

RESEARCH ARTICLE



Potato Plant Leaf Diseases Identification Using Transfer Learning

Birhanu Gardie¹, Smegnew Asemie¹, Kasahun Azezew¹,
Zemedkun Solomon^{1*}

¹ School of Computing and informatics, MizanTepi University, Ethiopia

 OPEN ACCESS

Received: 20.10.2021

Accepted: 05.01.2022

Published: 05.02.2022

Citation: Gardie B, Asemie S, Azezew K, Solomon Z (2022) Potato Plant Leaf Diseases Identification Using Transfer Learning. Indian Journal of Science and Technology 15(4): 158-165. <https://doi.org/10.17485/IJST/v15i4.1235>

* **Corresponding author.**

zemedk.2m@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2022 Gardie et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](https://www.isee.org/))

ISSN

Print: 0974-6846

Electronic: 0974-5645

Abstract

Background/Objectives: Agriculture is a major food source for Ethiopian population. Plant diseases contribute a great production loss, which can be addressed with continuous monitoring. Early plant disease identification using computer vision and Artificial Intelligence (AI) helps the farmers to take preventive course of action to increase production quality. Manual plant disease identification is strenuous and error-prone. **Methods:** In this study, we present a convolutional neural network architecture inception-v3 model to detect potato leaf diseases using a deep learning-based transfer learning technique. We used separable convolution in the inception block that can minimize the number of parameters by an outsized margin and to utilize resource efficiently. The inception-V3 model have a higher training accuracy and needs less training time than the main CNN architecture, as the used parameters are fewer. **Findings:** In this study, there is an improvement on the little noisy on sample images which leads to misidentification of diseases. In our experiment, we have used an RGB color channel image dataset to train model, which yields an overall accuracy performance of 98.7% on the held-out test set. **Novelty:** In order to identify potato leave diseases, we conducted transfer learning for high performance classification with pixel-wise operation to enhance the number of leaf images. A model based on inception-v3 transfer learning approach is presented in this study for disease identification of potato leave images, thus provide an effective computer-aided recognition model for potato disease classification in the absence of large data.

Keywords: Artificial intelligence; convolutional neural network; deep learning; leaf disease identification; Softmax

1 Introduction

Potato is an important food source across the world, and it can be stored to provide food security as a result it is a commercial and existence agriculture⁽¹⁾. In developing countries like Ethiopian, plant diseases are detected manually by trained experts scouting in cultivation field and inspecting potato foliage⁽²⁾⁽³⁾. This task is very monotonous, in some cases it is impractical due to the unavailability of professionals in remote regions⁽²⁾. However, advancements in image processing, and deep learning

in disease recognition of plant leaves using images can make the process far effective and timely. The common disease of potato plant is early blight and Late blight. Such kind of plant diseases is caused by pathogens like bacteria, fungus, parasites, virus, and unfavorable environmental conditions. Plant diseases in the leaf affects photosynthesis process, thereby leads to plant death⁽⁴⁾.

Disease identification in its early stage plays a vital role in the agriculture industry due to plant diseases are often unavoidable. It is essential to be treated the potato plant diseases as soon as possible to manage the degradation of the plant quality and its production quantity which will lead to a loss, when a certain potato plant is infected by diseases. Although, as simple as it may sound, identification of potato plant disease is not a light task to be carried out. To be able to detect potato plant leaf diseases, it needs trained personnel knowledge which is in fact, doesn't incorporate. As a result, the farmers would require to consult trained professionals in potato plant identification which can be rather expensive, boring, and sometimes results would be inaccurate. Nowadays, deep learning has become a significant field of interest of researchers, contributing an essential part in the farming industry for a various importance's from fruitlet grading to weed recognition.

In this study, we try to design a model for potato plant leaf disease identification based on images taken from kaggle plant village dataset using inception-v3 neural network architecture. In image identification, inception-v3 is powerful in understanding patterns from large numbers data, increase the depth of the network to achieve higher performance, reasonable computation cost, number of trainable parameters used in the network and resource utilization⁽⁵⁾⁽⁶⁾. Transfer learning is immensely and effective when we have small dataset, which has a pre-trained model. This model has been trained on an extremely large dataset⁽⁷⁾.

2 Literature review

In the past few decades, machine learning approaches were used in several fields such as face recognition, image processing and video processing, but feature engineering remain a difficult issue. However, the emergence of image processing and deep learning gains a substantial quality results in various benchmarks for plant pathologists without strenuous feature engineering. Plant disease identification has been the most important research field in which convolutional neural network and image processing methods have been widely used for accurate disease identification. Here we take some of research works conducted related to plant leaf disease detection and classification using different advanced approaches. Usama Mokhtar⁽⁸⁾ provides an efficient method that could identify a tomato leaf as healthy and diseased. The image that has been given to the model was pre-processed initially by removing unwanted background and gray level occurrence matrix is used for texture feature extraction. It has achieved an accuracy of 99.83% using the linear kernel function using support vector machine classifier. However, the gained performance accuracy is high, it is not adequate enough to identify the healthy and diseased tomato leaves. David Hughes⁽⁹⁾ presents image based plant disease detection using deep learning techniques in which authors used AlexNet⁽¹⁰⁾ and GoogleNet⁽¹¹⁾ models to train 54,306 images from the plant village dataset in that GoogleNet Achieves 99.35% of training accuracy. But, the accuracy downs to 31.4% while tested in images taken under various conditions to train the model. In the paper three training split distributions 75/25, 60/40 and 70/30 was used. In⁽¹²⁾ an artificial intelligence based banana diseases identification is presented using deep learning approach in which they trained in three different convolutional neural network architectures i.e. inception V2, ResNet50, and MobileNetV1 to recognize banana leave diseases and pests in transfer learning technique. Authors have used 18,000 banana leave images taken from different areas and labeled into 18 various categories. The experimental output shows that the model achieved 90% training accuracy from the dataset used in the experiment. In⁽¹³⁾ neural network based tea leave disease recognition is presented in which authors firstly, data enlargement and segmentation are used to process the tea leave images, and then input to network for training the model. Secondly, to achieve a higher identification accuracy of the network, the iteration and learning rate were adjusted and dropout was applied in the case of overfitting. Finally the experimental result indicates that the identification accuracy achieved 93.75% of training accuracy. However, in this study, high performance classification is employed by adopting the pre-trained inception-v3 based transfer learning weights which is extended by training models from ImageNet dataset to accelerate the learning efficiency of the new inception-v3 model to build better performance detection through by sharing the trained model parameters to the new model via transfer learning.

3 Materials and methods

A generalized overview of the classification of the potato plant leave disease is presented in the following Figure 1. To implement our new transfer learning model a dataset is taken from public database. The images are labeled according to their class category then pre-processing is conducted including resizing of images, filtering of images, applying various data augmentation techniques such as image rotation, flipping and shifting to maximize the size of the dataset. The training and validation images are fed into the pre trained inception-v3 model and features are extracted. Deep learning is a part of machine learning and artificial intelligence algorithm in which its layers are closely related⁽¹⁴⁾. The result of the first layer will be used as an input to

the next layer. In this work, we try to design an inception-v3 based transfer learning model for potato leaf disease detection to build high performance detection for small data using pre trained on large datasets. In plant disease identification experiment, convolutional neural network is an appropriate learning technique in deep learning approach in which it can accurately recognize plant diseases⁽¹⁵⁾. The main steps in this work are image acquisition, image preprocessing, segmentation, feature extraction and identification of potato diseases as depicted in the following Figure 1.

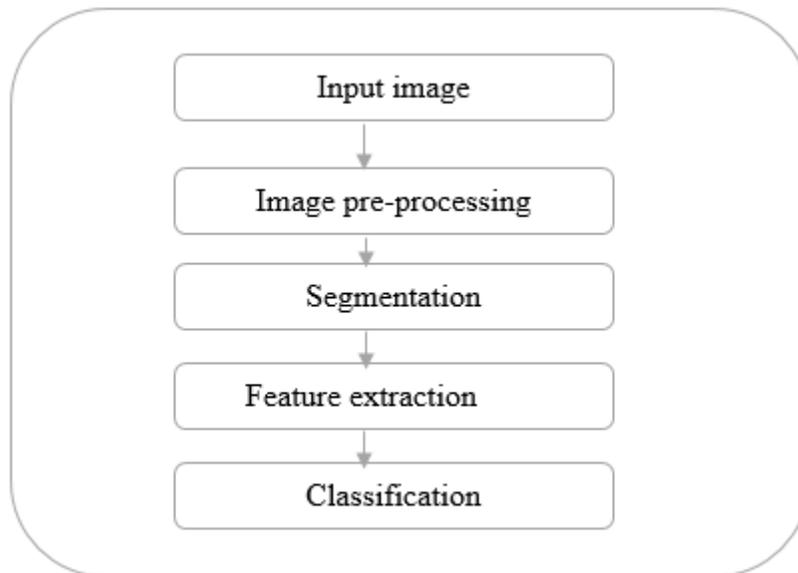


Fig 1. Methodology of the proposed model

3.1 Data acquisition

In our work we have analyzed 2152 potato leaf images taken from plant village dataset, which have three categories. We split the data into two sections, the training portion which is dedicated to train the proposed model and the testing part is used for validation purpose. The data is divided by 80/20 for training and testing respectively. In this approach, we have resized the image dimension into the standard deep learning approach models which is 256 x 256 x 3 pixels in order to train the inception v3 model and to make the training computationally feasible.

Table 1. Dataset details

| Class | Training data | Testing data |
|--------------|---------------|--------------|
| Early blight | 800 | 200 |
| Late blight | 800 | 200 |
| Healthy | 121 | 31 |
| Total | 1721 | 431 |

3.2 Image pre-processing

The contaminated plant leaves in an image produces noise. The noise here would be leaf sand, may be dust and other stuff. To get high training accuracy performance, it is significant to remove the noisy data from plant images. Image pre-processing methods then are used to remove noises from leave images. Many image pre-processing ways are accessible such as image clipping in which cropping the leaf image to get area of interest. The other technique is smoothing filter which is performed to achieve image smoothing. In image processing ZCA whitening, standardized rotation and translation were used for data augmentation.

3.3 Segmentation

Image segmentation is a technique of classifying each pixel in an image as belonging to specific class⁽¹⁶⁾. As for the various sizes of the potato plant leaves, it is imperatively essential to locate and segment the image to increase the performance of identification of potato diseases by reducing the background interference information of the leaf images to get the image's interest of region that is convenient for the inception v3 model to extract features. Image segmentation technique is conducted based on various intensity discontinuity and similarities among the pixels⁽¹⁷⁾. Image segmentation means partitioning the image into various parts with same features, or having rough resemblance which can be used to identify feature similarities in the gray levels between the pixels in an image region. We have achieved segmentation in this work through converting an RGB color mode images to HIS model.

3.4 Feature extraction

Feature extraction plays an important role in digital image analysis⁽¹⁸⁾. Different image pre-processing methods such as standardization, thresholding, binarization, etc. are applied to the sample digital images before gaining functionality. After this feature extraction technique is applied to acquire patterns that are useful in image identification. After all the depiction of the image significant patterns of the image is extracted using innumerable natures of feature extraction with respect to images, the similar features together form a feature vector to recognize and categorize an object. In this work feature extraction is performed using inception-v3 model.

3.5 Classification

The great variations in size, shape, color, texture, background, layout, shape and imaging illumination of plant diseases and pests in real time environment makes the detection task difficult. Because of the strong feature extraction capability, the adoption of convolutional neural network-based identification and classification network has become the most commonly applied pattern in plant leaf diseases and pests detection. Feature extraction part of neural network identification network consists of cascaded convolution layer plus pooling layer which is followed by fully connected layer plus softmax classification layer. Softmax classifier performs identification of outputs based on the given inputs.

4 Experiments and results

When we look the following plot in figure 2, in the first epoch the training accuracy is around 65% and validation accuracy is around 75%. Then both the training and validation accuracy automatically increases linearly and no decreasing, at the same time the validation loss is decreasing linearly and oscillates. There are much variations in between the training and validation accuracy. Therefore, there is an over fitting problem in the model while we train using the dataset. To address this issue, we used a dropout and augmentation methods.

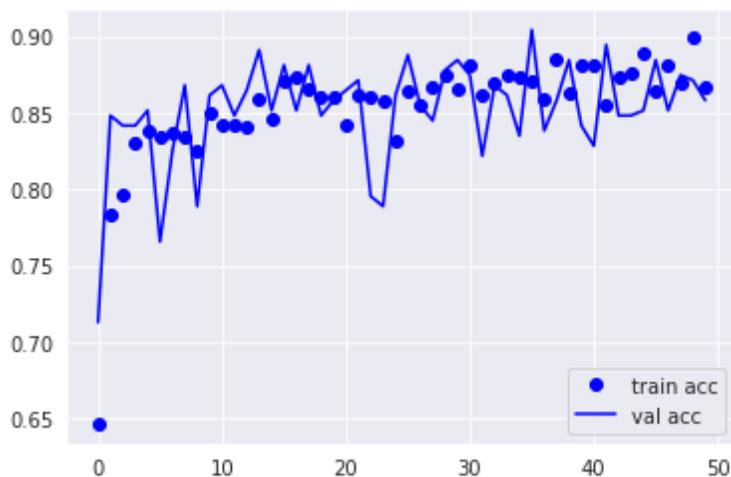


Fig 2. Training and validation accuracy of Inception v3 model

As clearly depicted in the training accuracy and validation curve in Figure 2, most of the time the training accuracy is higher than validation accuracy throughout the curve. However, the variation between the training accuracy and validation accuracy is lesser than compared to in the training phase.

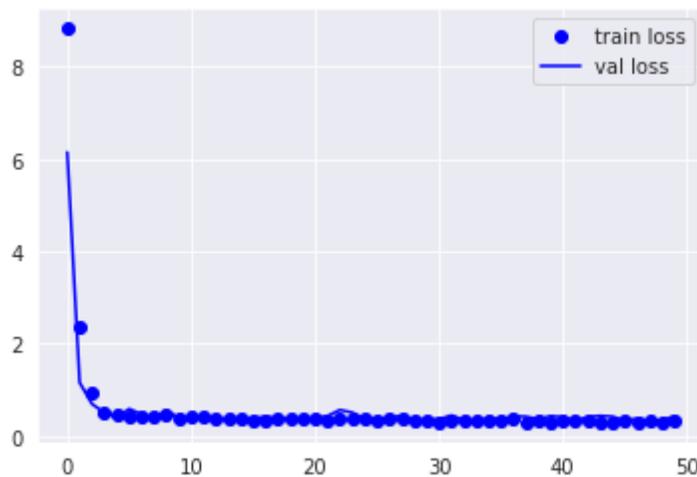


Fig 3. Training and validation loss of Inception v3 model

As clearly depicted in Figure 3, the training loss is far lesser than the validation loss throughout the curve in the plot. It is due to a random sample from our validation set: the validation set at each evaluation step is unrelated. Validation loss was very high initially, while it gradually decreases, over fitting happens in the validation phase as well. However, the amount is lower than compared to in the training phase. Training accuracy is augmented linearly and reaches 95%, while the validation accuracy oscillates up and down to reach 96% accuracy at some instant.

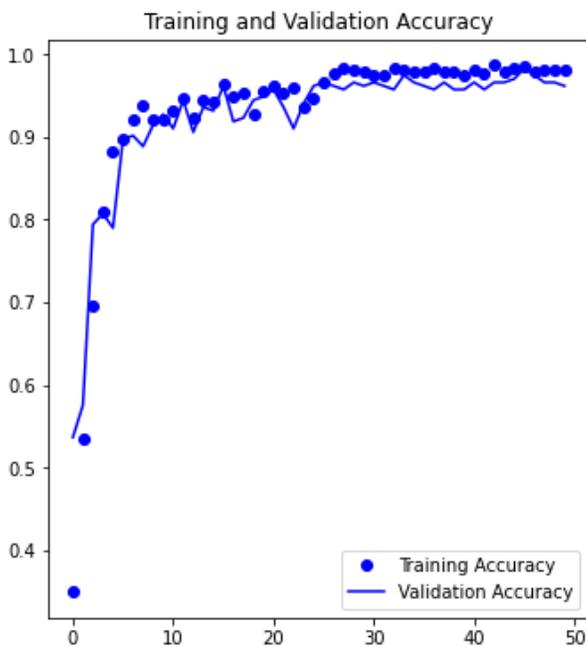


Fig 4. Training and Validation accuracy

As clearly shown in the above Figure 4, our model achieves 98.7% training accuracy and 97.3% validation accuracy. This identification accuracy is realized when the model is trained with the segmented or region of interest of the potato image which can increase the validation accuracy. As a result, segmenting an image and training the model with the segmented region of interest have an important impact on the model performance. We trained the model by applying augmentation, segmentation and dropout methods to increase the training and validation accuracy performance.

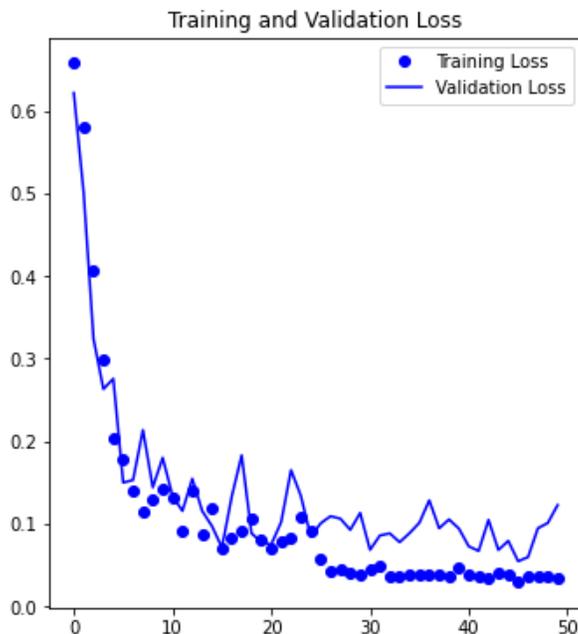


Fig 5. Training and validation loss

When we see the above plot in Figure 5, the training loss is much lesser than the validation loss across the curve. The curve in the training loss is more stable around in the final epochs as compared to the validation loss in the validation phase. However, the amount is decreases as compared in the training phase of the model. The validation loss somehow go oscillates towards in the validation phase. It was relatively small throughout the plot. Except at some instants such as epoch 25. The model here, performs out very well as compared in the validation and training phase.

In⁽¹⁹⁾ employed a multi-level deep learning model for potato disease categorization using YOLOv5 image segmentation technique. The dataset used are taken from Central Punjab and Pakistan which can represent a specific region and plant diseases have a significant environmental factor and the model would not outperform in datasets that collected in other regions. In⁽²⁰⁾ plant disease identification using transfer learning and CNN is conducted in which authors replace the standard convolution with separable convolution to minimize the number of parameters that is trained on 14 various plant species and 38 different disease classes. The implemented model achieves 99.56%, 98.42%, 99.11%, and 97.02% using EfficientNetB0, inception-v3, inceptionResNetV2, and MobileNetV2 respectively. They used three various representations which is color, grayscale and segmented image datasets taken from plant village. In⁽⁹⁾ performed a work by applying GoogleNet and AlexNet CNN architectures to train 54,306 image datasets taken from the kaggle public plant village database collected under in controlled environment, in which GoogleNet architecture realizes enhanced and reliable through a training accuracy of 99.35% and 85.53 % in case of AlexNet architecture. However, the performance accuracy decreases to 31.4% while tested in image datasets that taken under conditions varied from images used to train the model. In⁽²¹⁾ contributed a residual network CNN architecture based plant phenotyping and infection classification on an unequal imbalanced data. But in this study, we build a high-performance potato leave disease detection using inception-v3 transfer learning weights which is expanded by training the new model from ImageNet dataset to accelerate the learning efficiency of the new inception-v3 model to build a higher accuracy performance. In our experiment the datasets are taken from the public dataset in that we apply 80/20 splitting ration for the training and testing phases.

In our experiment, the overall accuracy on the training dataset is performed to evaluate the performance of the proposed model with other models using unseen data. Table 2, shows the performance result of each architecture, Inception-v3 achieves

Table 2. Comparison of models using the same dataset

| Model | Accuracy |
|----------------|----------|
| AlexNet | 93.75 |
| GoogleNet | 96.63 |
| Inception - V3 | 98.3% |

the highest test accuracy. Various parameters affect the performance of the pre-trained weight in deep learning approach. Metrics such as pre-trained weight size, number of parameters and depth of layers has significant impact on the CNN architectures accuracy performance. The accuracy of CNN models performance would enhance with the increase of layers.

5 Conclusion

Plant disease identification in its early stage plays a vital role in the agriculture industry. In this study, we attempt to design an inception-v3 transfer-learning model for potato plant leave diseases identification. The model is fine-tuned and trained to detect the healthy and diseased potato leave images. The achieved results indicate that the proposed model outperforms than the AlexNet and GoogleNet architectures. In our experiment work, the potato leave image from plant village dataset has three classes including the healthy leave images. The dataset we used for the experiment is a three-color channel image dataset by applying segmentation method. In the first experiment, the model achieves a training accuracy of 96.8%. However, after the augmented dataset, and applying segmentation on the images, the training accuracy is enhanced to 98.3% which is a higher performance. In the future work, potato leave disease identification would be further investigated with large number of datasets. We will conduct further research works using ensemble learning to analyze the diseases severity and to find higher performance.

References

- 1) Diro S. Assessment of the effect of plant density on yield of irrigated potato and analysis of input supply chains at Adea Berga district of west shoa zone, Oromia. 2016. Available from: <https://cgspace.cgiar.org/handle/10568/90423>.
- 2) Johnson J, Sharma G, Srinivasan S, Masakapalli SK, Sharma S, Sharma J, et al. Enhanced Field-Based Detection of Potato Blight in Complex Backgrounds Using Deep Learning. *Plant Phenomics*. 2021;2021:1–13. Available from: <https://dx.doi.org/10.34133/2021/9835724>.
- 3) Bhagwat R, Dandawate Y. A Review on Advances in Automated Plant Disease Detection. *International Journal of Engineering and Technology Innovation*. 2021;11(4):251–264. Available from: <https://dx.doi.org/10.46604/ijeti.2021.8244>.
- 4) Arivazhagan S, Ligi SV. Mango Leaf Diseases Identification Using Convolutional Neural Network. *Mango Leaf Diseases Identification Using Convolutional Neural Network*. 2018;120:11067–11079. Available from: <https://www.acadpubl.eu/hub/2018-120-6/8/731.pdf>.
- 5) Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016;2016:2818–2826. doi:10.1109/CVPR.2016.308.
- 6) Indolia S, Goswami AK, Mishra SP, Asopa P. Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. *Procedia Computer Science*. 2018;132:679–688. Available from: <https://dx.doi.org/10.1016/j.procs.2018.05.069>. doi:10.1016/j.procs.2018.05.069.
- 7) Irla T. Transfer Learning using Inception-v3 for Image Classification. 2021. Available from: <https://medium.com/analytics-vidhya/transfer-learning-using-inception-v3-for-image-classification-86700411251b>.
- 8) Mokhtar U, Bendary NE, Hassenian AE, Emary E, Mahmoud MA, Hefny H, et al. SVM-Based Detection of Tomato Leaves Diseases. *Advances in Intelligent Systems and Computing*. 2015;323:641–652. doi:10.1007/978-3-319-11310-4_55.
- 9) Mohanty SP, Hughes DP, Salathé M. Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science*. 2016;7:1419–1419. Available from: <https://dx.doi.org/10.3389/fpls.2016.01419>.
- 10) Zhang Y, Gao J, Zhou H. Breeds Classification with Deep Convolutional Neural Network. *Proceedings of the 2020 12th International Conference on Machine Learning and Computing*. 2020;p. 145–151. doi:10.1145/3383972.3383975.
- 11) Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015.
- 12) Selvaraj MG, Vergara A, Ruiz H, Safari N, Elayabalan S, Ocimati W, et al. AI-powered banana diseases and pest detection. *Plant Methods*. 2019;15(1):1–11. Available from: <https://dx.doi.org/10.1186/s13007-019-0475-z>.
- 13) Sun X, Mu S, Xu Y, Cao Z, Su T. Image Recognition of Tea Leaf Diseases Based on Convolutional Neural Network. *2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*. 2018;p. 304–309. doi:10.1109/SPAC46244.2018.8965555.
- 14) What is deep learning and how does it work. 2021. Available from: <https://www.techtarget.com/searchenterprisetalks/definition/deep-learning-deep-neural-network>.
- 15) Tiwari D, Ashish M, Gangwar N, Sharma A, Patel S, Bhardwaj S. Potato Leaf Diseases Detection Using Deep Learning. *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*. 2020;p. 461–466. doi:10.1109/ICICCS48265.2020.9121067.
- 16) New deep learning model brings image segmentation to edge devices – TechTalks. 2021. Available from: <https://bdtechtalks.com/2021/05/07/attendseg-deep-learning-edge-semantic-segmentation/>.
- 17) Getahun T, A. Automatic flower disease identification using image processing techniques. *Addis Ababa University*. 2015;p. 1–69. Available from: <http://213.55.95.56/bitstream/handle/123456789/1993/Getahun%20Tigistu.pdf?sequence=1&isAllowed=y>.
- 18) Kumar G, Bhatia PK. A Detailed Review of Feature Extraction in Image Processing Systems. *2014 Fourth International Conference on Advanced Computing & Communication Technologies*. 2014;p. 5–12. doi:10.1109/ACCT.2014.74.

- 19) Rashid J, Khan I, Ali G, Almotiri SH, AlGhamdi MA, Masood K. Multi-Level Deep Learning Model for Potato Leaf Disease Recognition. *Electronics*. 2021;10(17):2064–2064. Available from: <https://dx.doi.org/10.3390/electronics10172064>.
- 20) Hassan SM, Maji AK, Jasiński M, Leonowicz Z, zbieta Jasińska E. Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach. *Identification of Plant-Leaf Diseases Using CNN and*. 2021;10(12):1388–1388. Available from: <https://doi.org/10.3390/electronics10121388>.
- 21) Hati AJ, Singh RR. Artificial Intelligence in Smart Farms: Plant Phenotyping for Species Recognition and Health Condition Identification Using Deep Learning. *AI*. 2021;2(2):274–289. Available from: <https://dx.doi.org/10.3390/ai2020017>.