

ASSESSMENT OF PLANT DISEASE DETECTION BY DEEP LEARNING

Akan Alpyssov

Candidate of Pedagogical Sciences, Assistant Professor
Graduate School of Natural Science
Pavlodar Pedagogical University
Mir str., 60, Pavlodar, Republic of Kazakhstan, 140000

Nurgul Uzakkyzy

Corresponding author
Doctor of PhD, Associate Professor
Department of Computer and Software Engineering*
E-mail: nura_astana@mail.ru

Ayazbaev Talgatbek

Candidate of Physical and Mathematical Sciences, Associate Professor
Department of Information and Communication Technologies
International Taraz Innovative Institute
Zheltoksan str., 69B, Taraz, Republic of Kazakhstan, 080000

Raushan Moldasheva

Doctoral Student in the Specialty "8D06101 Big Data Analytics"
Department of Information Systems
S. Seifullin Kazakh Agrotechnical University
Zhenis ave., 62, Astana, Republic of Kazakhstan, 010011

Gulmira Bekmagambetova

Doctor PhD, Assistant Professor
Department of Information Technologies
Kazakh University of Technology and Business
K. Mukhamedkhanova str., 37A, Astana, Republic of Kazakhstan, 010000

Mnyaura Yessekeyeva

Candidate of Physical and Mathematical Sciences, Associate Professor
Department of Information Systems and Technologies
Esil University
Zhubanov str., 7, Astana, Republic of Kazakhstan, 010000

Dossym Kenzhaliev

Candidate of Physical and Mathematical Sciences,
Associate Professor of Physics
Department of General and Theoretical Physics*

Assel Yerzhan

Doctor in Radio Engineering, Electronics and Telecommunications
Department of Telecommunications and Innovative Technologies
Almaty University of Power Engineering and Telecommunications
named after Gumarbek Daukeyev
Baytursynuli str., 126/1, Almaty, Republic of Kazakhstan, 050013

Ailanysh Tolstoy

Master of Radio Engineering, Electronics and Telecommunications
Department of Radio Engineering, Electronics and Telecommunications*
*L. N. Gumilyov Eurasian National University
Satpayev str., 2, Astana, Republic of Kazakhstan, 010008

Plant disease and pest detection machines were originally used in agriculture and have, to some extent, replaced traditional visual identification. Plant diseases and pests are important determinants of plant productivity and quality. Plant diseases and pests can be identified using digital image processing. According to the difference in the structure of the network, this study presents research on the detection of plant diseases and pests based on three aspects of the classification network, detection network, and segmentation network in recent years, and summarizes the advantages and disadvantages of each method. A common data set is introduced and the results of existing studies are compared. This study discusses possible problems in the practical application of plant disease and pest detection based on deep learning.

Conventional image processing algorithms or manual descriptive design and classifiers are often used for traditional computer vision-based plant disease and pest detection. This method usually uses various characteristics of plant diseases and pests to create an image layout and selects a useful light source and shooting angle to produce evenly lit images.

The purpose of this work is to identify a group of pests and diseases of domestic and garden plants using a mobile application and display the final result on the screen of a mobile device. In this work, data from 38 different classes were used, including diseased and healthy leaf images of 13 plants from plantVillage. In the experiment, Inception v3 tends to consistently improve accuracy with an increasing number of epochs with no sign of overfitting and performance degradation. Keras with Theano backend used to teach architectures

Keywords: image processing, Inception v3, deep learning, classification, plant diseases, clustering

Received date 03.12.2022

Accepted date 13.02.2023

Published date 28.02.2023

How to Cite: Alpyssov, A., Uzakkyzy, N., Talgatbek, A., Moldasheva, R., Bekmagambetova, G., Yessekeyeva, M., Kenzhaliev, D., Yerzhan, A., Tolstoy, A. (2023). Assessment of plant disease detection by deep learning. *Eastern-European Journal of Enterprise Technologies*, 1 (2 (121)), 41–48. doi: <https://doi.org/10.15587/1729-4061.2023.274483>

1. Introduction

Deep learning has received a lot of attention recently with the goal of developing a fast, automatic, and accurate image identification and classification system. This work focused on fine-tuning and evaluating state-of-the-art deep

convolutional neural networks for image-based plant disease classification. Deep learning or deep structured learning or hierarchical learning is a branch of machine learning that attempts to model high-level abstractions in data based on a set of algorithms [1, 2]. Such algorithms create a multi-level hierarchical architecture of learning and data presentation.

Based on [2, 3], DL algorithms are useful when dealing with large amounts of unsupervised data. Problems that were considered unsolvable are now being solved with inhuman precision. Image classification is a prime example of this. The performance of traditional machine learning algorithms stabilizes when the training data limit is reached, while deep learning increases performance as the data grows.

Traditional data analysis methods such as linear regression [4] and linear discriminant analysis [5] are based on predefined distributions and model assumptions. These methods are applicable without loss of accuracy only for data that meets these requirements. In this study, let's focus on developing plant disease detection methods using traditional machine learning methods. Several methods are evaluated: support vector machines, decision tree, random forest, and Naïve Bayes. Usually, when examining plants with the naked eye, they use information about color (leaves with diseases usually change color) or about the presence of spots or rotten areas on the leaves. Therefore, let's investigate various image processing functions that capture color information or a local descriptor. The evaluating feature extraction methods include RGB, Scale Invariant Feature Transformation [6], Accelerated Reliable Features [7], Oriented FAST and Rotated BRIEF [8], Histogram Oriented Gradients [9]. The features are selected because they contain color, local features, or object detector information from the image data. Therefore, the scope of such analysis methods is limited.

In this work, machine learning methods such as Decision Tree, Random Forest, Naïve Bayes and deep learning methods such as the Inception v3 method and VGG19 were used to classify the data set. The existence of an automated computer system for the detection and diagnosis of plant diseases could be a valuable aid to plant growers who are asked to make such a diagnosis by optical observation of the leaves of infected plants [10]. If the system were easy to use and easily accessible through a simple mobile app, it could also be a valuable tool for houseplant growers in parts of the world that lack the infrastructure to provide phytopathological advice. Causes of diseases can be biotic (bacteria, fungi, viruses, and nematodes) and abiotic (effects of temperature, humidity, and nutrient deficiencies). Some of the causes of these diseases are interconnected with each other. For example, plants that lack nutrients are prone to diseases caused by bacteria.

Indoor and garden plants have a great influence on human health, some of them bear fruit, and some are used as medicines. Therefore, knowing and evaluating their current state, and knowing and predicting the type of disease is still an urgent problem.

2. Literature review and problem statement

Decision tree classifiers are regarded to be a standout of the most well-known methods to data classification representation of classifiers. Different researchers from various fields and backgrounds have considered the problem of extending a decision tree from available data, such as machine study, pattern recognition, and statistics. In various fields such as medical disease analysis, text classification, user smartphone classification, images, and many more, the employment of Decision tree classifiers has been proposed in many ways.

In paper [11] provides a detailed approach to the decision trees. Furthermore, paper specifics, such as algorithms/ap-

proaches used, datasets, and outcomes achieved, are evaluated and outlined comprehensively. In addition, all of the approaches analyzed were discussed to illustrate the themes of the authors and identify the most accurate classifiers. As a result, the uses of different types of datasets are discussed and their findings are analyzed. This work has been a thorough analysis of the decision tree method. But the authors did not consider the comparison between other machine learning methods.

In this [12] article, the author's considered methods for analyzing aerospace images and showed the state of crops in different growing seasons. In particular, methods of orthogonal transformations were applied to aerospace images. The authors did not determine the effectiveness of other methods other than those indicated in the article.

This article [13] proposes an efficient selective Naïve Bayes algorithm that uses only some attributes to build selective Naive Bayes models. These models are built in such a way that each is a trivial extension of the other. The most predictive selective Naïve Bayes model can be selected using stepwise cross-validation measures with one exception. As a result, attributes can be selected by efficient model selection. Empirical results show that the selective Naïve Bayes method demonstrates excellent classification accuracy, while at the same time maintaining simplicity and efficiency. This paper provides a complete overview of Naïve Bayes, but the authors did not consider more efficient methods of model selection.

This article [14] presents an ensemble deep learning approach for a defined classification of non-carcinoma and carcinoma histopathological images of breast cancer using our collected dataset. The authors trained four different models based on pre-trained VGG16 and VGG19 architectures. Initially, the authors performed a 5-fold cross-validation of all individual models, namely the fully trained VGG16. Then, by taking the average of the predicted probabilities, it was found that an ensemble of finely tuned VGG16s showed a competitive classification, especially for the carcinoma class. However, the dataset collected by the authors is relatively small compared to the datasets used in numerous current studies. Also, the dataset contains only images of two classes.

In this article [15], the authors explored the aspects of reducing the time and effort of the user by recommending the best product to it. To do this, this article proposes a Naive Bayes classifier that accurately labels reviews and combines reviews to give a final score to a product. The performance of the model is evaluated and the results are analyzed. But the effectiveness of other traditional methods of machine learning has not been considered.

In this paper [16], a transfer learning model is implemented to identify unmasked person detection processes. A model proposed by the authors to build an algorithm to refine the preprocessing of the modern InceptionV3 deep learning model. An image scaling technique used to solve problems with limited data availability for better evidence and model testing. The model outperformed other proposed methods, achieving 99.9 % probability during training and 100 % during testing. The work does not use large amounts of data and has not been expanded to classify the type of mask and implement a face recognition system deployed at various workplaces to support the identification of a person while wearing a mask.

In this article [17], the authors propose the Inception V3 subtitle generator model using convolutional neural net-

works and long-term memory modules. The InceptionV3 model was trained on 1000 different classes on the ImageNet dataset. The model was imported directly from the Keras application module. Remove the final classification layer of the dimension vector (1343) from the InceptionV3 model. The embedded array is used for dictionary links. The constructive matrix is a linear transformation of the original space with important relationships in real space. Subtitles are widely used and are important, for example, to implement human-computer interaction. But the authors did not consider the comparison of other deep learning methods with the Inception V3 method used.

A feature of this work is the use of deep learning methods to create a cross-platform application for plant growing enthusiasts. In particular, the effectiveness of these methods is to determine the state of the plant and the application of measures to destroy pests, which leads to plant productivity.

3. The aim and objectives of the study

The aim of the study is to evaluate the detection of plant diseases and pests using deep learning methods.

To achieve this aim, the following objectives are solved:

- to analyze traditional methods of machine and deep learning to identify diseases and pests of plants;
- to implement a mobile application for the detection of plant diseases for growers.

4. Materials and methods

In this work, the object of study is plants and their state in real time. The traditional approach is to use well-established SV methods such as feature descriptors (SIFT, SURF, BRIEF, etc.) to detect objects. Before the advent of DL, for tasks like image classification, there was a step called feature extraction. Features are small “interesting”, descriptive or informative spots on images. Several computer vision algorithms can be used in this step, such as edge detection, corner detection, or threshold segmentation. The maximum possible number of features are extracted from the images, and these features form a definition (known as a set of words) for each class of objects. During the deployment phase, these definitions are looked up in other forms. The image is classified as containing this particular object, i.e. abnormal values in the images, which are identified as plant diseases in this work.

The difficulty with this traditional approach is that you have to choose which features are important for each image. As the number of classes for classification increases, it becomes increasingly difficult to obtain estimates. Deciding which features best describe different classes of objects in deep learning depends on pre-trained data and a lengthy process of trial and error. Based on the collected data, the identification of deviations in plants by traditional methods was trained according to the model shown in the figure below (Fig. 1).

As shown in Fig. 1, the dataset was trained by traditional machine learning methods on the training set and the test dataset was validated according to the result of the trained model. For example, one of the traditional methods is shown – the decision tree architecture (Fig. 2).

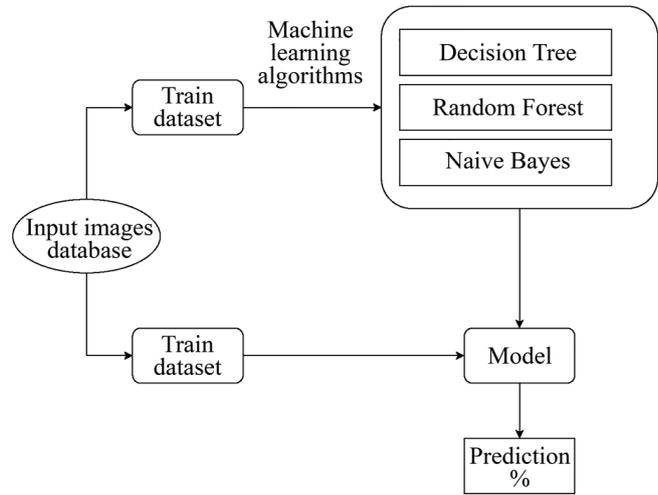


Fig. 1. Learning model based on traditional machine learning method

The decision tree builds classification or regression models in the form of a tree or hierarchical structure (Fig. 2). It splits the data set into smaller and smaller subsets, at the same time, the corresponding decision tree is gradually developed, that is, the final result of the classification will be in the form of leaves of the tree, and this will be our decision (Fig. 3).

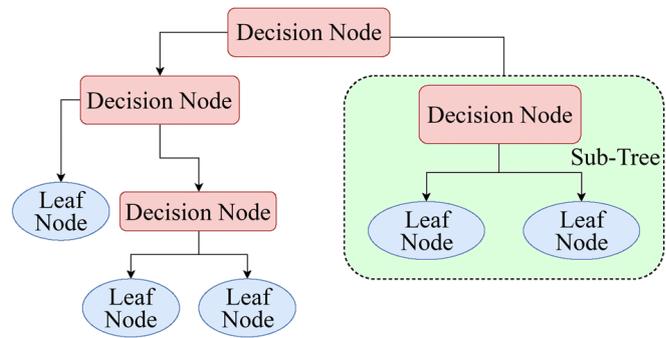


Fig. 2. Traditional Machine Learning Architecture Decision Tree

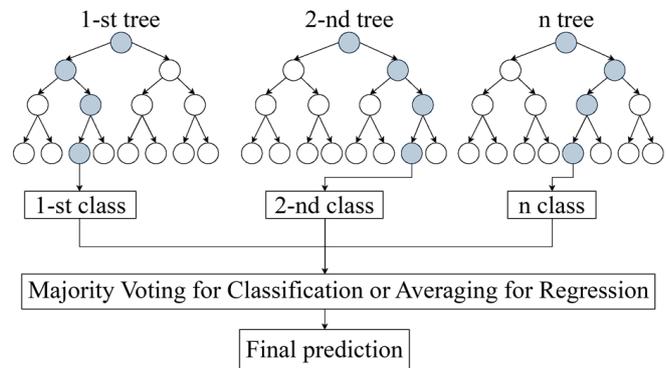


Fig. 3. Machine learning architecture based on traditional random forest method based on decision tree

A random forest is a classifier that consists of several decision trees, and when classifying, it is also divided into several subsets that ultimately form the result. But at the very end of the random forest, several answers are displayed, which pass the majority vote for classification or averaging for regres-

sion (Fig. 3). At the end, the answer with the highest vote value is formed. The Naïve Bayes classifier is a probabilistic machine learning model that is used for a classification problem. The essence of the classifier is based on the Bayes theorem (1) [19].

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

Using Bayes' theorem, it is possible to find the probability that *A* will happen, given that *B* has happened. Here *B* is the evidence and *A* is the hypothesis. The assumption here is that the predictors/features are independent. That is, the presence of one particular feature does not affect the other. When used for classification, as well as in general, machine learning requires pre-trained feature data. In this work, the deep learning architecture used is shown in Fig. 4.

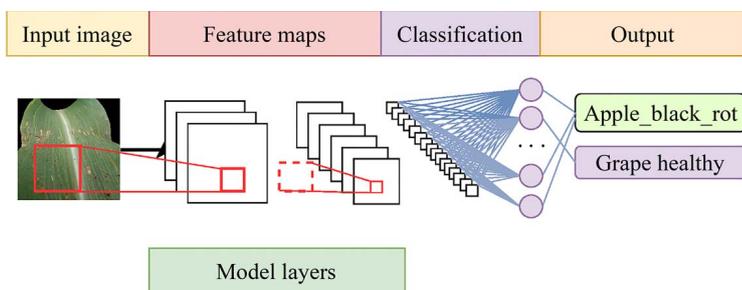


Fig. 4. Deep learning model with convolutional neural networks

Structured numerical data that helps to train the model for future use, while deep learning does not require any human intervention, when training data that has been divided for training, methods detect features on their own using layers and filters.

5. Results of the application of machine and deep learning methods for assessing the detection of plant diseases

5.1. Analysis of traditional methods machine and deep learning methods for the detection of plant diseases

In this work, VGG19 was used as a convolutional neural network, which is a good neural network architecture, but it cannot handle complex tasks well, since it is a simple set of convolutional and maximum pooled layers, following each other and finally completely connected layers (Fig. 5). Simply put, it cannot extract very complex features.

On the other hand, initial networks have initial modules consisting of 3x3 filters, also known as point convolutions, followed by convolutional layers with filters of different sizes applied simultaneously. This allows initial networks to learn more complex features. They have more hidden layers compared to VGG19. Hence, they are used for more complex tasks.

Inception v3 consists of symmetric and asymmetric building blocks, including convolutions, pooling of averages, pooling of maximums, concatenations, exclusions, and fully connected layers (Fig. 6). Batch normalization is widely used throughout the

model and applied to activation inputs. The Softmax activation function allows the model to make probabilistic predictions for each class, which is useful in classification tasks, that is, which class an image belongs to and makes it possible to determine plant species and their diseases.

In total, the initial V3 model consists of 42 layers, which is slightly higher than the previous initial V1 and V2 models.

The training dataset contains 23,102 pre-trained image sets. Including apples – 2215, cherries – 1717, corn – 1834, grapes – 2419, oranges – 1576, peaches – 1347, bell peppers – 1479, potatoes – 1793, raspberries – 1946, strawberries – 1983, soybeans – 1790, tomatoes – 1791, indoor plants – 2590 images. Healthy and diseased species of fruits, vegetables (Fig. 7) and houseplants (Fig. 8) were examined. For example, apple scab, black rot, cedar apple rust, corn leaf spot, gray leaf spot, common rust, northern leaf spot, grape black rot, esca (black beet) for grapes, grape leaf spot (isariopsis).

To treat the disease, fruit leaves can be processed in the first period of ripening. Fig. 7 shows healthy and diseased varieties of apples, the most common types of fruits in every garden.

This article reviews healthy and diseased houseplants for those interested in growing houseplants (Fig. 7). During the experiment, an analysis of the results obtained and a comparison between the methods of traditional machine learning and deep learning were carried out. Table 1 below shows the result of the average accuracy of determining diseases of indoor and garden plants (Table 1).

The performance of the model (Fig. 6) is really impressive. It has a deeper network compared to traditional machine learning methods, and its speed does not suffer. The result of the experiment (Table 1) showed a high accuracy of object classification compared to traditional machine learning methods.

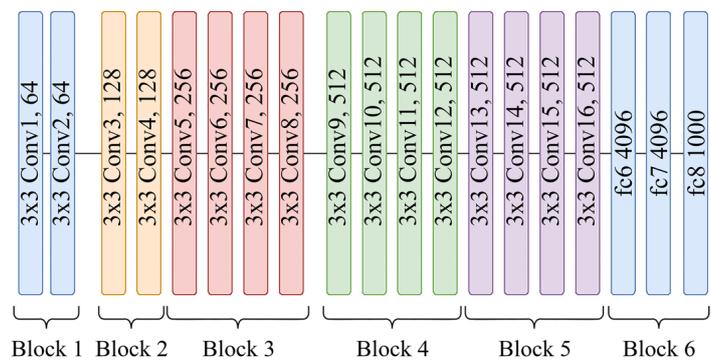


Fig. 5. VGG19 Convolutional Neural Network Architecture

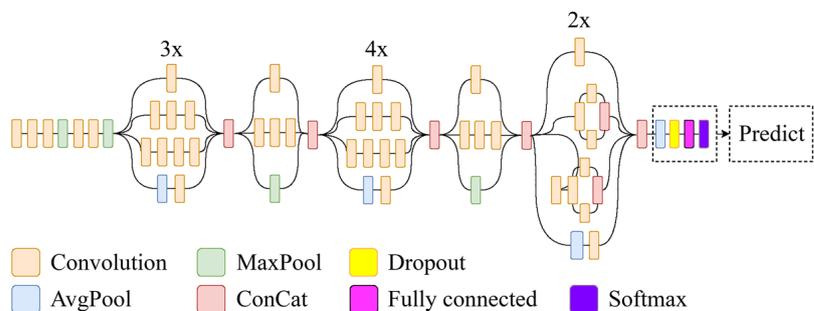


Fig. 6. Convolutional Neural Network Architecture Inception v3

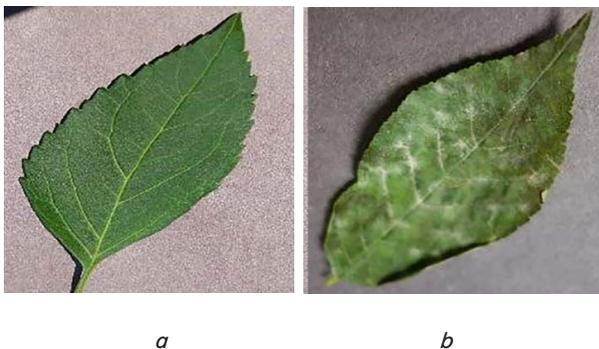


Fig. 7. Varieties of apple leaves: *a* – healthy; *b* – diseased

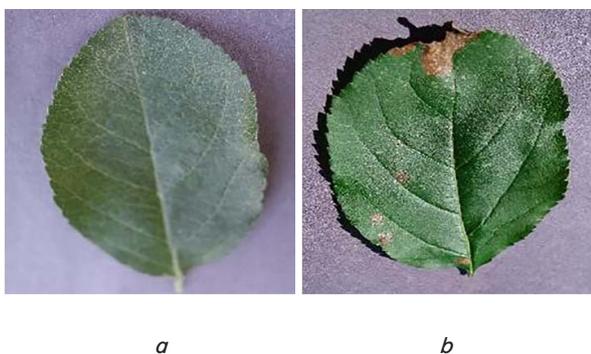


Fig. 8. Types of houseplant leaves: *a* – healthy; *b* – diseased

Table 1

The average accuracy of determining diseases of indoor and garden plants

Initial Images / Algorithms	Decision Tree	Random Forest	Naïve Bayes	VGG19	Inception v3
Apple	78 %	79 %	82 %	90 %	93 %
Sherry	77 %	80 %	84 %	92 %	95 %
Corn	79 %	78 %	82 %	90 %	90 %
Grapes	80 %	81 %	84 %	93 %	95 %
Orange	78 %	79 %	82 %	90 %	93 %
Peach	79 %	78 %	82 %	90 %	90 %
Bell pepper	80 %	81 %	84 %	93 %	95 %
Potatoes	78 %	79 %	82 %	90 %	93 %
Raspberry	77 %	80 %	84 %	92 %	95 %
Strawberry	79 %	78 %	82 %	90 %	90 %
Soy	80 %	81 %	84 %	93 %	95 %
Tomato	78 %	79 %	82 %	90 %	93 %

5. 2. Implementation of a mobile application for the detection of plant diseases for breeders

Based on the experiment, 20,000 sets of images were considered. To see the effectiveness of deep learning methods, they were trained against traditional machine learning classification methods. In this work, when developing a mobile application, TensorFlow was used for deep learning models and FireBase for data storage. Fig. 9 below shows the working architecture of the mobile application.

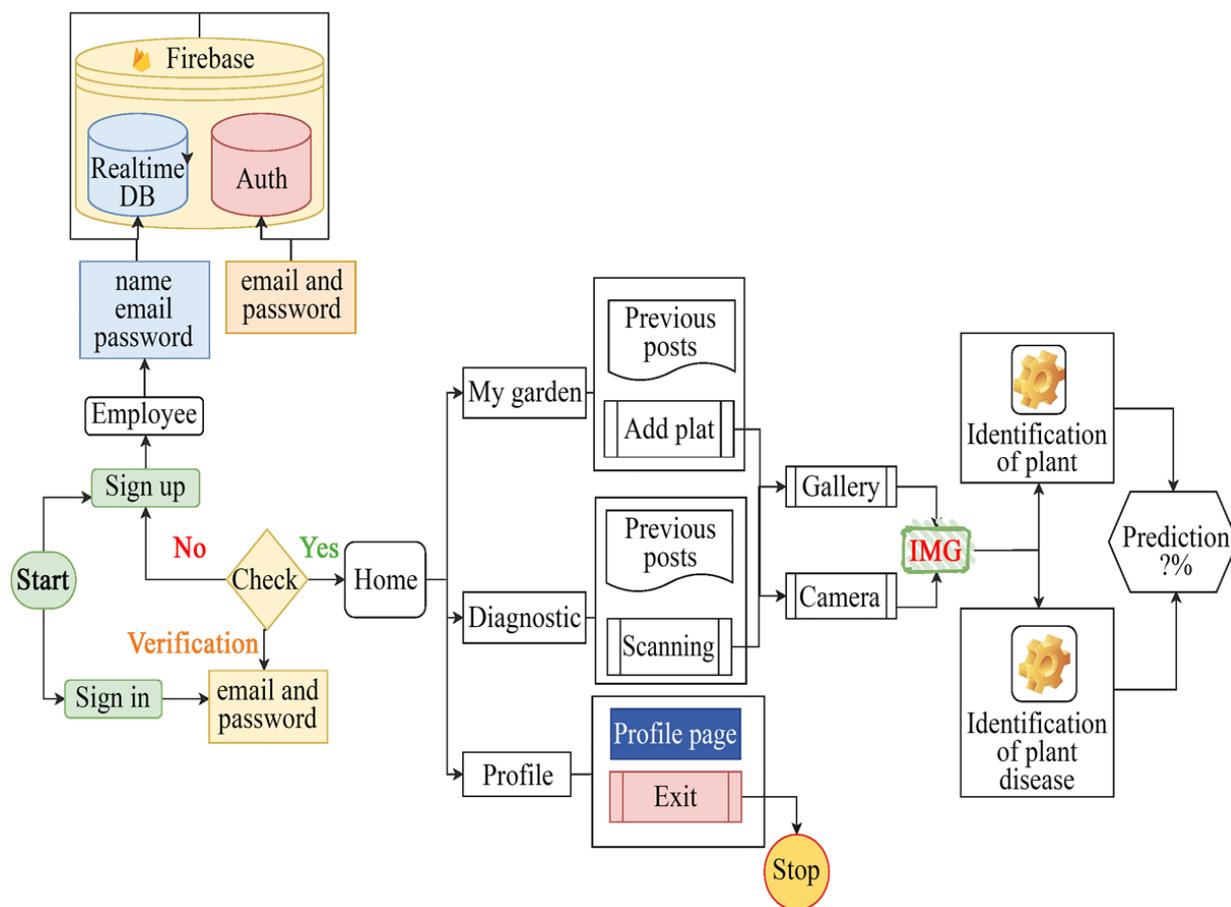


Fig. 9. The architecture of the mobile application

Each image was pre-processed according to the requirements of the methods. Fig. 10 below shows a graph of the accuracy and loss of the training and testing datasets considered in this paper.

below (Fig. 12) shows one of the apple tree diseases in the database.

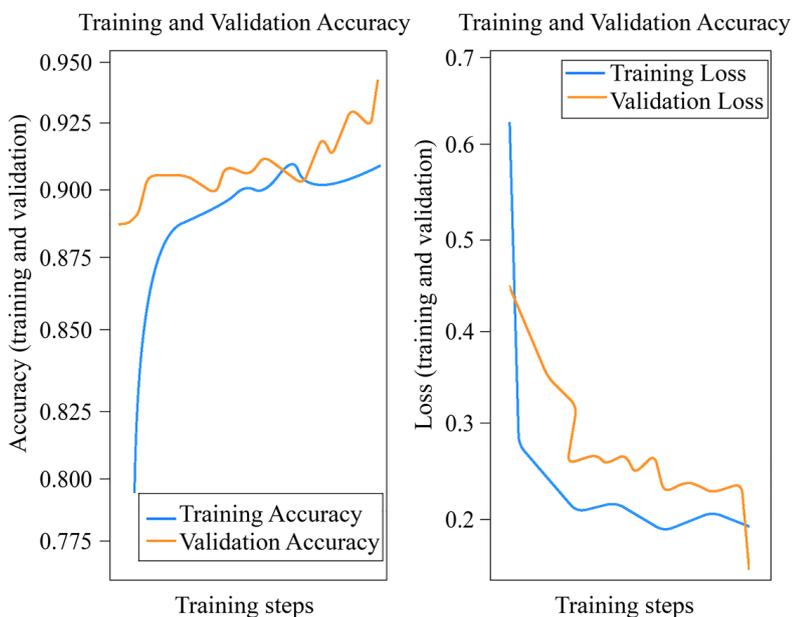


Fig. 10. Graph of accuracy and learning loss and Inception v3 testing

The accuracy of deep learning when training with the Inception v3 method was 92.5 %, and when testing it was 94 %. According to this model, the losses during training were about 8 %, and during testing – 6 %. Fig. 11 shows the interface of the developed mobile application.

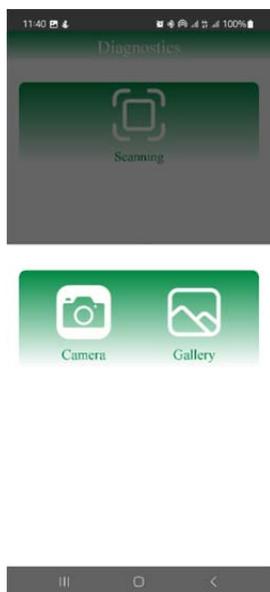


Fig. 11. Home page of the mobile application

As shown in Fig. 11, the user can check the status of the plants by uploading a picture from the gallery section of their mobile phone or by using the camera. The picture



Fig. 12. Definition of apple disease

The database covers the most common diseases of many garden fruits and vegetables. For example, one of them is apple cedar rust.

6. Discussion of the results of using algorithms and methods of machine learning for image processing

With the help of modern technologies, problems that one has to face daily are solved. One of them is monitoring the condition of plants growing in our garden. In this work, it is recommended to identify and prevent diseases during their growth in order to obtain a good harvest of fruits. To solve the tasks set, methods for classifying machine learning and deep learning methods were considered. In addition, several methods were compared with each other to check the accuracy of the methods (Table 1). The advantages and disadvantages of each method considered in this paper are taken into account. The advantages of decision trees are interpretability, fewer data preparation, generality, and nonlinearity. And the disadvantages of this method are overfitting, feature reduction, data resampling, and optimization.

Random Forest has the advantage of reducing overfitting in decision trees and helps improve accuracy, and flexibility for classification problems and for regression, works well with both categorical and continuous values and automates missing values present in the data. However, despite these advantages, the random forest algorithm also has some disadvantages, such as it requires a lot of computing power as well as resources, since it builds many trees to combine their results, it also takes a long time to train, since it combines many decision trees to class definitions.

Naïve Bayes methods have the advantage of being simple and easy to implement, do not require as much training data, handle both continuous and discrete data, and scale easily due to the number of predictors and data points. And the

disadvantages of the method are that all predictors (or signs) are independent and rarely occur in real life. This limits the applicability of this algorithm in real use cases.

This algorithm runs into the “zero frequency problem” when it assigns a probability of zero to a categorical variable whose category in the test dataset was not available in the training dataset. It would be better if you use dithering technique to solve this problem.

Taking into account the advantages and disadvantages of each considered method, the Inception v3 algorithm was built into the developed mobile application. An important limitation of this method is the need to include deep learning image processing methods to improve the accuracy of this method as a further modification of the application.

7. Conclusions

1. In this work, 23102 pre-trained datasets of many common garden fruits and vegetables were trained as normal plants and disease types. As image classification methods, the efficiency of machine learning methods Decision Tree, Random Forest, Naïve Bayes and deep learning methods such as VGG19, Inception v3 were analyzed. As a result of the experiment, the approximate accuracy of each model was determined.

2. The created mobile application is cross-platform, that is, it is possible to work on any gadget. Using a mobile application, the user can determine the type of plants and the presence of diseases in them from the photos stored in the gallery or using the camera in real time. The result of training the model using the Inception v3 algorithm was 93 %. It

is the most optimal among the considered methods. It can be treated in a timely manner according to the type of disease identified. The user can take appropriate action based on the result of the percentage determined in the mobile application. The implemented mobile application is simple and understandable to the user.

Conflict of interest

The authors declare that there is no conflict of interest regarding this research, including financial, personal nature, authorship or other nature that could affect the research and its results presented in this article.

Financing

The study was conducted without financial support.

Data Availability

The manuscript has no associated data.

Acknowledgments

For providing data on agricultural crops of Northern Kazakhstan in the preparation of this article, the author expresses gratitude to the Scientific and Production Center of Grain Farming named after A. I. Barayev.

References

1. Mathew, A., Amudha, P., Sivakumari, S. (2020). Deep Learning Techniques: An Overview. *Advanced Machine Learning Technologies and Applications*, 599–608. doi: https://doi.org/10.1007/978-981-15-3383-9_54
2. Ahmad, J., Farman, H., Jan, Z. (2018). *Deep Learning Methods and Applications*. SpringerBriefs in Computer Science, 31–42. doi: https://doi.org/10.1007/978-981-13-3459-7_3
3. Golenko, Y., Ismailova, A., Shaushenova, A., Mutalova, Z., Dossalyanov, D., Ainagulova, A., Naizagarayeva, A. (2022). Implementation of machine learning models to determine the appropriate model for protein function prediction. *Eastern-European Journal of Enterprise Technologies*, 5 (4 (119)), 42–49. doi: <https://doi.org/10.15587/1729-4061.2022.263270>
4. Tussupov, J., La, L., Mukhanova, A. (2014). A model of fuzzy synthetic evaluation method realized by a neural network. *International Journal of Mathematical Models and Methods in Applied Sciences*, 8, 103–106. Available at: <https://www.naun.org/main/NAUN/ijmmas/2014/a142001-069.pdf>
5. Orazayeva, A., Wójcik, W., Pavlov, S., Tymchenko, L., Kokriatska, N., Tverdomed, V. et al. (2022). Biomedical image segmentation method based on contour preparation. *Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2022*. doi: <https://doi.org/10.1117/12.2657929>
6. Burger, W., Burge, M. J. (2022). Scale-Invariant Feature Transform (SIFT). *Digital Image Processing*, 709–763. doi: https://doi.org/10.1007/978-3-031-05744-1_25
7. Yang, J., Huang, J., Jiang, Z., Dong, S., Tang, L., Liu, Y. et al. (2020). SIFT-aided path-independent digital image correlation accelerated by parallel computing. *Optics and Lasers in Engineering*, 127, 105964. doi: <https://doi.org/10.1016/j.optlaseng.2019.105964>
8. Awaludin, M., Yasin, V. (2020). Application of Oriented Fast and Rotated Brief (Orb) and Brute-force Hamming in Library OpenCV for Classification of Plants. *JISAMAR (Journal of Information System, Applied, Management, Accounting and Research)*, 4 (3), 51–59. Available at: <https://journal.stmikjayakarta.ac.id/index.php/jisamar/article/view/247>
9. Rahmad, C., Asmara, R. A., Putra, D. R. H., Dharma, I., Darmono, H., Muhiqqin, I. (2020). Comparison of Viola-Jones Haar Cascade Classifier and Histogram of Oriented Gradients (HOG) for face detection. *IOP Conference Series: Materials Science and Engineering*, 732 (1), 012038. doi: <https://doi.org/10.1088/1757-899x/732/1/012038>
10. Mohanty, A., Nam, A., Pozhitkov, A., Yang, L., Srivastava, S., Nathan, A. et al. (2020). A Non-genetic Mechanism Involving the Integrin $\beta 4$ /Paxillin Axis Contributes to Chemoresistance in Lung Cancer. *IScience*, 23 (9), 101496. doi: <https://doi.org/10.1016/j.isci.2020.101496>

11. Charbuty, B., Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2 (01), 20–28. doi: <https://doi.org/10.38094/jastt20165>
12. Yessenova, M., Abdikerimova, G., Adilova, A., Yerzhanova, A., Kakabayev, N., Ayazbaev, T. et al. (2022). Identification of factors that negatively affect the growth of agricultural crops by methods of orthogonal transformations. *Eastern-European Journal of Enterprise Technologies*, 3 (2 (117)), 39–47. doi: <https://doi.org/10.15587/1729-4061.2022.257431>
13. Chen, S., Webb, G. I., Liu, L., Ma, X. (2020). A novel selective naive Bayes algorithm. *Knowledge-Based Systems*, 192, 105361. doi: <https://doi.org/10.1016/j.knosys.2019.105361>
14. Hameed, Z., Zahia, S., Garcia-Zapirain, B., Javier Aguirre, J., Maria Vanegas, A. (2020). Breast Cancer Histopathology Image Classification Using an Ensemble of Deep Learning Models. *Sensors*, 20 (16), 4373. doi: <https://doi.org/10.3390/s20164373>
15. Vangara, R. V. B., Thirupathur, K., Vangara, S. P. (2020). Opinion Mining Classification using Naive Bayes Algorithm. *International Journal of Innovative Technology and Exploring Engineering*, 9 (5), 495–498. doi: <https://doi.org/10.35940/ijitee.e2402.039520>
16. Jignesh Chowdary, G., Punn, N. S., Sonbhadra, S. K., Agarwal, S. (2020). Face Mask Detection Using Transfer Learning of InceptionV3. *Lecture Notes in Computer Science*, 81–90. doi: https://doi.org/10.1007/978-3-030-66665-1_6
17. Degadwala, S., Vyas, D., Biswas, H., Chakraborty, U., Saha, S. (2021). Image Captioning Using Inception V3 Transfer Learning Model. 2021 6th International Conference on Communication and Electronics Systems (ICCES). doi: <https://doi.org/10.1109/icces51350.2021.9489111>
18. Berrar, D. (2019). Bayes' Theorem and Naive Bayes Classifier. *Encyclopedia of Bioinformatics and Computational Biology*, 403–412. doi: <https://doi.org/10.1016/b978-0-12-809633-8.20473-1>