Autonomous Classifier of Medicine with NVIDIA Deep Learning

Arnab Paul

Abstract—The general purpose of this project is to classify a particular medicine vs other medicines for an autonomous classifier device . There could be huge application of this project in the material handling industry where medicine material handling requires high level of precision to reduce human error in identifying a particular medicine and supplying right medicine to its consumer. Two different neural networks are used to demonstrate classification methods that could help fellow developer of the project. The raw data for the first one is supplied by Udacity Robotic Inference training team and the raw data for the second network was collected by individual/s worked for this project.

Index Terms—Robot, IEEEtran, Udacity, classify, deep learning.

1 INTRODUCTION

It is a well-known fact that machine vision has come a long way since the optical character recognition using CNN by LeCun et all[1]. Object recognition has achieved sub-pixel accuracy. Today it can be used to solve many problems of the most basic human need- healthcare . From a safety perspective, a variety of errors can occur across the pharmacological chain. Such errors can result in adverse drug events e.g. injuries resulting from medication use. Identification and confirmation of prescribed pills can mitigate dispensing errors. While dispensing error rates are low, pharmacies continue to incorporate technologies (e.g. bar code systems) to further reduce them[2]. A study showed that in Mexico selfmedication problems were observed mostly in women and in adult population with less access to health services[3]. Even when self-medication is a common practice, potential risks exist, for example an incorrect self-diagnosis, adverse reactions, incorrect administration or dosage, risk of dependence and abuse, delaying medical advice, and in the worst-case scenario, death[4]. In the past, several attempts have been made to mitigate the problem of accurately identifying a pill/medicine using machine vision and providing a details knowledge to its user.In January 2016 the U.S. National Library of Medicine announced a challenge competition which was motivated by the need to easily identify unknown prescription pills both by health-care personnel and the general public[5]. Different groups from all over the world participated in this competition to address the challenge using various methods. Team castelo used a Convolutional Neural Network (CNN) implemented with Googles TensorFlow open source library. Team msumpf used a combination of features obtained using deep learning and the SIFT descriptor. This entry used a CNN implemented with University of California Berkeleys Caffe open source framework; image similarity was defined as the weighted sum of the two similarity scores with the weights being 0.8 and 0.2 for CNN and SIFT. In the proceedings of CESCG,2010 Andreas Hart and his team presented a mobile computer vision system where a single input image of pills on a special marker based target is processed by an efficient method for object segmentation on structured background[6]. A prototype application is constructed using the Studierstube ES framework, which allows to perform pill recognition on off-the-shelf mobile phones. Eduardo et al. presented[7] a classifier model that uses machine learning and computer vision techniques to classify counter medications based on an image. This classifier was implemented inside a web application that allows users to take a picture of a medication's box in order to receive information related to its content, such as its chemical composition, possible secondary effects, and general comments. The main motivation for this study was to present a tool that can be used to make better decisions about which counter medication to buy. The work presented in this paper is largely motivated by the techniques used in the last solution described above. Here robotic inference method is being chosen to address the problem of automatic identification of medicine with the help of DIGITS lab provided by NVIDIA. The idea is to reduce data latency and improve data accuracy for fast distribution of medicines to the suppliers.

2 BACKGROUND / FORMULATION

Robotic inference helps real data coming from a camera to be classified by the network in real time. It simplifies some common deep learning task, such as managing data, designing a network and rapid prototyping. The prototype model classifies 3 types of medicine container i.e. 4 categories in total: Metformin, Vitamin B12, Vitamin D3 and no pill. A LeNet training model was used on 28x28 pixel size color images having RGB as color channels. A Stchastic Gradient Descent(SGD) technique with a base learning rate of 0.001 is used to solve the optimization problem for minimizing the mean square error between sample data and the hypothesis. 15 training epoch is used with snapshot interval of 1 in each epoch. A graph of epoch vs loss confirmed the values of training and validation loss decreasing from epoch to epoch.

On the other hand, GoogLeNet training network was used for the supplied data on 256x256 pixel size color images having RGB as color channels. An Adaptive Gradient Descent(AGD) technique with a base learning rate of 0.01 is used to minimise error between the hypothesis and training data. 8 training epoch is used with snapshot interval of 1 in each epoch. The respective training models for different data set of the given project and the robotic inference project were chosen to achieve better validation accuracy.

3 Data Acquisition

Data for the project was collected in the form of video and it proved to be a easier method to collect data. A python code was used to extract images out of the video and then about 400 best selected images per category of the data were uploaded into provided DIGITS work space. It compresses and pre-process image data and make suitable for the training (fig 1). Following are some technical details of the data acquisition

- Video recording: Android device (Samsung Galaxy note 8, image resolution 1440x1440)
- Training Images: 1542
- Validation Images: 411(20 per cent)
- Test Images: 102(5 per cent)
- Image Dimension: 28x28
- Image Type: Color(RGB)
- Resize Transformation: Squash
- Image Encoding: PNG

The provided data have been pre-processed with following specifications (fig 2)

- Training Images: 7570
- Validation Images: 2524(25 per cent)
- Image Dimension: 256x256
- Image Type: Color(RGB)
- Resize Transformation: Squash
- DB Backend: lmdb
- Image Encoding: PNG

4 Results

The training model accuracy could be identified by the graph of validation accuracy vs epoch(fig 3,4) and learning rate vs epoch(fig 5,6) for both of the training model. Two instances of successful classification are shown at fig 7. Now before deploying the model architecture and learned weights into a hardware platform it is possible to understand the inference time and accuracy using 'evaluate' command in the DIGIT workspace and job ID. Fig 8 shows average inference time and accuracy of the project model to be 5.33 mS and 75.4 per cent respectively.

5 Discussion

It is clear from the above result that the project training model is successfully able to classify a new image with above 98 per cent accuracy. However, inference accuracy is not that high and there is scope for improvement. For the purpose of this project inference accuracy is more important than time, since the customer reliability on a machine deployed would mostly depend on the accuracy. More training data with clearer image and less noisy background on the image data could significantly improve the training model performance.

Fig. 1. Inference Project Data Set

Fig. 2. Given Project Data Set

6 Conclusion / Future work

The project mainly focuses on building the training model and evaluating the inference accuracy and time. There are a lot of scope for improvement on both of them. Due to time and funding constraints this work could not be taken any further. However, The next step for the purpose of this work would be to build a model with all different types of medicines as category and implement not only classification, but also object detection. After the model is built, one needs to deploy the model at a powerful embedded platform such as NVIDIA Jetson TX2, which is energy efficient for deep learning inference. A detailed analysis of how does the training model avoids bandwidth constraints and improves latency is required, which the

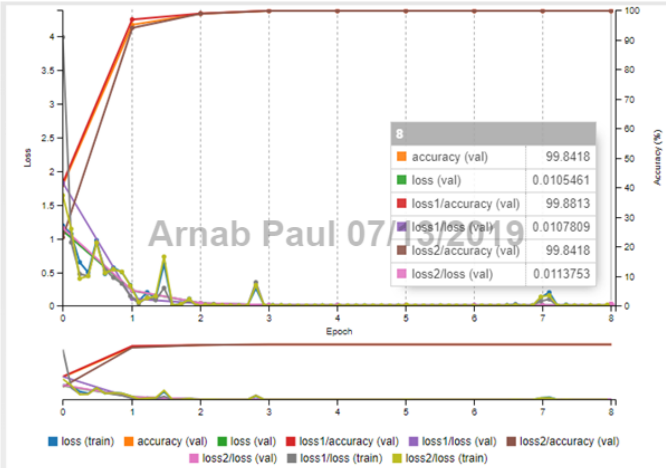

Fig. 3. Given Project Accuracy

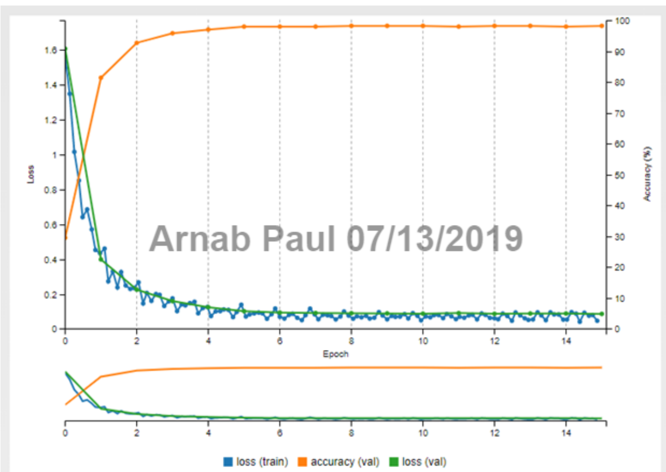

Fig. 4. Inference Project Accuracy

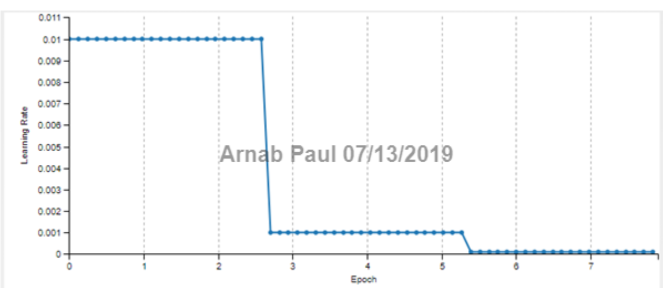

Fig. 5. Given Project Learning

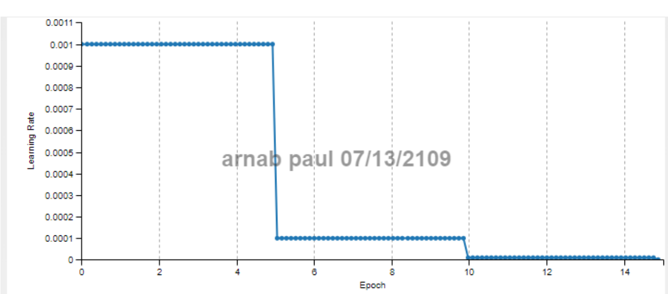

Fig. 6. Inference Project Learning

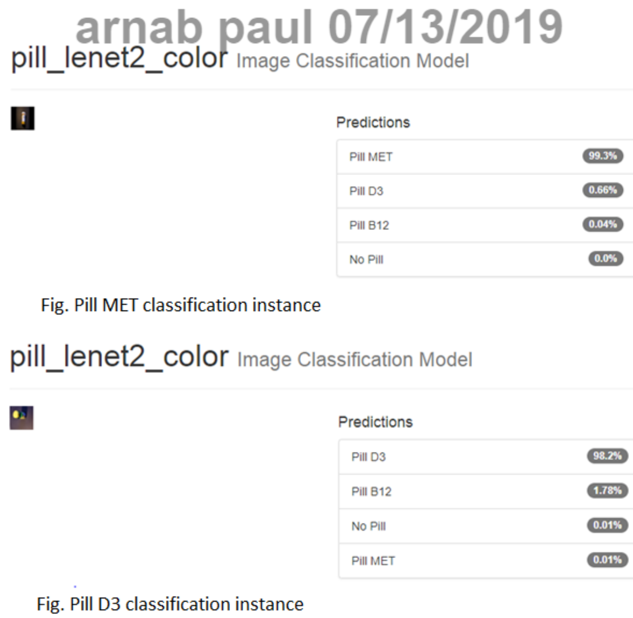

Fig. 7. Inference Project Classification

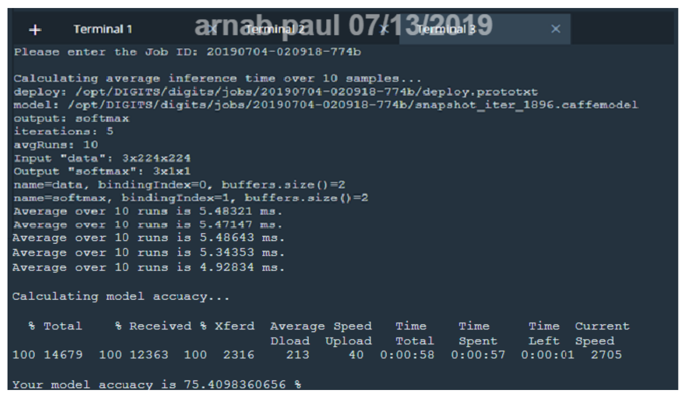

Fig. 8. Inference Time and Accuracy

current work did not address. Finally, a prototype could be build to understand its performance in the real world and commercial viability. The prototype could be improved further using the feedback from human user.

References

[1] LeCun, Yann; Lon Bottou; Yoshua Bengio; Patrick Haffner (1998). "Gradient-based learning applied to document recognition" (PDF). Proceedings of the IEEE. 86 (11): 22782324.doi:10.1109/5.726791. Retrieved November 16, 2013.

[2] Cheung K-C, Bouvy ML, De Smet PAGM. Medication errors: the importance of safe dispensing. British Journal of Clinical Pharmacology. 2009; 67(6):676680. [PubMed: 19594537]

[3] Spiller HA, Griffith JRK. Increasing burden of pill identification requests to US Poison Centers. Clinical Toxicology. 2009; 47(3)

[4] Krenzelok EP, Mrvos R. The use of an automated interactive voice response system to manage medication

identification calls to a poison center. Clinical Toxicology. 2009; 47(5):425429. [PubMed: 19504736]

[5] IEEE Appl Imag Pattern Recognit Workshop. 2016 October ; 2016,The National Library of Medicine Pill Image Recognition Challenge: An Initial Report,Ziv Yaniv1,2, Jessica Faruque1,3, Sally Howe1, Kathel Dunn1, David Sharlip1, Andrew Bond4, Pablo Perillan4, Olivier Bodenreider1, Michael J. Ackerman1,2, and Terry S. Yoo11

[6] Andreas Hart, Clemens Arth.Computer-Vision based Pharmaceutical Pill Recognition on Mobile Phones.In proceedings of CESCG 2010: The 14th Central European Seminar on Computer Graphics

[7] A Counter Medication Classifier Using Machine Learning and Computer Vision Techniques.Eduardo Ceh-Varela1, Gandhi Hernndez-Chan1, Marisol Villanueva-Escalante2, Giner Alor-Hernndez3, Jos Luis Snchez-Cervantes4,pages 5-10