# Robotic Inference Project Writeup

pook, chun

**Abstract**—This Report presents the result of using NVIDIA DIGIT to prototype the implementation of Real time object classification. A CNN is trained to classify objects on conveyor belt. And a second CNN is trained to classify object which is either plastic bottle, can or others.

**Index Terms**—Object Classification, NVIDIA DIGITS, CNN.

✦

## 1 INTRODUCTION

CLASSIFYING objects in real time has a lot of practical application. However classification task typically requires a lot of computational power in order to achieve the desired accuracy and timing constraint.Using GPU make this possible.

Returning items to be recycled, usually require sorting objects into different categories, e.g. plastic bottle, can, paper etc. Being able to classify recycle items to its proper category, is the necessary step to automate the recycle work flow.

## 2 BACKGROUND / FORMULATION

For the supplied data set, based on the requirement of having the inference time of less than 10ms and accuracy greater than 75%. Googlenet is chosen, with the learning rate of 0.001, SDG optimizer and 6 epoch, after some trial and error.

The same set of parameters is used on the collected data set except this time it is ran with 30 epoch.

## 3 DATA ACQUISITION

### 3.1 Supplied Data Set

The Data Set is provided by udacity, sample images is as shown below.There are 3 types of objects provided, which are candy boxes, bottle and nothing.

### 3.2 Collected Data Set

The sample data is collected using a USB web cam, each object are taken with different angles, pose, and the size of the image is 256x256, stored in png format. Three Datasets has been collected, they are bottle, can and others. Because of the limited amount of data, the dataset is augmented by rotating, shifting, shearing the original image to create about 400 images per category.

## 4 RESULTS

### 4.1 Supplied Data Set

As Shown in the figure below, the accuracy of the model after training is almost 100% after 6 epoch.

The accuracy reported by evaluation command is 75.4%, and the inference time is around 5ms, which satisfies the requirement to be accurate on or above 75% with an inference time less than 10ms.



Fig. 1. Sample of Supplied Data Set.



Fig. 2. Sample of collected Data Set.

### 4.2 Collected Data Set

As Shown in the figure below, the accuracy of the model on the collected data set after training is almost 100% after 30 epoch.however the inference time has not been measured.
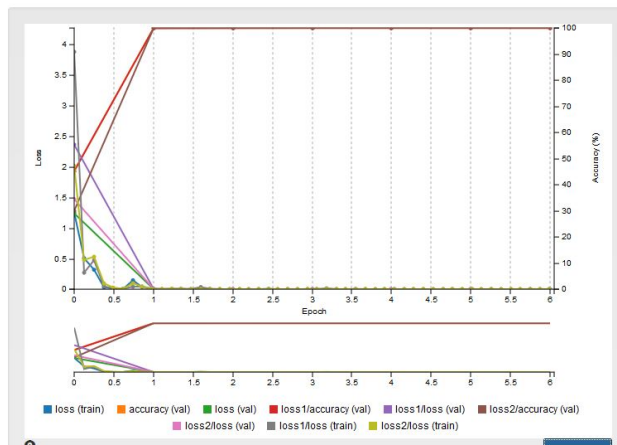
Fig. 3. loss and accuracy on supplied data set.



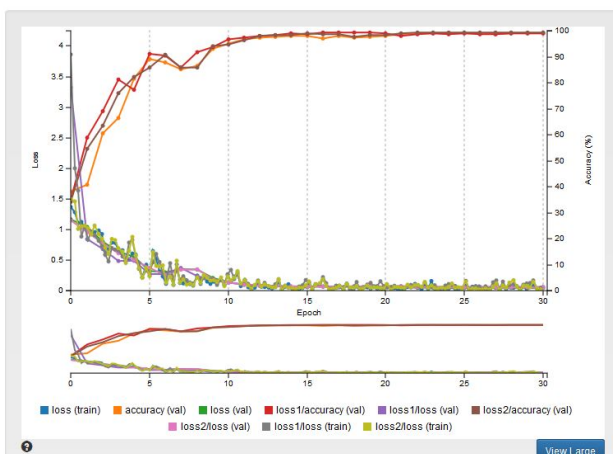Fig. 4. evaluation on supplied data set.



Fig. 5. loss and accuracy on collected data set.

## 5 DISCUSSION

### 5.1 supplied Data set

The model is able to achieve an inference time of around 5ms and accuracy 75.4% satisfying the requirement. However the accuracy for both training and valuation accuracy is around 100%. It may be the model is not able to generalize well. Including more variety of images from different source should help alleviate the problem.

## 6 COLLECTED DATA SET

The accuracy on the collected data set seems to be too good, achieving almost 100%. But probably the performance will drop for unseen photos like deformed plastic bottle or can Or bottle with liquid inside.

## 7 CONCLUSION / FUTURE WORK

Prototypes for object classification for objects on conveyor belt, and classifier for recycle object has been built. Further improvements will be to include more object types such as paper, glass etc in the classification process. Also inference time is important to real time application, an improvement will be to make sure the resulting model satisfy real time constraint, e.g. the inference time has to be less than 10ms. In Real life, bottles or cans are usually slightly deformed, sometimes still have liquid inside, images should include those kind of images to deal with real life situation.