CUSTOM IMAGE CLASSIFIER

Group No: 2

Amrutha K V Roll No: 10

Jithin K M Roll No: 29

Jithin T Mathews Roll No: 30

Prasila P Roll No: 45

S_8 B.Tech CSE

Government Engineering College, Wayanad

Guided By: Prof. Nikesh P, Assistant Professor CSE

September 20, 2020

Overview

INTRODUCTION

- Image recognition refers to technologies that identify places, logos, people, objects, buildings, and several other variables in images.
- Users are sharing vast amounts of data through apps.
- The large volume of digital data is being used by companies to deliver better and smarter services to the people accessing it.
- Image recognition is a part of computer vision and a process to identify and detect an object or attribute in a digital video or image
- Image recognition using Convolutional neural networks is capable of dealing with millions of parameters and provides high accuracy.

LITERATURE REVIEW

1. Rapid Object Detection using a Boosted Cascade of Simple Features.

Introduced by Paul Viola and Michael Jones[1], this is a machine learning approach for visual object detection which is capable of processing images extremely rapidly and achieving high detection rates. This approach has three key areas such as,

- Integral image : allow quick detection of features.
- Adaboost algorithm : selects critical visual areas of the image.
- Cascade : multiple classifiers are combined.

limitations- restricted to face data, Experiments on such a large and complex dataset are difficult and time-consuming.

2. Histograms of Oriented Gradients for Human Detection.
Navneet.Dalal,Bill.Triggs[2] studied the the question of feature sets for robust visual object recognition. They use the concept of histogram for detection of humans. They showed that experimentally the grids of Histograms of Oriented Gradient (HOG) descriptors significantly outperform existing feature sets for human detection. This includes :

- Compute gradients of input image
- Construct reflections of the image
- Apt descriptors are chosen
- Construct HOG
- Identify human/non-human

limitations-Again restricted to human detection.

LITERATURE SURVEY CONT..

Similarities between Rapid and Histogram approaches:

- Both have high detection rates.
- Both are limited to human detection.

Dissimilarities:

RAPID	**HISTOGRAM**
Limited to facial features.	Can detect a human completely.
Integral image concept.	Uses Histogram concept.

LITERATURE REVIEW CONT..

Rapid Object Detection using a Boosted Cascade of Simple Features

MERITS	DEMERITS
High detection rates.	Fails to identify identical twins.
Fully Automated.	Experiments on large dataset is difficult
Improved Security.	Multiple camera angle required.
Continuous Improvement.	Time consuming.

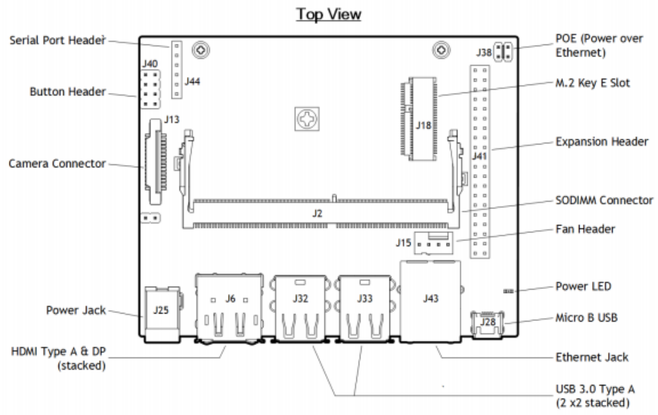
Histograms of Oriented Gradients for Human Detection.

MERITS	DEMERITS
High detection rates.	Does not provide specfic classification.
Fully Automated.	Large amount of data required.
Can handle with large amount of data	Use only continuous data.
Easy decision making.	Issues About Reliability and Efficiency.

PROBLEM STATEMENT

Lack of specific image classifiers to classify the vast amount of digital data being shared by users. The available image classifiers are capable of classifying inputs that are trained by neural networks generally. We don't need to recognize some generic objects, like cats and dogs and airplanes. We want to recognize something specific for example, identifying a breed of the dog for its automatic pet door, or a plant species for sorting.

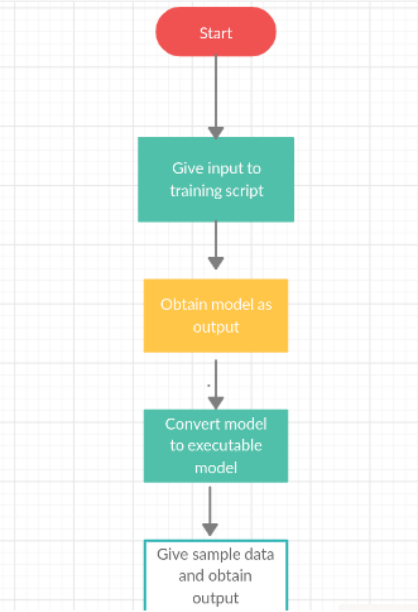
ARCHITECTURE



BOARD SPECIFICATION

- GPU: 128-core NVIDIA Maxwell architecture-based GPU
- CPU: Quad-core ARM® A57
- Video: 4K @ 30 fps (H.264/H.265) / 4K @ 60 fps (H.264/H.265) encode and decode
- Camera: MIPI CSI-2 DPHY lanes, 12x (Module) and 1x (Developer Kit)
- Memory: 4 GB 64-bit LPDDR4; 25.6 gigabytes/second
- Connectivity: Gigabit Ethernet
- OS Support: Linux for Tegra®
- Module Size: 70mm x 45mm
- Developer Kit Size: 100mm x 80mm

FLOW CHART


Start
Give input to training script
Obtain model as output
Convert model to executable model
Give sample data and obtain output

PROPOSED SOLUTION

Create a custom image classifier with transfer learning(re-training of last few layers of convolutional neural network) to identify specific classes of objects using Nvidia jetson nano and PyTorch.

WORK DONE IN PROJECT PRELIMINARY

- Setting up an environment manager for python and the creation and maintenance of multiple environments.
- Collection of the dataset.
- Installing necessary packages such as Keras, TensorFlow, NumPy etc.
- Setting up the initial training script using python and performing a sanity check of the initial model.
- Setting up PyTorch for transfer learning.

WORK DONE IN MAIN PROJECT

- Running the training script on the collected data set.
- Change variation of the number of neurons and epochs according to sanity checks.
- Saving the model constructed using the training script.
- Convert the constructed model to a model that can be run on the development board.
- Perform real-time classification of image inputs obtained through the camera sensor on the development board.
- Retrain the model until perfect classification is obtained.
- Try to improve confidence score of the model by performing hyper-parameter learning and feature extraction

WORK PLAN

Work Split

- AMRUTHA KV - managing environment for training model and feature extraction.
- PRASILA P - dataset collection, crossing(optional).
- JITHIN KM - pytorch for the development board.
- JITHIN T MATHEWS - training script and necessary packages.

WORK PLAN CONT...

Action plan

Building Initial Model	20/1/2020
Implementing on board	20/2/2020
Retraining of model	10/3/2020
Precise model	20/3/2020

Table: PLAN OF ACTION

DATA SETS

Tools used for Data collection:
Image Batch assistant

- Ad-on that can be installed on your browsers.
- Useful for downloading all the images found in a web page as a batch file rather than downloading them individually.
- This was used to collect data set of pens,markers etc... as readily available data set could not be found.

As for the plant and cat-dog data set,they were collected as a subset from these sources:
PLANTS - PLANTCLEF DATA SET 2017-
https://www.imageclef.org/lifeclef/2017/plant
CAT DOG - subset of ILSCRV12 -
https://drive.google.com/file/d/1LsxHT9HX5gM2wMVqPUfILgrqVIGtqX1o/

DATA SET CONT..

The collected data need to be set up before it can be fed as input for training. The data need to be split into three separate sets:

- Training set-This will contain the majority of the data set and will be used for training.
- validation set- This set will be used for testing the model accuracy after each epoch/training rounds.
- Test set- This will be used by the user to test the performance of the constructed model.

DATA SET CONT..

Feature extraction

- To avoid noise and perform pre-processing feature extraction where we try to highlight the important parts of the image that were tried out.
- However doing that on a data set that contains more than 10,000 images manually was thought to be infeasible.

Deep learning was the second option but these could only be used to separate human-like faces efficiently. Therefore it could not be done on the data set of pens, plants.

SETUP

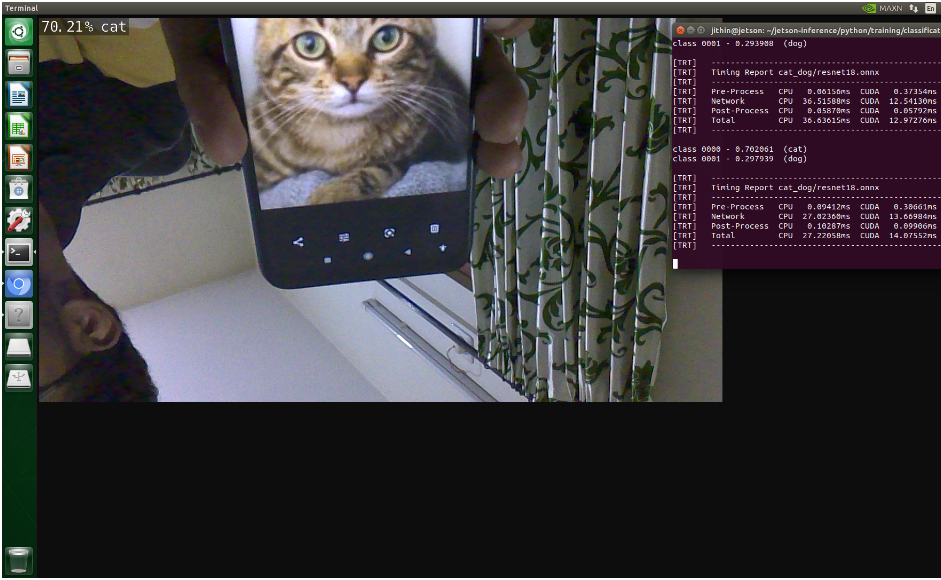
Important Tools used in the project include:

- Pytorch-PyTorch is an open-source machine learning library.
- Nvidia Jetson nano kit-NVIDIA® Jetson Nano™ Developer Kit is a small, powerful computer that lets you run multiple neural networks in parallel for applications like image classification, object detection, segmentation, and speech processing.
- Waveshare imx219-77- camera used for the development kit.
- Miniconda- Environment manager.
- Resnet-18 - ResNet-18 is a convolutional neural network that is 18 layers deep.

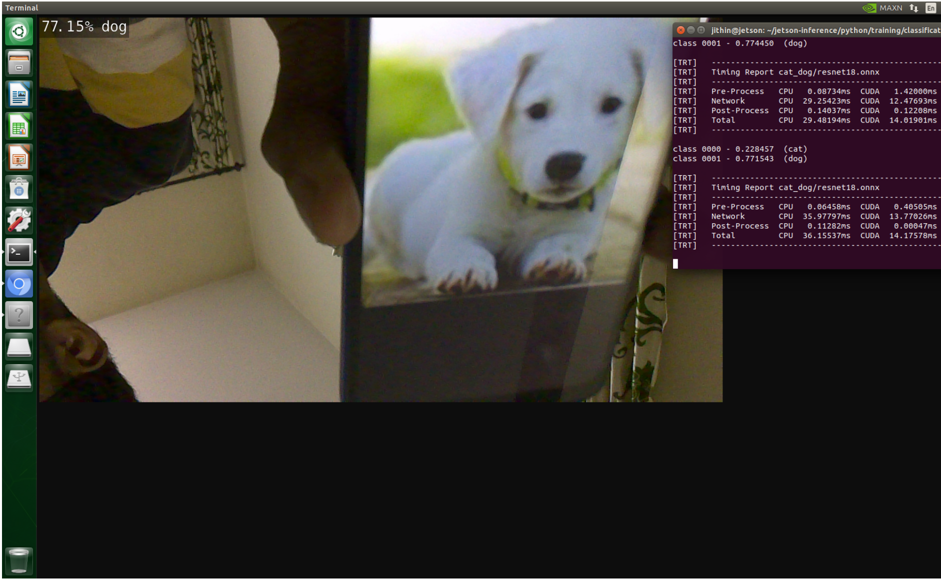
WORK PROGRESS

- Initial decision was to train a model that will tell apart different breeds of dog and cats but due to size constraints of the system(32 GB disk space), the model will be telling apart from 20 species of plants.
- The model will also classify a custom dataset ordered by the coordinator.
- Firstly, the model was trained to classify simple classes of cats and dogs.
- The model was able to classify among the two classes with a reasonable confidence score.

WORK PROGRESS



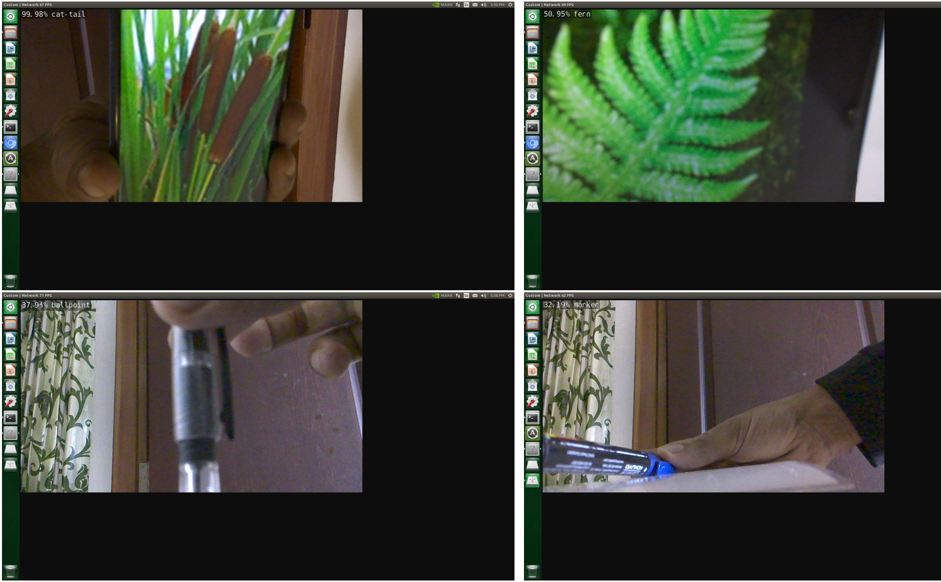
WORK PROGRESS



WORK PROGRESS

- Trained model to classify among different species of plants,training took around 36 hrs.
- Also trained model to classify the custom dataset.
- Model was able to classify the plants with good confidence score but the score for the custom dataset was relatively low,this was because lack of training data and the data that could be collected was of low quality.

WORK PROGRESS



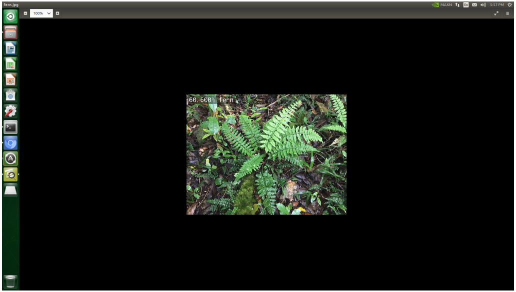
WORK PROGRESS

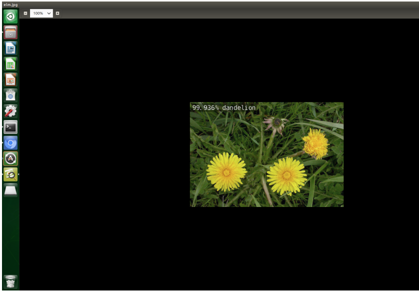
- Next feature extraction was planned to mark the important parts of the images that needed to be trained on.
- Owing to the large size of the dataset, a program or deep neural network was considered to do the feature extraction.
- However these could only separate and highlight human faces or similar so these could not be adopted to extract the custom dataset.

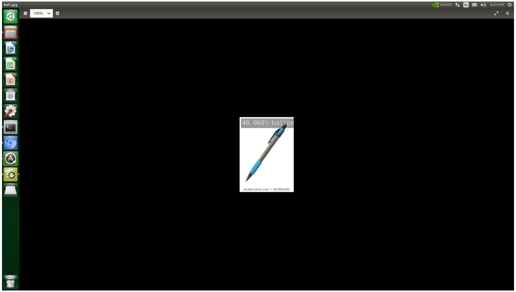
WORK PROGRESS

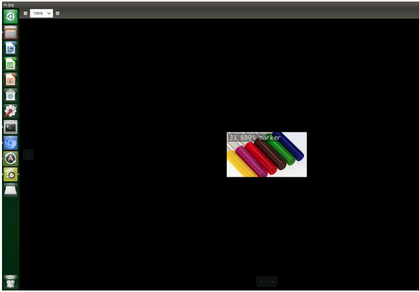
- For further research work, Hyperparameter learning was tried out, Here we modify the layers in the neural network and make slight variations in parameters and see how this affects the model accuracy.
- Minor improvements were made during hyperparameter learning for the custom dataset owning to poor training data.
- Notable improvements were found in the plant dataset.
- Ongoing work to further the project. The major parameters that affected the accuracy was an epoch, picture size, and the number of neurons or classes.

WORK PROGRESS

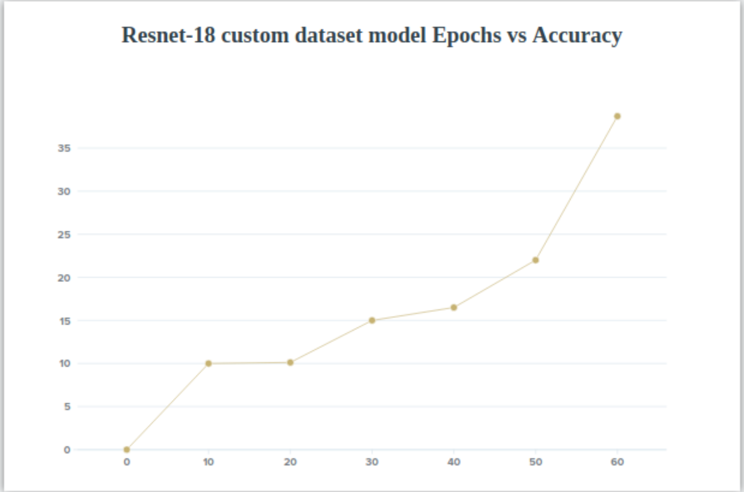




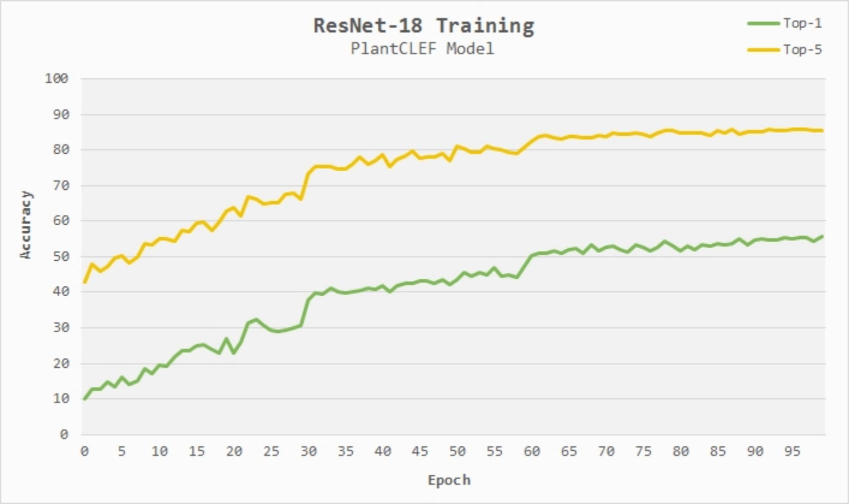




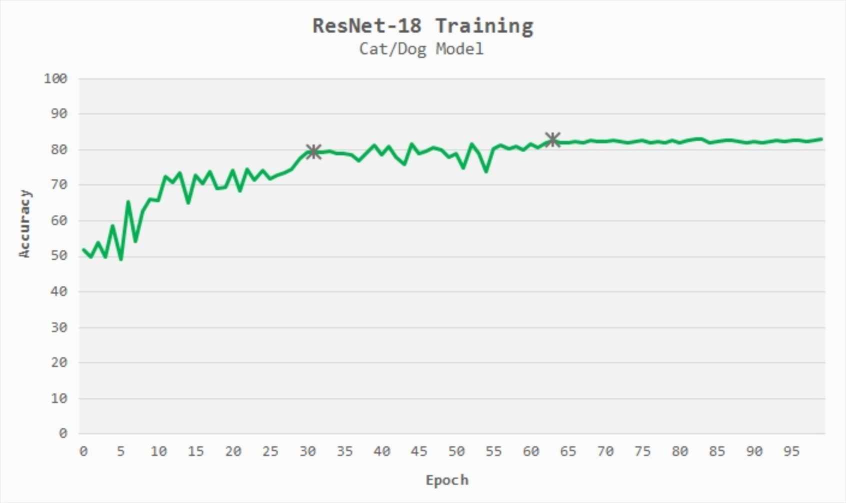
COMPARISONS



COMPARISONS



COMPARISONS



LIMITATIONS FACED DURING PROJECT AND ASPECTS NOT CONSIDERED

Limitations:

- The major limitation was the lack of availability of data set for pen, marker, etc....This made the project hard because data size is important in machine learning.
- Camera used had poor autofocus performance.
- Lack of time to do more training rounds or epochs

Aspects not considered:

- Mechanical actions-once a class has been identified, the development board could be connected to the motor to cause an action to occur.eg.. opening a door, raising an alarm, etc..
  These were not considered due to a lack of facilities at the time and is considered as a possible future upgrade.

ADVANTAGES

- More precise classification with better accuracy.
- System can be integrated into machines so that it can do a specific action once a specific class is recognized.
- Can identify more classes than the original model.
- The retrained model will be able to identify the new classes of the target images.
- Fully Automated.

LIMITATIONS OF PROPOSED SOLUTION

- Since there are lots of image classes that can be more deeply and specifically classified, it would be impossible to create a perfect one for all models that can classify everything it is given as input to. Therefore we have to choose things the model will classify and train the model on those sample images.
- A new model will have to be constructed for each instance of new user demand.
- Sometimes the accuracy of the model is deeply affected by the quality of the camera attached to the development board.
- The model may be quick to detect the noise in the test data image's environment and classify it accordingly.
- Large amount of data required.

CONCLUSION AND FUTURE WORKS

- Proposed system retrains the last few layers of the traditional CNN to classify its inputs more deeply.
- System is easier to implement with the use of transfer learning where a pre-trained model is used.
- Proposed system will provide real-time classification along with its confidence score.
- A lot of tasks can be automated using Image Recognization.
- It is a futuristic and relatively unexplored field, with wide areas of practical applications, including industrial, scientific, and medical applications.

CONCLUSION AND FUTURE WORKS

- During the research phase of the project, it was found out that the accuracy of the model depends on:
  - The size of training data up to a point.More the data, the better the accuracy.
  - The number of training rounds or Epochs done. On average 35 epochs produce good results and accuracy starts to converge at around 70 epochs.
  - The number of input layers used. More layers provide better top class predictions.
  - Hyperparameter learning where we manually reshape the neural network was able to provide an accuracy boost of 2-10 percent provided the training data was sufficiently large.

REFERENCES I

[1] Paul Viola Michael Jones.

Rapid object detection using a boosted cascade of simple features.

Mitsubishi Electric Research Labs Compaq CRL 201 Broadway, 8th FL One Cambridge Center, 14(2):131–142, 2001.

[2] Navneet Dalal and Bill Triggs.

Histograms of oriented gradients for human detection.

INRIA Rhone-Alps, 2005.

[3] Andrew G. Howard Menglong Zhu Bo Chen Dmitry Kalenichenko Weijun Wang Tobias Weyand Marco Andreetto Hartwig Adam.

Mobilenets: Efficient convolutional neural networks for mobile vision applications.

Google Inc., 9(3), 2015.

REFERENCES II

[4] Ferhat Culfaz.

transfer learning using mobilnet and keras.

towardsdatascience., 9(3), 2018.

[5] Zepan.

Train, convert, run mobilenet on sipeed maixpy and maixduino.

bbs.sipeed, 9(3), 2019.

Thank You