Comparative Analysis of PyTorch And Caffe Frameworks

[1]N Kanakapriya ,[2] Dr. Anala M R

[1]*Student, M.Tech Department of CSE R V College of Engineering Bengaluru-560059, India*

[2]*Associate Professor ,Department of CSE ,R V College of Engineering ,Bengaluru-560059, India*

Abstract

Deep learning is a model of machine learning whose demand is rapidly increasing. It can change the view of artificial intelligence and achieve the goal significantly. With many deep learning frameworks available, the choice of selection becomes difficult. In this paper, two deep learning frameworks namely PyTorch and Caffe are analysed, with various parameters and deep learning architectures. The frameworks are implemented for specific applications and results are compared to make the better choice.

Keywords—*Deep Learning, PyTorch, Caffe.*

I. INTRODUCTION

Deep learning is a subset of machine learning. Deep learning is a neural network method and provides neural network architectures in many languages which can be used in any platform [1]. Deep learning frameworks are easier to build. The framework is also called as "toolkit" or "library". The frameworks can be selected based on several factors such as style, core development architecture, neural network architecture, application areas, hardware and deployment. There are two types of frameworks such as imperative framework and symbolic framework. Imperative framework execute as the program flows and are more flexible as it is near to the language. Symbolic framework executes using symbol and later combined which gives a computational graph. Symbolic frameworks are efficient in case of memory and speed and less flexible since it is written in domain specific language and deep learning solutions can be built quickly. Deep learning framework allows developers to easily build deep learning models without getting in depth with technicality of algorithm. A clear deep learning model is build using pre-built components. Theses frameworks reduces complexity of deep learning and makes it more accessible to developers [2]. The models can be designed, trained and validated using deep neural networks on a programming interface. Widely used deep learning frameworks such as Caffe and PyTorch depend on GPU such as CUDA, to provide high-quality training.

Section 2 describes the related work with respect to comparison of deep learning frameworks, Section 3 describes the comparative analysis of deep learning frameworks considered for study, Section 4 describes the methodology carried out, Section 5 describes results and conclusion, Section 6 gives conclusion and Section 7 describes the future work.

II. RELATED WORK

In the arena of machine learning, the growth of deep learning is gaining more scope [3]. The view towards Artificial Intelligence is advancing along with their goals for which survey of various deep learning software tools such as Torch, Caffe, Theano, Cuda-Convert and Pylearn2 were discussed. Various algorithms for evaluation and comparison of deep learning tools is studied. Attributes such as extensibility, hardware, performance on CPU or GPU machines, open source, interface etc. Torch and Theano work well on CPU. Theano obtains good performance on GPU. Cuda-convert gains optimization on many objects.

Deep learning is a machine learning model that is based on human brain. With developments in recent times of hardware such as GPU and software such as cuDNN, Torch, Caffe, Theano and new training models have made the neural networks fast and easy. Comparison of deep learning models based on attributes like interfaces, platforms supported, modelling capability, documentation quality, pre-trained models availability and training techniques supported [4].

Deep Learning methods improves performance of applications. To implement the deep learning models, a set of software frameworks are developed and is in use. Selection of framework is important when resources are limited. Comparative study of three frameworks such as Tensorflow, Theano and Torch is done. The deep learning architectures such as convolutional and recurrent architectures to classify images based on two databases: MNIST and CIFAR-10 is done. Computation cost that is gradient computation time, forward time and memory consumption are documented for both CPU and GPU settings [5].

Deep learning is gaining popularity in many fields such as computer vision and natural language processing. Deep learning has strong learning ability and extracts features from dataset. Due to its practicability, deep learning is powerful for researchers to do their research work. Advanced deep learning models such as convolutional neural network, recurrent neural network etc are discussed. Also, deep learning frameworks such as Caffe, Tensorflow, Torch and Theano is studied along with their applications [6].

III. DEEP LEARNING FRAMEWORKS

In this work, two kinds of frameworks such as Caffe and PyTorch are considered for study. The following section discusses the need to select these frameworks for comparative analysis. Fig.1 depicts the relation between framework, libraries and program. The libraries are imported from the considered framework to obtain the result by running the code.

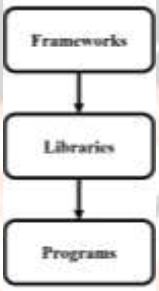

Fig. 1. Flow of Framework.

A. PyTorch

PyTorch is developed from Torch deep learning framework which is used to model deep neural networks and run computations. PyTorch is open source library. PyTorch runs on Python and trained well on a GPU. PyTorch uses computation graph and allows users to build framework by themselves and also change them during runtime. PyTorch is imperative style deep learning framework as it is flexible and closer to language. This is valuable since memory allocation is done as per user requirement. It supports CUDA implementation and is fast, portable and easy usage. It has many libraries for deep learning algorithms [7]. It includes packages for neural networks, image processing and energy-based models. It has applications which include speech, video and image applications. Networks such as Recurrent, convolution etc are trainable using PyTorch. PyTorch can be embedded into iOS and Android. Drawback of PyTorch is documentation is inconsistent. It is used by Google, Twitter, Nvidia and so on.

B. Caffe

Caffe is an open source deep learning framework which has good architecture, speed and modularity. It is cross platform which is written in C++, Python and runs on Ubuntu, Android and other platforms. It is a popular framework used for image classification. Caffe supports CPU and CUDA for GPU computation and OpenCL. Caffe is symbolic style deep learning framework and is efficient in terms of speed and memory. Caffe supports recurrent and convolution neural network. Caffe provides pre-trained models for learning and demo purpose. It has better speed in training models and obtaining accuracy of results. It is used in simple regression, speech applications, robotic applications and so on. Collection of layers is called blob in Caffe and these layers are classified into vision layer, loss layer, activation layer, data layer and common layers. Advantage of Caffe is instead of writing code, models are optimized and defined as plaintext schemas [8]. For cloud environments, Caffe allows platform switching. Caffe provides libraries for training, testing and tuning models. Caffe is easier to understand, hence many beginners choose Caffe framework. Caffe is used in academic, startups and some companies like Yahoo. Fig.2 shows the hardware and applications of deep learning interaction. Table. 1 shows the comparison of deep learning frameworks: PyTorch and Caffe, on the basis of various parameters.

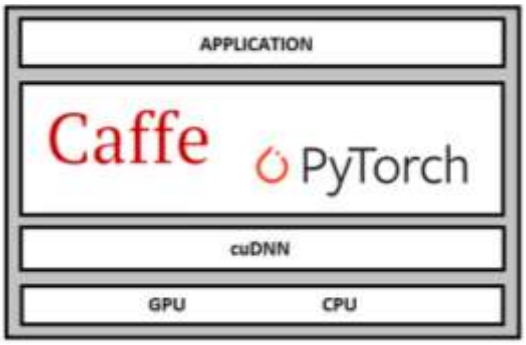

Fig. 2. Framework Architecture

Table 1: Comparision of PyTorch and Caffe Framework

Parameters	PyTorch	Caffe
Creator	Ronan Collobert, Koray Kavukcuoglu, Clement Farabet	Berkeley Vision and Learning Center
Platform	Linux, Android, iOS, Win- dows, Mac OS	Ubuntu, OS X, AWS, unofficial port for windows and android
Interface	C, C++	C++, Command line, Python, MATLAB,PyCaffe
Written in	C	C++
Parallelizing technique support	CUDA, OpenMP, OpenCL (third party implementation)	CUDA, OpenMP (third party implementation)
Modeling capability	RNNs, CNNs, RBMs, DBNs	RNNs, CNNs, RBMs, DBNs
Has pre- trained models	Yes	Yes
Extensibility	Easier	Not easy
Hardware Utilization	CPU, GPU	CPU, GPU
Architecture	Well design with modular interface	Standard layer wise design
Platform	Linux based	Cross Platform
Open source	Yes	Yes
Advantages	Easy to write own layer types and run on GPU, Lots of pre-trained models	Train models without writing any code, feedforward networks and image processing
Disadvantages	Not good for recurrent neural networks	Need to write C++ / CUDA for new GPU layers, Not extensible, bit of a hairball

Parallel execution	Yes	Yes
Application	NLP, Speech Recognition0	Image classification with convolutional nets

IV. METHODOLOGY

The block diagram consists of input data that can be image or video frames. The input dataset is decided as per user requirement. To obtain the trained model, these input data is trained using deep learning architectures such as convolutional neural network or recurrent neural network [9] on any deep learning framework such as Caffe or PyTorch. This model is analysed using any language preferably Python. The detection and classification results of dataset is obtained. The methodology is explained in Fig.3.

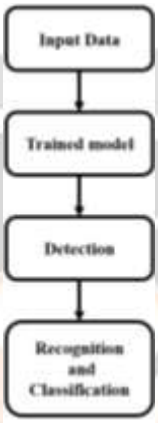

Fig. 3. Block diagram of methodology.

V. RESULTS AND DISCUSSION

The deep learning frameworks such as Caffe and PyTorch were used to train the input dataset considered for applications such as Object detection and recognition and Speech recognition.

A. Speech Recognition

Analysis of audio leads to speech recognition to provide text as outputs. Speech search assistance is mainly used in smart phones, which help the user by giving speech/voice as input data to the system. The input data is trained using convolutional neural network on deep learning frameworks considered. The graph between the two frameworks determine that PyTorch works well for speech recognition compared to Caffe framework. From fig. 4 , it is observed that PyTorch recognition is 92% and Caffe is at 78%.

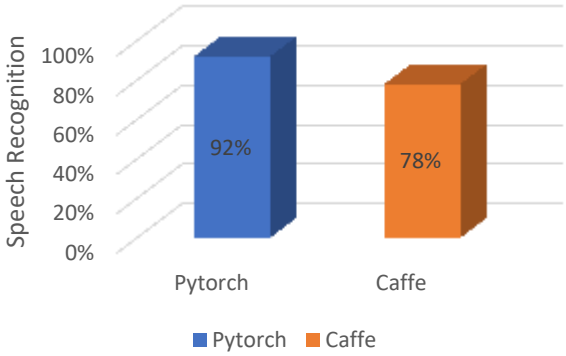

Fig. 4.Graph for Speech Recognition

B. Object Detection and Recognition

Object detection is defined as detection of objects present in an image or set of images. Object recognition is classification of these objects into respective classes such as person, vehicle etc. This, type of application is used in image processing, robotics etc. The example image is executed on PyTorch and Caffe frameworks using convolutional neural network [10]. The results show that accuracy is more in Caffe compared to PyTorch. Fig. 5 shows the result obtained from Caffe is 96% and PyTorch is 85%.

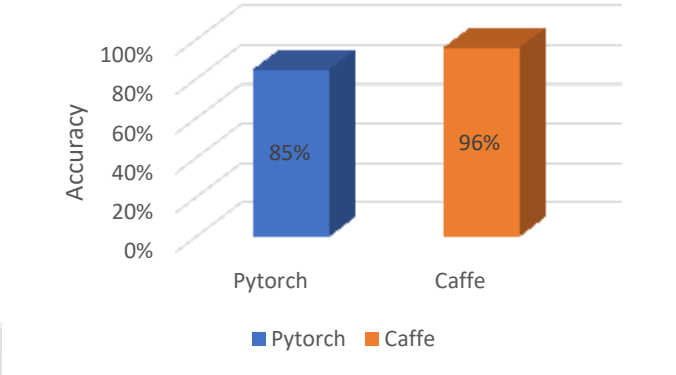

Fig. 5. Graph for Object Detection and Recognition

VI. CONCLUSION

The deep learning frameworks namely PyTorch and Caffe is studied. Deep Learning frameworks gives significant performance improvement in many applications. Observations made are PyTorch has lot of pre-trained models available. It is fast and hence used in speech recognition application which comes under natural language processing. PyTorch documentation needs update. Caffe is useful in image processing application such as object detection and recognition. Caffe has pre-trained convolutional models hosted on their website. If bulk data needs to be trained, Caffe can be preferred as it trains around 60 thousand images per day. Thus, comparative analysis of deep frameworks with respect to its architecture and applications is carried out.

VII. FUTURE WORK

The enhancement of this of paper work can be carried out by considering various input dataset for more applications area and train the dataset using Caffe and PyTorch frameworks and obtain the trained model. Other deep learning frameworks can be compared and analysed with respect to architecture and application.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol.521, no. 7553, pp. 436–444, 2015.

[2] S. Bahrampour, N. Ramakrishnan, L. Schott, M. Shah, "Comparative Study of Deep Learning Software Frameworks", *Research and Technology Center, Robert Bosch LLC.*

[3] Priyanka Sherkane, Deepali Vora, "Survey of Deep Learning Software Tools", *International Conference on Data Management, Analytics and Innovation*, 2017.

[4] Anirudha Parvat, Jai Chavan, Siddesh Kadam, "A Survey of Deep-learning Frameworks", *International Conference on Inventive Systems and Control*, 2017.

[5] Ruben D Fonnegra, Bryan Blair, Gloria M Diaz, "Performance Comparison of Deep Learning Frameworks in Image Classification Problems using Convolutional and Recurrent Networks", *IEEE Colombian Conference on Communications and Computing*, 2017.

[6] Xuedan Du, Yinghao Cai, Shuo Wang, Leijie Zhang, "Overview of Deep Learning", *31st Youth Academic Annual Conference of Chinese Association of Automation*, 2016

[7] R. Collobert, K. Kavukcuoglu, and C. Farabet, "PyTorch: A matlab-like environment for machine learning," in *BigLearn, NIPS Workshop, no. EPFL-CONF-192376*, 2011

[8] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia. ACM*, pp.675–678, 2014

[9] P. Liu, X. Qiu an X. Huang, " Recurrent Neural Netwok for Text Classification with Multi-Task Learning", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 2016

[10] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks", *European Conference on Computer Vision*, Zurich, pp. 818-833, 2014