A Survey On Deep Learning Methods And Tools In Image Processing

B.Meena ,Prof.K.Venkata Rao ,Prof.Suresh Chittineni

Abstract: Deep Learning brings provocative light on various fields and research areas, such as image processing, Artificial Intelligence, and natural language processing. it is an essential method applied to different applications in future. In recent years, the image processing and has been attracting and increasing attention as it is one of the emerging areas of research. There are various applications for image processing in various domains such as agriculture, space agencies, medical field, forensics and many others. As there is a progress in deep learning, many modern methods in deep learning are recommended to improve image processing and image analysis performance. This document provides a broad overview on deep learning methods used in image processing. Initially, an introduction on deep learning and image processing is given. Then, architectures of deep learning are introduced. Then, the applications of deep learning on image processing like the image detection, image segmentation, and classification image classification are explained. Thereafter, the benefits and weaknesses of deep learning tools common tools are mentioned along with the deep learning tools used in image processing applications.

Index Terms: Artificial Intelligence, Convolution Neural Networks, Deep learning, Image Processing, Image segmentation, Machine Learning, Tools.

1. INTRODUCTION

Artificial Intelligence [1] is one of the branches of computer science and Information Technology that creates computers or machines that behave as intelligent as humans. John McCarthy, who is the father of artificial intelligence said that "The Science and engineering of creating intelligent machines particularly artificial intelligence attempts to simulate intelligence of humans and results with a new intelligent machine that shall be able to produce information as human awareness and behavior . Artificial Intelligence has the application in various fields such as natural language processing, Neural Networks, robotics and image processing. The heart of artificial intelligence is machine learning [2] and the basic approach for designing intellectual systems .Machine learning includes a many practices on statistics and probability theory, convex analysis and approximation as well as algorithm complexities. the concepts of machine learning essentially adopts intuition and integration, in order to create the contemporary learning by reproducing the human nature and performance ,Then restructures the new knowledge to improve the performance regularly. Machine learning has been extensively applied in various areas such as space agencies, automation, industry Finance, Industry Healthcare, Government Organization, Marketing, etc. The concept of deep learning [3] has been proposed with the deepening of artificial neural network[3],[4],[5] Deep learning uses Deep neural networks and it's a software that simulates human brain neural network. Hence Deep learning is a subdivision of Machine learning which is bringing the machine learning and artificial intelligence closer to create innovative applications. The human brain is an interdependence of network of neurons , this scheme has lead to the emergence of artificial neural networks. Artificial neural network has a dissimilarity with human brain, that the human brain neurons are connected through physical path especially the neural network generates data, connections and discrete layer propagation direction . In 2011, the first-generation learning systems DistBelief was launched by Google. Google has been able to analyze thousands of their data from various data centers Google is broadly using Deep learning in processing its products such as Google search and Google photos etc. .One of the important aspects in image processing and analysis is feature representation. The following are the two advantages that deep learning yields: (1) Given a dataset, deep learning finds features naturally for a definite application. Usually semi automatic learning method is a classical feature extraction method that uses a prior knowledge to extract the features (2) Deep learning can find new features that are relevant to the suitable applications. According to traditional extraction methods , they are always defined by previous knowledge, which only collect certain features associated with an appropriate application. In addition to this, there are two other components that alter image processing results . they are image acquisition and interpretation : Image Acquisition: It is demonstrated experimentally by researchers that the higher the image quality, the more the results obtained in image processing and analysis. Nonetheless, image quality depends on the image acquisition

Image interpretation: It is a procedure of reviewing an aerial image or a digital remote sensing image and manually determines image features. This approach is very trust worthy and broad features can be collected from an image. some of them can be riparian features ,anthropogenic features etc . Yet it takes lot of time and only expert people like image analyst having sound knowledge can do it. Although image interpretation works best on the built in attributes of image which are composed of seven aspects such as texture, size, association, tone, pattern and shape . These aspects are used to know the information about the objects present in the image.

However to get better results during image interpretation, it is very important to design and develop a system that would interpret the images automatically. This kind of systems should incorporate abundant operations like image classification[20,27], image registration, image segmentation[20,27] and image detections. This paper focuses on deep learning by providing a broad inspection about the architectures and applications of deep learning over

- *B.Meena ,(Research Scholar, Department of CS &SE ,AUCE, Visakhapatnam) ,Associate Professor Department, Raghu Engineering College,Dakamarri,Visakhapatnam, E-mail: meenacserec@gmail.com*
- *Prof.K.Venkata Rao,Professor,Department of CS & SE,Andhra University,Visakhapatnam.E-mail: professor_venkat@yahoo.com*
- *Prof.Suresh Chittineni ,Professor,CSE Department ,GITAM University,Visakhapatnam.E-mail: sureshchittineni@gmail.com*

image processing. Later the pros and cons of various tools are deliberated along with the tools that are applied to image processing. Finally the objective of deep learning in image processing is expressed along with forth coming tools

2. ARCHITECTURES OF DEEP LEARNING

2.1 Artificial neural networks (ANN) :

ANN is important because just like human brain network it grasps the information which is processed to establish various models pertaining to neural network. The concepts of artificial neural network evolved during 1980's. These models of artificial neural network are calculation models having huge finite nodes and links among them. every node in an distinguished ANN present an image of particular output function . The link between the nodes communicate information between them . This information is represented as a weight . Hence ANN is an algorithm whose output rely upon weights and activation function (output function) . In accordance with, ANN algorithms are mainly classified into feedback neural networks and feed forward neural networks. Few of them are described below : Feed Forward neural networks: It is graph which is directed and having no cycles and no feedback in the network. It is applied for the conversion the input space information to output space. As it is combined with various simple nonlinear functions; it has got a good capability of processing information. its network topology is simple and easy to apply.

Feedback neural networks: It is a complete graph having a feedback and no directions .it has state transitions while processing information. In order to handle the information processing it uses dynamic system theory. the network cohesion is firmly connected to the function of associative memory. It's a decade that the research on artificial neural network is getting deepened and a lot of advancement has been made in various domains in biology, smart robots, finance and medical. In addition to these, artificial neural networks have been strongly used to solve countless problems by showing many intellectual characteristics. Commonly the attributes of artificial neural networks are primarily mirrored on associative memory, self-learning capability and quick fixing the solution using optimization.

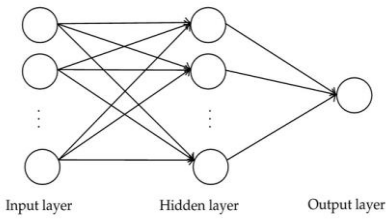

Figure 1 : *Sample deep feed forward networks*

2.2 Deep feed forward networks

The traditional deep learning models are deep feed forward networks. deep feed forward networks are also called as feed forward neural networks. deep feed forward networks aim to train the network for computing the equivalent objective functions. A deep feed forward network is described as $y= f (x; \theta)$. This mapping function y acquires the value of θ settings to compute the best equivalent functions. Usually, a network of direct action in extent consists of an input layer, different hidden layers and an output layer. In addition to that the information flow in deep feed forward network occurs in one direction and does not move backward . This is depicted in the sample figure . In this example the deep feed forward network has few input layers, hidden layers and one output layer. Hence the deep feed forward network is one the ancient architecture of deep learning

2.3 Auto encoders Stacked

A simple deep feed forward network which is simple and having input layers , hidden layers and an output layer is referred as an auto encoder[6],[7],[8],[9] . A decoder and an encoder are two parts of an auto encoder according to its functionalities .An encoder is notated as f(x) that produces a decreased characteristic description from x,h where x is a primary input h is a hidden layer. the notation of decoder is g(f(x)) . the decoder is familiar with reorganizing the basic input from the encoder output. the decoder minimizes the loss function during reorganizing process:

$$L(x,g(f(x))$$

The high-dimensional data is converted to low-dimension by encoding and decoding technique. Hence auto encoder is much beneficial in image classification and various identical tasks. The auto encoder is derived from sparse auto encoder, denoising auto encoder and contractive auto encoder. sparse auto encoder is dissimilar to auto encoders. the sparse constraint $\Omega(h)$ is added to the hidden layer. Hence its restoration error is computed as

$$L(x,g(f(x)))+ \Omega(h)$$

De noising auto encoder : the objective of denoising auto encoder is to minimize the loss function , they are dissimilar to sparse auto encoder

$$L(x,g(f(x)))$$

Contractive auto encoder: Similar to sparse auto encoders, they decrease the specific regularizer by accumulating $\Omega(h)$ to hidden layer .The explicit regularizer is computed as as :

$$L(x,g(f(x)))+ \Omega(h)$$

where $\Omega(h)$ is the squared Frobenius norm of the Jacobian partial derivative matrix of the encoder function f(x) and λ is a free parameter. The neural network with many auto encoder layers is referred as stacked autoencoder . In addition to that lateral layer input is taken from the former layer output in stacked auto encoders . An auto encoder usually consists of ternary layers only without having any stacked auto encoders like the architecture of deep learning . It is much noted that stacked auto encoder can only finish the task when they are trained . As an example in order to train a n→m→k network , the stacked auto encoder need to be trained with n→m→n for obtaining a conversion n→m and then train m→k→m network to obtaining a conversion of m→k . At the end it stacks the transformations (conversions) to form n→m→k auto encoder . This series of actions are referred as unsupervised pre-training in layers.

2.4 Deep belief networks

Unlike stacked auto encoders , deep belief networks are likewise referred as neural networks and is a special form of Boltzmann machine[10],[11],[12] with some restrictions . Furthermore , the input of next layer comes from output of previous layer in deep belief networks. In order to pre-train every restricted Boltzmann machine hierarchically, the deep belief Boltzmann machine accustoms an unsupervised greedy method which is pre trained . the output of this used as an basic input for supervised learning probability model. Boltzmann machine is a method of modeling that uses functions of energy and has been derived from statistical physics. This characterizes the higher interactions among the variables .when compared it is complex having total tangible analysis and is supported by strict statistics concepts. It is symmetrically joined by arbitrary binary unit neural network with feedback having visible and hidden units. A special form of Boltzmann machine is referred as restricted Boltzmann machine that includes only a visible layer and a hidden layer . In restricted Boltzmann machine the connections between the hidden layer nodes and visible layer are bi directional. the restricted Boltzmann machine have faster calculation speed when compared to Boltzmann machine as it has only one hidden layer .The restricted Boltzmann machine has two main functions : (1) Used to reduce the dimensions of data. (2) A weight matrix is acquired that is a basic input to other neural networks.

2.5 Convolutional neural networks

Convolutional neural networks [24][25] are broadly used in tasks recognition like image classification, handwriting recognition and document recognition. The dissimilarity in convolution neural networks and fully connected feed forward is their connection type among the adjacent layers. In feed forward neural networks some nodes are connected where as in convolution neural network, all the nodes are completely connected. Convolution neural networks consists of two kernel layers namely the Convolutional and pooling layers. Convolutional Layer: the input of each node is only a tiny area of the previous layer and the size of this tiny patch is usually 3 X 3 or 5 X 5. the convolution layer analyzes deeply each small portion of the neural network and results a better feature representation. Pooling Layer: The layer followed after the conventional layer is pooling layer .The purpose of this layer is to decrease the matrix size from convolution layer, the result of this layer is also to decrease number of parameters in neural network The advantage of this pooling layer is to prevent over fitting problem and speed up the calculation. Usually the convolution neural network architecture are categorized into two types basing on their connection modes .first type connects various convolutions layers in an order .As an example LeNet-5,AlexNet and ZfNet , while the second one connects in parallel like Inception and later versions

3. APPLICATION OF LEARNING IN DEPTH IN THE IMAGE PROCESSING

There are many deep learning methods that have been proposed for applying in Image analysis and Image processing areas, like Image classification, Image registration, Image segmentation and Image detection. The problems in the above mentioned areas can be solved using feature representation in collaboration with Deep learning techniques. In this portion of the paper we provide an analysis on the advancement of putting deep learning architecture in all the above areas mentioned.

3.1.Deep learning in image segmentation

One of the important and interesting concepts in image processing is image segmentation

The major categories of computer vision are:

1. Image Classification
2. Image Detection
3. Image segmentation

The classification of images by identifying an existence of an image in a picture is known as image classification. Image detection follows image classification and localization. the concept localization finds the area if a given object is located . A boundary is drawn around it to highlight the region. The difficult and the most beneficial type among the above three is image segmentation. This is related to classify every pixel into a group of objects in a given picture.

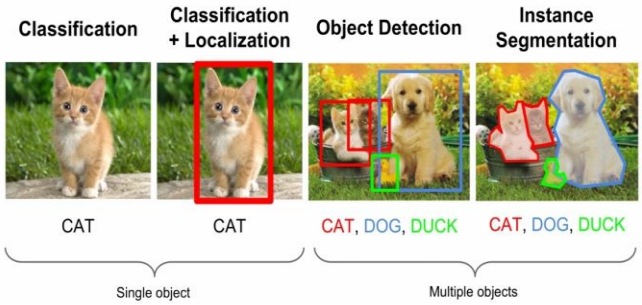

Figure 2 : *Image Classification of single object and multiple objects*

Image segmentation is classified into simultaneous detection and semantic segmentation. Both of the type of image segmentation seems to be similar but the dissimilarity is simultaneous detection identifies exclusive objects where as semantic segmentation deals with pixel categorization. The above figure is considered to quote the dissimilarity of these two types of segmentation. as per simultaneous detection which is also known as instance aware segmentation find each cat image separately where as the semantic segmentation deals with categorizing all cats as a single instance .These kind of problems can be solved easily using deep learning techniques.

3.2. Image Segmentation using Convolution Neural Networks (CNN):

Here in this procedure the pieces of images which are marked are inserted as an input of convolution neural network. This procedure is done again and again on the complete array of pixel. This architecture has a disadvantage of performing this set of operations only over well consistent image. Passing complete images to CNN: In this architecture a complete image is given to the network where the pixels are labeled in a single step rather than many steps. It is an improved architecture than traditional one . here the mask of segmentation is decreased in size due to the convolution and pooling technique . if the input of an image is 512 X 512 , the result may be 28 X 28. Fully CNN: The next improved architecture than the earlier one is fully convolution neural

network. Here the final tier of CNN follows Deconvolutional neural network . Given a picture, image segmentation is very essential to figure out varied things and relationship. The utilization of image segmentation differs from simple applications in image recognition to spatial images as they are exemplary generous to know the texture . they are applied to medical images to identify cancer cells and even in surgery preparations . This is an essential constituent in searching the images in Fashion and retail enterprises

3.3. Classification of images using Deep Neural Networks

Image classification is a process of drawing out the information of classes from a multiband raster image . Here the computer is taught to identify images and categorize into the some of the given classes. In this example some images are given to the computer and it could classify into 10 categories are as follows

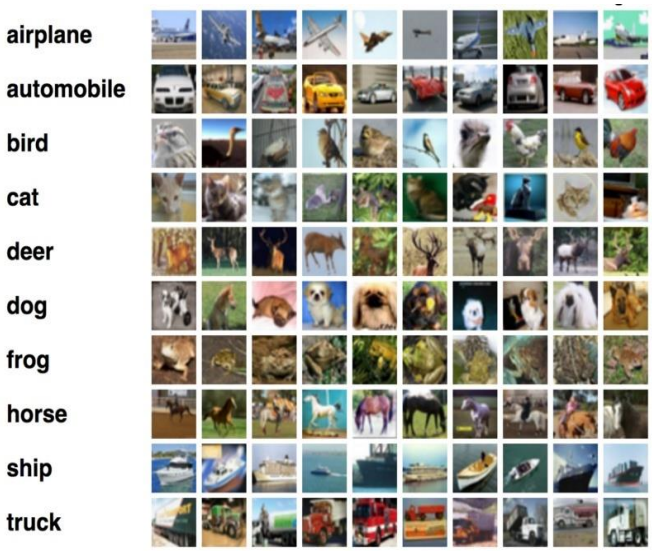

Figure 3 : *Image Classification*

Here the computer is taught with the behavior of bird, dog, cat etc, in order to find out the objects. The more cats the computer gets trained, it recognizes the more cat objects. This kind of training is said to be supervised learning. Here the images are labeled to accomplish such task. Google has given an open source framework for deep learning called as Tensor Flow. The developer would have a better control at micro level over each node (neuron) in Tensor Flow. To gain the best performance the weights have to be set. For image classification Tensor Flow has built in APIs. Convolution neural networks is one of the well know techniques to improve image classification accuracy. In this technique the image is broken up into several parts where the machine tries to identify each part rather than sending the image array as an input . this technique of sending the parts to machine allows to accomplish the task parallel and identify the objects irrespective of the object location.

3.4 . Sample data

In this example a dataset CIFAR-10[20] having images calibrated 60 000 in number each of size 32 X 32 pixels is taken. this dataset is categorized into ten classes that do not occur simultaneously accompanying every class that has images in number 6000.the images in this dataset has no noise , small in size and are definitely labeled . Hence this is an ideal dataset to finish the task significantly with a fewer pre processing steps. the following are some of the pictures collected from the above dataset:

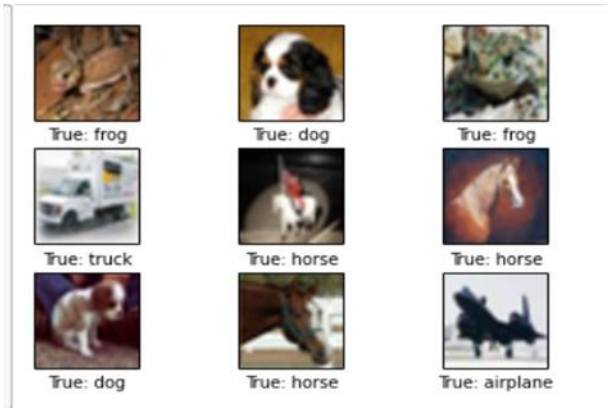

Figure 4 : *Pictures from cifar-10 data set and image*

Here are the steps to apply image classification upon the above data set

1. Pre-processing
2. Splitting the dataset
3. Building a CNN

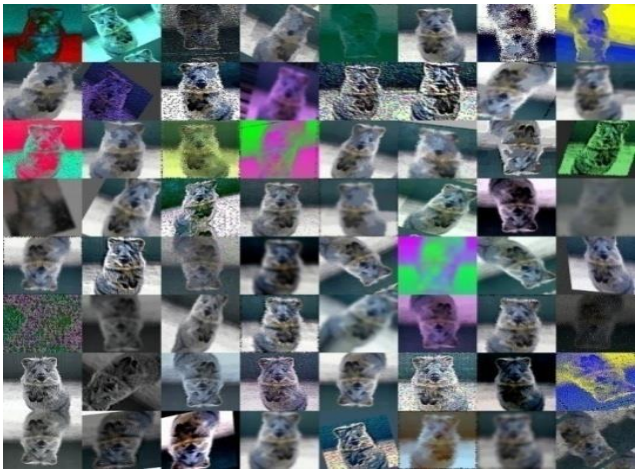

Figure 5 : *Image after applying above three steps*

Step 1: Pre-treatment: In first step of preprocessing, the data does not have any noise and it is organized so variance is added to the data. The noise can be added using python library named imgaug. Then a random combination of cropping the image, flipping and applying saturation and hue adjustments is done.

Step 2: Splitting our data set: Computing gradient for the given large dataset model is a time taking process . Hence the images are processed in tiny batches at every optimizer iteration. Generally the size of the batch is considered as 32 or 64. In this example 64 is the batch size as it has big number of image. Further the dataset is broken up into training set each having 50000 image and test set having 10

000 images

Step 3: Building a convolution neural network : After pre processing and splitting up the neural network , the implementation begin with three convolution layers having a max pooling of 2 X 2 .max pooling - a technique that reduces the image dimensions . It collects maximum value of a pixel in the grid . Hence it makes the model generic and over fitting is reduced . Below is an example that show the working of 2 X 2 max pooling . Here is a max pool with 2 X 2 filters and stride 2.

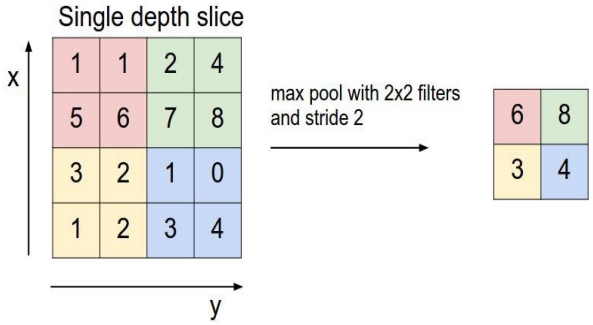

Figure 6: *Two cross two max pooling*

Now the neural network is trained and ready to use . An artificial convolution neural network can recognize the image more accurately than previous approach using tensor flow .

4. DEEP LEARNING TOOLS :

4.1. Common tools of Deep Learning

Various tools [13],[14],[15],[16],[17],[18],[19],[20],[21],[22],[23],[24],[25] of deep learning have been evolved that are convenient for analyst and researchers. this advancement helps a lot for all those working in deep learning. The table below introduces various deep learning tools.

TABLE 1 : *SOME OF THE DEEP LEARNING TOOLS AND REFERENCES*

Tool Name	URLs
DeepLearnToolbox	https://github.com/PierreHao/Deep-Learning-ToolBox-CNN
Caffe	http://caffe.berkeleyvision.org/
Torch	http://torch.ch/
Theano	http://deeplearning.net/software/theano
Pylearn2	http://deeplearning.net/software/pylearn2/
Keras	https://github.com/EderSantana/keras
tensorflow	https://www.tensorflow.org/
CNTK	https://www.microsoft.com/en-us/research/product/cognitive-toolkit/
MXNet	https://github.com/dmlc/mxnet
chaining	http://chainer.org/
Deeplearning4j	https://deeplearning4j.org/
MatConvNet	http://www.vlfeat.org/matconvnet/
maxDNN	https://github.com/eBay/maxDNN

TABLE 2: *DEEP LEARNING TOOLS FOR IMAGE PROCESSING AND REFERENCES*

Tool name	URL references
OpenCV	http://www.opencv.org/
BoofCV	http://www.boofcv.org/
NASA Vision Workbench	http://t1.arc.nasa.gov/tech/asr/intelligent-robotics/nasa-vision-workbench/
SimpleCV	http://simplecv.org
Tesseract	http: //github.corn/tesseract-OCR / Tesseract
SLIC superpixels	http://ivrl.epfl.ch/supplementary_material/RK_SLICSuperpixels/
OpenMVG	https://github.com/openMVG/openMVG
LIBVISO	http://www.cvlibs.ne/software/libviso
VisualSFM	http://hornes.cs.washington.edu/~ccwu/vsfm/
MeshLab	http://meshlab.sourceforge.net/
Bundler	http: //phototour.cs/washington.edu/bundler/
Vid.stab	https://github.com/georgmartius/vid.stab

5. CONCLUSION:

The aim of this paper is to present relevant information for investigators and researchers, with respect to the deep learning applications in image processing research area. In order to acquire better results in image processing research area, the architectures of deep learning are extensively used in sub domains like image segmentation, image classification, image analysis, image recording and image sensing. Deep learning has faced various challenges and proved that these methods give improved and more appropriate results for researches working in image processing. To achieve the best results the large data sets are to be given to the network, limited data set would lead to imbalances in classes. The large the data set, deep learning learns in depth. Deep learning with image processing would present outstanding achievements in future.

REFERENCES :

[1] RABrooks Intelligence without representation. Artif.Intell., Vol 47 139-159 1991

[2] B.Bishop, Recognition and Machine Learning.Springer-Verlag, 2006

[3] Y.Bengio, learning deep Architechtures for AI, Found. Trends Mach. Learn., Vol.2 no.1,2009, (1-127)

[4] B.Yegnanaryana, artificial neural networks. New Delhi, India: PHI Learning Pvt. Ltd. 2009

[5] H.Bourlard and Y.Kamp, Self-organization by multi preceptrons and singular value decomposition, Biol.Cybern .. PP 291-294

[6] RS GEHinton and Zemel, Autoencoders, minimum description length and Helmholtz free energy in Proc. 6 Int.Conf.Newural 1994 Information processing systems

[7] G.Alain and Y.Bengio, this regularized self-learning encoders from the distribution data generating, Mach.Learn J. Res., Vol 15, No. 1, pp 3563 to 3593.2014

[8] MARanzato, C.Pultney, S.Chopra and Y. LeCun, effective learning of the few performances with an energy-based model in the 20th annual conference of information processing systems Neural. 2006

[9] P. Vincent H. Larochelle, I. Lajoie, Y. Bengio and PA Manzagol, autoencoders denoising Stacked: Learning useful representations in a deep network with a criterion of local

noising, J. Mach. Learn. Res., Vol. 11, pp. 3371-3408, 2010.
[10] Yi Pan,Min Li,Ziyue Chen,Lu Tang,Chengqian Lu. and Jianxin Wang Applications of deep learning to MRI images : A survey Jin Liu . ISSN 2096-0654 01/06 pp 1– 18 Volume 1, Number 1, March 2018
[11] SE Fahlman, GE Hinton and TJ Sejnowski, Massively parallel architectures for AI: NETL, thistle and Boltzmann machines, in Proc. AAAI 3rd Conf. Artificial intelligence, Washington, DC, USA, 1983.
[12] RB Palm, forecasting as a candidate for learning deep hierarchical data models, master thesis, Technical University of Denmark, Denmark, in 2012.
[13] Y. Jia, Shelhamer E., J. Donahue, S. Karayev, J. Long, R. Girshick S. Guadarrama and T. Darrell, Caffe Architecture for embedding fast convolution function, in Proc. 22 ACM Int. Conf. Multimedia, Orlando, FL, USA, 2014, pp. 675-678.
[14] R. Collobert, K. and C. Kavukcuoglu Farabet, Torch7: Matlab type learning environment of the machine, BigLearn, Workshop NIPS, No. EPFL-CONF-192376.2011.
[15] R. Al-Rfou, Alain G., A. Almahairi, C. Angermueller, D. Bahdanau N. Ballas F. Bastien, J. Bayer, A. Belikov, A.Belopolsky, et al., Theano : A python framework for rapid calculation of mathematical expressions, arXiv preprint arXiv: 1605.02688 2016.
[16] IJ Goodfellow, D-Warde Farley, P. Lamblin, V. Dumoulin, Mr. Mirza Pascanu R., J. Bergstra F. Bastien and Y. Bengio, Pylearn2: A machine learning research library , arXiv preprint arXiv: 1308.4214 2013.
[17] F. Chollet, Keras: Théano-libraryCode learning base long, https://github.com/fchollet. Documentation: http: //keras.io 2015.
[18] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, ZFChen, C. Citro, GS Corrado, A. Davis, J. Dean, M. Devin, et al, tensorflow :. Large scale machine learning on heterogeneous distributed systems, arXiv preprint arXiv:
[19] D. Yu, A. Eversole, Mr. Seltzer, KS Yao, O. Kuchaiev, Y. Zhang, F. Seide, ZH Huang, B. Guenter, HM Wang et al., Introduction to computer networks and computer network transmission tools, technical report
[20] https://medium.com/@tifa2up/image-classification-using-deep-neural-networks-a-beginner-friendly-approach-using-tensorflow-94b0a090ccd4
[21] TQ Chen, Li, Li YT, Lin, NY Wang, J. Wang, TJ Xiao, B. Xu, Z. Zhang and Zhang CY, MXNet: A flexible efficient Machine learning library for heterogeneous distributed systems, arXiv preprint arXiv: 1512.01274 2015.
[22] S. Tokui, K. Oono, S. Hido, J. and Clayton, Chainer: A new generation open source framework for in-depth learning, in Proc. Workshop on Machine Learning Systems (LearningSys) in the Twenty-ninth Annual Conf. Neural Processing Systems (PNM), in 2015.
[23] Deeplearning4j development team, Deeplearning4j: Open-source distributed deep learning for JVM, Apache Software Foundation License 2.0, http://deeplearning4j.org 2017.
[24] BC Ooi, KL Tan, S. Wang, W. Wang, Cai QC, G. Chen, Gao JY, ZJ Luo, AKH Tung, Y. Wang, et al, SINGAPORE. A learning platform distributed depth, Proc. 23 ACM Int. Conf. Multimedia, Brisbane, Australia, 2015, pp. 685-688.
[25] A. and K. Vedaldi Lenc, MatConvNet: convolutional neural networks for Matlab, in Proc. 23 ACM Int. Conf. Multimedia, Brisbane, Australia, 2015, pp. 689-692.