Transfer Learning-Based Deep Learning Model for Corn Leaf Disease Classification

Justin An[1], Nian Zhang[2(✉)], and Wagdy H. Mahmoud[2]

[1] Department of Mechanical Engineering, University of the District of Columbia, Washington, D.C 20008, USA
justin.an@udc.edu

[2] Department of Electrical and Computer Engineering, University of the District of Columbia, Washington, D.C 20008, USA
{nzhang,wmahmoud}@udc.edu

Abstract. Corn, also recognized as maize, is one of the most popular food grains in the world, accounting for 80% of the total volume of global trade of feed grains [1]. Corn leaf diseases pose a significant threat to global agricultural production, impacting crop yield, quality, and economy. Rapid and accurate detection of these diseases is crucial for reducing the impact to the agricultural production. In this study, we propose a transfer learning (TL) based model for the classification of corn leaf diseases. The proposed TL model integrates a convolutional neural network (CNN) known as AlexNet as it has been pre-trained on a large dataset containing more 1000 images [2]. The model is then trained on a dataset from PlantVillage [3] which consists of images of three types of corn leaf diseases, including common rust, gray leaf spot, northern leaf blight, and healthy corn leaves. To ensure better performance, the TL model, entitled NewNet, uses data augmentation techniques such as rotation, scaling, and flipping during training, effectively enhancing the model's ability to handle variations in image appearance and diseases. Transfer learning was achieved by fine-tuning the pre-trained AlexNet CNN and leveraging its learned features. Experimental results show that the TL model achieves a 99.2% accuracy for classifying common rust disease, 93.5% for gray leaf spot, 100% for healthy corn leaves, and 90.0% for northern leaf blight. Therefore, by using the TL technique, we can improve the accuracy of detecting plant diseases, specifically corn leaf diseases.

Keywords: Convolutional neural network (CNN) · Plant disease · Deep learning · Machine learning · Transfer learning

1 Introduction

The agricultural industry plays a significant role within the world economy. It employs more than one billion people, while serving the entire world's population with numerous crops and meats. The security of these crops is of great importance, and it can be easily negatively influenced by a plethora of natural and human factors. Some of these include natural disasters [4], chemical influences [5], plant diseases [6], and many others. Plant

X. Le and Z. Zhang (Eds.): ISNN 2024, LNCS 14827, pp. 163–173, 2024.
https://doi.org/10.1007/978-981-97-4399-5_16

disease is one of the most common factors for the decrease of agricultural yield. It affects both global food security and farmers who depend on healthy crops and high yield. Corn is one of the most popular food grains in the world, accounting for 80% of the total volume of global trade of feed grains [1]. The United States alone produces approximately 95% of total feed grain production and use in the world. However, in other developing countries, farmers produce approximately 80% of their agricultural production [7], and reports have shown that more than 50% loss of crops due to disease and pests [8]. With the expected increase in global population, food security is a concern and must be prioritized. Therefore, rapid, and accurate techniques to classify plant diseases should be investigated.

To aid towards this issue, our goal of this research is to develop a CNN model that can detect and classify corn leaf diseases within a given dataset. This model will aid in the detection of plant diseases within developing countries while being low-cost and easily applicable. This can lead to higher yields of corn in such countries, thus enhancing food security. Early methods of detection of plant diseases required agricultural experts and farmers to physically visit the location and manually determine the disease. This was limited by various reasons, such as the number of agricultural experts available to oversee crop lands and number of locations. Thus, researchers have been looking into automated and smart methods for plant disease detection. Computer-aided detection systems have become a popular method in reducing human dependency and aiding in plant disease detection by using a combination of artificial intelligence (AI), machine learning (ML) and deep learning (DL). Machine learning (ML), in combination with image processing approaches, presents an opportunity to benefit the agricultural scene. Deep learning (DL) is a type of machine learning technique that has been actively used in various fields outside of agriculture, such as medical [9] and climate change [10]. Its popularity comes from its various capabilities, which includes feature extraction, pattern discernability, and data classification. Convolutional neural networks (CNNs) are a branch of deep learning renowned for their proficiency in extracting and learning features from images. The efficacy of CNN models is based on factors such as the dimensions of the training dataset and the architecture of the network layers. Due to the time-consuming nature of collecting agricultural data, such as plant diseases in various crops, an approach in deep learning is investigated. Transfer learning is a suggested suitable approach for training the CNN model for detecting plant disease.

The remainder of this paper is organized as follows. In Sect. 2, we provide a literature review of previous similar studies. Section 3 then explains how we prepared the data for training and testing our transfer learning model, along with its structure and algorithm. Next, in Sect. 4, we present and analyze the results obtained from our model, comparing them with those of other methods and models. Finally, Sect. 5 concludes the paper.

2 Literature Review

Deep learning techniques have been used in many applications, such medical, automotive, climate change, and networking. Transfer learning techniques have been used in such applications for their ability to detect and classify data. In the medical field for examples, Justin et al. [9] applied transfer learning to brightness-mode (B-mode) and ultrasound elastography images to distinguish malignant and benign breast tumors. This was done by using a pre-trained network of AlexNet, fine-tuning it to fit their needs, and feeding the medical images to it. This resulted in a classification accuracy of 88%.

Mishra et al. [11] investigated combining smart devices and machine learning in diagnosing corn leaf diseases automatically to reduce crop loss. They developed a real time method based deep CNN to detect and classify corn crop diseases. They also used the corn leaf dataset from the PlantVillage dataset [3] for their model. They fine-tuned the model's hyperparameters and used a system equipped with GPU to enhance the performance of the proposed model. They then optimized for real-time performance. The pre-trained deep CNN model was then deployed on a Raspberry Pi 3 using the Intel Movidius Neural Compute Stick, which features dedicated hardware blocks for CNNs. The testing results showed that the deep learning model achieved an accuracy of 88.46% in identifying corn leaf diseases, highlighting the effectiveness of this approach. This approach shows that it can be applied to other smart devices.

Bai et al. [12] uses a similar AlexNet approach as our research. They developed and used CNN to detect food plant diseases of rice, wheat, and corn. They adjusted the CNN's hyperparameters for optimization of the datasets. Bai et al. used four datasets, one from PlantVillage [3], one containing wheat leaf diseases, and two containing rice leaf diseases. They fine-tuned the CNN by using different types of layers compared to the traditional AlexNet, along with adjusting the number of layers. By doing so, they were able to achieve classification accuracies of 99.34%, 95.15% and 92.47% for wheat, corn, and rice respectively.

In the paper by Panigrahi et al. [13], they investigate different supervised machine learning methods like Random Forest (RF), Naive Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Decision Tree (DT) to detect diseases in maize plants using images of the plants. The various classification techniques are evaluated and compared to find the most accurate model for predicting plant diseases. Among them, the Random Forest (RF) algorithm demonstrated the highest accuracy at 79.23%, outperforming the other classification methods.

Agarwal et al. [14] developed a CNN model to identify corn crop diseases. The model developed consists of three convolution and max pooling layers, followed by fully connected layers. They used the PlantVillage dataset [3], and after running 1000 epochs, they were able to obtain an accuracy of 94%. They compared their model with state-of-the-art CNN models, VGG-16 [15] and Inception V3 [16] and found that they had a higher classification accuracy compared to the Inception V3 architecture, but lower accuracy compared to the VGG-16 architecture.

3 Proposed Methodology

Several state-of-the-art machine learning techniques that have been used in image classification include convolutional neural networks (CNNs), transfer learning, ensemble learning, deep learning architectures and self-supervised learning. Transfer learning is chosen due to knowledge transfer from a pre-trained CNN, and accelerated model training as it takes considerably less time and resources to start with a model that has already learned image features than starting from scratch.

3.1 Dataset

The dataset used for fine-tuning the pre-trained model is from PlantVillage dataset [3]. The PlantVillage dataset consists of images of leaves of various crops, such as apples, corn, and tomatoes. The dataset expands on these crop leaves by also including crop leaves afflicted by diseases or pests. These images are sorted and classified correctly already. As we are focusing on corn leaf diseases, only the folders containing images of corn leaves were used in our proposed model as our dataset. The corn leaf classes consist of common rust, gray leaf spot, northern leaf blight and healthy.

Difference in image sizes poses a serious challenge in image detection, especially from public datasets. This inconsistency poses a hindrance to feeding deep learning (DL) models effectively, as these models usually require homogenized image sizes for optimal performance. Without preprocessing the data from diverse sets to standardize image dimensions, leveraging DL for disease detection becomes impractical. Sample images of common rust, gray leaf spot, healthy, and Northern leaf blight classes in the dataset are shown in Fig. 1. In this research, the dataset was pre-processed using MATLAB. The dataset is placed into a shared folder, where we then use MATLAB's resizing feature to homogenize the image sizes. The input format for the pre-trained AlexNet CNN model consists of batches of RGB images, each sized at 227-by-227-by-3, where 227-by-227 denotes the image dimensions and 3 represents the number of color channels (red, green, blue). Following the homogenization of image sizes and conversion to a readable format, the images are poised for training and testing within the NewNet model architecture.

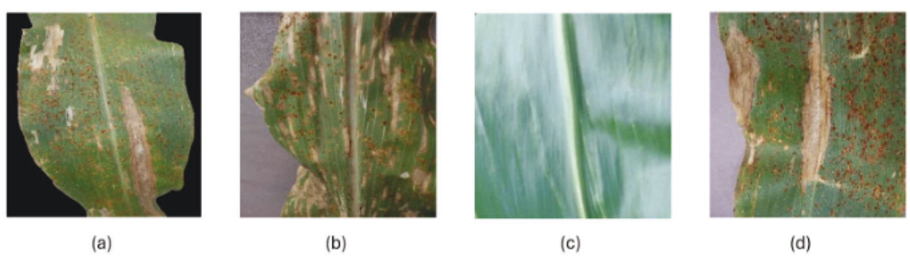

Fig. 1. Sample images of each class in the dataset. (a) Common rust, (b) Gray leaf spot, (c) Healthy, and (d) Northern leaf blight

3.2 Transfer Learning Architecture

Transfer learning is a technique in machine learning which involves leveraging knowledge obtained from training a model on one task to a related but distinct task [17]. In terms of neural networks, transfer learning typically occurs in two stages. In our research, the pre-trained CNN of AlexNet is used as the pre-trained model. AlexNet has been trained on the ImageNet dataset, which encompasses over 1000 images of numerous classes. Afterwards, fine-tuning occurs, where the pre-trained network is further refined using the specific dataset of interest. This fine-tuned model becomes the transfer learning model, which we entitled NewNet. Typically, the initial layers of the model are tasked with learning basic features such as edges, corners, and curves, while the final layers focus on capturing more intricate features such as complex shapes and textures.

We used the pre-trained AlexNet [2], a widely recognized convolutional neural network. AlexNet was used as leveraging GPU acceleration during training significantly enhances its computational efficiency compared to alternative models. Originally designed by Alex Krizhevsky, AlexNet was primarily engineered to categorize large quantities of images into a thousand object classes, including animals, household items, and clothing items. Thus, the network underwent extensive training to discern a spectrum of features across a diverse set of images from the ImageNet database. The AlexNet model used in this research comprises 25 layers, consisting of convolution layers, rectified linear unit (ReLU) layers, max pooling layers, cross channel normalization layers, dropout layers, fully connected (FC) layers, a softmax layer, and an output layer. To address potential overfitting issues, pooling layers are integrated. Our proposed model, NewNet, uses transfer learning, which is a technique that capitalizes on a pre-trained model, such as AlexNet in our case, to discern patterns and features. Leveraging AlexNet's training on over 1000 augmented images, we use this acquired knowledge and adapt it to our dataset. This adaptation process involves substituting the final layer of the pre-trained model with a new layer. Additionally, we replaced the training and testing images used by AlexNet with corn leaf images from our dataset after resizing them. The fully connected (FC) layer, originally used in AlexNet to categorize images into 1000 distinct labels, was changed to accommodate 4 classes: common rust, gray leaf spot, healthy, and northern leaf blight. The architecture of the proposed NewNet transfer learning model is illustrated in Fig. 2.

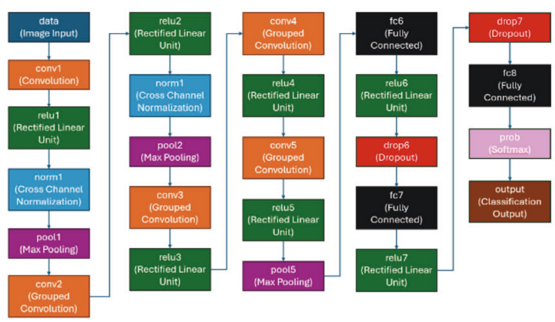

Fig. 2. Proposed transfer learning CNN architecture

In AlexNet, fully connected layer 8 (fc8) in the figure refers to the final fully connected layer that consists of 1000 neurons. In AlexNet, this layer is responsible for producing the final output for classification across 1000 distinct labels, such as pen, keyboard, or cat. In the proposed NewNet, this layer is changed to house 4 neurons, as we have 4 classification classes, or labels. This was done as it is more efficient to have the output consist of 4 neurons rather than 1000, as we only have 4 classes. This adjustment lessens the model's computational resources, making it more appropriate for the task.

3.3 Proposed Transfer Learning Algorithm

The proposed CNN, NewNet, works in a step-by-step manner. Initially, images or data are fed into the network, passing through each layer one after the other. As it moves through these layers, the model gradually identifies the unique traits of each class, helping it become better at accurately classifying new images over time. The dataset is divided into training and testing sets, constituting 80% and 20% of the data, respectively. The proposed architecture relies heavily on convolutional layers, which play a crucial role in extracting features from input images. These layers apply filters to the images, moving across them to calculate dot products between the filter weights and corresponding pixel values. The resulting output highlights patterns and features present in the image.

To train convolutional neural networks effectively, it's essential to incorporate batch normalization layers between convolutional layers and nonlinear activation functions. To achieve this, we integrated ReLU (Rectified Linear Unit) layers into our architecture. This not only speeds up processing but also reduces the network's dependence on initialization conditions. The ReLU layer utilizes the ReLU activation function, transforming negative values into zero while preserving positive values unchanged. This process enhances the network's ability to learn and make predictions. Following the ReLU layer, the rectified feature map undergoes pooling, a procedure that reduces the dimensionality of the feature map. In our network, we utilize max pooling, which selects the maximum value within a filter and transfers it to the new pooling feature map. After the convolutional and pooling layers, fully connected (FC) layers are commonly included in neural network architectures. These layers establish connections between every neuron in the layer before, linking the features extracted by preceding layers across the entire image. This interconnectedness aids the neural network in identifying additional features. The final FC layer combines these features to effectively classify the CXR images.

When training a neural network like NewNet, a common issue that can arise is overfitting. Overfitting occurs when the model learns to perform well on the training data but struggles to generalize to new, unseen data. To combat overfitting, a technique called "Dropout" can be employed. The Dropout layer randomly deactivates a portion of input units during each training step, preventing the network from relying too heavily on specific features and reducing the risk of overfitting. The SoftMax layer is another crucial component in neural network architectures, often used as the final layer for classification tasks. Its main function is to convert the raw output values, known as logits or scores, into probabilities for each class in a set of mutually exclusive classes. This enables the network to make predictions by identifying the class with the highest probability as the predicted label. In the NewNet convolutional neural network, the final layer is the ClassOutput layer, also known as the classification layer. This layer computes the probabilities of the output classification obtained from the SoftMax layer, assigning a probability value to each class for every input image. Using these probabilities, the ClassOutput layer determines the class for each input image and calculates the associated loss.

4 Experimental Results

The corn leaf dataset used consists of four classes of images, including common rust, gray leaf spot, healthy, and northern leaf blight. The training and testing data is shown in Table 1.

Table 1. Corn Leaf Dataset

Corn Leaf Class	Number of images used for training (80%)	Number of images used for testing (20%)	Total number of images
Common rust	400	100	500
Gray leaf spot	410	103	513
Healthy	400	100	500
Northern leaf blight	400	100	500
Number of images	1610	403	2013

Throughout the training phase, we use the cross-entropy loss function in conjunction with the Stochastic Gradient Descent Optimizer (SGDM), configured with a learning rate of 0.0001. The training procedure extends over 20 epochs, with a MiniBatchSize of 124. Additionally, the training images undergo resizing to dimensions of 227-by-227-by-3 and data augmentation to feed the dataset. This augmentation is done by random vertical flipping and random translations of up to 20 pixels in both horizontal and vertical directions. The integration of data augmentation helps in mitigating overfitting by deterring the network from memorizing specific details of the training images. During the training phase of the network, we were able to achieve a training accuracy of 96.07%. Figure 3 represents the training progress and accuracy throughout each iteration. Figure 4 shows the corresponding loss curve associated with the training progress.

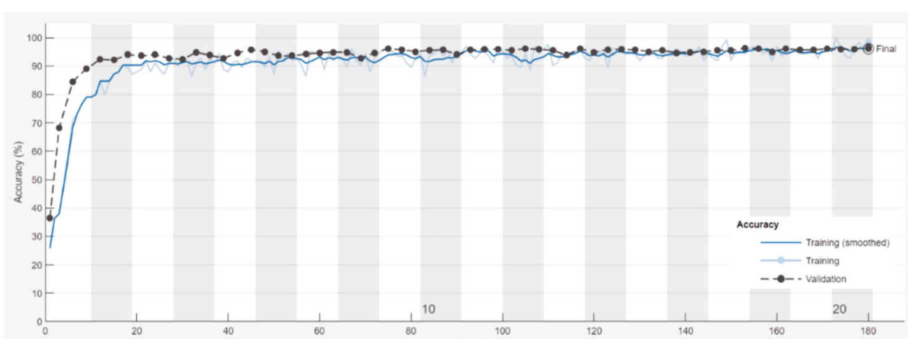

Fig. 3. Accuracy of trained model

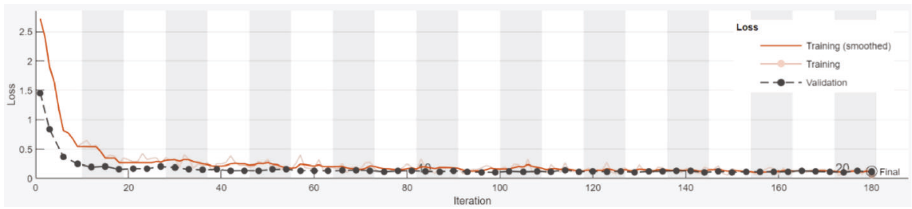

Fig. 4. Loss curve

In terms of feature visualization, Fig. 3 shows that neurons in the beginning layers are where the CNN is able to recognize simple features and patterns. This occuers around the Conv1 and Conv 2 layers. Figure 4 shows that the model accuracy becomes stable after around 18 iterations. This trend is similar in the loss curve, as it starts to stabilize early after 10 iterations. The early stabilization of the accuracy curve of our model indicates that the model's accuracy on the training dataset reached a steady state early in the training process. Therefore, the model has quickly learned to make accurate predictions and has converged towards its optimal performance. Similarly, the early stabilization of the loss curve of our model means that the loss value reached a steady state early in the training process, indicating that the model has quickly converged towards an optimal solution. This signifies that our model has effectively learned the underlying patterns in the training data and will not make significant improvements in performance with further training iterations.

Figure 5 shows the confusion matrix presenting the classification accuracy of the transfer learning model on the test images. It demonstrates an average accuracy of 95.70% in accurate predictions. Cells along the diagonal, in both numerical and percentage terms, depict correctly classified observations, while cells off the diagonal signify instances that are classified incorrectly. The column on the far right displays the percentages of

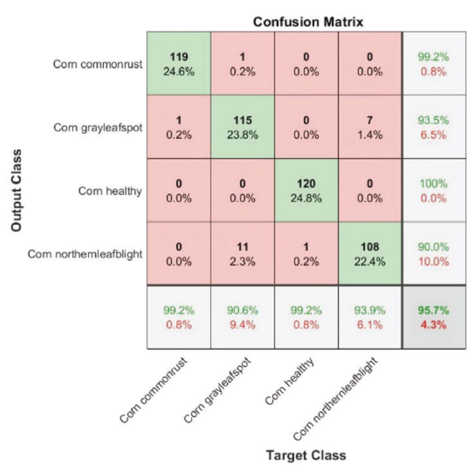

Fig. 5. Confusion matrix for the NewNet transfer learning model

accurately and inaccurately classified instances for the predicted class. Similarly, the bottom row indicates the percentages of each class that are correctly and incorrectly classified.

As shown in Fig. 5, the classification accuracy of common rust is 99.2% with 119 images correctly classified out of 120, gray leaf spot is 93.5% with 115 images correctly classified out of 123, healthy is 100% with 120 out of 120 images correctly classified, and northern leaf blight has 90% of images correctly classified with 108 images out of 120 images.

Figure 6 shows the testing results of the NewNet TL model on the corn leaf dataset provided. The labels are either common rust, gray leaf spot, healthy, or northern leaf blight.

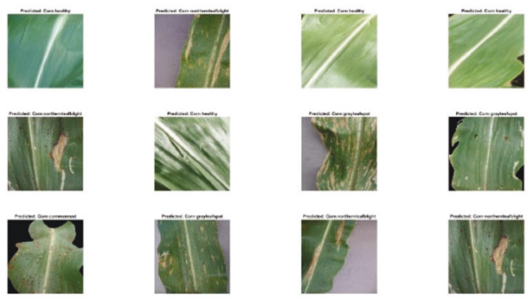

Fig. 6. NewNet classification of corn leaf images

Evaluation of our model was done by comparing the overall classification accuracy of our NewNet model with the works of Mishra et al. [11], Bai et al. [12], Panigrahi et al. [13], and Agarwal et al. [14], Yu et al. [15], and Szegedy et al. [16]. The overall classification accuracy comparison of models is shown in Table 2.

Table 2. Overall Classification Accuracy Comparison of Models

Model	Overall Classification Accuracy (%)
NewNet (SGDM Optimizer)	**95.70**
NewNet (Adam Optimizer)	**95.00**
Mishra et al. NCS Model [11]	88.66
Bai et al. AlexNet [12]	95.15
Panigrahi et al. Random Forest [13]	79.23
Agarwal et al. [14]	94.00
VGG-16 [15]	95.00
Inception V3 [16]	93.50

Additional data augmentation techniques were implemented to compare the accuracy of our model. Solely changing the optimizer to "Adam" (Adaptive Moment Estimation) instead of "sgdm" resulted in an overall accuracy of 95.0%, which is lower than using the "sgdm" optimizer. Another augmentation technique used was to add additional translations in both the X and Y axes directions. This resulted in a similar result of 95.7% as the original result.

5 Conclusions

This paper proposes a transfer learning model called NewNet, based on the AlexNet architecture, to identify corn leaf disease images accurately and effectively. The proposed NewNet model extracts features from corn leaf images and compares them with features learned from training images. To evaluate its performance, we used a dataset comprising 500 common rust infected corn leaf images, 513 Gy leaf spot infected corn leaf images, 500 healthy corn leaf images, and 500 northern leaf blight infected corn leaf images. The results demonstrate that NewNet achieved an overall classification accuracy of 95.70% for the corn leaf images used in testing. Comparisons of accuracy were done with other developed CNN models and state-of-the-art CNN models, which showed promising results. These findings suggest that the NewNet model can significantly aid in the agricultural industry in detecting crop diseases. Future work will seek to integrate our NewNet model into real-time monitoring systems to aid farmers in the field. An option discussed has been to develop mobile applications for smart devices, so that it is easily accessible for all farmers. Another option is to develop a device that can process images and training in real-time.

Acknowledgment. This work is supported by the National Science Foundation (NSF) grants #2401880 and #2011927, NIH grant #1R25AG067896–01, and USGS/DCWRRI FY2023 grant.

References

1. Economic Research Service. Feed Grains Sector at a Glance, https://www.ers.usda.gov/topics/crops/corn-and-other-feed-grains/feed-grains-sector-at-a-glance (2024)
2. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: 25th International Conference on Neural Information Processing Systems, vol. 1, 1097–1105 (2012)
3. PlantVillage Dataset, https://www.kaggle.com/datasets/emmarex/plantdisease (2024)
4. Natural Events and Disasters, U.S. Environmental Protection Agency, https://www.epa.gov/agriculture/agriculture-and-natural-events-and-disasters#:~:text=loss%20of%20harvest%20or%20livestock,systems%20and%20other%20agricultural%20infrastructure (2024)
5. Vasfilov, S.P.: Possible ways of negative influence of acid gases on plants. J. Gen. Biol. **64**(2), 146–159 (2003)
6. Strange, R.N., Scott, P.R.: Plant disease: a threat to global food security. Annu. Rev. Phytopathol. **43**, 83–116 (2005)
7. Smallholders, Food Security and the Environment.: Int. Fund Agric. Dev. (2013)
8. Harvey, C.A., et al.: Extreme vulnerability of smallholder farmers to agricultural risks and climate change in Madagascar. Phil. Trans. Roy. Soc. B, Biol. Sci. **369**(1639), (2014)

9. An, J., Abdus-Shakur, T., Denis, M.: Quantitative assessment of tissue stiffness using transfer learning ultrasound elastography: a breast cancer phantom study. IEEE Sens. Lett. **7**(10), 1–4 (2023)
10. Zhang, N., Dai, X., Ehsan, M.A., Deksissa, T.: Development of a drought prediction system based on long short-term memory networks (LSTM). In: Han, M., Qin, S., Zhang, N. (eds.) ISNN 2020. LNCS, vol. 12557, pp. 142–153. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-64221-1_13
11. Mishra, S., Sachan, R., Rajpal, D.: Deep convolutional neural network based detection system for real-time corn plant disease recognition. Procedia Comput. Sci. **167**, 2003–2010 (2020)
12. Bai, F., Mahmoud, W., Zhang, N.: Deep learning-based plant disease image recognition for cyber-physical systems. In: 13th International Conference on Information Science and Technology (ICIST 2023), Cairo, Egypt, 1–8 (2023)
13. Panigrahi, K.P., Das, H., Sahoo, A.K., Moharana, S.C.: Maize leaf disease detection and classification using machine learning algorithms. In: Das, H., Pattnaik, P.K., Rautaray, S.S., Li, K.C. (eds.) Progress in Computing, Analytics and Networking. AISC, vol. 1119, pp. 659–669. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-2414-1_66
14. Agarwal, M., et al.: A convolution neural network based approach to detect the disease in corn crop. In: IEEE 9th International Conference on Advanced Computing (IACC), pp. 176–181 (2019)
15. Yu, W., et al.: Visualizing and comparing AlexNet and VGG using deconvolutional layers. In: 33rd International Conference on Machine Learning (2016)
16. Szegedy, C., Vanhoucke, V., Ioffe, et al.: Rethinking the inception architecture for computer vision. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826 (2016)
17. An, J., Zhang, N., Mahmoud, W., Denis, M.: A COVID-19 CXR image recognition method based on deep transfer learning. In: The IEEE World Congress on Computational Intelligence (IEEE WCCI 2024), Yokohama, Japan, June 30-July 5 (2024)