Development Environment of Machine Learning for Image Classification

Irwan Jani Tarigan\textsuperscript{1}, Poltak Sihombing\textsuperscript{2} and Pahala Sirait\textsuperscript{2}

\textsuperscript{1}Postgraduate, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Padang Bulan, Medan, Indonesia
\textsuperscript{2}Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Padang Bulan, Medan, Indonesia

ijanitarigan@gmail.com, poltak@usu.ac.id, sirpahala@yahoo.com

Abstract. Machine learning is an interesting topic nowadays, it is also a subarea of artificial intelligence. An interesting thing that needs to be done in the beginning is how to configure the Development Environment (DE). In this paper, DE is configured using Acer-Notebook with jupyter notebook and Smartphone with google colab for image classification using convolution neural network method. The configuration process is easier to do on Smartphone Realme-C2 than Acer-Notebook and the best level of accuracy is epoch 100 in training and testing data obtained by Smartphone Realme-C2 are 0.70 and 0.66, while Acer-Notebook is 0.70 and 0.63.

Keyword: Development Environment, Jupyter Notebook, Google Colab, Classification, Machine Learning, Convolution Neural Network

1. Introduction

Development environment is needed in terms of machine learning to apply the problem solving methods in classification cases, how to apply them and understand about utilize hardware, software, installation process and programming language to used [1]. Machine learning is a subarea of artificial intelligence, a concept that is more devoted to providing the ability of the system to learn from a collection of data by applying algorithms in problem solving [2].

Machine learning techniques have several categories, depending on what issues will be handle, including:

1.1. Supervised Learning
Learning models that can be used to find the right pattern of data that will be applied in the analysis process, the data will be modeled and labeled for identification, also known as classification.

1.2. Unsupervised Learning
Learning models that can be used for very large data problems and do not have labels, which have the ability to find patterns in the classification and cluster.
1.3. Reinforcement Learning
Learning models that lead to system behavior in making the best decisions with the right sequence of steps, usually found in the field of robotics or simulation devices (games and self-driving).

1.4. Deep Learning
Known as deep learning, a more specific method is by combining artificial neural networks for more complex problems. This model is often used in applications such as image recognition, speech and computer vision [3].

Image classification is a very complex process and is impacted by many factors, it is necessary to choose a method or algorithm that supports to improve the accuracy [4].

![Artificial Intelligence Machine Learning](image-url)

**Figure 1.** Subarea Artificial Intelligence and Machine Learning [3]

In this paper, we will discuss how to implement Development Environment in machine learning for image classification using Jupyter notebook and Google Colab based on the python programming language.

2. Related Work
Phan Thanh Noi and Martin Kappas, 2018, in their paper used three parameters of the classification method, such as: Random Forest (RF), K-Nearest Neighbor (KNN) and Support Vector Machine (SVM), the aim is to classify 30 x 20 Km² land using the sentinel-2 dataset located in Vietnam. From the performance of the proposed method obtained an overall accuracy of 90% - 95% [5-6].

Rohit Verma and Dr. Johid Ali, 2019, in their paper outlining the KNN, SVM, ANN and CNN methods to compare methods that might be beneficial in assignments and image classification studies [7].

In this paper, we will use a Machine Learning algorithm with Convolutional Neural Network method for the classification of flower images in development environment such as Notebook / Laptop and Smartphone.

3. Methodology
DE of Machine Learning will be divided into several parts, such as:

3.1. System Specification
The initial stage is the preparation process such as hardware and software, can be seen in the Table 1 below.

| Development Environment | System Spesification         |
|-------------------------|------------------------------|
| **Hardware**            |                             |
| Processor Intel(R) Core(TM) i3 M 370 | Processor Eight Core RAM 2 GB |
| Acer-Notebook           | Smartphone Realme-C2        |

**Table 1.** DE System Specification
Table 1. DE System Specification

| Development Environment | System Specification |
|--------------------------|----------------------|
| **Acer-Notebook**        | **Smartphone Realme-C2** |
| RAM 4 GB                 |                       |
| Operating System :       | Operating System :    |
| Linux Deepin 15.11 Desktop, 64-Bit | Android 9 ColorOS Version 6.0 |
| Software                 |                       |
| Interface :              | Interface :           |
| Jupyter Notebook, Python3, Keras | Google Colab Web, Python 3, Keras |
| Account Type             |                       |
| Personal                 | gmail                 |
| Network Technology       | localhost, internet   |
| internet                 |                       |

3.2. Dataset
Image classification using 17 Flower Categories dataset with contain 1360 images, which will be divided into a training and testing data [8].

3.3. Flowchart Process

![Flowchart Process of Image Classification](image)

In the flowchart above, experiment are carried out on Acer-Notebook and Smartphone Realme-C2 with several stages of the process, as follows:

3.3.1. Load image data, this process reads the image dataset stored to do the classification stage.
3.3.2. Pre-processing, the image will be resized.
3.3.3. Split dataset, devide the set of data into training and testing data.
3.3.4. Classification, Processing a classification using the convolution neural network.
3.3.5. Result, print the accuracy results in the classification are displayed.
The model can be seen in Figure 3. below

![Figure 3. a CNN Model](image)

### 3.4. Configuration

#### 3.4.1. Acer-Notebook

3.4.1.1. **Installation Operating System**, software that can communicate with hardware that can run the applications needed to run [9].

3.4.1.2. **Working Environment**, allows to have an isolated space in the work of each machine learning project, so it does not interfere with other projects [10].

3.4.1.3. **Python**, an interpreter programming language for object programming, has a simple syntax and supports modules and packages [11].

3.4.1.4. **Library**, required a package installer for python to start machine learning [12].

![Figure 4. Diagram of Notebook-Acer](image)

![Figure 5. Interface of Operating System](image)

3.4.2. **Smartphone Realme-C2**

3.4.2.1. **Web Browser**, application software to access information from web pages, can be utilized from desktops, notebooks, tablets and smartphones [13]

3.4.2.2. **Email**, a method of exchanging messages between humans and electronic devices [14].

3.4.2.3. **Google Colab**, a free cloud service that supports graphics processing units for python programming in machine learning [15].

![Figure 6. Interface of Jupyter Notebook](image)
4. Experiment and Result

After configuring the development environment, it is continued at the experimental stage to get the expected results from this paper.

4.1. Notebook-Acer

| Epoch | Train_Accuracy | Testing_Accuracy | Times (s) |
|-------|----------------|-----------------|----------|
| 10    | 0.39           | 0.44            | 162      |
| 20    | 0.47           | 0.42            | 321      |
| 50    | 0.62           | 0.55            | 806      |
| 70    | 0.63           | 0.55            | 1118     |
| 100   | 0.70           | 0.63            | 1616     |

The classification process in the table above shows the best accuracy results on epoch 100 where training data is 0.70 and the testing data is 0.63, the graph can be seen in Figure 9.
4.2. Smartphone Realme-C2

| Epoch | Train_Accuracy | Testing_Accuracy | Times (s) |
|-------|----------------|------------------|-----------|
| 10    | 0.39           | 0.45             | 88        |
| 20    | 0.49           | 0.51             | 175       |
| 50    | 0.61           | 0.58             | 450       |
| 70    | 0.63           | 0.62             | 600       |
| 100   | 0.70           | 0.66             | 895       |

The results of the best tests in the table above are on Epoch 100 where the training data is 0.70 and the testing data is 0.66, the graph can be seen in Figure 10.

![Training Loss and Accuracy](image)

**Figure 10.** Training accuracy of Smartphone Realme-C2

5. Conclusion

The result of configuration Development Environment for Machine Learning, the configuration process is easier to use on smartphone with classification process obtained the best level of accuracy on epoch 100 for training data and data testing, where Acer-Notebook is 0.70 and 0.63, where as for Smartphone Realme-C2 0.70 and 0.66. It is hoped that further research can be developed to obtain accuracy in predictions and image recognition. Hopefully this paper can be a reference source for machine learning studies.

References

[1] Brownlee J., 2018, Machine Learning Development Environment, Machine Learning Mastery, online: https://machinelearningmastery.com/machine-learning-development-environment/
[2] Julin F., Vision Based Facial Emotion Detection Using Deep Convolutional Neural Network, page 9, Sweden, 2019.
[3] Hurwitz J. and Kirsch D., Machine Learning for Dummies, pages 13-18, United State of America, 2018, online: https://www.ibm.com/downloads/cas/GB8ZMQZ3
[4] D. Lu and Q. Weng, A Survey of image classification methods and techniques fro improving classification performance, International Journal of Remote Sensing 28(5), 2007.
[5] Noi P. T. and Kappas M., Comparison of Random Forest, K-Nearest Neighbor and Support Vector Machine Classifier For Land Cover Classification Using Sentinel-2 Imagery, MDPI, 2018.
[6] Jonny HP, Rajagukguk D, Rajagukguk J. Computational Modelling Based on Modellus to Improve Students’ Critical Thinking on Mechanical Energy. In Journal of Physics: Conference Series 2020 Jan (Vol. 1428, No. 1, p. 012042). IOP Publishing.

[7] Verma R. and Ali J. Dr., Comparative Analysis of Advanced Classification Techniques For Multilingual Orc System, International Journal Of Scientific & Technology Research (IJRST), Vol. 8(11), 2019.

[8] Nilsback M. and Zisserman A., 17 Category Flower Dataset, online : http://www.robots.ox.ac.uk/~vgg/data/flowers/17/index.html

[9] TechTerms, Operating System, 2020, online : https://techterms.com/definition/operating_system

[10] Tagliaferri L., Digital Ocean, How To Install Python 3 and Set Up a Programming Environment on an Ubuntu 18.04 Server, 2020, online : https://www.digitalocean.com/community/tutorials/how-to-install-python-3-and-set-up-a-programming-environment-on-an-ubuntu-18-04-server

[11] Python, What is Python ? Executive Summary, 2020. online : https://www.python.org/doc/essays/blurb/

[12] Python Software Foundation, 2020. online : https://pypi.org/project/pip/

[13] Wikipedia, Web Browser, 2020. online : https://en.wikipedia.org/wiki/Web_browser

[14] Wikipedia, Email, 2020. online : https://en.wikipedia.org/wiki/Email

[15] Fuat, Google Colab Free GPU, 2018. online : https://medium.com/deep-learning-turkey/google-colab-free-gpu-tutorial-e113627b9f5d