

**AI Methods – Cover Sheet**

ASSIGNMENT

Instruction:

 Marks will be awarded for good presentation and thoroughness in your approach.

 Complete this cover sheet and attach it to your assignment.

|  |  |  |
| --- | --- | --- |
| Student declaration: | | |
| *I declare that:* | *I understand what is meant by plagiarism*  *The implications of plagiarism have been explained by our lecturer*  *This assignment is all my work and I have acknowledged any use of the published or unpublished works of other people in documentation.* | |
| Signature: | | Date: |

|  |  |  |
| --- | --- | --- |
| Name and ID | Zhang Zi Teng | TP052096 |
| Name and ID | Tan Wei Xian | TP048417 |
| Name and ID | Leong Soong Jun | TP050994 |
| Name and ID | Lyu Chen Yang | TP052020 |
| Intake | UC2F1809IS |  |
| Research Topic | Phone Recognition System |  |

# Table of Content

[Table of Content 2](#_Toc26537395)

[Abstract 3](#_Toc26537396)

[Introduction 4](#_Toc26537397)

[Materials Method 5](#_Toc26537398)

[Literature Review 5](#_Toc26537399)

[Tan Wei Xian TP048417 5](#_Toc26537400)

[Leong Soong Jun TP050994 12](#_Toc26537401)

[Zhang Zi Teng TP052096 14](#_Toc26537402)

[Lyu Chen Yang TP052020 16](#_Toc26537403)

[Algorithm implements 18](#_Toc26537404)

[Results and Discussion 19](#_Toc26537405)

[Discussion on implementation 19](#_Toc26537406)

[Install 19](#_Toc26537407)

[Downloading image data 19](#_Toc26537408)

[Loading and viewing data 20](#_Toc26537409)

[Creating a model and initial training 22](#_Toc26537410)

[Cleaning data 24](#_Toc26537411)

[Interpretation 24](#_Toc26537412)

[Result 25](#_Toc26537413)

[Improvement / Suggestion 26](#_Toc26537414)

[Conclusions 27](#_Toc26537415)

[References 28](#_Toc26537416)

# Abstract

Convolutional Neural Network (CNN) is one of the best approaches to identify and recognize objects from and image. Since the late 1980s, Visual studies is applied to Convolutional Neural Network (CNN) but due to the lack of big data and computer power. This technique was dormant until the mid-2000s and started a boom since 2010s due to the remarkable performance and huge potential of business opportunities. (Rawang, Wang, 2017). In this review, we will talk about the meaning of CNN, the techniques to identify specific objects from a picture and recognize the objects based on the data. Its goal is to give a brief introduction to the reader to understand how CNN works and why CNN has a huge potential in some automated vision projects.

# Introduction

In the era of 21st century, every country in the world has slowly risen its research & development department to improve the nation in various ways. However, we believe that technology is the key factor to develop the country at a fast pace. Undoubtedly, Malaysia is a developing country, which is doing well in this field. We are the student of Asia Pacific University, would like to follow the steps of our country, then we received this task as our assignment to deeply immersed ourselves into the field of artificial intelligence. Upon the requirement of this assignment, we have studied various artificial intelligence algorithms that we see the potential of future value and future improvement, etc. Artificial intelligence can be understanding the simulation of human intelligence processes by machines, especially computer systems. The types of artificial intelligence included expert systems, speech recognition, and machine learning. Nevertheless, we will further explain the algorithm we choose in the next following statement by providing enough evidence to support our algorithm.

# Materials Method

## Literature Review

### Tan Wei Xian TP048417

The usage of CNN in object detection

Convolutional Neural Network is the greatest in learning object and detect an object. Researchers have found that CNN has the highest accuracy when it comes to object detection. In this study, The method based on the foreground target modelling is designed by establishing the apparent model of the target. The appropriate classifier is used to classify and detect the targets in the video. The researcher used Hog method to perform the object detection, but the Hog method has low accuracy in object detection. This paper proposes a model based on R-CNN(Girshick R, Donahue J, Darrell T, et al,2014) and SPP-Net (He K, Zhang X, Ren S, et al.)network structure, which fully exploits the advantages of deep convolutional neural networks, and transforms the target detection problem of images into classification problems, which has achieved proper results. In conclusion, Convolution Neural Network has a remarkable result in Computer Vision field because of it high capability of learning the pattern of image

Transfer Learning in CNN

Transfer Learning has greatly helped the researcher to construct their own neural network and greatly improve the accuracy of the model. Researcher has proved that the usage of Transfer Learning greatly increases the accuracy of the model. In (Giuseppe Amato, Paolo Bolettieri, Vinicius Monteiro de Lira, Cristina Ioana Muntean, Raffaele Perego, and Chiara Renso. 2017.) project, Md Tohidul Islam, Sagidur Rahman, B.M. Nafiz Karim Siddiqueand Taskeed Jabid used food-11 dataset and they used Transfer Learning for their project. Transfer Learning is a method which is used a pre-trained model from other sources, unfrozen the model and train with our data. They used pre-trained model Inception V3 model in detail (Giuseppe Amato, Paolo Bolettieri, Vinicius Monteiro de Lira, Cristina Ioana Muntean, Raffaele Perego, and Chiara Renso. 2017.). The datasets had been pre-processed by researchers to have better performance, reduce complexity and increase accuracy. As the result, their model has higher accuracy than the model without Transfer Learning. In conclusion, Transfer Learning is a great way to improve your model.

### Leong Soong Jun TP050994

Deep-learning based automatic recognition network is broadly implemented in the field of agriculture in China. Previously, to chase back the past, traditional image processing technology is classified into a few specific features. For instance, including color, shape, and graining. Due to the massive amount of data generated by the mobile Internet and the development of large-scale computing devices and technologies, the deep learning algorithm (LeCun et al., 2015) has seen a breakthrough in terms of image recognition technology. Beijing Engineering Research Centre for Intelligent Agricultural Machinery developed an agricultural machinery operation supervision service system (Yin et al., 2018) that consists of a vehicle-mounted monitoring terminal, Global Navigation Satellite System (GNSS) positioning sensor, machine operation monitoring sensor, vehicle-mounted waterproof camera, and machine recognition sensor. Image processing systems then classify the images through methods such as an artificial neural network (Roffman et al., 2018) and support vector machine (Thanh Noi and Kappas, 2018). Nonetheless, the neural network and specified algorithms are introduced to help in recognizing the image in different types of problems. Hence, there is a must to further investigate in the field of image recognition regarding machinery by using the methods below; image classification, processing, cropping, color adjustment, then motion blur elimination, image denoising, collect and filter the image dataset.

Video image recognition on obstacle detection in an intelligent vehicle is debating in the technology industry towards the era of industrial 4.0. Previously, the elite only focuses on the performance of the intelligent vehicle. With the rising of development of the intelligent vehicle, Mohammed Shakeel et al. (2018) study about multi-model background detection, illumination, change subtraction and shadow removal methods. Lan et al (2017) had tested intelligent vehicle control and external sensors. Next, further exploration of image recognition technology yet obstacle detection. Yong et al. (2018) used radar technology to detect crop identification and measurement in farmland. However, this multi-sensor technology is implemented to detect and identify the target in color images. Hence, there are a few methods is to experiment in further investigation. For instance, intelligent vehicle technology, obstacle detection technology, machine vision technology. However, the various experiment is conducted to test the real-time detection.

### Zhang Zi Teng TP052096

Deep learning is a technology that simulates the human brain function by automatically analyzing data to identify unstructured data through an artificial neural network. Nowadays, it is popularly used in the fields of data analysis and prediction, such as image classification. A deep learning convolutional neural network based on Keras and Tensorflow is deployed using python for binary image classification by (Karan Chauhan, Shrwan Ram, 2018). In this study, (Karan Chauhan, Shrwan Ram, 2018) compared four different structures of CNN on the CPU system, with different combinations of classifiers and activation functions. With experiments, (Karan Chauhan, Shrwan Ram, 2018) observed results for each combination and observed that for binary image classification, Relu activation function and Sigmoid classifier combination gives better classification accuracy (90.54%) than any other combination of activation function and classifier. Therefore, it is shown that the combination of the sigmoid classifier and the Relu activation function for binary object classification gives greater categorization accuracy than any other classifier and activation function combination, especially on the CPU system.

Face recognition is big research in the field of artificial intelligence. With this technology, the computer can recognize a person's face without any human assistance. In the olden days, many studies and practical performances of the face-detection were not satisfied until Viola and Jones proposed work. (2000) The Viola and Jones are the first people who applied rectangular boxes for the face, but, it had a lot of disadvantages as its feature size was large (Manik Sharma, J Anuradha, H KManne and GS CKashyap, 2017). Garcia et al. (2002) introduced a neural network to find the semi-frontal human faces in the complex images. Osadchy et al. (2005) trained a convolutional neural network for face detection. After many experiments, (Manik Sharma, J Anuradha, H KManne and GS CKashyap, 2017) found A DBN can work globally which affects a model that should improve performance. Like the camera lenses slowly focus on the picture the reason DBN works better is highly technical and a stack RBM will work as a single unit. Therefore, compared to other models, the DBN model is successfully implemented and provided very good results. In the still images, videos, paintings, and webcam capture, it can recognize faces.

### Lyu Chen Yang TP052020

Nowadays, Deep Learning in the AI field has achieved significant development in many fields for several years, and when we need to process a large number of pictures, we can use convolutional neural networks to process them. The researchers found that after training the convolutional neural network and then using the ANN binary classifier for classification, the classification accuracy of 88.31% was finally achieved. Hopfield et al. proposed the Hopfield network in the early 1980s. (Santisudha Panigrahi&Anuja Nanda&Tripti Swarnkar,2018) This had invigorated the artificial neural network. Further, Hinton et al. proposed the Boltzmann machine by using a simulated annealing algorithm (D.H Ackley&G.E. Hinton&T. J Sejnowski,1985). For CNN implementation a large volume of data set is required for training the model (Simonyan&Zisserman,2014). Otherwise, we can use the pre-trained CNN i.e. called transfer learning to fit our requirement (M. Oquab & L. Bottou & I. Laptev&J. Sivic,2014). Natural images are the photos having sharp edges and flat areas of color. Many researchers have developed new algorithms, new optimization methods in the field of computer vision using deep learning (Santisudha Panigrahi&Anuja Nanda&Tripti Swarnkar,2018). Hence, after various levels of optimization have been carried out to improve the performance of convolutional neural networks, the best classification accuracy using deep learning has been achieved.

Manual identification of objects and facilities in satellite images is subject to human resources and scope resulting in inaccuracies and inaccuracies, hence the need for intelligent algorithms. Deep learning is a family of machine learning algorithms that have shown promise for the automation of such tasks. It has achieved success in image understanding by means of convolutional neural networks (Mark Pritt & Gary Chern, 2017). It has achieved astonishing success in object detection and classification by combining large neural network models, called convolutional neural networks (CNNs), with powerful graphical processing units (GPUs). Since 2012, CNN-based algorithms have dominated the annual ImageNet Large Scale Visual Recognition Challenge for detecting and classifying objects in photographs (Mark Pritt&Gary Chern, 2017). On the IARPA fMoW dataset of one million images in 63 classes, including the false detection class, the system achieves an accuracy of 0.83 and an F1 score of 0.797. It classifies 15 classes with an accuracy of 95% or better (Mark Pritt&Gary Chern,2017). Hence, by combining this deep learning system with the detection component, the speed of post-natural disaster processing is improved.

## Algorithm implements

Our system will be a Phone recognition system based on Python and involved in the field of image recognition technology. Image classification is one of the most basic goals in deep learning computer vision. However, image classification is not that simple every day. Recognize and classify most images in similar shapes and visual representations. For example, our system will recognize different models of smartphones. Just as people can see an animal's tail and claws to tell what animal it is.

In order to achieve the above system purpose, the system includes the following steps:

1. Enter the image of the mobile phone that has been trained and to be recognized.

2. Segment the training samples in the database to obtain an image template set.

3. Analyze the resulting image template.

4. Match and process this image with the image template set.

5. Finally, get its model from the representation data of this image.

# Results and Discussion

## Discussion on implementation

### Install

Firstly, we can install the FastAI library using either conda or pip. Such as the following image: (Figure 1):

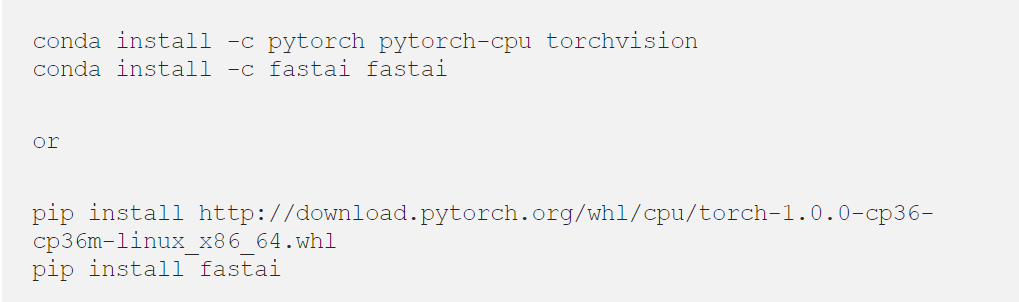


Figure 1 “conda” or “pip” to install FastAI

Next is to import the version module by typing Such as the following image: (Figure 2):



Figure 2 import fastai library

### Downloading image data

The FastAI library provides a lot of specific datasets that can be loaded directly, but it also provides features to import images that provided a file containing the URLs of these images. FastAI will navigate to Google Images in order to get the URLs, searching for sufficient images for the same category.

In our project, we create a phone classifier in this project that can differentiate between ' Samsung S10 plus, '' iPhone 11, '' Redmi K20 plus ' and ' Huawei P30 pro. ' To do this, we searched for all the images of the four modern of the phone and used the command to save a linked CSV file.

After downloading the CVS files, using the "download images method", we can download the data and check if there are any corrupted images using the "Verify images method". Such as the following image: (Figure 3):

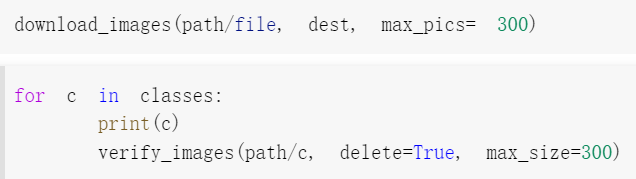


Figure 3 download images method and Verify images method

It will try if it is possible to open any image in the given folder and have n-channels. If there is no one of these conditions, the image will be removed.

### Loading and viewing data

FastAI has specific data objects that are called “data bunches” to train a prototype. It is possible to create such “data bunches” in two main ways.

The first solution is to use specific problem solutions such as “ImageDataBranch”. This approach is simple but can't be used for others. Such as the following image: (Figure 4):

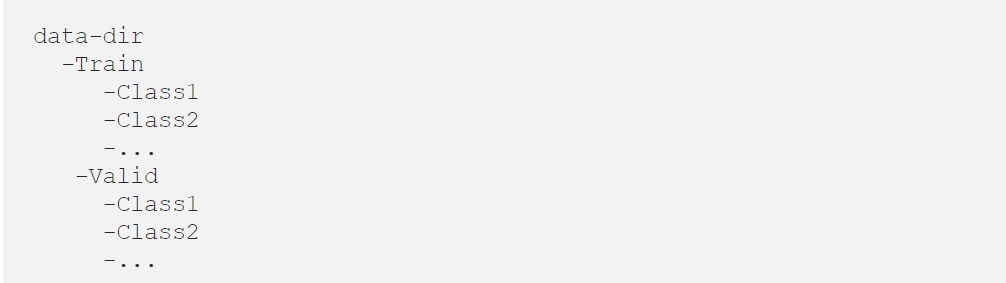


Figure 4 ImageDataBranch

So, in our project, we use another way to load data called the “data block api”, which provides you with more options by isolating the underlying parts of that process into separate blocks. Such as the following image: (Figure 5):

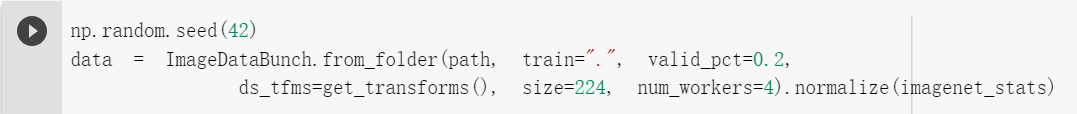


Figure 5 data block api

Next, we can use the object of “data” to collect further knowledge or data. Such as the following image: (Figure 6):

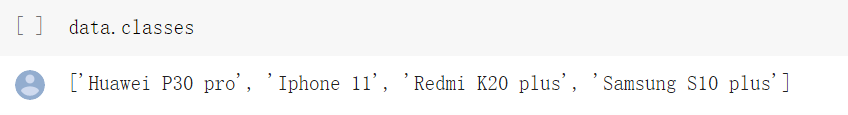


Figure 6 use object “data” to get more data

Using the “show batch method” we can view a random batch of pictures. Such as the following image: (Figure 7):

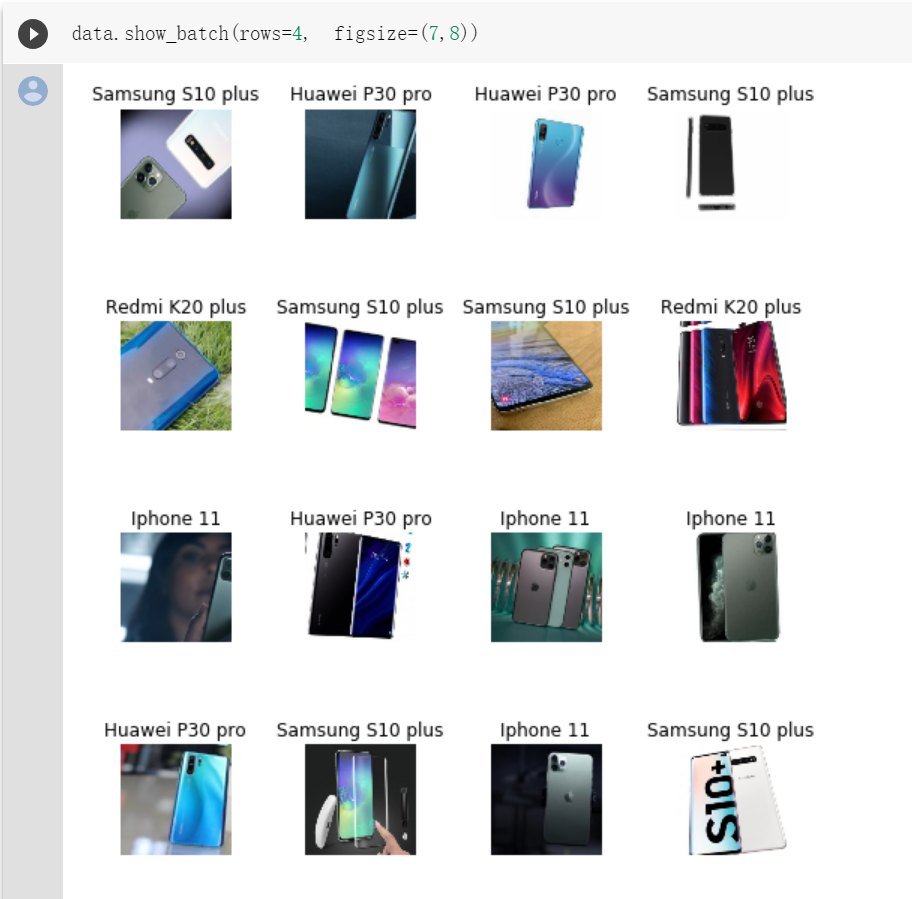


Figure 7 Random batch of images

### Creating a model and initial training

The FastAI library is designed to allow you to build models with just a few lines of code. It provides a method called "build cnn" that can be used to create a neural network that is convolutional. The method requires two arguments, "data" and "architecture", but it also supports many other parameters that can be used to customize the model for a particular problem. Such as the following image: (Figure 8)



Figure 8 build cnn method

The "fit" method is the "standard" way to train a neural net with a constant rate of learning, while the "fit one cycle" method uses something called the 1 cycle rule, which effectively adjusts the rate of learning over time to achieve better results. Such as the following image: (Figure 9)

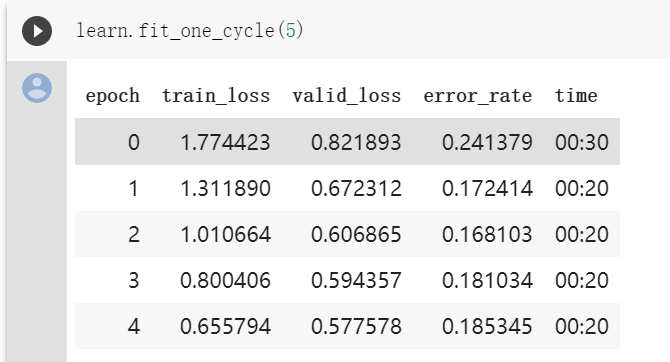


Figure 9 Training output

FastAI offers another technique to improve learning transfer called differential learning rates, enabling us to set different learning rates for different parts of the network. We can use the "lr find" and "recorder.plot" methods to find the perfect learning levels, and generate a plot that relates the learning level to the loss. Such as the following image: (Figure 10)

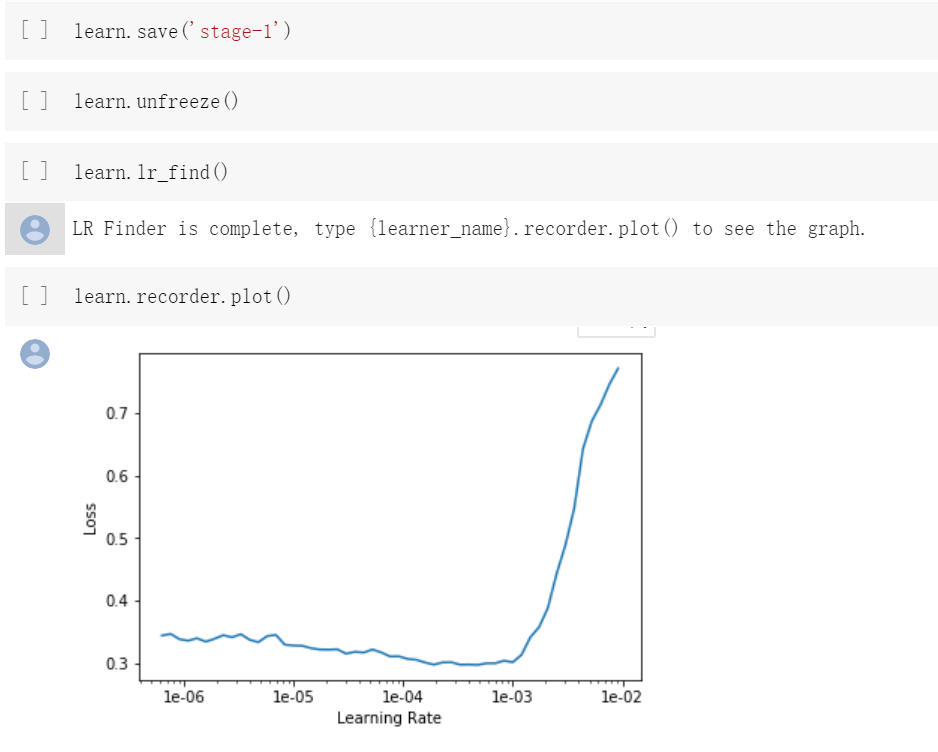


Figure 10 Learning rate plot

Next, we need to move the “max\_lr” argument to the “fit\_one\_cycle” method to train the model using differential learning rates. Such as the following image: (Figure 11):

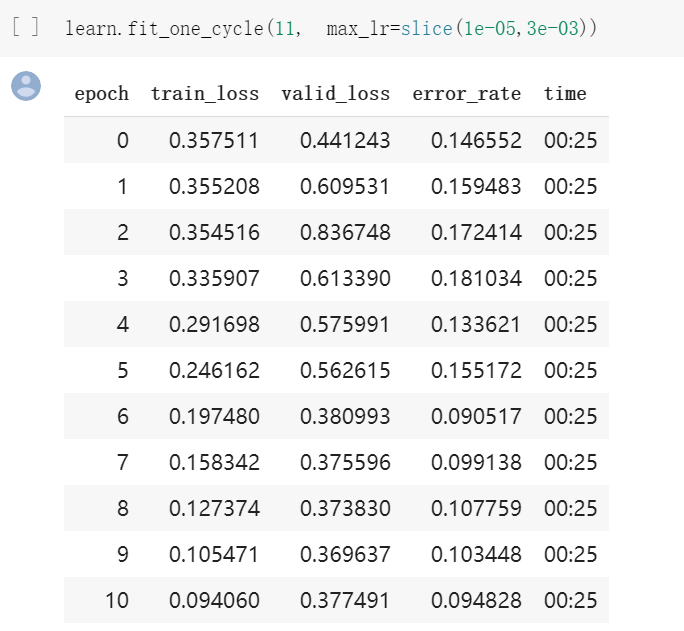


Figure 11 Training output

Lastly, it can be saved using the “save” method. Such as the following image: (Figure 12):



Figure 12 save method

### Cleaning data

FastAI also offers the functionality to use Jupyter widgets to clean the data. The ImageCleaner class shows relabeling or removal images and stores route changes as “cleaned.csv”.

To use “ImageCleaner”, first use “DatasetFormatter()” from top losses to get the suggested image indices. Such as the following image: (Figure 13):

****

Figure 13 ImageCleaner method

### Interpretation

Ultimately, to analyze our tests, we can use FastAI's “ClassificationInterpretation” class. We need to call the “from\_learner” method and move it on to our learner/model to construct an interpretation entity. Instead strategies such as “plot\_confusion\_matrix”, “plot\_top\_losses” or “most\_confused” can be used. Such as the following image: (Figure 14):

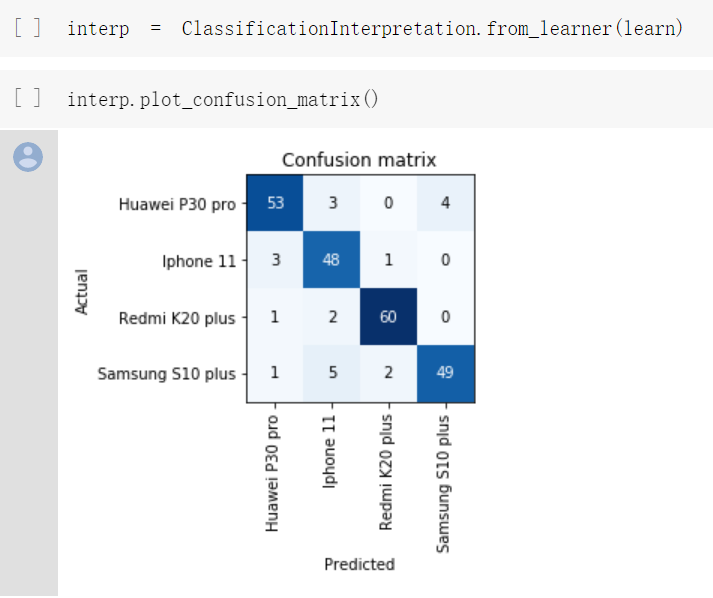


Figure 14 Interpretation

# Result

A general description of the algorithm that we have chosen, it is a very simple and easy algorithm that is able to perform smartphones recognition function. We have executed the algorithm on Google Colab to train and test it. As a result, we expect the algorithm to identify the different types of phones in severe types of color, a different angle of pictures taken, the quality of image and encounter obstacles in the picture. Nevertheless, we found out severe issues. For example, we feed other picture which is a totally different object compare to the smartphone, the algorithm obviously will reveal its recognition but end up with a wrong result and failed to detect different objects. We only able to access limit resources since our computer device could not afford higher requirement, therefore, we feed 300 pictures to 4 smartphones to train the algorithm. In results, we saw a high percentage results to recognize these 4 smartphones; iPhone 11 Pro, Samsung S10 Plus, Huawei P30 Pro and Redmi K20 Plus.

# Improvement / Suggestion

We have noticed our mistakes and ways to improve our algorithm. We found out the importance of a database. A larger database allows us to develop or train our algorithm to perform seamlessly and provide a better result. Nevertheless, instead of a larger database, we also need a computer device with higher specifications. A higher specification computer device can resist the physical requirement of our algorithm needs. It certainly resolves the major issues that we faced. For instance, we would depend on the graphic card, RAM, memory storage of the computer device to execute the algorithm and the process of executing the code is seamless. In general, we suppose to feed more data to train our algorithm to reduce the failure rate, unfortunately, we have hardly provided more data since our computer device does not meet the requirements. Hence, we have identified one bug which is recommended to repair. Our algorithm will identify the error, but it failed to report a result that an object is different from a smartphone. However, it will report a smartphone even the pictures are shown in other objects instead. We will work on to repair this bug to display a message, “Identify different types of the object”, or even recognize it and report the exact object’s name. We have observed severed some bugs in our algorithms, then we will repair it to achieve higher goals.

# Conclusions

Phone Classification is a challenging task because of our technical issue and so much variation on the same task (example of these are the color of the phones, the size difference between each phone model, etc). Despite this, CNN has shown a great result without any rule-based system. This paper also proved that Transfer learning can greatly improve the model accuracy. Even though our model is still far from capable to be used for business since there are still lots of wrong predictions, I believe this paper has proved Convolutional Neural Network (CNN) is one of the most effective methods to develop a model that learns the pattern of image and predict the label of the image.

# References

AI (2018). What is AI (artificial intelligence)? - Definition from WhatIs.com. [online] SearchEnterpriseAI. Available at: <https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence> #:~:targetText=Artificial%20intelligence%20(AI)%20is%20the,conclusions)%20and%20self%2Dcorrection.‌[Accessed 4 Dec. 2019].

Pritt, M. and Chern, G. (2017). *Satellite Image Classification with Deep Learning*. [online] Computer.org. Available at: https://www.computer.org/csdl/proceedings-article/2017/aipr/13xI8A66zF1/13xI8A8WyWY [Accessed 4 Dec. 2019].

Panigrahi, S., Nanda, A. and Swarnkar, T. (2018). *Deep Learning Approach for Image Classification*. [online] Computer.org. Available at: https://www.computer.org/csdl/proceedings-article/2018/icdsba/17D45VtKirR/17D45VN31g8 [Accessed 3 Dec. 2019].

Medium. (2019). *FastAI Image Classification*. [online] Available at: https://towardsdatascience.com/fastai-image-classification-32d626da20 [Accessed 5 Dec. 2019].

Docs.fast.ai. (2019). *Installation | fastai*. [online] Available at: https://docs.fast.ai/install.html [Accessed 5 Dec. 2019].

GitHub. (2019). *fastai/fastai*. [online] Available at: https://github.com/fastai/fastai/blob/master/README.md#installation [Accessed 5 Dec. 2019].

Medium. (2019). *10 New Things I Learnt from fast.ai v3*. [online] Available at: https://towardsdatascience.com/10-new-things-i-learnt-from-fast-ai-v3-4d79c1f07e33 [Accessed 5 Dec. 2019].

Fast.ai. (2019). *Practical Deep Learning for Coders 2019 · fast.ai*. [online] Available at: https://www.fast.ai/2019/01/24/course-v3/ [Accessed 5 Dec. 2019].

Ke Ning, W. (2019). *Five techniques of computer vision*. [online] Cloud.tencent.com. Available at: https://cloud.tencent.com/developer/article/1109237 [Accessed 3 Dec. 2019].