

# Aquaponics Handbook: A Deep Convolutional Neural Network helper tool for Aquaponics

Lanz Vincent T. Vencer

Lyceum of the Philippines University – Manila, lanz.vencer@lpunetwork.edu.ph

**Abstract** - In recent times, the use of artificial intelligence(AI) in aquaculture and hydroponics is rapidly becoming the norm as more data is being acquired and generated. The advancements in Internet of Things(IoT) and Machine Learning amplified the data gathering process and extended the development of new approaches in image classification and pattern recognition. However, the field of aquaponics, which encompasses both mentioned agricultural sciences, also have the same potential especially when portrayed through an economic or scientific perspective. The main objective that this research is trying to achieve is to provide a helper tool for non-practitioners of aquaponics through a data-driven web application that uses the concepts of deep learning as a subset of AI. Two image datasets were acquired composing of fish and plant diseases which will be used for image classification. Fish and plants play a core element in every aquaponic system, and it is important to have a tool that would ensure accurate decisions when handling both these components. The MobileNetV2 Deep Convolutional Neural Network was used as a transfer learning base model for the image classification features of the website as it fits in many platforms and lightweight for mobile devices. The performance measure parameters, i.e., precision, recall, F1-score, and test accuracy were also calculated and monitored. The designed model achieved a 99.78% test accuracy for the fish classification feature and a 96.35% test accuracy for the plant-disease classification feature. The proposed research directly supports the intersection of artificial intelligence and aquaponics as it is beneficial in terms of biodiversity and economics.

*Index Terms* – Aquaponics, Convolutional Neural Networks, Fish Classification, MobileNetV2, Plant Disease Classification, Transfer Learning.

## INTRODUCTION

Aquaponics is a typical bio-integrated method of production that links the cyclic nature of aquaculture with hydroponics, which shares a sustainable agricultural system [1]. It merges fish cultivation(aquaculture) and plant farming(hydroponics), without the presence of soil at the base of the setup(although modern versions of aquaponics may sometimes include soil). The mechanism works by combining these two agricultural sciences to sustain the recirculating system by natural

biological cycles (nitrification) to supply nitrogen and reduce water inputs together with the non-renewable fertilizers [2]. Several systems, techniques, design, scales and nomenclature within aquaponics have been developed throughout the years and it has already been commercialized and used in many scientific studies [3].

Producing plants hydroponically and farming fish using aquaculture requires special methodologies to make sure that the respective systems are implemented and managed properly [2]. The complexity of each system models is proportional to the parameters being considered and monitored across varying scales and purpose of implementation. Problems may occur when these parameters do not align with the acceptable thresholds that were initially set or external factors like the presence of disease-causing agents for the crops and unwanted chemicals in the waters are seen upon the actual implementation. In the real world, it is very difficult for fish farmers to recommend the perfect fish species for aquaculture in a specific aquaponics system [4]. Land farmers may also be unfamiliar with the merger since growing crops through a land-based system is significantly different compared to an aquaponics environment. Whatever the case may be, investing into the production of an aquaponic system whether home-based or commercialized, is an expensive and risky venture if not equipped with the necessary guides and practices to not just make it work properly, but also sustainably maintaining its mechanisms.

Within the emergence of IoT based smart systems and the advancements in Artificial Intelligence, more innovative approaches would be developed to further advance the field of aquaponics and minimize the trade-offs enabling it to become a household or commercial venture. In this new era of science and technology, computer-aided diagnosis (CAD) systems are being used and developed for agricultural sciences along with the applications of Machine Learning and automation [5]. There are many applications of AI (Deep Learning and Machine Learning) in the aquaculture and hydroponics field and several journal articles have been published to provide empirical methods and sound practices that are based on gathered data and insights made through data mining [4]–[8]. Although both these fields share many common intersection with aquaponics, the implementation and scale are relatively different. In a general view, aquaponics encompasses both these fields and merges them

to create a feedback and two player system that enables survival of various living agents within the model [9].

Proper recommendations and guidelines for aquaponic farming are still lacking and requires more research investments focusing on possible production systems (type of fish  $\times$  type of plant  $\times$  densities  $\times$  filtration system  $\times$  hydroponic system  $\times$  aquaculture system) and different varieties of crops [2]. There are studies that integrate Internet of Things (IoT) devices within their aquaponic systems to gather data and make sense about the patterns that occur [10]. This intersection is commonly referred as “smart aquaponics” where data and information are core elements. However, measured data are oftentimes affected by complicated environmental factors that are usually nonlinear and various, which increases the complexity and difficulty of maintaining an accurate control system [11]. The massive amount of data in smart aquaponics imposes a variety of challenges, such as multiple sources, data inconsistencies, data volatilities, and structural complexities. To find a way to convert this data into insights and imperatives, various sophisticated Machine Learning (ML) algorithms are being used. For image data, for example, Deep Learning can be utilized because it involves multilevel data representations, from low to high levels, in which high-level features are built on the low-level features and carry rich semantic information that can be used to recognize and detect targets or objects in the image [11]. Deep Learning is commonly used together with other Machine learning models to makes sense of unstructured data like images or text. There are also other Machine Learning algorithms that can be used for structured or tabular datasets that can be utilized for regression or classification problems without the use of sophisticated Deep Learning techniques. By combining IoT, Data Mining, Artificial Intelligence and Data Science, it is possible to create a framework or standard that would aid beginning agriculturists or the common people to implement their own aquaponics system.

Transforming data from previously acquired analytics or creating an image classification deep learning algorithm are just some of the possibilities that can be implemented through Artificial Intelligence that would help non-experts of the field have a data-driven decisions for implementing their own aquaponics system. As of the writing of this paper, there are only handful of AI-based papers that focuses solely on aquaponics and to the best of my knowledge, this is the first paper to provide a handbook for aquaponics covering several preliminaries using both Machine Learning and Deep Learning from the gathered datasets.

## METHODOLOGY

The underlying step-by-step process followed when building the machine learning models are presented on Appendix 1. The processes within the workflows were thoroughly discussed and the models were specifically created for each respective feature present at the Aquaponics Handbook website.

### I. The dataset

The varying datasets were all gathered from Kaggle.com. The largescale fish dataset was adapted from the fish segmentation and classification study by Ulucan, et al. The dataset is composed of raw fish images separated through nine different classes namely: Black Sea Sprat, Gilt Head Beam, Horse Mackerel, Red Mullet, Red Sea Bream, Sea Bass, Shrimp, Striped Red Mullet, and Trout. Images were collected via 2 different cameras, Kodak Easyshare Z650 and Samsung ST60. Therefore, the resolution of the images are 2832 x 2128, 1024 x 768, respectively [12]. For each class, there are 1000 augmented images and their pair-wise augmented ground truths. However, this paper only considered the raw 1000 augmented images and completely disregarded the others. The sample images from randomly chosen classes are present in Figure I.



FIGURE I  
IMAGES OF FISH WITH LABES FROM THE DATASET

The image dataset for plant disease classification was taken from a publicly known dataset called Plant Village from Pennsylvania State University that is available in many machine learning repositories. In this data-set, 38 different classes of plant leaf and background images are available.

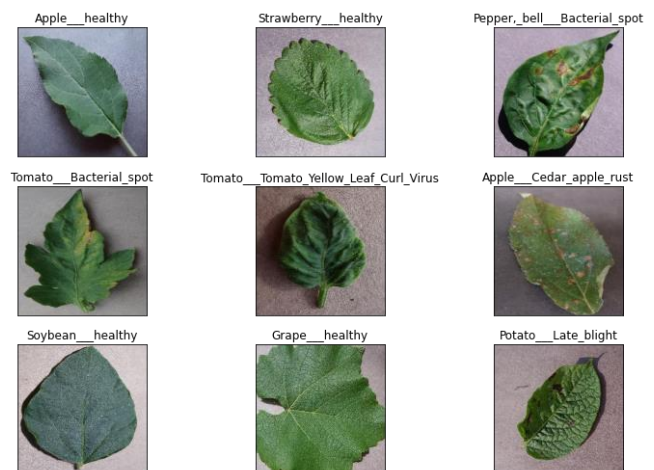


FIGURE II  
IMAGES OF PLANT WITH LABES FROM THE DATASET

These classes include 13 different plant types some with many available images of different plant diseases. In this study, the proposed deep learning model utilized the augmented version of the dataset by Geetharamani and Arun Pandian in their paper “Identification of plant leaf diseases using a nine-layer deep convolutional neural network” [13]. Overall, the data-set contained 54, 305 images. The sample images from randomly chosen classes are presented in Figure II.

## II. Experimental Setup

The Image classification features of the aquaponic handbook website utilized the pre-trained Deep Convolutional Neural Network(CNN) model called MobileNetV2 as base model for transfer-learning [14], [15]. This was implemented through the TensorFlow, Keras, and pandas libraries which use the Python programming language. The source code was written on Google Colaboratory which offers a free Graphics Processing Unit(GPU) that makes training and testing the models faster. The trained weights were then downloaded locally and integrated to a Streamlit backend which was used to create the website. The overall web application was then uploaded at GitHub and hosted at Streamlit sharing.

## III. Image Classification

The schematic in Figure III depicts a potential view for the image classification and analysis of the available features in the aquaponics handbook website. Initially, plant leaf disease and market fish images are collected and classified into

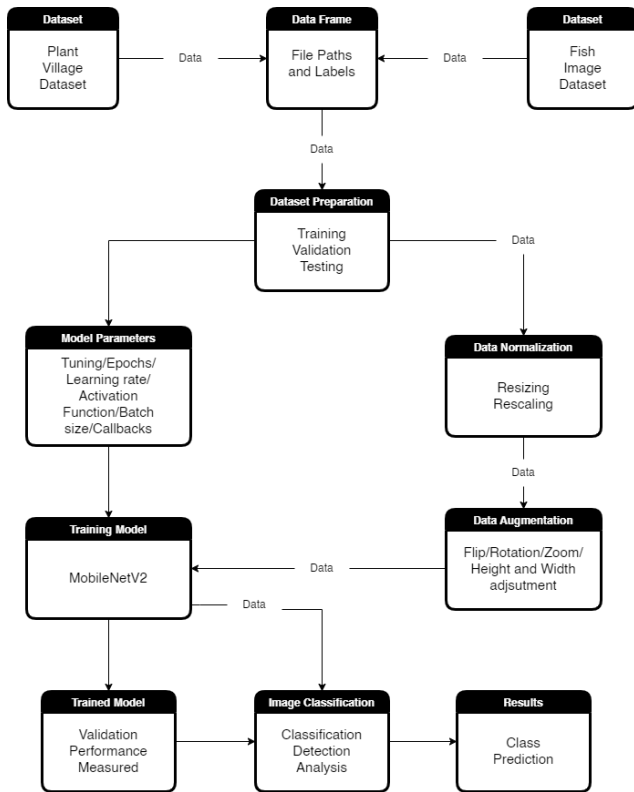


FIGURE III  
IMAGE CLASSIFICATION SCHEMATIC DIAGRAM

several categories. The image file paths were then arranged with its corresponding labels through a pandas Data frame to structure the data into two columns(file paths and labels) with the individual images as its rows. The data was then split into three categories namely: training, validation, and testing. Each category has a target size of 224 x 224, color mode is 'rgb', class mode is categorical, and the batch size is 32. The individual images were also preprocessed by scaling them to range from 0 to 1. Random flip, random rotation, random zoom, height adjustment, and width adjustment are some of the data augmentation techniques used. By using data augmentation methods, new sample photos are created from available photos to enhance and prepare the dataset [8]. In this case, only the images from the plant village dataset undergone data augmentation preprocessing. The photos are then used as input to the suggested approach for training the model in the following stage. After undergoing through transfer learning, the data from the base model is then connected to two Dense layers with 256 hidden fully connected(FC) layers and the activation function used was Rectified Linear Unit (ReLU). In the fish classification feature, only 128 fully connected layers are used within the two Dense layers. Both the FC layers are separated by a Dropout function with a value of 0.2 respectively [16]. The dropout function was only applied to the images in the plant village dataset since there is a tendency for the model to overfit in the training data. The newly trained architectural model will be used to anticipate previously unseen images. In this case, the model was evaluated on the test data from the previous split. Eventually the findings of the classifications are achieved.

## IV. Transfer Learning

The Deep Learning model's optimization and training is a computationally intensive and time-consuming operation. As mentioned earlier, a powerful graphics processing unit(GPU) is required for training the model, as well as large amounts of data. However, transfer learning, which is deployed in deep learning, solves these problems [8]. The pre-trained Convolutional Neural Network (CNN) used in transfer learning is optimized for one task and transfers knowledge to different modes [8], [17]. The images from the gathered datasets were compromised of different file sizes. Due to this, it was resized to a size of 224 x 224 with three channels to cater its 'rgb' type. The pre-trained CNN used is MobileNetV2 was used to find patterns within the input images and it's corresponding final layers have to be connected to a dense layer. The final layers before the softmax is a 11 x 11 Dense layer for the Fish classification, and a 38 x 38 Dense for the plant disease classification. The basic picture preparation is necessary for the transfer learning considerations with the data augmented images. Figure IV shows the whole Convolutional Neural Network architecture used to train the images.

## V. Fully Connected Layers

Fully connected layers are needed for extracted images classification features because of their special purpose. The

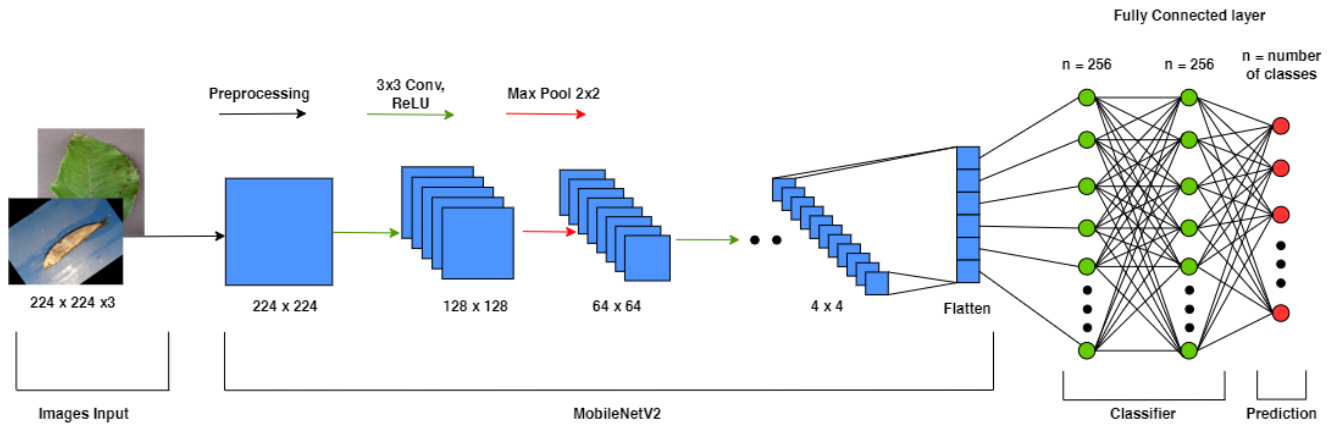


FIGURE IV  
CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE

softmax function predicts earlier extracted image attributes from preceding layers. Softmax is a multiclass classification activation function in the output layers. The neural network layer uses a multilayer perceptron model (MLP) as a classifier for two-class classification. The model with nonlinearity, which is introduced in the full vectors using rectified linear unit (RELU) activation [8].

#### VI. Model Evaluation

The test dataset gathered from the splitting of the dataset during the data preparation stage was used to evaluate the trained model. Classification metrics from scikit-learn library were also used. This includes the classification reports and confusion matrix. Throughout the training of the model, the training accuracy, training loss, validation accuracy, and validation loss were also tracked to determine if the model is overfitting or underfitting.

#### VII. Aquaponics Handbook Website

The saved weights from the trained models were downloaded locally. The image classification features were made through the Streamlit library that use the Python programming language. The finished website was then uploaded to GitHub pages together with the images, weights, and helper python files.

## RESULTS AND DISCUSSION

#### I. Metrics

The metrics used to evaluate how well the model performs are precision, recall, F1 score and the confusion matrix. Higher precision leads to less false positives, higher recall leads to less false negatives, f1-score is the combination of precision and recall, usually a good overall metric for classification model. When comparing predictions to truth labels, a confusion matrix can be used to see where the model gets confused.

#### 1. Precision(P)

The fraction of true positives (TP, correct predictions) from the total amount of relevant results, i.e., the sum of TP and false positives (FP). For multi-class classification problems, P is averaged among the classes [7]. Equation 1 shows the formula for computing the precision.

$$P = TP / (TP + FP) \quad (1)$$

#### 2. Recall(R)

The fraction of TP from the total amount of TP and false negatives (FN). For multi-class classification problems, R gets averaged among all the classes [7]. Equation 2 shows the formula for computing the recall.

$$R = TP / (TP + FN) \quad (2)$$

#### 3. F1 score(F1)

The harmonic mean of precision and recall. For multi-class classification problems, F1 gets averaged among all the classes [7]. It is mentioned as F-measure in [18]. Equation 3 shows the formula for computing the F1 score.

$$F1 = 2 * (TP * FP) / (TP + FP) \quad (3)$$

#### II. Fish Classification Evaluation

The images from the fish classification was trained using the proposed Deep Convolutional Neural Network. The model scored a test accuracy of 99.78% and a test loss of 0.00722. The timeline for the accuracy and loss of the training and validation datasets are presented on Figure V and Figure VI respectively. The corresponding classification reports are presented on Table I and the confusion matrix on Appendix A. Sample predictions on the test data are presented on Figure VII.

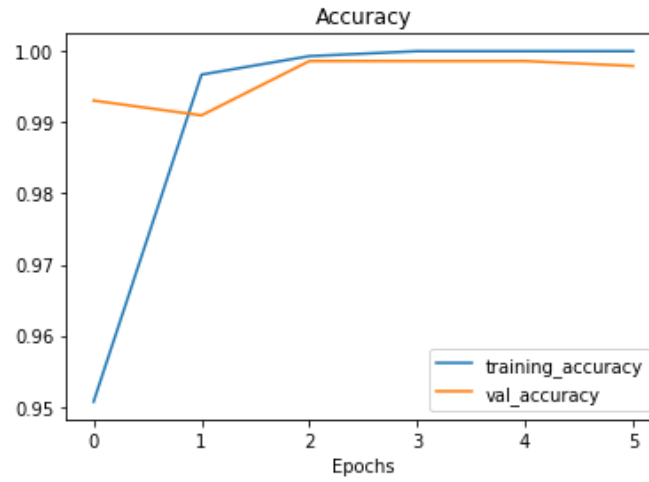


FIGURE V  
FISH CLASSIFICATION MODEL TRAINING AND VALIDATION ACCURACIES PER EPOCHS

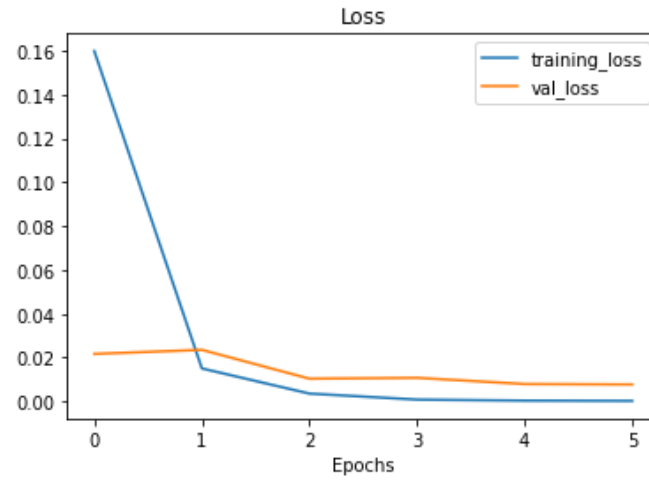


FIGURE VI  
FISH CLASSIFICATION MODEL TRAINING AND VALIDATION LOSS PER EPOCHS

TABLE I  
CLASSIFICATION REPORTS OF THE FISH DATASET AFTER EVALUATING THE MODEL

Class Name	Precision	Recall	F1-score	Support
Black Sea Sprat	1.00	1.00	1.00	206
Gilt-Head Bream	0.99	0.99	0.99	188
Horse Mackerel	1.00	1.00	1.00	202
Red Mullet	1.00	0.99	1.00	187
Red Sea Bream	0.99	1.00	1.00	196
Sea Bass	1.00	1.00	1.00	225
Shrimp	1.00	1.00	1.00	208
Striped-Red Mullet	0.99	0.99	0.99	183
Trout	1.00	1.00	1.00	205





FIGURE VII  
FISH CLASSIFICATION MODEL SAMPLE PREDICTIONS ON TEST DATA

### III. Plant Disease Classification Evaluation

The images from the plant disease classification was trained using the proposed Deep Convolutional Neural Network. The model scored a test accuracy of 96.35% and a test loss of 0.11303. The timeline for the accuracy and loss of the training and validation datasets are presented on Figure VIII and Figure IX respectively. The corresponding classification reports are presented on Table II and the confusion matrix on Appendix B. Sample predictions on the test data are presented on Figure X.

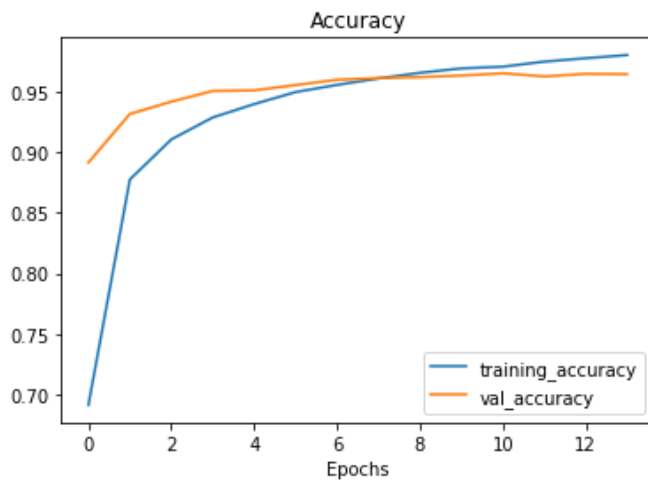


FIGURE VIII  
PLANT DISEASE CLASSIFICATION MODEL TRAINING AND  
VALIDATION ACCURACIES PER EPOCHS

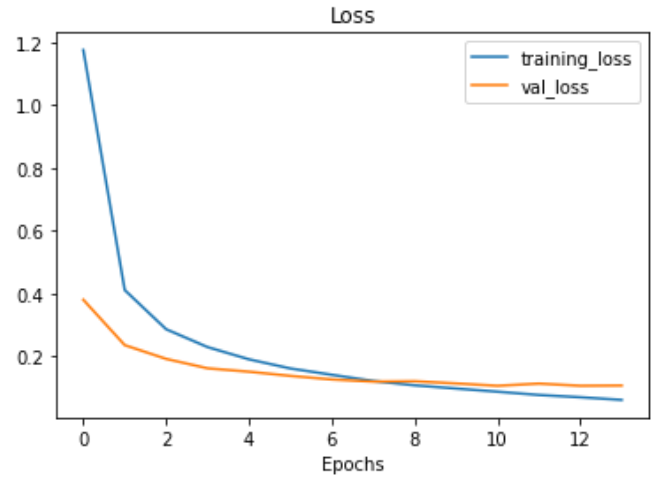


FIGURE IX  
PLANT DISEASE CLASSIFICATION MODEL TRAINING AND  
VALIDATION ACCURACIES PER EPOCHS

TABLE II  
CLASSIFICATION REPORTS OF THE PLANT VILLAGE DATASET AFTER EVALUATING THE MODEL

Class Name	Precision	Recall	F1-score	Support
Apple scab	0.97	0.81	0.89	124
Apple Black rot	0.99	0.98	0.99	128
Apple Cedar - apple rust	0.96	0.92	0.94	60
Apple healthy	0.97	1.00	0.98	336
Blueberry healthy	0.99	1.00	0.99	326
Cherry(including sour) Powdery mildew	1.00	0.99	0.99	227
Cherry(including sour) healthy	1.00	0.99	1.00	180
Corn(maize) Cercospora leaf spot - Gray leaf spot	0.83	0.81	0.82	110
Corn(maize) Common rust	0.98	1.00	0.99	255
Corn(maize) Northern Leaf Blight	0.90	0.90	0.90	201
Corn(maize) healthy	1.00	1.00	1.00	237
Grape Black rot	0.98	0.97	0.97	228
Grape Esca (Black Measles)	0.97	0.98	0.98	256
Grape Leaf blight (Isariopsis Leaf Spot)	1.00	0.97	0.98	232
Grape healthy	0.99	1.00	0.99	82
Orange Haunglongbing (Citrus greening)	1.00	1.00	1.00	1085
Peach Bacterial_spot	0.99	0.99	0.99	469
Peach healthy	1.00	0.97	0.98	59
Pepper, bell – Bacterial spot	0.95	0.94	0.95	212
Pepper, bell – healthy	0.98	0.99	0.98	302
Potato Early blight	0.96	0.98	0.97	202
Potato Late blight	0.93	0.95	0.94	204
Potato healthy	1.00	0.85	0.92	39
Raspberry healthy	0.98	1.00	0.99	89
Soybean healthy	0.99	1.00	0.99	980
Squash Powdery mildew	1.00	0.99	1.00	366
Strawberry Leaf scorch	0.97	1.00	0.98	235
Strawberry healthy	1.00	0.99	0.99	86
Tomato Bacterial spot	0.90	0.96	0.93	417
Tomato Early blight	0.85	0.67	0.75	189
Tomato Late blight	0.93	0.95	0.94	380
Tomato Leaf Mold	0.91	0.91	0.91	201
Tomato Septoria leaf spot	0.87	0.89	0.88	360
Tomato Spider mites Two-spotted spider mite	0.90	0.93	0.91	316
Tomato Target Spot	0.86	0.88	0.87	259
Tomato Yellow Leaf Curl Virus	0.99	0.99	0.99	1028
Tomato mosaic virus	0.94	0.82	0.88	73



FIGURE X  
PLANT DISEASE CLASSIFICATION MODEL SAMPLE PREDICTIONS ON TEST DATA

#### IV. Website deployment

The website was deployed through the Streamlit sharing feature associated with GitHub. Figure XI shows the interface of the website.

### CONCLUSION

In this paper, two datasets were collected, prepared, analyzed, and implemented as a feature for an Aquaponics helper website. These datasets are fish images and plant disease images taken from previous similar studies available in public repositories. The techniques of data augmentation, dataset pre-processing, training, and testing are applied to the convolutional neural network-based MobileNetV2 model. The proposed model is built and tested to improve the performance measured and compared it throughout different parameters. The evaluation metrics like precision, recall, and f1-score were utilized to evaluate the model as the hyperparameters are adjusted to determine the best model for the image classification. After the training and evaluation, two classifier weights were eventually achieved. The first model which was trained with the fish images achieved a test accuracy of 99.78% and loss of 0.00722. On the other hand, the second model which was trained with the plant disease images achieved a test accuracy of 96.35% and a loss of 0.11303. Always improving the performance of our models for fish and plant disease classification and analysis is a critical step, but our model achieved a satisfactory performance, which hopefully will support aquaponics

practitioners. The major focus of this research is to provide a helper tool that would help non-experts in the field aquaponics to have a guide for data driven decisions about creating and maintaining an aquaponic system. The collection and preparation of genuine datasets and its applications to aquaponics to add more features to the website is a future target. Using more sophisticated deep learning models and gathering more datasets for better classification and analysis are anticipated. My work encourages and stimulates aquaponics practitioners and hobbyist alike, which ultimately helps raise their respective incomes and promotes a bio-friendly way of agriculture that could be implemented by anyone.

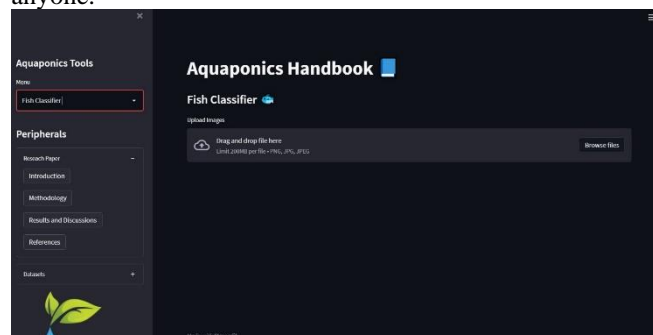


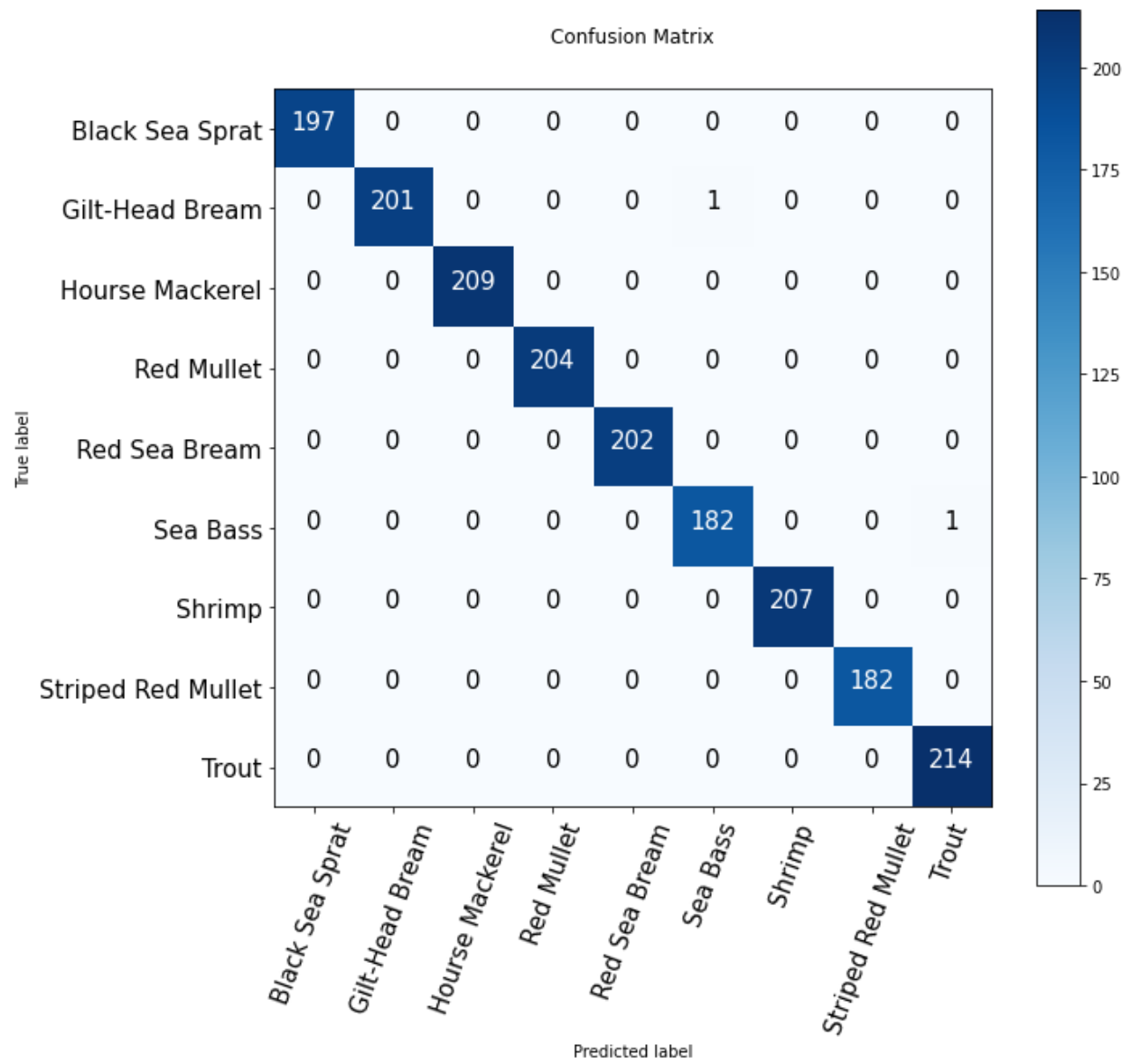
FIGURE XI  
AQUAPONICS HANDBOOK WEBSITE INTERFACE



## REFERENCES

- [1] P. G. Panigrahi, S. Panda, and S. N. Padhi, "Aquaponics: An innovative approach of symbiotic farming," *IJB*, vol. 5, no. 09, p. 4808, Aug. 2016, doi: 10.21746/ijbio.2016.09.005.
- [2] R. V. Tyson, "Aquaponics—Sustainable Vegetable and Fish Co-Production," p. 5.
- [3] H. W. Palm *et al.*, "Towards commercial aquaponics: a review of systems, designs, scales and nomenclature," *Aquacult Int*, vol. 26, no. 3, pp. 813–842, Jun. 2018, doi: 10.1007/s10499-018-0249-z.
- [4] Md. M. Islam, M. A. Kashem, and J. Uddin, "Fish survival prediction in an aquatic environment using random forest model," *IJ-AI*, vol. 10, no. 3, p. 614, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp614-622.
- [5] V. K. Shrivastava, M. K. Pradhan, S. Minz, and M. P. Thakur, "RICE PLANT DISEASE CLASSIFICATION USING TRANSFER LEARNING OF DEEP CONVOLUTION NEURAL NETWORK," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. XLII-3/W6, pp. 631–635, Jul. 2019, doi: 10.5194/isprs-archives-XLII-3-W6-631-2019.
- [6] S. Dilmi and M. Ladjal, "A novel approach for water quality classification based on the integration of deep learning and feature extraction techniques," *Chemometrics and Intelligent Laboratory Systems*, vol. 214, p. 104329, Jul. 2021, doi: 10.1016/j.chemolab.2021.104329.
- [7] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, Apr. 2018, doi: 10.1016/j.compag.2018.02.016.
- [8] A. S. Paymode and V. B. Malode, "Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG," *Artificial Intelligence in Agriculture*, vol. 6, pp. 23–33, 2022, doi: 10.1016/j.aiia.2021.12.002.
- [9] D. Karimanzira, C. Na, M. Hong, and Y. Wei, "Intelligent Information Management in Aquaponics to Increase Mutual Benefits," *IIM*, vol. 13, no. 01, pp. 50–69, 2021, doi: 10.4236/iim.2021.131003.
- [10] C. S. Arvind, R. Jyothi, K. Kaushal, G. Girish, R. Saurav, and G. Chetankumar, "Edge Computing Based Smart Aquaponics Monitoring System Using Deep Learning in IoT Environment," in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, Canberra, ACT, Australia, Dec. 2020, pp. 1485–1491, doi: 10.1109/SSCI47803.2020.9308395.
- [11] X. Yang, S. Zhang, J. Liu, Q. Gao, S. Dong, and C. Zhou, "Deep learning for smart fish farming: applications, opportunities and challenges," *Rev. Aquacult.*, vol. 13, no. 1, pp. 66–90, Jan. 2021, doi: 10.1111/raq.12464.
- [12] O. Ulucan, D. Karakaya, and M. Turkan, "A Large-Scale Dataset for Fish Segmentation and Classification," in *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Istanbul, Turkey, Oct. 2020, pp. 1–5, doi: 10.1109/ASYU50717.2020.9259867.
- [13] G. Geetharamani and J. Arun Pandian, "Identification of plant leaf diseases using a nine-layer deep convolutional neural network," *Computers & Electrical Engineering*, vol. 76, pp. 323–338, Jun. 2019, doi: 10.1016/j.compeleceng.2019.04.011.
- [14] Y. LeCun, Y. Bengio, and T. B. Laboratories, "Convolutional Networks for Images, Speech, and Time-Series," p. 15.
- [15] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *arXiv:1801.04381 [cs]*, Mar. 2019, Accessed: May 07, 2022. [Online]. Available: <http://arxiv.org/abs/1801.04381>
- [16] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," p. 30.
- [17] P. Nevavuori, N. Narra, and T. Lipping, "Crop yield prediction with deep convolutional neural networks," *Computers and Electronics in Agriculture*, vol. 163, p. 104859, Aug. 2019, doi: 10.1016/j.compag.2019.104859.
- [18] D. H. T. Minh *et al.*, "Deep Recurrent Neural Networks for mapping winter vegetation quality coverage via multi-temporal SAR Sentinel-1," *arXiv:1708.03694 [cs]*, Aug. 2017, Accessed: May 08, 2022. [Online]. Available: <http://arxiv.org/abs/1708.03694>

# APPENDIX A FISH CLASSIFICATION MODEL CONFUSION MATRIX



### PLANT DISEASE CLASSIFICATION MODEL CONFUSION MATRIX

