

Crop Leaf Disease Detection Using Convolution Neural Network

Chaskar Prasad¹, Dhage Shubham², Vaidya Shubham³, Chavan Abhijeet⁴, Dr. Swati Bhavsar⁵

Student Department of Computer Engineering^{1,2,3,4}

Professor, Department of Computer Engineering⁵

Matoshri College of Engineering and Research Centre, Nashik

Savitribai Phule Pune University, Nashik, India

Abstract: It has become inevitable in recent years to use technology to promote agricultural awareness. Seasonal climatic conditions are also altering in respect to important assets like land, water, and air, causing food insecurity. Crop production rates are always falling short of meeting demand in this climate, necessitating the development of a smart system capable of dealing with the problem of diminishing agricultural output. To solve this problem, we propose a system that provides accurate plant disease diagnostics and crop selection based on economic and environmental parameters to maximise production for farmers, therefore assisting in meeting the country's growing food supply need. The proposed technique makes use of machine learning to provide predictions. In order to maximise production for farmers, the system will give many types of disease detection and crop selection based on soil weather factors suited for the crop. Crop production is forecasted using parameters such as rainfall, temperature, area (in hectares), season, and so on. The approach also helps in identifying whether a specific fertiliser is suitable for use. Crop disease detection and yield prediction are major agricultural concerns. Every farmer wants to know how much yield will be produced and whether it will meet their expectations. Previously, yield prediction was calculated by looking at a farmer's previous experience with a certain crop. Weather, pests, and harvest operation planning all have a large impact on agricultural productivity and crop health. Accurate disease and crop yield history is crucial for making agricultural risk management decisions..

Keywords: Image Classification, Crop Recommendation, User authentication, Classification

I. INTRODUCTION

Within the scope of this investigation, the application of convolutional neural networks and deep learning algorithms will take place with the intention of determining the presence of disease in plant leaves (CNN). Learning by machine is a significant tactic that may be used to give answers to this challenge that are not only practically applicable but also effective. This method has the potential to be used. India has the second-highest position internationally in terms of the quantity of agricultural products it produces. Agriculture and the industries that are closely related with it, such as forestry and fisheries, accounted for 16.6 percent of GDP in 2009 and employed about half of the whole population. Other closely associated industries include horticulture. Animal husbandry and horticulture are two more industries that are very closely tied to one another. It is expected that this declining trend would continue for the foreseeable future in India's agricultural sector's contribution to the country's GDP, which has been continuously decreasing. The harvest of plant crops is affected by a broad variety of variables, including climatic, geographical, organic, political, and economic difficulties. These elements all interact with one another to create a complex system. These many aspects all have an effect on one another. When farmers do not have a solid grasp of how much their harvests are worth on the market, it is

extremely difficult for them to raise more than one crop at the same time. According to Wikipedia, the rate of farmer suicides in India fluctuated from 1.4 to 1.8 per 100,000 of the overall population during the ten-year period that started in 2005 and ended in 2005. This period of time began in 2005 and completed in 2005. 2005 was the year that marked the conclusion of this historical period. The year 1995 served as the starting point for this time period. In 2014, 5650 farmers were found to have taken their own lives; the next year, however, that number surpassed 8000. In recent years, there has been no other option for expanding one's knowledge of agriculture other than to make use of the many technological instruments that are available. The meteorological conditions throughout the seasons are also fluctuating, which is producing insecurity with regard to essential resources such as land, water, and air, which in turn causes insecurity about the supply of food. As a result of the fact that crop production rates in this environment are never sufficient to meet demand, it is essential to devise a complex system that is capable of resolving the issue of diminishing agricultural output. This is because of the fact that crop production rates in this environment are never enough to meet demand. This might be done by concentrating on the following, among other things: We have come up with a solution to this issue in the form of a system that not only allows for the precise diagnosis of plant diseases but also allows for the selection of crops based on features that are significant to both the economics and the environment. The purpose of this method is to maximise agricultural production to the greatest extent possible, which, in turn, will assist the country in getting one step closer to meeting the ever-increasing need for food supply. The use of machine learning, which has been suggested as a method of doing things, is used to provide predictions and support their accuracy. Farmers will be able to utilise the system to recognise a broad range of illnesses and choose crops depending on the parameters of the soil and weather that are ideal for the crop. Because of this, the output that the farmers are able to achieve their maximum potential for will be achievable. It is possible to make reasonable estimations regarding crop yields by taking into account a number of different elements, including the amount of precipitation that actually falls, the temperature, the area (measured in hectares), and the season. This method is also useful in determining whether or not a certain kind of fertiliser need to be administered to a given region, which is an important consideration. Agriculture is one of the most important contributors to India's gross domestic product, which is a reflection of the fact that India is mostly an agricultural economy (GDP). The growing prevalence of suicide among farmers, which may be connected to crop failure, served as the inspiration for the formation of this effort. This initiative was established in order to address this problem. The possibility of farmers killing themselves by taking their own life is one of the goals of the programme. The agricultural industry is faced with a significant challenge in the shape of a climate that is always shifting, in addition to other aspects of the environment that are undergoing change. Learning by machine is a significant tactic that may be used to give answers to this challenge that are not only practically applicable but also effective. This method has the potential to be used. A prognosis of the yield of the crop that takes into consideration historical information such as the condition of the soil, the climate, and the quantity of precipitation, in addition to the yield of the crop in previous years. The use of a method for machine learning allowed us to realise this objective with flying colours. We investigated and compared a wide variety of machine learning techniques, such as ANN, K Nearest Neighbor, Random Forest, SVM, and Linear Regression, among others. In the end, we decided to go with the Random Forest Algorithm due to the fact that it had an accuracy rate of 95 percent. It is becoming increasingly important to conduct research that is relevant to the expansion of environmentally friendly agriculture in light of recent developments in agricultural technology as well as the implementation of artificial intelligence in the process of diagnosing plant diseases. This is because it is becoming increasingly difficult to diagnose plant diseases without the assistance of artificial intelligence. The quality and quantity of potatoes are significantly impacted by a number of diseases, including early and late blight, and the manual interpretation of these leaf diseases is a labor-intensive process that takes a significant amount of time and can be challenging. Early and late blight are two examples of these diseases. The early blight and the late blight are both types of this disease. An online tool was developed during the course of this study with the purpose of forecasting crop yields

by taking into account a broad spectrum of soil properties as well as diseases that might potentially damage the leaves of potato and rice plants. This study was carried out in order to investigate the relationship between soil properties and crop yields. In addition to this, it reveals whether or not the usage of a certain kind of fertiliser is authorised. Within the scope of this investigation, the application of convolutional neural networks and deep learning algorithms will take place with the intention of determining the presence of disease in plant leaves (CNN). Learning by machine is a significant tactic that may be used to give answers to this challenge that are not only practically applicable but also effective. This method has the potential to be used. A prognosis of the yield of the crop that takes into consideration historical information such as the condition of the soil, the climate, and the quantity of precipitation, in addition to the yield of the crop in previous years. The use of a method for machine learning allowed us to realise this objective with flying colours. It has been shown that the strategies of deep learning are particularly adaptive to a broad spectrum of agricultural picture datasets. [Citation needed] [Citation needed] [Further citation is required] Because there are so few labelled medical pictures accessible, in order to perform an accurate computer-assisted diagnosis, a substantial number of data augmentation (DA) approaches are required. Among these methods is the practise of modifying the geometric structure or intensity of the original pictures. When added to the data that were used for training, these data are helpful in addressing the restricted medical picture collection that was assembled from a variety of various sources. These data may be added to the data that were used for training. One of the many distinct types of DA approaches that are accessible to users is known as a Generative Adversarial Network (GANs). A GAN that has been trained on pictures has the potential to produce new images that incorporate a range of real-world qualities and give human viewers the impression that they are looking at real-world settings. As a direct result of this, the primary focus of this investigation is centred on the application of deep convolution GANs in an effort to solve the issue of datasets that do not contain sufficient labels. A visual Turing test was carried out with the assistance of trained medical professionals in order to verify the model that was presented.

II. OBJECTIVE OF PROJECT

In order to achieve the system's goals, we must first minimise the amount of physical labour required. to include it in the computer so that it not only captures photographs but also processes them. Leaf disease may also be detected. Detection of plant leaf diseases at an early stage To verify the correctness of the ML model's predictions.

Project Idea:

Farmers Early on, one of the hardest things to do is figure out what leaf diseases are. Pathogens are things that can make you sick. Pests and diseases are most likely to be found on the plant's leaves and stems. For effective crop cultivation, it is important to know how to identify plants, leaves, symptoms, pests, and diseases, as well as the percentage of pests and diseases and the signs of an attack by a pest or disease. What can I do to fix my plant's problem once I know what's wrong with it? Even if it's too late to help the plant in question, a good diagnosis could help stop the problem from happening to other plants or in the future. To come up with effective ways to control diseases, it is important to know what causes them and how they spread. Because of this, a plant pathologist must learn a lot about diagnosis as part of their education. Plants may die if disease management methods are used before the disease has been correctly diagnosed and the disease-causing agent has been found. Think about many different things, both natural and cultural. Learn to tell the difference between a sickness and what caused it. Before coming up with a final diagnosis, do more lab tests to rule out all but a few of the possible causes, or Recognize the symptoms and signs that help you figure out who the person is. It might be hard to describe a person's specific symptoms in a clear way. Because of this, it might be hard or even impossible to figure out what's wrong with a plant based on a phone call from a worried person. To see if this is true, take a sick plant and have three different people write down the symptoms they see.

Software Requirement

Hardware Requirements

Processor A computer's processor is an electrical circuit incorporated into the motherboard. An operating system instructs a processor to execute arithmetic, logical, input/output (I/O), and other fundamental commands (OS). Almost all other processes rely on a processor's actions. We suggest using a CPU with at least 1 GHz of processing power, however we would like S2GHz or more.

- **Wi-Fi** In wireless local area networking (WLAN), Wi-Fi refers to a group of radio technologies based on the IEEE 802.11 family of standards. Wi-Fi technology may be used by a wide variety of devices, including desktops and laptops, smartphones and tablets, TVs and printers, digital music players, digital cameras, and even automobiles and drones. Devices that are compatible with one other may share a Wi-Fi access point, as well as Ethernet connections, to access the Internet. One of these access points (or hotspots) may reach a distance of around 20 metres (66 feet) inside, and a larger distance outside. One room with walls that restrict radio waves may have hotspot coverage that extends to many square kilometres, if numerous overlapping access points are used.

- **Hard Drive** An electro-mechanical data storage device, known as a hard drive, employs magnetic storage to store and retrieve digital information using one or more rigid spinning discs covered with magnetic material. Data is read and written to the platters using magnetic heads attached to a moving actuator arm, which move in tandem with the platters. As the name implies, data is accessible randomly, which means that blocks of information may be stored or retrieved randomly, rather than sequentially. Data may be kept on a hard drive even after the computer is turned off since it's non-volatile. For the suggested system, a memory capacity of at least 32 GB is recommended.

- **Memory (RAM)** The data and machine code now being utilised are stored in this kind of computer data storage. Allows data objects to be read or written in almost the same amount of time regardless of where they are physically located in a random access memory device Random-access memory (RAM) is now often found in integrated circuits (ICs). Although nonvolatile RAM (NVRAM) has been created, RAM is often linked with volatile forms of memory (such as DRAM modules), where recorded information is lost if power is disconnected. The suggested system should have at least 2 GB of RAM.

Requirements for software The MySQL database server MySQL is a free and open-source database management system for relational databases (RDBMS). A well-known SQL data language is used in a standard version that is fast and performs well even with large datasets. Language: Python Python is a widely used programming language. Developed by Guido van Rossum, it was launched in 1991.:

(a) web development (server-side).

(b) software development. MCERC 7 Crop Leaf Disease Detection Using Convolution Neural Network

- **Operating System:** As with Windows 8 and above, Windows is the most extensively used operating system for desktop and laptop computers. Windows pre-production runs on x86-based PCs developed by Microsoft. Windows offers Graphical User Interface and desktop Environment in which programme presented in resizable, movable windows on screen.

Notebook for coding in Jupyter Using the Jupyter Notebook, you may create and share documents with live code, equations, visualisations and text. The Jupyter Notebook is an open source online application. Many additional kernels are available, including the IPython kernel, which is included with Jupyter, although there are presently more than 100 available.

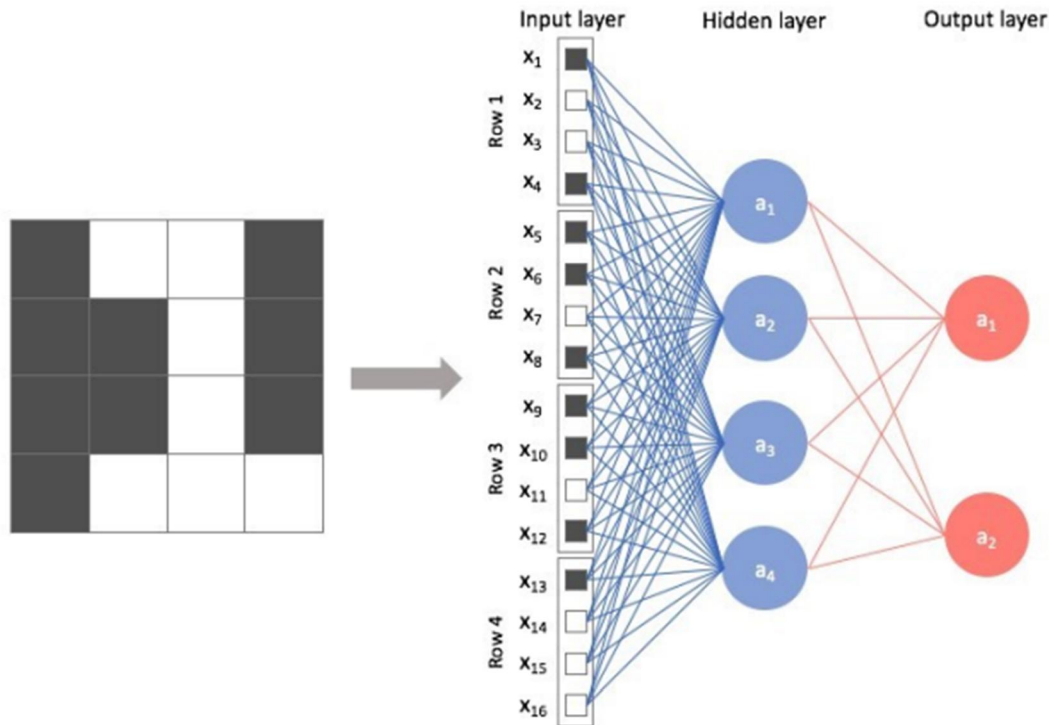
Affirmatively Supportive In addition to superb code completion and inspection with a powerful debugger and support for web development and multiple frameworks, VsCode is the most popular Python IDE. With the help of VsCode's intelligent code editor, Python programmers may produce better code. Color schemes, automatic indents on new lines, coding style preferences, and context-aware code completion are only some of the features of the editor.

As of this writing, the Django framework is supported. Fast development and clean, pragmatic design are encouraged by Django, a Python web framework. Because it was created by seasoned programmers, it alleviates a lot of the burden of web development so you can concentrate on building your app instead. It's open source and free.

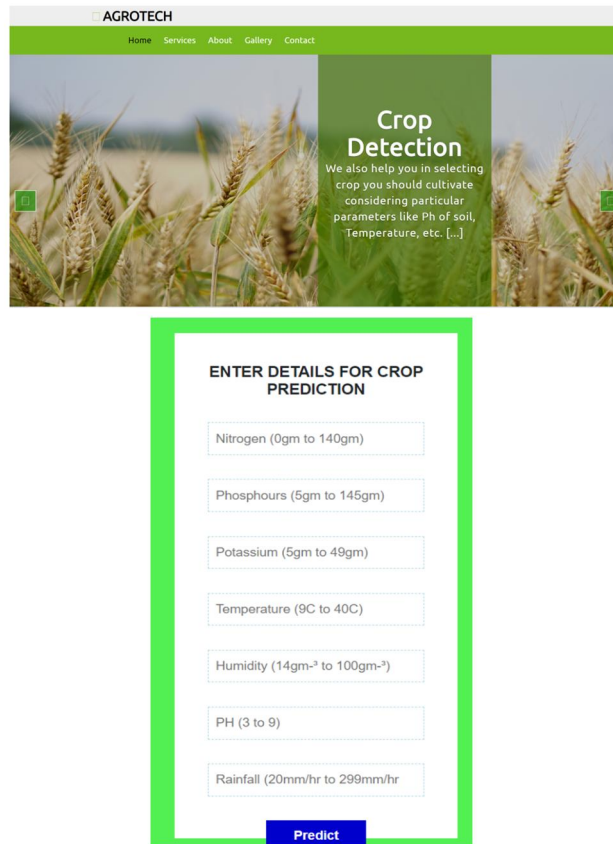
III. RESULT AND DISCUSSION

The concept of a graphical user interface, more often referred to as a GUI, is something that is introduced to us for the very first time in this chapter. This is the very first and most important thing that this chapter does for us, therefore it's important to pay attention to it. This idea outlines the many different ways in which users will be involved in the development of the project, in addition to the many different methods in which the first module will be constructed. In addition to this, it describes in great depth the procedure that will be followed in order to deliver the project to the users after it has been finished. The fully connected layer is always the very last layer that is included in a neural network, regardless of whether or not the network also contains convolutional layers. This is the case even when the network has both types of layers. Even if there are a lot of convolutional layers in the network, this remains true. This is the case despite the fact that the network may have a different number of convolutional layers. This is because the layer that has retained all of its connections has also preserved all of the connections that were present in the layers that came before it. This is because the layer that has preserved all of its connections has preserved all of its connections. As a direct consequence of the aforementioned characteristic, this specific layer is not exclusive to a CNN; rather, it is capable of being used in a broad variety of other types of neural networks as well. This is because of the fact that it is not exclusive to a CNN. This is due to the fact that it is not very noteworthy to a network like CNN. This layer takes as its input a vector, which is then used in the creation of another vector, which is then output by this layer. The operation of multiplication will be carried out by this layer as its responsibility. As a result of the fact that this layer can generate vectors at both its input and output locations, it can perform the functions of both receiving and transmitting vectors. In order to do this, first the input values are subjected to a linear combination, and then maybe an activation function is applied to them. This is done in order to get the desired result. This action is taken in order to get the wished-for outcome. The desired result can only be achieved by carrying out this particular course of action. Only by following this precise plan of action will it be possible to accomplish the outcome that has been envisioned. The desired end result can be achieved, but only if the specific steps outlined in this plan of action are carried out in the correct order. The last layer, which links everything together, generates a vector with a size of N , where N is the number of categories that need to be allocated to the photos that need to be categorised. This layer is fully connected. The size of the resulting vector is N , which is due to the fact that this layer has full connectivity throughout. Each component of the vector in the vector space provides an indicator of the degree to which an image is likely to correspond to a certain category. This indicator may be used to help determine whether or not an image belongs to the category. Utilizing this information, it is feasible to determine which categories a photo falls under. This indication, which is derived from the things that the vector gives, has its origin in the vector itself. When computing the probabilities, the fully-connected layer multiplies each input element by weight, adds up the total, and then applies an activation function (logistic if $N=2$, soft-max if $N>2$) to the sum. If there are two input elements, the fully-connected layer uses the soft-max activation function. The soft-max activation function is used by the fully-connected layer if there are two items serving as inputs. When there are two different things functioning as inputs, the fully-connected layer will employ the soft-max activation function to activate itself. The fully-connected layer will use the soft-max activation function in order to activate itself if there are two distinct items operating as inputs. [Case in point:] [Case in point:] If there are two separate things functioning as inputs, the fully-connected layer will employ the soft-max activation function in order to activate itself. This is because it is more efficient. [Here's a good example:] [Here's a good example:] When N equals 2, the activation function changes in such a way that it converts into a soft-max function, as opposed to the hard-max function that it had been in the past. This is in contrast to

the previous state of the activation function, which was a hard-max function. This is a substantial departure from the original. When we say that the input vector is multiplied by itself, we are essentially expressing the same thing that we say when we say that the input vector is multiplied by the weights matrix. When we say that the weights matrix is multiplied by the input vector, we are essentially saying the same thing. When we state that the weights matrix is multiplied by the input vector, we are essentially expressing the same thing as when we say that the matrix is multiplied. When we describe the same idea by saying that the input vector is multiplied by the weights matrix, we are, in fact, explaining the same thing. This is because both explanations refer to the same thing. This is because when we say that the input vector is multiplied by itself, we are saying the same thing as when we say that the weights matrix is multiplied by the This is because when we say that the input vector is multiplied by itself, we are saying the same thing as when we say that the weights matrix is multiplied by the This is because when we say that the input vector is multiplied by itself, we are saying the same thing as when we say that the weights matrix is multiplied by When we speak about anything being "fully-connected," what we mean is that each and every value that is inputted into it is linked to each and every value that is outputted from it. This is what we mean when we talk about something being "fully-connected." When we speak about anything being "fully-connected," this is exactly what we mean by that phrase. When we discuss anything except this, we are particularly referring to the preceding matter. When we talk about anything else outside this, we are specifically referring to the topic that was discussed before. The current layer that we are on is the layer that connects to every other layer in its structure in every conceivable manner. This is the layer that we have just entered. This is the layer that is the focus of our investigation at the moment.



Rectified Linear Units



Crop Prediction Interface

the user must enter all the necessary details such N,P,K,PH etc values using above form and after submitting all information he will get recommended crop as result



Some Images of Diseases and soils

These are the sample images, this type of images user can input from the device or can capture from the disease leaf for prediction of the disease.

IV. CONCLUSION

By using a web-based programme, it has become feasible to both identify illnesses that affect agricultural leaves and take prophylactic steps against them. The programme offers disease prediction modules, suggestions for the kind of soil and fertiliser that should be used for specific crops, and an estimate of the crop yield. All of these features may be found inside the application. This is a web-based application that may be utilised whenever it is most convenient for the user.

FUTURE SCOPE

Machine learning can be accomplished at the present level of Agrotechnology technology, as shown in this experiment. This project is not the only one that may benefit from neural network training; it can also be used to other agricultural difficulties such as the identification of weeds or the categorization of animal breeds. Neural networks perform better when trained with several layers of the deep network on datasets including tens of thousands of pictures. Deep learning can be used to enhance I.O.T. systems in agriculture as well.

REFERENCES

- [1] Brahimi, M., Arsenovic, M., Laraba, S., Sladojevic, S., Boukhalfa, K., Moussaoui, A., 2018. Deep learning for plant diseases: Detection and saliency map visualisation, in: Human and Machine Learning. Springer, pp. 93–117.
- [2] Brahimi, M., Boukhalfa, K., Moussaoui, A., 2017. Deep learning for tomato diseases: classification and symptoms visualization. Applied Artificial Intelligence 31, 299–315.
- [3] DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E.L., Yosinski, J., Gore, M.A., Nelson, R.J., Lipson, H., 2017. Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. Phytopathology 107, 1426–1432.
- [4] Fujita, E., Kawasaki, Y., Uga, H., Kagiwada, S., Iyatomi, H., 2016. Basic investigation on a robust and practical plant diagnostic system, in: 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE. pp. 989–992.
- [5] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 .
- [6] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700–4708.
- [7] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014. Caffe: Convolutional architecture for fast feature embedding, in: Proceedings of the 22nd ACM international conference on Multimedia, ACM. pp. 675–678.
- [8] Kawasaki, Y., Uga, H., Kagiwada, S., Iyatomi, H., 2015. Basic study of automated diagnosis of viral plant diseases using convolutional neural networks, in: International Symposium on Visual Computing, Springer. pp. 638–645.
- [9] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, pp. 1097–1105.
- [10] Mohanty, S.P., Hughes, D.P., Salathe, M., 2016. Using deep learning for image-based plant disease detection. Frontiers in plant science 7, 1419.
- [11] Moriones, E., Navas-Castillo, J., 2000. Tomato yellow leaf curl virus, an emerging virus complex causing epidemics worldwide. Virus research 71, 123–134.

- [12] Navas-Castillo, J., Sanchez-Campos, S., D'íaz, J.A., Saez-Alonso, E., Moriones, E., 1999. Tomato yellow leaf curl virus-is causes a novel ' disease of common bean and severe epidemics in tomato in Spain. *Plant Disease* 83, 29–32.
- [13] Pico, B., D'íez, M.J., Nuez, F., 1996. Viral diseases causing the greatest economic losses to the tomato crop. ii. the tomato yellow leaf curl virus—a review. *Scientia Horticulturae* 67, 151–196.
- [14] Rangarajan, A.K., Purushothaman, R., Ramesh, A., 2018. Tomato crop disease classification using pre-trained deep learning algorithm. *Procedia computer science* 133, 1040–1047.
- [15] Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [16] Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning, in: *Thirty-First AAAI Conference on Artificial Intelligence*.
- [17] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826.
- [18] Tan, W., Zhao, C., Wu, H., 2016. Intelligent alerting for fruit-melon lesion image based on momentum deep learning. *Multimedia Tools and Applications* 75, 16741–16761.
- [19] Too, E.C., Yujian, L., Njuki, S., Yingchun, L., 2018. A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*.
- [20] Wang, J., Chen, L., Zhang, J., Yuan, Y., Li, M., Zeng, W., 2018. Cnn transfer learning for automatic image-based classification of crop disease, in: *Chinese Conference on Image and Graphics Technologies*, Springer. pp. 319–329.
- [21] Yamamoto, K., Togami, T., Yamaguchi, N., 2017. Super-resolution of plant disease images for the acceleration of image-based phenotyping and vigor diagnosis in agriculture. *Sensors* 17, 2557.