Analysis of Image Classification Models for Use on Low-Resource IoT Devices

A preliminary analysis report for a data science project for course LTAT.02.002 at the University of Tartu.

Villem Susi

2023

This project lives at https://github.com/villemsusi/IDS_project.

1. Background

This project seeks to delve into image classifiers and their usability on hardware with low computing power. The author, a student at the University of Tartu, aims to complete this project with the goal of using the achieved results as the basis for his thesis on pet identification with image classifiers on a Raspberry Pi. There are a multitude of pre-trained image classifiers available with varying performance and efficiency metrics and therefore it is daunting to choose one at random without any previous insight. Previous research conducted by the author on articles and research papers on similar topics has revealed different takes and results without many similarities and therefore it is deemed necessary by the author to conduct his own analysis for the best fitting results. The focus point of this project is finding the classifiers which got the best result metrics while calculating in the most efficient manner.

2. Business Goals

The main goal of this project is to simplify further work done by the author. The results from this project can and will be used to design the system behind the author's thesis.

3. Business Success Criteria

The results will be measured strictly on a per metric basis. This means that two models, working on the same data with the same parameters, are objectively comparable by their result performance and time usage metrics. I.e. better accuracy and lower amount of FLOPs signifies a better model.

4. Inventory of Resources

The data-mining aspect of this project will be run on a Raspberry Pi 4 with 4GB of memory. The codebase will be running mainly on PyTorch libraries and the pre-trained models will be taken from the TIMM library. The data to be used is taken from Oxford III-t Pet Dataset. A secondary source of data that may or may not be used is the Stanford Dogs Dataset.

5. Requirements, Assumptions, and Constraints

This project is to be completed by December 11th 2023 and to be presented on December 14th 2023. The project will be deemed complete when a definitive decision can be made about the best model. This requires running multiple variations of parameter-data combinations on many models and analysing the results.

6. Risks and contingencies

The project could be delayed most probably by one of two reasons. Firstly, if the software implementation deems to be overtly difficult then no progress can be made until it is completed. This can be averted by utilising outside help for working on the codebase. Secondly, if the hardware stops working, then also no progress can be made as the whole operation relies on hardware based performance. This can be averted by obtaining extra hardware as possible replacements.

7. Terminology

Image Classification - Process of categorising a given image into a set list of labels.

FLOPs (Floating Point Operations per second) - Measure of how many operations a computer makes in one second.

Accuracy - Number of correct predictions divided by number of all predictions .

Precision - Number of correct positive predictions divided by number of all positive predictions.

8. Costs and benefits

This does not have relevance to this project on a monetary basis. The only cost/benefit ratio relevant is time spent on further work. This is also irrelevant because the same amount of time would be put into this research whether or not this project is started/completed.

9. Data-mining goals

The goal regarding data-mining is to find all the necessary metric values for all the used image classifier models. This results in a dataset with all the gathered data in an analysable structure.

10. Data-mining success criteria

Data-mining has been successful if enough data has been collected of sufficient detail so that it can then be analysed. This is to be assessed after the completion of data-mining by the author. When not enough data was collected, further data-mining needs to be completed.

11. Gathering Data

● Data Requirements

For this data-mining project, a lot of images of pets are required. These images need to be of definitive classes and labelled with the names of those classes. There needs to be a sufficient number of images belonging to each class. The images need to be of distinct variations of one another for there to be enough varying data for the model to learn. The images need to be of type .jpg or .png.

● Data Availability

All the datasets mentioned previously are freely available and accessible. It is downloadable for anyone on the internet. Looking into the data shows that there are no problems with it. Therefore there needs to be nothing done to substitute or fix the currently selected datasets.

● Selection Criteria

The datasets to be used were mentioned earlier in the Inventory of Resources section. From these datasets, two categories of data will be used. Firstly, the images of pets containing their full body shots will be used to train the models for grasping the focus of the image. Secondly, for each of the full body images, there exists data for a region of interest (ROI). This will be used to train the models for classifying on data with minimal amount of unnecessary information. This is crucial for time performance optimization.

12. Describing Data

The Oxford III-t Pet Dataset consists of 16 000 images of 37 classes of pets, both dogs and cats. For each image there exists a breed label, ROI coordinates around the face of the pet, and a trimap segmentation which will not be used. This is all that is needed for the data-mining.

13. Exploring Data

Observing the data closely reveals that 1 image seems to be corrupt (“Abyssinian_34.jpg”). This image will be left out of the dataset for the mining purposes. The rest of the data in the dataset seems to be of sufficient quality.

14. Verifying Data Quality

Overall the data is of great quality. Disregarding the 1 corrupt image which plays an insignificant role considering the amount of data left.

15. Project Plan

1. Setup the necessary hardware

This task consists of acquiring a Raspberry Pi and setting it up to run PyTorch models. This task should take approximately 12 hours. It is necessary during this task to foolproof the underlying infrastructure on the hardware so that further work won't be delayed by problems with the base system.

2. Modify the original data

For better image classification performance, the model should be trained on augmented data. The original images will be transformed, coloured etc to enhance the amount of variation in the original data. This should take approximately 4 hours.

3. Train the models on the chosen dataset

Model training will take the most time, as it consists of firstly setting up the software base required to run training and then the passive time that training takes. It is impossible to gauge the time spent on this task as it relies heavily on training speed.

4. Import the models to the Raspberry Pi along with test data

This task should take approximately 4 hours.

5. Test the models on the Raspberry Pi and measure the performance with the chosen metrics

This task should take approximately 5 hours.

6. Analyse the result data

This task consists of delving into the results and making meaningful conclusions. The analysis will be made using Python 3.9 and supporting external libraries i.e Scikit-Learn. This should take approximately 15 hours of work.