Classification of Medicinal Plant using an Optimized Deep Learning Method

Dominador M. Acasamoso Jr.
Department of Information Technology
Polytechnic University of the Philippines - Ragay Campus
Ragay, Philippines
dmacasamosojr@pup.edu.ph

***Abstract*—Convolutional Neural Network or CNN is a popular approach in deep learning commonly used in Image Recognition because of its high and optimal potential in deep learning. In this study, CNN is used to classify common backyard medicinal plants that can be found in the Philippines which can be further integrated into mobile applications for use by people in remote locations where medicinal plants serves as first aid remedies. There are training images belonging to three categories of medicinal plants which are distributed in training, validation, and testing and are fed to a CNN with 3 Convolution Layer with Rectified Linear Function (ReLU) Activation, 3 Max Pooling Layer, Flatten, and Dense Layer with Softmax Activation, and Adam as the set optimizer. The CNN provides outstanding performance in training so as to validation which makes it a good model to implement in such a context.**

***Index Terms*—Deep Learning, Plant Classification, Medicinal Plants, CNN, Image Recognition**

I. Introduction

The power of computing has immersed and become useful in different aspects and disciplines. Years ago, the sole function of computers was just for mathematical computation tasks, but, with the brilliant innovations contributed by different researchers to improve the capabilities of machines, computing has become a multidisciplinary necessity. With advanced innovations, artificial intelligence is a field that most researchers and scientists are fascinated by. In agriculture and smart farming as cited in the study of [4], artificial intelligence has brought drastic improvement in harvests by optimization through guided decision-making based on the qualities and best practices that make increased produce. This increase in production is only one of the numerous reasons why machine learning is integrated even into agriculture and farming.

Another example of the vast applications of artificial intelligence is in the field of biology like what is applied in the study of [5] where Convolutional Neural Networks (CNN) for Gray Scale Normalization in image processing. In their study, they applied CNN for grayscale normalization to lessen the lighting differences of every image for classification purposes, hence, improving the speed of its classification process. In this study, CNN, an approach in Artificial Intelligence will be used in image classification for backyard medicinal plants so that people will be assisted in identifying herbal plants that can be taken for medicinal or dietary purposes. The researcher implemented CNN for image processing due to its optimal power in image recognition just like what has been attested by [8] which provides high accuracy for recognition. It is inevitable that due to the limited medical and healthcare supplies in remote areas, locals often rely on herbal medicinal plants most especially for minor illnesses and discomfort. With the help of this research, people can have a reference for herbal plants that they are not familiar with since they can submit a photo of an unfamiliar plant to the trained predictive model through CNN. In this study, the researcher included only five types of herbal medicinal plants for experiment purposes only to test the accuracy of implementing Convolutional Neural Networks in herbal plant classification.

II. Related Works

A. Image Recognition Approaches

Image recognition is a method where a machine tries to recognize images based on the information and patterns that can be found on training images fed to a predictive model powered by machine learning. During this time, biometric-based recognition is already in substitute for non-biometric-based recognition in a variety of systems, for example, face recognition [3]. In their study, there are algorithms that have been put into an experiment to identify their accuracy level compared to one another and these algorithms are Principal Component Analysis (PCA), Artificial Neural Network (ANN), and Signal Value Decomposition (SVD). Researchers did an image segmentation/classing from the image database before fitting it to the predictive model.

Their experiment shows that Artificial Neural Network shows the best accuracy by calculating the percentage of correctly identified images making it the best algorithm for the selected training images by the researchers. In this experiment, ANN has won the challenge in terms of accuracy scoring 85%.

Another approach in image recognition is Principal Component Analysis (PCA), a powerful and popular approach in image recognition and classes which have been implemented by [1]. The primary role of PCA is to lessen the dimension of the data with a large quantity of correlated variables. Their study includes the implementation of PCA for feature extraction in recognizing human faces. In the training and

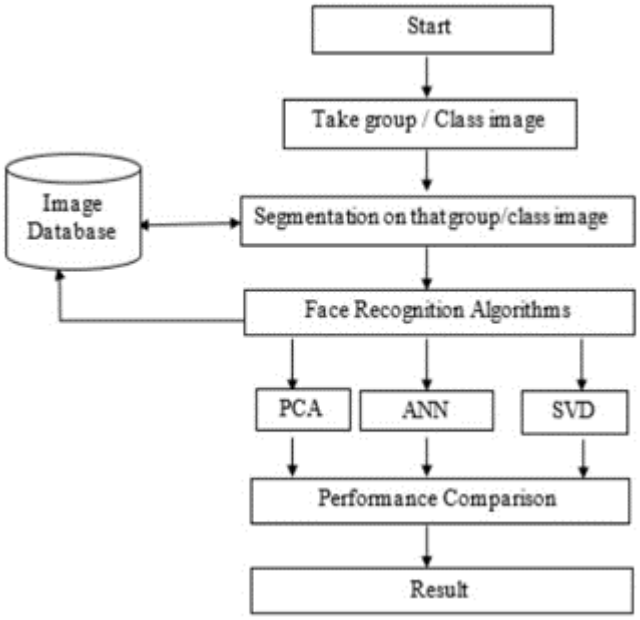

Fig. 1. Architectural Diagram retrieved from the study of (Fahad et al., 2017) [3]

testing phase, a total of ten images per individual has been acquired.

A total of three methods of classifier which are Euclidian Distance Method, Squared Euclidian Distance Method, and City-Block Distance Method which provide acceptable accuracy in recognizing human faces.

Support Vector Machine categorized as pattern detection algorithm is an approach under supervised learning used for regression and classification. It is primarily used for pattern detection for classification or prediction. In the same objective for image recognition, SVM has been studied and experimented by [14], whereas PCA and SVM have been combined for the recognition process. In this study, PCA was implemented for feature extraction and transformation that lessens the dimensions of feature space while the classification task is given to SVM which produced high accuracy because of the two algorithms combined.

Finally, the most popular approach in image recognition, the Convolutional Neural Networks (CNN) which is a type of Neural Network primarily used in image processing and recognition. This type of neural network use Convolution, Pooling, and Dense Layers for prediction or classification. There is a study about the implementation of CNN in Image Character Recognition where they train a model to do feature extraction from images through number of layers and detect patterns that can be utilized in recognition of unknown text or characters. By using this approach through an experiment, they gathered a test accuracy of 97.93% [9]. Figure 2 shows the structure of CNN they applied.

In this study, Convolutional Neural Network will be used as an approach to classify backyard medicinal plants because

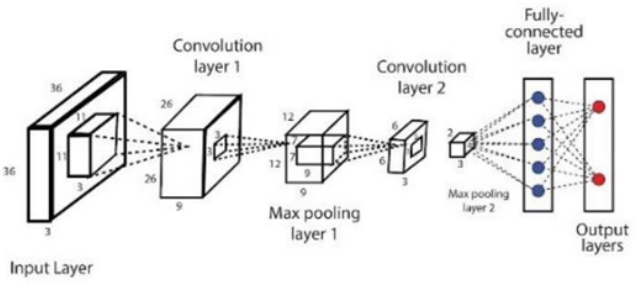

Fig. 2. CNN Structure as implemented in the study of (Narayan & Muthalagu, 2021) [9]

of its optimal potential and learning capabilities. It has been chosen because of the opportunities it can discover in the learning process.

B. Convolutional Neural Network

The convolutional neural network algorithm or CNN is a supervised learning approach which utilize learning samples to use the backpropagation method. It is an approach which have a concept of neural networks that has input, convolution, pooling, and output layers (Li, 2021). It is widely used in different purposes like image processing [6] , image recognition [11], networking [7] [12], and prediction [13]. In a comparative study conducted by [2] regarding the implementation of CNN, different models in face recognition have been put into test including AlexNet, VGG, and MobileNet. Their study shows that VGG19 performs best in providing maximum accuracy through CNN.

Similar study as to CNN application is the work of [10] about plant leaf disease detection but with the help of Learning Vector Quantization (LVQ). They gathered and pre-processed 500 images of several plants with symptoms in leaf disease for training. The CNN functions as feature extractor among the training images where leaf diseases are present and LVQ is fed to the CNN model output feature vector.

III. METHODS

As Artificial Intelligence broaden its capabilities like in biology and agriculture, different opportunities for studies have been unleashed. In the field of botany and biology, AI can be used through image recognition and classification, plant disease detection, and genome sequencing in bioinformatics. This research sought to provide a Deep Learning model that will classify backyard medicinal plants that can be found in the Philippines. CNN will be applied to training images gathered in raw by the researcher for the model to learn the classification model and CNN is applied because of its capabilities like feature extraction and patter recognition which is best for plants dataset. In addition to the objectives of this study, CNN will be put into test as to how it will perform in terms of accuracy in classifying medicinal plants in the Philippines.

Prior to classification, the following are the procedures conducted.

A. Preparation of Training Images

The dataset used in this study was completely gathered by the researcher. It is composed of raw images with various images of the medicinal plants belonging to three (3) categories. For the purpose of experimentation, a total of 1,199 images have been gathered that belongs to different class with different angles.

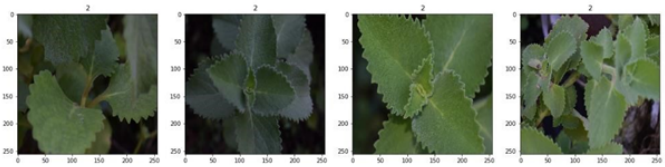

Fig. 3. Sample Training Images

B. Partition of Dataset for Training, Validation and Testing

The training images collected and pre-processed will be divided into two three allocations. The 70% of training images will be used for training, 20% will be for validation, and the remaining 10% will be used for testing.

C. Data Pre-processing

After the dataset has been divided for training, validation, and testing, the dataset has been divided into four batches which is essential for model training.

D. Model Training and Evaluation

Convolutional Neural Network or CNN is an approach in deep learning more particularly in image recognition, this method is also used in the experiment of classification in this study. There are three convolution layers in the CNN model implemented by the researcher with Rectified Linear Function (ReLU) that comes with Max Pooling Layers in different sizes, this is an optimization conducted to gain better classification performance without any compromise on efficiency. After convolution and max pooling layers, there is a flattening and dense layer for final classification. The model is evaluated using Accuracy Score which is available in the TensorFlow Library in Python.

IV. Simulation Results

The model developed has been put into an experiment by using Adam optimizer and to check the losses of in every step using 20 epochs. Results in Figure 5 below shows how the model performs in classification.

It can be observed that the model is gradually minimizing the loss both in training and validation as the number of epochs increase. The lowest loss is in training is about 35%, and in validation at 39% which is in epoch 20. On the opposite, it is also observed that model is learning progressively through 20 epochs by looking at the accuracy rate in training and validation. It can be observed that the accuracy improves gradually.

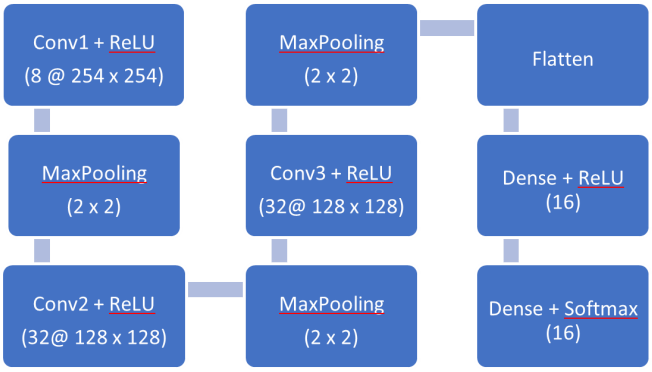

Fig. 4. Convolutional Neural Network Configuration for Classification

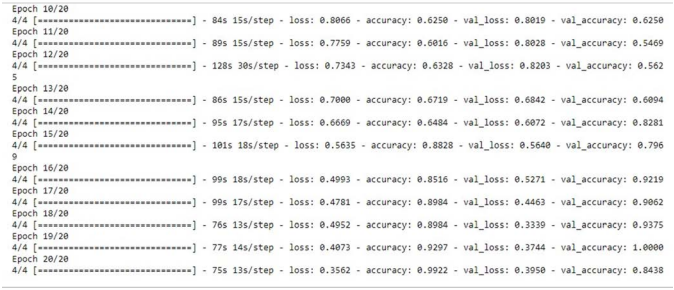

Fig. 5. Convolutional Neural Network Configuration for Classification

This is because of the combination of Conv3 and ReLU as optimizer. Detailed visualization of the model performance is shown in Figure 6 and Figure 7.

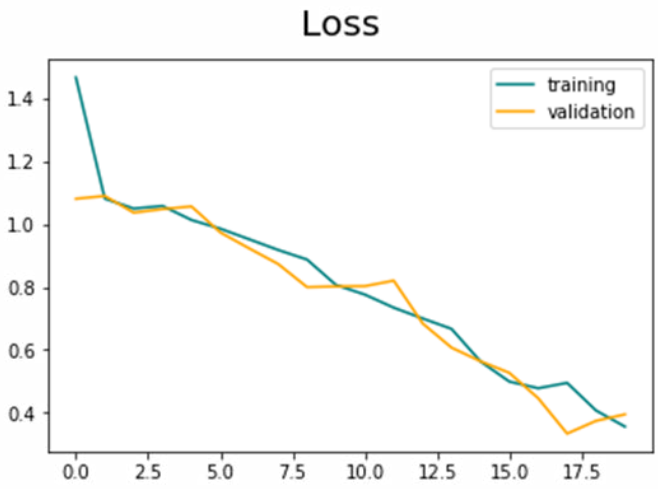

Fig. 6. Loss Value in Training and Validation

V. Conclusion

After thorough experiment in model training and testing, the model developed performs outstandingly with an accuracy rate of 99.22% for training, and 84.38% for validation. It can be attested that the model developed is capable of classifying backyard medicinal plants efficiently, given that it only has

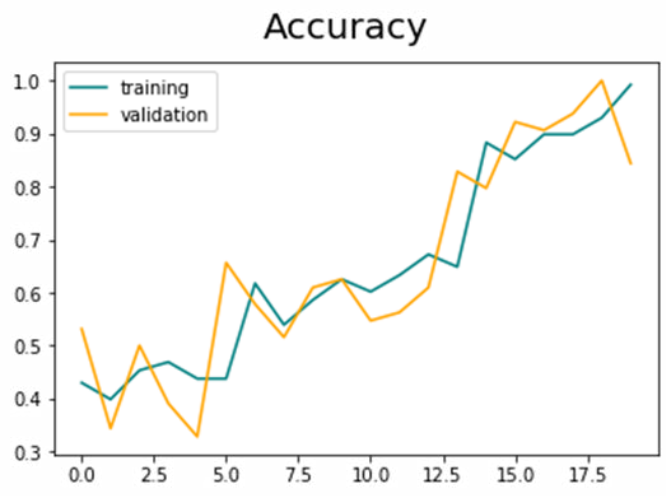

Fig. 7. Accuracy Value in Training and Validation

three categories to classify from. For future works, it is recommended to add more categories of medicinal plants to be fitted in the model and classify more variety of plants and geo-tagging to specify the locality of every medicinal plants that will be used by the model. Integration of the model in a mobile application can also be considered for portability and ease of use.

ACKNOWLEDGMENT

The author extends its sincerest gratitude to the Polytechnic University of the Philippines for its support on the conduct of this study.

REFERENCES

[1] Abbas, E. I., Safi, M. E., Rijab, K. S. (2017). Face recognition rate using different classifier methods based on PCA. International Conference on Current Research in Computer Science and Information Technology, ICCIT 2017, 37– 40. https://doi.org/10.1109/CRCSIT.2017.7965559

[2] Ahmed, T., Das, P., Ali, F., Mahmud, M.-F. (2020). A Comparative Study on Convolutional Neural Network Based Face Recognition; A Comparative Study on Convolutional Neural Network Based Face Recognition. In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT).

[3] Fahad, S., Rahman, S. ur, Khan, I., Haq, S. (2017). An Experimental Evaluation of Different Face Recognition Algorithms Using Closed Circuit Television Images. IEEE 2nd International Conference on Signal and Image Processing, 51–54. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=8124504

[4] Jaybhaye, N., Tatiya, P., Joshi, A., Kothari, S., Tapkir, J. (2022). Farming Guru: - Machine Learning Based Innovation for Smart Farming. Proceedings - 4th International Conference on Smart Systems and Inventive Technology, ICSSIT 2022, 848–851. https://doi.org/10.1109/ICSSIT53264.2022.9716287

[5] Kothari, D., Patel, M., Sharma, A. K. (2021). Implementation of Grey Scale Normalization in Machine Learning Artificial Intelligence for Bioinformatics using Convolutional Neural Networks. Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021, 1071–1074. https://doi.org/10.1109/ICICT50816.2021.9358549

[6] Lee, M., Kim, J., Yu, S. C. (2019). Robust 3D Shape Classification Method using Simulated Multi View Sonar Images and Convolutional Nueral Network. OCEANS 2019 - Marseille, OCEANS Marseille 2019, 2019-June. https://doi.org/10.1109/OCEANSE.2019.8867438

[7] Li, H. (2021). Computer network connection enhancement optimization algorithm based on convolutional neural network. Proceedings - 2021 International Conference on Networking, Communications and Information Technology, NetCIT 2021, 281–284. https://doi.org/10.1109/NetCIT54147.2021.00063

[8] Lou, G., Shi, H. (2020). Face image recognition based on convolutional neural network. China Communications, 17(2), 117–124. https://doi.org/10.23919/JCC.2020.02.010

[9] Narayan, A., Muthalagu, R. (2021, March 25). Image Character Recognition using Convolutional Neural Networks. Proceedings of 2021 IEEE 7th International Conference on Bio Signals, Images and Instrumentation, ICBSII 2021. https://doi.org/10.1109/ICBSII51839.2021.9445136

[10] Sardogan, M., Tuncer, A., Ozen, Y. (2018). Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm; Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm.

[11] Tun, N. M., Gavrilov, A., Tun, N. L., Paing, P. P. (n.d.). Face Identification System using Deep Learning Method at Low Lighting Resources. Retrieved December 2, 2022, from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=9039321

[12] Xin, R., Zhang, J., Shao, Y. (2020). Complex Network Classification with Convolutional Neural Network (Vol. 25, Issue 4). https://unstats.un.org/unsd/trade/sitcrev4.htm

[13] Yu, J., Peng, Y., Deng, Q. (2021, June 7). Remaining Useful Life Prediction Based on Multi-scale Residual Convolutional Network for Aero-engine. 2021 IEEE International Conference on Prognostics and Health Management, ICPHM 2021. https://doi.org/10.1109/ICPHM51084.2021.9486524

[14] Zhang, Y., Zhang, F., Guo, L. (2021). Face Recognition Based on Principal Component Analysis and Support Vector Machine Algorithms. Chinese Control Conference, CCC, 2021-July, 7452–7456. https://doi.org/10.23919/CCC52363.2021.9550727