical architecture of learning and presenting information, where higher-level (more dynamic) highlights are characterized in terms of lower-level (less unique) highlights.

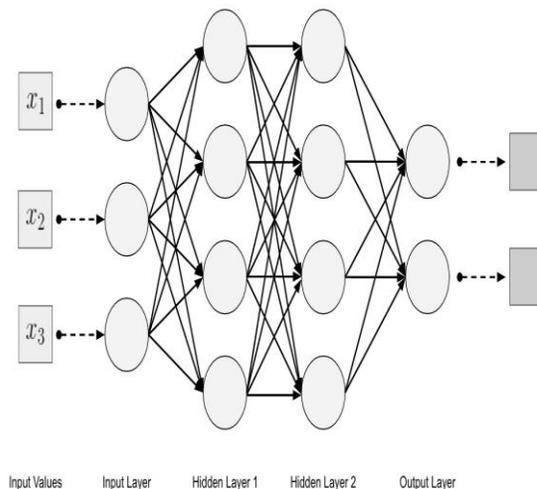


Figure 1 Deep neural network

The progressive learning engineering of Deep Learning procedure is inspired by artificial intelligence imitating the deep, layered learning procedure of the essential sensorial zones of the neocortex in the human brain, which consequently separates highlights and reflections from the basic information [2, 3]. Deep Learning calculations are very helpful when managing learning from a lot of unsupervised information, also, ordinarily learn information portrayals in a greedy layer-wise design[4,5].

*Author for correspondence

2.Applications of deep learning

Deep learning has occupied an importance place in almost every field. Its applications can be viewed in important areas of industry. Applications of deep learning are as following:

2.1Image colorization of black and white images

Image colorization is the issue of adding color to black and white photos. Deep learning can be utilized to utilize the items and their setting inside the photo to shading the picture; much like a human administrator may approach the issue.

2.2Naturally adding sounds to silent movies

In this assignment the framework must combine sounds to coordinate a noiseless video. The framework is prepared using 1000 cases of video with sound of a drum stick striking diverse surfaces and making distinctive sounds. A deep learning model links the video outlines with a database of pre-recorded sounds keeping in mind the end goal to choose a sound to play that best matches what is going on in the scene.

2.3Programmed machine translation

This is the work where given words, expression or sentence in one language; consequently make an interpretation of it into another language.

Deep learning is achieving top results in two specific areas:

- Automatic Translation of Text.
- Automatic Translation of Images.

2.4Image classification and detection in photographs

This requires the grouping of items inside a photo as one of an arrangement of already known articles. A more composite variety of this task called object recognition includes particularly recognizing at least one object inside the scene of the photo and drawing a crate around them.

2.5Automatic handwriting generation

This is where given a corpus of handwriting cases, produce new handwriting for a given word or expression. From this corpus the connection between the pen movement and the letters is found out and new cases can be created ad hoc.

2.6Automatic image caption generation

Programmed picture inscribing is the process where given a picture the framework must produce an inscription that portrays the content of the image. The

frameworks include the utilization of extensive convolutional neural systems for the object identification in the photos and afterward an recurrent neural network to transform the marks into an intelligent sentence.

3.Image classification

Image classification frames an essential part of image processing. The goal of image classification is the programmed allotment of image to topical classes [6]. Two sorts of classification are supervised classification and unsupervised classification.

The process of image classification includes two steps; training of the framework took after by testing. The training process implies, to take the characteristic properties of the images (frame a class) and shape a special description for a specific class. The process is done for all classes relying upon the sort of classification problem; binary classification or multi-class classification [7-10]. The testing step intends to classify the test images under different classes for which system was trained. This assigning of class is done based on the partitioning between classes in view of the training features.

4.Image classification procedure

It has been seen that the task in Image Classification undertaking is to take an array of pixels that represents a single image and assign a label to it [11-20]. Complete pipeline can be formalized as follows:

Input: Input comprises of a collection of N images, each marked with one of K distinctive classes. This data is referred to as the training set.

Learning: Task is to utilize the training set to realize what each one of the classes resembles. This step is referred to as training a classifier, or learning a model.

Evaluation: In the end, the nature of the classifier asking it to predict labels for a new set of images that it has never seen before is evaluated. True labels of these images are then compared to the ones anticipated by the classifier.

Different estimations of those elements used for the CNNs are considered [7]. From the results of experiments, it can be discussed as follows:

With a series of control comparisons, we progressively modify a baseline model as

1. Network width in a stage-a phase characterizes as those layers between two adjacent pooling layers.
2. Different convolutional filter size.
3. Different depth in a stage.
4. Phenomenon in deeper convolution with different width of network in a stage.
5. Phenomenon in deeper convolution with different size of filter in a stage.

5.Related work

CIFAR dataset is utilized for investigating impact of different CNNs structure in all the cases.

CIFAR-10 is a collection of natural color images of 32x32 pixels. It comprises of 60,000 images in 10 classes, with 6,000 images for each class. There are 50,000 training images and 10,000 testing images.

The classes are totally fundamentally unrelated and marked by hand.

Deeper neural networks are harder to train [8]. A residual learning system to facilitate the training of networks was presented. At the point when deeper networks are able to start converging, a debasement problem has been uncovered: with the system depth expanding, precision gets immersed (which may be obvious) and afterward corrupts quickly. Surprisingly, such degradation isn't caused by overfitting, and adding more layers to an appropriately deep model prompts higher training error as depicted in [9,10]. *Table 1* shows the comparative study.

Table 1 A comparative study of various deep learning networks, datasets and their accuracy

| Paper name | Dataset | DL Network | Learning technique | Accuracy |
|-----------------|--------------|--------------------|---------------------|----------|
| Zhang K (2017) | BSD68 | DnCNN ¹ | Batch Normalization | 82% |
| Kolsch A (2017) | TOBACCO-3482 | ELM ² | Transfer Learning | 83.25% |
| He K (2016) | CIFAR-10 | DnCNN | Residual | 85% |
| Liu P (2015) | CIFAR-10 | CNN ³ | Supervised | 88% |
| Jaswal D (2014) | UCML | CNN | Unsupervised | 93.18% |

¹Denoising Convolutional Neural Network (DnCNN)

²Extreme Learning Machine (ELM)

³Convolution Neural Network (CNN)

In [19] a method has been proposed for document classification that relies on codewords derived from patches of the document images. The code book is learned in an unsupervised way on the documents. To do that, the approach recursively partitions the image into patches and models the spatial relationships between the patches using histograms of patch-codewords. Two years later, the same authors presented another method which builds a codebook of SURF descriptors of the document images [20].

In [21] and [22] showed a great improvement in the accuracy by applying transfer learning from the domain of real-world images to the domain of document images, thus making it possible to use deep CNN architectures even with limited training data. With their approach, they significantly outperformed the state-of-the-art at that time. Furthermore, [22] introduced the RVL-CDIP dataset which provides a large-scale dataset for document classification and allows for training CNNs from scratch.

In terms of image classification, PASCAL Visual Object Classes has long been a popular benchmark. Deep learning systems from different research groups all reported impressive performance on the classification task [23-26].

6.Learning techniques

Any aspect of an agent can be improved through learning. Learning methodology can improve the results which depend on the following factors:

- What component to be improved
- Prior knowledge of the agent
- Representation used for data & the component
- What feedback is available for learning?

Two methods related to DnCNN were briefly reviewed in [11], i.e., residual learning and batch normalization.

Residual learning: Residual learning [5] of CNN was initially proposed to tackle the performance degradation problem, i.e., even the training precision starts to decrease alongside the expanding of network depth. By expecting that the residual mapping is easier to learn than the first unreferenced mapping, residual system explicitly takes in a residual mapping for a couple of stacked layers. With such a residual learning methodology, to a great degree deep CNN can be effortlessly trained and accuracy for image classification and object detection can be improved [5].

Batch normalization: Batch normalization [12, 13] is proposed to ease the internal covariate shift by consolidating a normalization step and a scale and a shift step before the nonlinearity in each layer. For batch normalization, just two parameters for each activation are included, and they can be updated with back-propagation. Batch normalization has a few benefits, for example, quick training, better execution, and low affectability to initialization.

Mini-batch stochastic gradient descent (SGD) has been generally utilized as a part of the training of deep CNNs. The training efficiency of deep CNNs is extremely lessened by inner covariate shift [14], which is change of internal input distributions during training. Batch normalization [14] is proposed to lighten interior covariance shift by changing internal inputs with a scale and shift step before nonlinear activation. Mean and variance of every activation are calculated to normalize features [15] by equations (1) and(2).

$$\bar{x}_f = \frac{1}{m} \sum_{i=1}^m x_{i,f} \quad (1)$$

$$\sigma_f^2 = \frac{1}{m} \sum_{i=1}^m (x_{i,f} - \bar{x}_f)^2 \quad (2)$$

where m is the measure of a mini-batch, and $x_{i,f}$ is the f th highlight of the i th test in the mini-batch. Utilizing mini-batch mean and variance, we can normalize each element as follows:

$$\widehat{x}_f = \frac{x_f - \bar{x}_f}{\sqrt{\sigma_f^2 + \epsilon}} \quad (3)$$

where ϵ is a small positive constant to enhance numerical dependability. However, the normalization of input features decreases the presentation capacity of the data sources. To take care of this issue, batch normalization presents two learnable parameters γ_f and β_f for each component f . Batch normalization is characterized by transforming normalized features with a scale and shift step:

$$\text{BN}(x_f) = \gamma_f \widehat{x}_f + \beta_f \quad (4)$$

The system can recoup the distribution of the actual contributions by the scale and shift step. Batch normalization has been generally utilized for CNN-based image classification [16, 17].

7. Conclusion and future work

The paper reviews some current advance in deep learning based image classification. Deep learning technique is an approach to improve the training

accuracy yet the cost gets multiplied by utilizing all the more training parameters. The streamlined training process is required. It has been observed that till date, there has been no order technique that works best on any given issue.

1. In future our concentration will be to improvise learning systems in deep learning and supervised learning in the vision group.
2. Batch normalization can be used to classify Cifar10 dataset in order to improve accuracy.
3. The substitution of a large filter size with stacked 2x2 convolutional neural networks likewise can enhance the recognition accuracy and reduce the time complexity.

Acknowledgment

None.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

- [1] Neterer JR, Guzide O. Deep learning in natural language processing. Proceedings of the West Virginia academy of science. 2018; 90(1).
- [2] Bengio Y, LeCun Y. Scaling learning algorithms towards AI. Large-scale kernel machines. 2007; 34(5):1-41.
- [3] Arel I, Rose DC, Karnowski TP. Deep machine learning-a new frontier in artificial intelligence research [research frontier]. IEEE Computational Intelligence Magazine. 2010; 5(4):13-8.
- [4] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. Neural Computation. 2006; 18(7):1527-54.
- [5] Bengio Y, Lamblin P, Popovici D, Larochelle H. Greedy layer-wise training of deep networks. In Advances in Neural Information Processing Systems 2007 (pp. 153-60).
- [6] Lillesand T, Kiefer RW, Chipman J. Remote sensing and image interpretation. John Wiley & Sons; 2014.
- [7] Jaswal D, Vishvanathan S, Kp S. Image classification using convolutional neural networks. International Journal of Scientific and Engineering Research. 2014; 5(6):1661-8.
- [8] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In proceedings of the conference on computer vision and pattern recognition 2016 (pp. 770-8). IEEE.
- [9] Liu PH, Su SF, Chen MC, Hsiao CC. Deep learning and its application to general image classification. In informative and cybernetics for computational social systems (ICSS), International Conference on 2015 (pp. 7-10). IEEE.
- [10] He K, Sun J. Convolutional neural networks at constrained time cost. In conference on computer vision and pattern recognition 2015 (pp. 5353-60). IEEE.

- [11] Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*. 2017 ; 26(7):3142-55.
- [12] Chan TH, Jia K, Gao S, Lu J, Zeng Z, Ma Y. PCANet: A simple deep learning baseline for image classification. *IEEE Transactions on Image Processing*. 2015; 24(12):5017-32.
- [13] Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E. Deep learning applications and challenges in big data analytics. *Journal of Big Data*. 2015; 2(1):1-21.
- [14] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*. 2015.
- [15] Yin Z, Wan B, Yuan F, Xia X, Shi J. A deep normalization and convolutional neural network for image smoke detection. *IEEE Access*. 2017; 5:18429-38.
- [16] Druzhkov PN, Kustikova VD. A survey of deep learning methods and software tools for image classification and object detection. *Pattern Recognition and Image Analysis*. 2016; 26(1):9-15.
- [17] Gao R, Grauman K. On-demand learning for deep image restoration. In *proceeding IEEE conferences. Computer Vision and Pattern Recognition 2017* (pp. 1086-95).
- [18] Zhao B, Feng J, Wu X, Yan S. A survey on deep learning-based fine-grained object classification and semantic segmentation. *International Journal of Automation and Computing*. 2017; 14(2):119-35.
- [19] Kumar J, Ye P, Doermann D. Learning document structure for retrieval and classification. In *Pattern Recognition (ICPR), 21st International Conference on 2012* (pp. 1558-61). IEEE.
- [20] Kumar J, Ye P, Doermann D. Structural similarity for document image classification and retrieval. *Pattern Recognition Letters*. 2014; 43:119-26.
- [21] Afzal MZ, Capobianco S, Malik MI, Marinai S, Breuel TM, Dengel A, et al. DeepDocClassifier: document classification with deep convolutional neural network. In *document analysis and recognition (ICDAR), 13th international conference on 2015* (pp. 1111-5). IEEE.
- [22] Harley AW, Ufkes A, Derpanis KG. Evaluation of deep convolutional nets for document image classification and retrieval. In *document analysis and recognition (ICDAR), 2015 13th international conference on 2015* (pp. 991-5). IEEE.
- [23] Chen Q, Song Z, Hua Y, Huang Z, Yan S. Hierarchical matching with side information for image classification. In *computer vision and pattern recognition (CVPR), IEEE Conference on 2012* (pp. 3426-33). IEEE.
- [24] Dong J, Xia W, Chen Q, Feng J, Huang Z, Yan S. Subcategory-aware object classification. In *computer vision and pattern recognition (CVPR), IEEE Conference on 2013* (pp. 827-34). IEEE.
- [25] Chen Q, Song Z, Dong J, Huang Z, Hua Y, Yan S. Contextualizing object detection and classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2015; 37(1):13-27.
- [26] Kölsch A, Afzal MZ, Ebbecke M, Liwicki M. Real-time document image classification using deep CNN and extreme learning machines. *arXiv preprint arXiv:1711.05862*. 2017.



Vishali Aggarwal received her B.Tech degree in Computer Science and Engineering from RBIEBT, Mohali, with 80% in 2015. She is currently pursuing M.Tech from Punjabi University, Patiala. Her area of research is Neural Networks and Image Processing.

Email: vishalisingla30@gmail.com



Dr. Gagandeep received her Ph.D degree in Computer Science from Punjabi University, Patiala in 2012 on the topic "Analysis and Development of Testing Techniques for Component-based Software Systems". Currently, she is working as Associate Professor in the Department of Computer Science, Punjabi University, Patiala. She has more than 20 years of teaching and research experience. She is currently involved in research work on data security, virtualization techniques and computational offloading mechanism for cloud and mobile computing.