Diagnostics system for plant leaf diseases using photo images

Dilnoz Muhamediyeva[1*] and *Nozir* Tukhtamurodov[2]

[1]"Tashkent Institute of Irrigation and Agricultural Mechanization Engineers" National Research University, Tashkent, Uzbekistan
[2]Research Institute for the Development of Digital Technologies and Artificial Intelligence, Tashkent, Uzbekistan

Abstract. The subject under consideration is the issue of detecting agricultural crop diseases. The preliminary data for determining the phytosanitary status of cultivated plants are the images of their leaves. To tackle this issue, a model of diagnostic algorithms has been proposed, which involves creating a set of preferred characteristics and making diagnostic decisions based on comparing these features. The process of establishing the model of diagnostic algorithms has been outlined. The effectiveness of the proposed model has been demonstrated in diagnosing wheat diseases based on leaf images.
A novel technique for detecting plant diseases has been introduced, relying on digital RGB photographs of leaves in the visible spectrum. The novelty of this method lies in diagnosing visible symptoms of plant leaf diseases from photographic images. There is a significant number of plant diseases that can reduce productivity, resulting in economic and environmental damage. Consequently, the accurate and timely diagnosis of plant diseases is paramount. One widely used approach for detecting plant pathologies involves using convolutional neural networks to diagnose plant leaf diseases from photographic images.

1 Introduction

According to UN estimates, agricultural sustainability is a critical issue due to the rapid growth of the global population, which is projected to reach 8.5 billion people by 2030. With most suitable agricultural land already in use, there is limited potential for expansion [1]. Thus, the challenge is to maximize crop yield while preventing the spread of pests and diseases [2]. Many plant diseases exhibit visible symptoms in the leaves. In most cases, humans make the initial diagnosis visually [1-3]. However, monitoring crops over large areas is difficult. To address this issue, unmanned aerial vehicles (UAVs) can be utilized to capture digital photographs for analysis using image processing, pattern recognition, and automatic classification techniques, as previously developed. Textural characteristics of images are commonly employed in leaf image analysis [4]. This study aimed to develop a software and hardware system for detecting and diagnosing eucalyptus leaf diseases from

* Corresponding author: dilnoz134@rambler.ru

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

photographic images. Since 2012, convolutional neural networks (CNNs) have been the top performers in the renowned ImageNet image recognition competition. CNNs' success is largely attributed to their shared weight concept, allowing for fewer configurable parameters than their predecessor, the neocognitron [5,6]. The mosaic convolutional neural network partially discards associated weights and employs the same error distribution-based learning algorithm as neocognitron-like variants of a flat neural network [7-8].
The human mind receives vast quantities of data through different images, and recognizing individuals by their faces or describing different creatures and events is a simple task for us. However, this is a complex task for computers. In recent years, machine learning has significantly addressed this challenge [9-10]. Specifically, researchers have discovered that a neural model called a deep convolutional neural network (CNN) can achieve reasonable results in recognizing and classifying complex visual images [11-12].

2 Methods

The iterative method restores the blurry and noisy image [8].

Step 1 entails reading the RGB image and selecting a portion of it.

Step 2 simulates a practical image with motion blur or poor focus and noise. The blur is created by convolving the image with a Gaussian filter that is represented by a point spread function. Gaussian noise with a variance of V is added to the blurred image to introduce noise into the simulated image. The value of the variance V serves as the attenuation parameter.

Step 3 entails restoring the blurry and noisy image using the iterative point spread function. The resulting array is in the same format as the original image.

Step 4 involves the interactive analysis of the restoration process. The image changes with each iteration. The function is executed in a step-by-step mode to examine the image restoration process, allowing the result to be analyzed at each iteration. The resulting array includes four numeric arrays, the first of which represents a blurred image, the second is the same image in double format, the third is the outcome of the penultimate operation, and the fourth contains processing parameters. Thus, the preliminary results are employed in the subsequent iteration.

Step 5 involves the processing of the noise component. A processing option is employed to control the noise component, which is described through processing the noise component.

Step 6 is creating a binary image by modeling a simple image.

Step 7 entails simulating the blur. The simulation is obtained by applying the Gaussian filter and point spread function and convolving them with the original image. To process this, the image is created by a weighting function represented by an array and divided into parts. The blurred image is positioned in the central part, whereas undistorted pixels are situated at the edges of the image behind the dotted line [9-10].

Step 8. Generating a weighted array. A method for assigning a specific weight to each pixel is employed to restore the image. It should be noted that the precision of the point placement is crucial when executing this technique.

Step 9. Generating the point extent function. The reconstructed image is produced using the modeled point extent function.

Convolutional Neural Networks (CNNs) are an advanced model architecture utilized for image classification problems. To extract and learn higher-level features, CNNs employ a series of filters that are applied to the raw pixel data of an image, and the resulting model can be used for classification. CNN comprises three primary components [21-22]:

In general, a CNN comprises a series of convolutional modules responsible for extracting features. Each module includes a convolutional layer followed by a pooling

layer. The final convolutional module is succeeded by one or more dense layers that carry out the classification process (refer to Figure 1). In the last dense layer of the CNN, there is a single node dedicated to each target class in the model, and an activation function is used to generate a value ranging between 0 and 1 for each node. These activation function values for a given image can be interpreted as relative indications of the likelihood of the image belonging to each target class [23].

To expedite the training process of CNNs, the sequential machine can be run at a higher speed, and the cortical process can be parallelized across each map, along with the reverse cortex of error distribution throughout the network. This parallelization leads to faster CNN training times [3-5].

Model of neuron

$$NET = \sum_{i=1}^{n} w_i * x_i + w_0, \tag{1}$$

here w_i is *i*-weight of neuron;

x_i is neuron output;

w_0 is auxiliary parameter, offset;

n is number of synaptic connections included in the neuron.

Determining the network topology is a task that should be accomplished based on research data and personal expertise [6]. Several steps influence the selection of the topology, such as defining the problem to be solved by the neural network (classification, forecasting, modification), identifying constraints on the problem (response speed, accuracy of the answer), and identifying incoming (e.g., image type, sound, size: 100x100, 30x30, format: RGB, grayscale) and outgoing (number of classes) data.

The task to be addressed using a neural network is classifying fundus images. The network's constraints are response speed (maximum of 1 second) and recognition accuracy (minimum of 70%). Input data consists of color JPEG images.

$$f(p, \min, \max) = \frac{p - \min}{\max - \min}, \tag{2}$$

here *f is* normalization function*;*

p is value of the transparent color of the pixel from 0 to 255*;*

min is minimal value of the pixel – 0*;*

max is maximal value of the pixel – 255.

A layer of maps is comprised of a collection of character maps, which are essentially basic matrices. Each map is equipped with a synaptic core, which is also referred to as a scan kernel or filter, depending on the source [8]. The dimensions of all maps in the convolutional layer are uniform and determined using the formula [7]:

$$(w, h) = (mW - kW + 1, mH - kH + 1) \tag{3}$$

here*(w, h)* is calculated size of the convolutional map;

mW is weight of the previous map;

mH is height of the previous map;

kW is weight of the core;

kH is height of the core.

Performs a scrolling operation according to the following formula used in image processing [24-25]:

$$(f * g)[m,n] = \sum_{k,l} f[m-k,n-l] * g[k,l], \tag{4}$$

here

f is specified image matrix;

g is convolution kernel.

During the scanning process of the previous layer's map with the sublayer's core (filter), there is no intersection between the scanned core and the layer's map. Usually, each map consists of a 2x2 core, which reduces the previous layer's maps by 2x. All the character maps are divided into cells of 2x2 elements, from which the maximum value is selected [26-27]. The activation function commonly used at the sub-selection level is the Rectified Linear Unit (ReLU). Additionally, the maximum value is selected using an operation known as Max Pooling. The layer can be mathematically expressed by using the following formula [7]:

$$x^l = f(a^l * subsample(x^{l-1}) + b^l), \tag{5}$$

here x^l is layer output;

$f()$ is activation function;

a^l, b^l are layer displacement coefficients;

Subsample ()–operation of selecting the local maximum value.

Fully glued layer

The final layer in the CNN architecture is a fully connected layer, also known as a multi-layer perceptron layer. This layer aims to classify a complex nonlinear function and improve recognition quality by optimizing it.

The training process involves selecting weights that minimize the difference between the predicted output and the object's actual properties. This is achieved by performing neuro-approximations and minimizing the difference between the predicted and actual outputs using various optimization techniques [27].

$$y = f_j(x_1, x_2, ..., x_n).$$

The recurrent relations system used to minimize the criterion is commonly used in neural network theory for training.

$$E = \frac{1}{2}\sum_{j=1}^{M}\left(y_j - \hat{y}_j\right)^2 \to \min \tag{7}$$

$$w_{jp}(t+1) = w_{jp}(t) - \mu\frac{\partial E_t}{\partial w_{jp}(t)}, \tag{8}$$

$$c_i^{jp}(t+1) = c_i^{jp}(t) - \eta\frac{\partial E_t}{\partial c_i^{jp}(t)}, \tag{9}$$

$$b_i^{jp}(t+1) = b_i^{jp}(t) - \eta\frac{\partial E_t}{\partial b_i^{jp}(t)}, \tag{10}$$

$$j = \overline{1,m}, i = \overline{1,n}, p = k_j, \tag{11}$$

here:

$\hat{y}_j$ and y_j are theoretical and empirical representations of the object (6) at the j-th stage of training;

$w_j^p; c_i^{jp}, b_i^{jp}$ are weights of the rules (w) and parameters of the corresponding functions at the t-step of training (b, c);

η is training parameter that can be selected according to the recommendations;

$\overline{d}_j$ - $d_j \in \left[\underline{y}, \overline{y}\right]$ is center of the class.

Trained rule weights:

$$w_{jp}(t+1) = w_{jp}(t) - \mu\left(y_t - \hat{y}_t\right) \frac{\overline{d}_j \sum_{j=1}^{m} \mu^{d_j}(y) - \sum_{j=1}^{m} \overline{d}_j \mu^{d_j}(y)}{\left(\sum_{j=1}^{m} \mu^{d_j}(y)\right)^2} w_{jp} \prod_{i=1}^{n} \mu^{jp}(x_i). \quad (12)$$

Similar to the rule-based system, the training process for a fuzzy neural network also involves two steps. The first step involves computing the object's output as a model value corresponding to the given network architecture.

3 Results and Discussion

The dataset underwent offline augmentation to recreate it from the original dataset. It comprises around 87,000 RGB images of diseased and healthy crop leaves categorized into 38 distinct classes. The dataset is segregated into training and validation sets at an 80/20 ratio while keeping the directory structure unchanged. A new directory consisting of 33 test images is created later for prediction purposes.

Classes:

1. Apple, Apple scab
2. Apple, Cedar apple rust
3. Apple healthy
4. Blueberry healthy
5. Cherry (including sour), Powdery mildew
6. Cherry (including sour), healthy
7. Corn (maize), Cercospora leaf spot Gray leaf spot Corn (maize), Common rust
8. Corn (maize), Northern Leaf Blight
9. Corn (maize) healthy
10. Grape, Blackrot
11. Grape, Esca (Black Measles))
12. Grape, Leaf blight (Isariopsis Leaf Spot)
13. Grape healthy
14. Orange, Haunglongbing (Citrus greening)
15. Peach, Bacterialspot
16. Peachhealthy
17. Pepper, bellBacterialspot
18. Pepper, bellhealthy
19. Potato, Early blight
20. Potato, Late blight
21. Potato healthy
22. Raspberry healthy
23. Soybean healthy
24. Squash, Powdery mildew
25. Strawberry, Leaf scorch

26. Strawberry healthy
27. Tomato, Bacterialspot
28. Tomato, Early blight
29. Tomato, Late blight
30. Tomato, LeafMold
31. Tomato, Septoria leaf spot
32. Tomato, Spider mites Two-spotted spider mite,
33. Tomato, Target Spot
34. Tomato, Tomato Yellow Leaf Curl Virus
35. Tomato, Tomato mosaic virus
36. Tomato, healthy

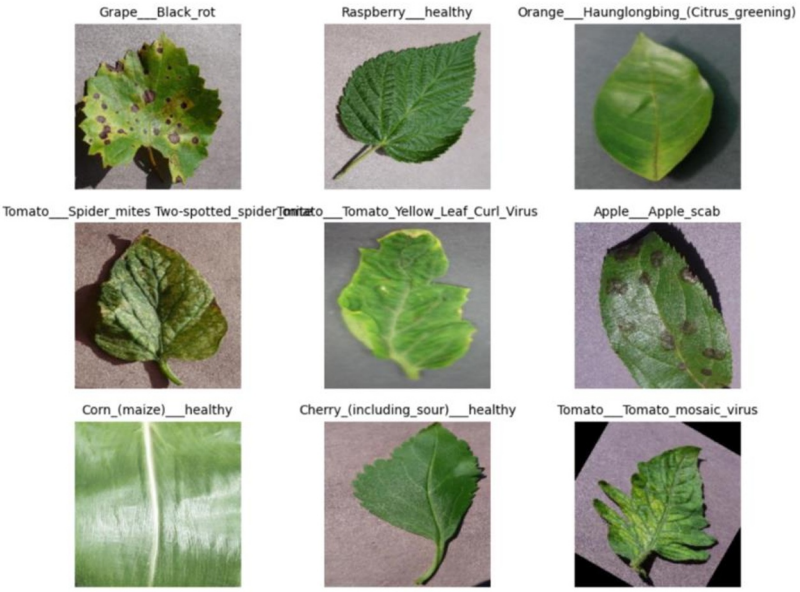

Fig. 1. Examples from dataset.

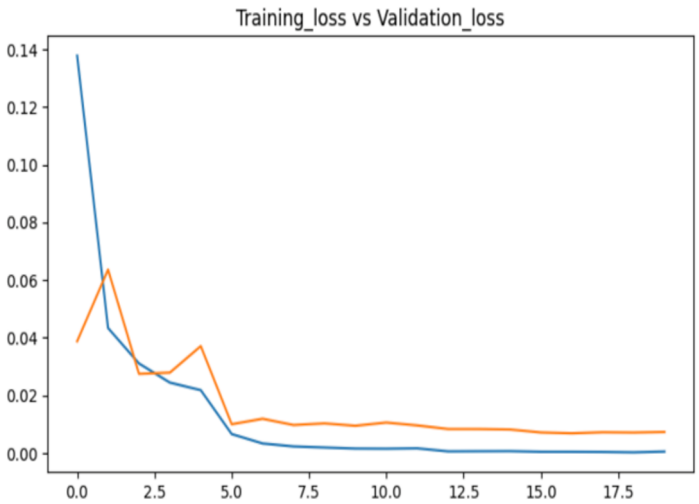

Fig. 2. Examples of results obtained.

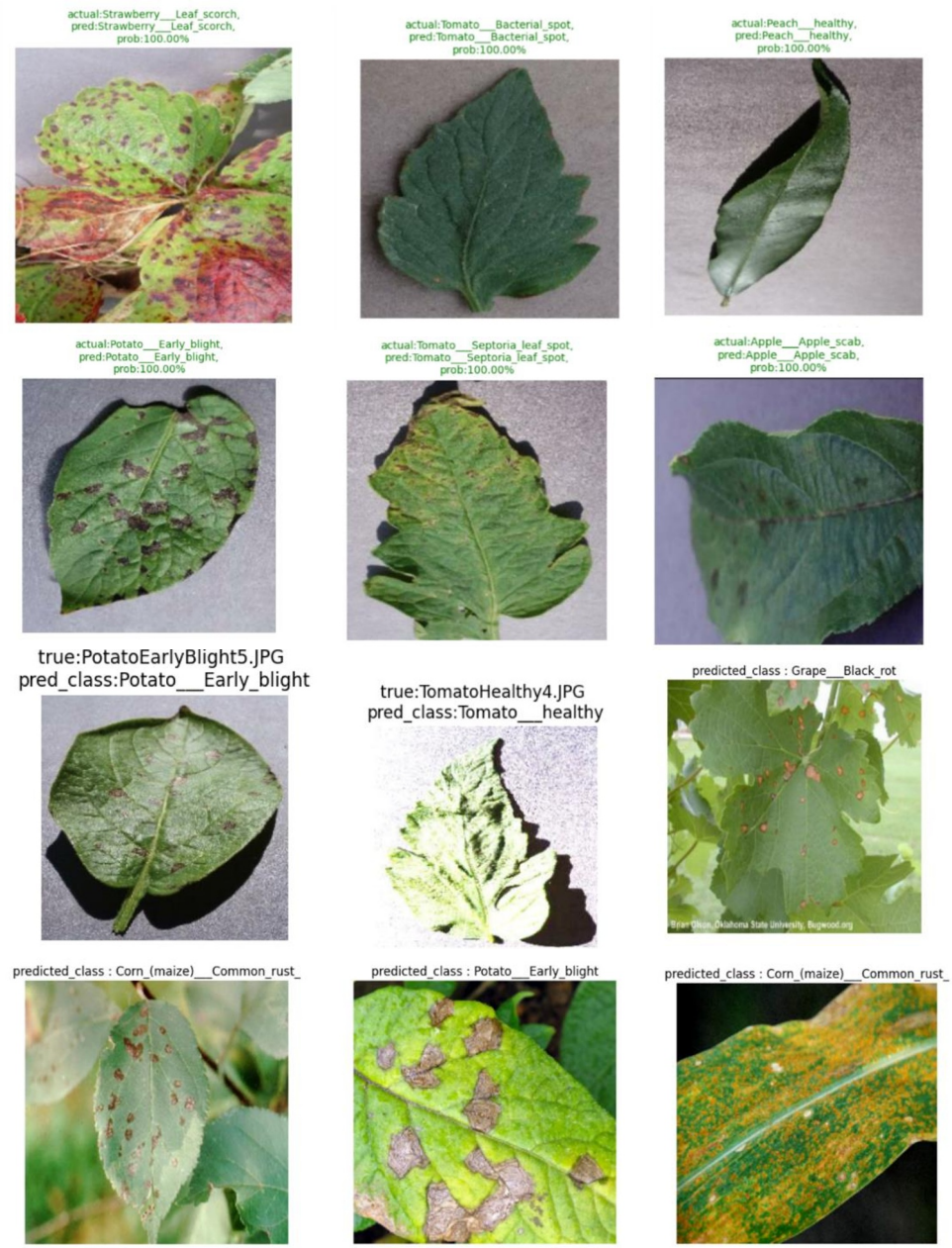

Fig. 3. Examples from dataset.

4 Conclusion

Timely identification of problems in a field is crucial to avoid wasting time and resources. Visual inspection of plants is a crucial aspect of plant health management, and growers rely on symptoms to make decisions about their crops. However, this process is time-consuming and often leaves growers uncertain. Laboratory tests are impractical for daily diagnosis due to their high costs and turnaround times. Crop consultants can assist growers in decision-making and ensure minimal errors, but not all growers can access such assistance. In such cases, AI-based plant disease detection applications can be helpful tool. With the advancements in AI research, these tools can quickly identify problems and provide growers with instant diagnoses. The recorded images can also be used as a reference in the future, making the tool a useful guide for growers who find it difficult to make accurate diagnoses. This technology is particularly important in countries where the ratio of producers to agronomists is high, and other sources of advice are limited.

References

1. P. Narayanasamy, Microbial Plant Pathogens Detection and Disease Diagnosis (Fungal Pathogens. Springer), 291, (2011)
2. M. B. Riley, M. R. Williamson, O. Maloy, Plant disease diagnosis (PHI), (2002)
3. M. McNeil, A. M. Roberts, V. Cockerell, V. Mulholland, Real-time PCR assay for quantification of Tilletia caries contamination of UK wheat seed (Plant Pathology), **53**, 741, (2004)
4. N. P. Cherepanova, Sistematika gribov: ucheb. posobie (2-e izd. SPb.: Izd-voS.-Peterb. un-ta), 344, (2005)
5. Smut bunt diseases of cereal biology, identification and management. Government of Western Australia Paplomatas E J Molecular diagnostics of fungal pathogens (Arab. J. Pl. Prot.) **24**, 147, (2006)
6. S. I. Bityukov, A. V. Maksimushkina, V. V. Smirnova, Comparison of histograms in physical research (Dated: 24 May 2016), (2016)
7. E. Duveiller, R. P. Singh, P. K. Singh, A. A. Dababat, M. Mezzalama, Wheat diseases and pests: a guide for field identification (Mexico: CIMMYT), 138, (2012)
8. B. Francesco, H. Richard, S. Paul, F. Antonio, Theoretical and experimental comparison of different approaches for colour texture classification (Dated: July 11, 2011), (2011)
9. G. Jayme, B. Arnal, Digital image processing techniques for detecting, quantifying and classifying plant diseases (Barbedo Springer Plus 2013, 2:660), (2013)
10. A. Vadivel, S. Shamik, A. K. Majumdar, An Integrated Color and Intensity Co-occurrence Matrix Communicated by R. Manmatha (29 March 2005), (2005)
11. K. Matkovic. et al., Global Contrast Factor – a New Approach to Image Contrast Computational Aesthetics, 159, (2005)
12. S. A. Kadnichanskiy, Assessment of the contrast of digital aerial and satellite images Geodesy and Cartography (No 3), 46, (2018)
13. V. V. Starovoitov, Refinement of the index of structural similarity of images SSIM Informatics (No. 3) **15**, 7, (2018)
14. S. A. Golestaneh, D. M. Chandler, No-reference quality assessment of JPEG images via a quality relevance map IEEE Signal Processing Letters (No 2) **21**, 155, (2014)

15. A. Graves, et al., Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks Proceedings of the 23rd international conference on Machine learning (ACM), 369, (2006)
16. Th. Bluche, Deep Neural Networks for Large Vocabulary Handwritten Text Recognition PhD thesis. Université Paris Sud-Paris XI, 268, (2015)
17. P. Doetsch, K. Michal, N. Hermann, Fast and robust training of recurrent neural networks for offline handwriting recognition Frontiers in Handwriting Recognition (ICFHR), 14th International Conference on. IEEE, 279, (2014)
18. A. Graves, S. Jürgen, Offline handwriting recognition with multidimensional recurrent neural networks Advancesinneural information processing systems, 545, (2009)
19. Th. Bluche, N. Hermann, K. Christopher, A Comparison of Sequence-Trained Deep Neural Networks and Recurrent Neural Networks Optical Modeling for Handwriting Recognition International BIBLIOGRAPHY 153 Conference on Statistical Language and Speech Processing (Springer), 199, (2014)
20. Th. Bluche et al., The asiaarabic handwritten text recognition system at the open hart2013 evaluation Document Analysis Systems (DAS), 11th IAPR International Workshop on. IEEE, 161, (2014)
21. Z. Sh. Juraev, D. T. Muhamediyeva, D. M. Sotvoldiev, IOP Conf. Journal of Physics: Conf. (Series 1546(1)012083), (2020)
22. D. Muhamediyeva, D. Sotvoldiyev, S. Mirzaraxmedova, M. Fozilova, Approaches to handwriting recognition International Conference on Information Science and Communications Technologies, ICISCT 2020, (9351505. 4-6 november. DOI: 10.1109/ICISCT50599.2020.9351505), (2020)
23. Sh. R. Farmonov, T. F. Bekmuratov, D. T. Muhamedieva, About the dodges plans of the continuous selective control International Conference on Information Science and Communications Technologies, ICISCT 2020 (9351415. 4-6 november. DOI: 10.1109/ICISCT50599.2020.9351415), (2020)
24. K. Mirzayan, M. Dilnoz, S. Barno, The Problem of Classifying and Managing Risk Situations in Poorly Formed Processes Aliev R.A., Yusupbekov N.R., Kacprzyk J., Pedrycz W., Sadikoglu F.M. (eds) 11th World Conference "Intelligent System for Industrial Automation" (WCIS-2020) (WCIS 2020. Advances in Intelligent Systems and Computing, Springer, Cham. https://doi.org/10.1007/978-3-030-68004-6_36) **1323**, 280, (2021)
25. D. T. Mukhamedieva, Journal of Physics: Conference Series (1546(1) 012091), (2020)
26. D. K. Muxamediyeva, IOP Conf.Journal of Physics: Conf. Series (1210), (2019)