

Deep Learning Concepts and Libraries Used in Image Analysis and Classification

K. Sai Prasad¹, Dr. S Pasupathy²

¹Department of Computer Science and Engineering, MLR Institute of technology, Dundigal, Hyderabad.

²Department of Computer Science and Engineering, Annamali University, Annamalai Nagar, Chidambaram, Tamilnadu.

Article Info

Volume 82

Page Number: 7907 - 7913

Publication Issue:

January-February 2020

Abstract

Deep learning has gained much importance in today's digital world. There are many algorithms in the industry which talk about how Convolutional neural networks or deep learning algorithms can be used to parse, analyze and predict type of an image and on how to extract image features or properties. Sometimes if the algorithm is not properly written and maintained then we may get wrong predictions and incorrect output results. It depends on the type of the data used in training the model.

To avoid such misconceptions, we here describe the hyper parameters and concepts that will help in writing better and optimized algorithm. Parameters like strides, epochs, pooling or number of layers are explained with examples. These parameters play key role in feature extraction. Here we also explain about the different technologies available in the industry till 2018, to write a deep neural network code along with different python libraries that can be used in image processing. Any researcher who want to do a statistical analysis in deep learning need to undergo various metrics in understanding the concepts. To make the task of researchers much easy we have grouped various Deep learning concepts and libraries as a study.

Article History

Article Received: 18 May 2019

Revised: 14 July 2019

Accepted: 22 December 2019

Publication: 05 February 2020

Keywords: Convolutional Neural Networks, Deep Learning, Image classification, Neural networks, Medical Image.

1. INTRODUCTION

Generally, Pathologist collects tissue samples of human body or plants on the glass slides as part of disease tests. These slides are examined under the microscope to find the characteristics of existence of disease. The study of tissues to find the presence of disease is called histopathology. There would be pathological reports generated as part of histopathological examination.

1.1 How does deep learning work?

A deep learning model is designed to continually analyze data with a logic structure similar to how

a human body would draw conclusions based on the knowledge. To achieve this, deep learning uses layered structure of algorithms called an **Artificial Neural Network(ANN) and Convolutional Neural Network(CNN)**.

Biological neural network structures of the human brain are the inspirational models for a typical artificial neural network which makes ANN to understand and perform than many of machine learning algorithms.

1.2 Best image processing libraries to use in python:

1. NumPy

Python NumPy provides support for large, multi-dimensional arrays. Using NumPy, we can easily represent images as multi-dimensional arrays.

2. PIL (Python Image Library) now is PILLOW

This is a free python library can be used to open, edit or save an image in different formats. This library can be used to filter a blurred image or to smooth an image or find edges.

3. OpenCV (Open Source Computer Vision Library)

This library has huge base for inbuilt optimized algorithms around 2500+. These algorithms can be used to distinguish different objects, detect and recognize faces and also classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects and many more.[8]

4. SimpleCV

This package does the similar jobs as OpenCV library. The learning curve is smaller than OpenCV [9]

5. Scikit-image

This package contains a collection of algorithms for image processing. It includes algorithms for segmentation, feature extraction from text or image datasets, prediction methods, geometric transformations, color space manipulation, analysis, filtering, morphology and more [10] [11].

This package is designed on top of other libraries NumPy, SciPy, Matplotlib, Pandas.

1.3 Top Libraries used in Deep Learning:

As per the study we have found 16 Top libraries used in deep learning for various purposes.

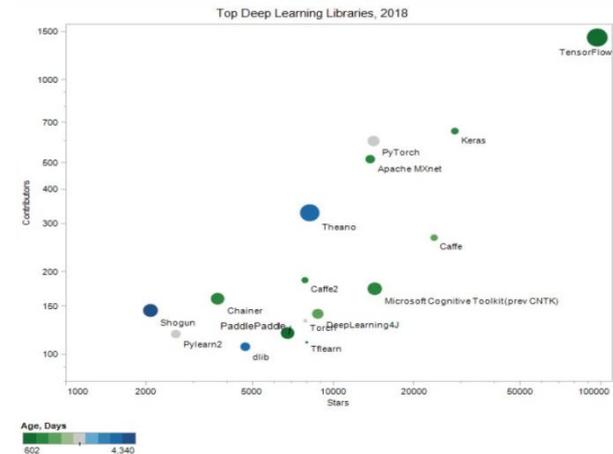


Fig1.2018 top deep learning libraries list

Figure 1 explains about the stars given on github and contributors for various libraries. TensorFlow is the top most used deep learning library in the industry, Keras and Caffe are the next in the list.

Details about few mostly used libraries:

1. TensorFlow

TensorFlow is open source and free software developed by google research team. This library has powerful functions to support in designing the network models with ease, provides lite versions of TensorFlow based on the need and sophisticated machine learning models. TensorFlow can be used with python, java, R and many more technologies.

2. Keras

Keras is an open source library. The main intension behind developing this library is to speed up the experiments with small sized prototypes. This library is more user friendly as its designed for faster prototypes. Keras can work with TensorFlow or Theano with inbuilt APIs. The error handling is well maintained with Keras.

3. PyTorch :

PyTorch has strong GPU acceleration techniques.

4. Theano:

Theano helps to work with multi-dimensional arrays along with NumPy library more efficiently. This library has inbuilt APIs to design, define mathematical expressions with ease and also supports several optimization techniques.

2. Basic concepts to understand in neural network and deep learning technologies:

Epoch:

The number of epochs is the number of times the learning algorithm has to go through the data set to understand the data. Each time the algorithm reads the training data, it minimizes the error and improves the predictability and accuracy of the model.

Number of epochs parameter in the algorithm is generally configured to a large number. There is no specific number. It can be 10, 100, 500 or 1000 even larger as well.

One epoch means each sample in the dataset gets an opportunity to update the internal model parameters.

This can be visualized as a for loop over the number of epochs where each loop proceeds over the dataset. Within this for loop another nested loop will be there that iterates each batch of the samples.

Weights:

Each node or neuron is a basic computational unit in the neural network. Each node takes more than one input values and produces one output by applying a formula on the input values. Input is basically a weight. The formula if applied on all these weights a biased value. These weights are not fixed. The adjustments to the weights are made during the model learning process. The

weights vector and biased value represents some of the features of the input image.

Following image talks about what happens at each neuron.

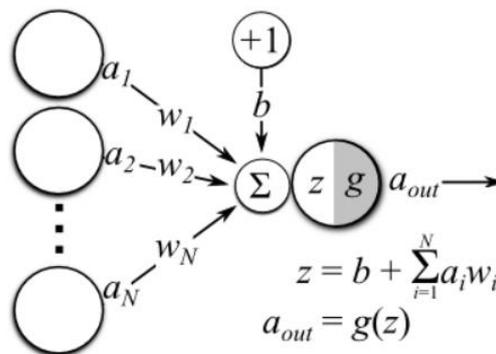


Fig 2: Weights and bias value representation on the nodes

Activation Function:

Activation function play an important role in the artificial neural network. Activation functions basically applied on the intermediate neurons and they decide whether to activate a neuron or not based on the inputs received on that node, weights and bias value. The following formula is one type of activation function.

$$Y = \text{Activation}(\Sigma(\text{weight} * \text{input}) + \text{bias})$$

If there are no activation functions used in the neural network, then it will become a linear regression and complex problems can't be solved by linear regression. Activation function basically makes the nonlinear transformation of the inputs such that model can learn new characteristics so that complex tasks can be performed. There are wide variety of activation functions available: Binary activation function, linear regression function, Sigmoid, Tanh and ReLU. The most used activation function is ReLU(Rectified Linear Unit).

Learning Rate:

Each model in the neural network starts understanding the input dataset by analyzing that dataset. This process happens back and forth in the network. During each run a weight is assigned or updated to each node in the network based on the data. The process is called back-propagation technique as the weights are assigned back to the nodes.

The amount of weights that are updated during the training of the model is called learning rate or step size.

For example, if the learning rate is set as 0.1, then the weights in the network are updated as $0.1 * (\text{estimated weight error value})$

The speed at which the model understands the input dataset depends on the learning rate. If the learning rate is set to high, the model will learn faster and there would be high chances of not assigning optimal weights as the number of back and forth runs in the training would be less. The resultant model may not provide accurate results. If the learning rate is set to very low, the model may take more time to train and sometimes the model may get stuck during the training process. Researchers have proven we may not get proper model when we use large learning rate or smaller learning rate. So we should consider using average set of weights.

Categorical Data:

In general data can be segregated into different groups. Each group represents a category of data. In the programming world, variables used to represent different groups are called categorical variables. For ex, gender or age are variables that can be used to categorize people. These variables can contain numerical values or string labels.

The data has to be represented in numerical format or in array format for a machine learning algorithm to understand the data easily. If incase

the dataset contains labels, then that data has to be converted into numerical format. It's because extending the numerical data format would be easier than extending the data which in string label format.

Integer or label encoding and one-hot encoding are most used data transformation techniques.

Integer or label encoding:

As the name indicates, an integer or a label is assigned to each of the input data category. If the data is assigned with integers, then if the new data comes it will be easy to increment the integer and assign to that new data. This is easy to implement but it would be difficult for the algorithms while prediction. Algorithms may consider the natural ordering of the integer numbers as feature of the input data which may not be true and prediction may go wrong. So integer encoding may not help algorithms most of the times in learning as well as prediction.

One-Hot encoding:

This encoding solves most of the problems that are present in integer type encoding. Each of the data is represented in binary sequence format instead of integers. These binary sequence would be easy to represent in array and can be understood easily by the algorithms.

Below table talks about how three different colors are represented. Color red will be represented as 100, green as 010 and blue as 001.

Red	Green	Blue
1	0	0
0	1	0
0	0	1

If more number of color values are added to the list of possible colors, then It would be easy to extend the one-hot encoding strategy for those new colors.

Strides:

Generally, computer understands an image by parsing the image from left to right direction and then from top to bottom direction. In the similar way if a person is reading the newspaper with a magnifier glass, then that person will move the magnifier from left to right direction first and then from top to bottom.

Here the person is using magnifier glass to read the newspaper. Similar computer uses a filter (an array of size $m \times n$ which would be lesser in size than the input image). This filter matrix is called a **stride**.

The stride is moved over the input image array to produce the shirked output array. This way the features present in the input image are extracted and stored in output array.

The process of moving the filter stride over the input array is repeated by changing the filter arrays.

The number of pixels the filter array needs to be moved is defined a stride parameter.

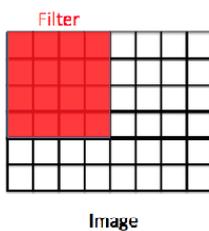


Fig 3: Input image array representation. Slots in red color shows the input filter

If the stride is set to 1 then the filter will move one pixel right as below

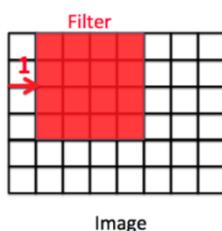


Fig 4: Representation of Filter moved by one stride

If the stride is set to 2 then the filter will move one pixel right as below

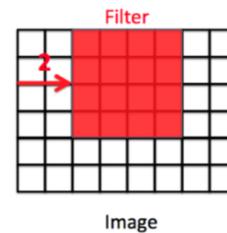


Fig 5: Representation of Filter moved by two stride

Here is an example:

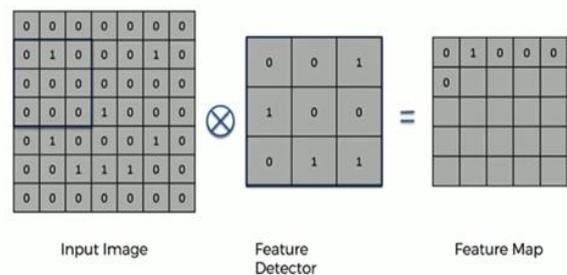


Fig 6: Creating feature map after moving the filter by one stride horizontal and vertical direction

These strides or feature detectors are picked by dividing the original image only, they are also called patches. [4]

Padding

Padding talks about adding extra pixels around the image. Zero padding means, we would be adding **zero's** as extra pixels around the image. Padding will help when the filter moves through the input image.

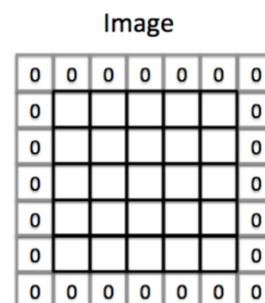


Fig7: Padding an image with zeros

Flattening:

Flattening is the concept of converting the multidimensional feature map array to a one dimensional array that will be used in convolution layer processing.

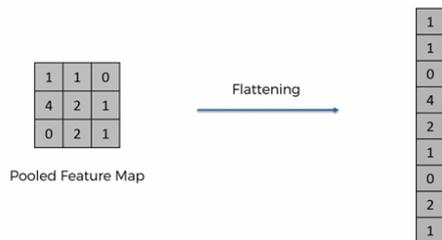


Fig 8: Data flattening into one dimensional array.

Max Pooling

Max pooling concept is used to reduce the size of the input feature map array by retaining the features. This concept is also called down sampling the image. The output of the max pooling is a reduced array so that computations in the next layers applied on this array would be reduced.

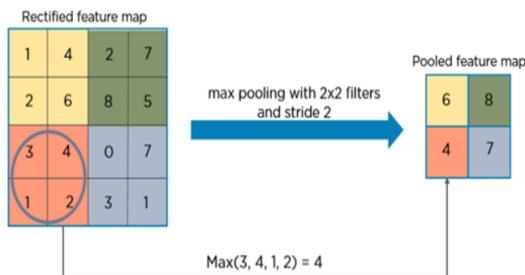


Fig 9: Max pooling using 2X2 filter with a stride of 2

In the above example, we have 4X4 feature map. Max pooling calculates the max of each 2X2 matrix and the max value will be inserted in the array. This was we are retaining the feature and also reducing the input feature map size. Some of the algorithms replaces the max by average value. This is called average pooling.

In the process of convolution, a filter of size 2X2 is applied on the feature map with a stride of 2. The output is the reduced pooled feature map.

Data Augmentation:

We feed several images to train neural network algorithm model. Sometimes we may not get so much of data to train the model. But whatever the data we use to train the model should be very much informative to differentiate the types of images. Input data size is not important. But the informative data is very important. For example, consider we are trying to train a model to differentiate two different brands of cars. Suppose if we have the brand A car images facing left and brand B cars facing right. We use these images to train the model. Then we want to test the model by providing the input as brand B car facing left. The model will predict that the car is of brand A, which is wrong. This is because the algorithm has understood in the training that all cars facing left are brand A. So it's important that the images set should contain different orientations of the images, different sizes of the images and there are many more.

Data augmentation basically talks about increasing the number of images in the dataset as well as several other concepts of image Flipping, Rotation, Scaling, Cropping, translation (moving image in X or Y direction) and more. There are many frameworks in the industry provide these features by default.

Supervised versus unsupervised learning:

Supervised learning is the concept of analyzing constructing a model on the data which is already labelled. Data labelling is nothing but the characteristics of the data are already known. So it would be easy for the algorithm to learn and create a model from such labelled data.

If the data is not labelled, means you don't know the data what you are trying to categorize and you start to create the model. With this type of data, the algorithm has to first find out features that should be able to categorize the images and it's

going to be difficult and time taking process for the algorithm.

3. Challenges in the image classification using deep learning techniques:

- Evolution of digital pathology: In the initial days' pathological images are collected on glass and storing and maintaining these glass samples for analysis is very difficult. Digital capture of these image characteristics has evolved and given lot of scope to research on the digital data. There are two main challenges of digital data: storage space and parsing and analyzing the giga byte data. [1]
- Choosing different hyperparamters for algorithms is a challenge and it need to be improved after several training runs.

4. References:

- [1]. Brady Kieffer, Morteza Babaie, Shivam Kalra and H.R. Tizhoosh "Convolutional Neural Networks For Histopathology Image Classification: Training Vs. Using Pre-Trained Networks" in 7th Intern. Conf. on Image Processing Theory, Tools and Applications(IPTA 2017)
- [2]. Mehdi Habibzadeh Motlagh, Mahboobeh Jannesari, HamidReza Aboulkheyr, Pegah Khosravi, Olivier Elemento, Mehdi Totonchi and Iman Hajirasouliha, "Breast Cancer Histopathological Image Classification: A deep Learning Approach"
- [3]. Tim Dettemers, "Understanding Convolution In deep Learning". [<http://timdettmers.com/2015/03/26/convolution-ddep-learning/>]
- [4]. Hafemann, L.g., Oliveira, L.S., & Cavalin, P.)2014, August). Forest species recognition using deep convolutional neural networks. In *Pattern Recognition (ICPR), 2014 22nd International Conference on* (pp. 1103-1107). IEEE.
- [5]. <https://machinelearningmastery.com/learning-rate-for-deep-learning-neural-networks/>
- [6]. <https://www.kdnuggets.com/2018/04/top-16-open-source-deep-learning-libraries.html>
- [7]. <https://www.kaggle.com/raoulma/cancer-image-tensorflow-cnn-80-valid-acc/notebook>
- [8]. Kundan Kumar, Annavarapu Chandra Sekhara Rao "Breast Cancer Classification of Image using Convolutional Neural Network" in 4th Int'l Conf on Recent Advances in Information Technology | RAIT – 2018 |
- [9]. Sai Prasad, K. and Miryala, Rajender (2019). Histopathological Image Classification Using Deep Learning Techniques. *International Journal on Emerging Technologies*, 10(2): 467–473.
- [10]. Kashi Sai Prasad, Dr. S Pasupathy, " Real-time Data Streaming using Apache Spark on Fully Configured Hadoop Cluster", *J.Mech.Cont.& Math. Sci.*, Vol.-13, No.-5, November-December (2018) Pages 164-176
- [11]. Sai prasad Kashi; K. Sai Prasad; O. Subhash Chander; G. Prabhakar Reddy; S. Gururaj, " Artificial Intelligence approach for Classifying Molecular Dataset using Density based technique with appropriate Euclidean Distance measure" *MATERIALS TODAY-PROCEEDINGS*, Volume 4, Issue 8, 2017, Pages 8827-8836. Elsevier.
- [12]. G. Arul Dalton, G. Prabhakar Reddy , K. Sai Prasad, "Fuzzy logics associated with neural networks in intelligent control for better world", *International Journal of Reasoning-based Intelligent Systems*, Volume 10, Issue 3-4 November 2018.