Improving Potato Diseases Classification Based on Custom ConvNeXtSmall and Combine with the Explanation Model

Huong Hoang Luong
Information Technology Department
FPT University, Can Tho, Viet Nam

***Abstract*—Potatoes are short-term crops grown for harvesting tubers. It is a type of tuber that grows on roots and is the fourth most common crop after rice, wheat, and corn. Fresh potatoes can also be used in an incredible variety of dishes by baking, boiling, or frying them. Moreover, the paper, textile, wood, and pharmaceutical industries also make extensive use of potato starch. However, soil and climate pollution are highly unfavorable for potato growth and lead to a lot of diseases such as common scab, black scurf, blackleg, dry rot, and pink rot. Thus. several types of research in medicine and computers were started for the early detection, classification, and treatment of potato diseases. In this study, transfer learning and fine-tuning were applied to potato disease classification based on a custom ConvNeXtSmall model. In addition, Gradient-weighted Class Activation Mapping (i.e., Grad-CAM) is provided for visual explanation in the final result after classification. For potato illness segmentation, k-means clustering was used to enable the difference between healthy and diseased sections based on color and texture. The data was collected from numerous websites and validated by the Bangladesh Agricultural Research Institute (i.e., BARI), including six types of potato diseases and healthy images. With a Convolutional Neural Networks (i.e., CNN) model from the Keras library, our study reached the unexpected validation accuracy, test accuracy, and F1 score in seven classifications of 99.49%, 98.97%, and 98.97%, respectively. Concerning four-class classification, high accuracy values were obtained for most of the models (i.e., 100%).**

Keywords*—Potato disease; classification; fine-tuning; transfer learning; Convolutional Neural Network (CNN); k-means clustering; Gradient-weighted Class Activation Mapping (Grad-CAM)

I. Introduction

Potatoes have been a crucial food source for humans for thousands of years. Originating in South America, they have spread worldwide due to their ability to grow in various climates and soil types. Potatoes are not only adaptable with hundreds of cooking ways but also packed with essential nutrients like carbohydrates, fiber, and vitamins. They are used in numerous dishes, from mashed potatoes to fries, and are also crucial in industries like starch production and biodegradable plastics. In the world, potatoes play a significant role, supporting about income of millions of farmers, and are one of the main food sources for billions of people.

In modern agriculture, potato production has become a cornerstone of global food systems. Based on statistical data, China holds the first spot in terms of potato production [1] [2]. In 2019, 370,436,581 tons of potatoes were produced globally and global potato production rose by 2.59% compared with 2012 [2]. 61.5% of the total EU production of potatoes is produced annually in Germany, France, the Netherlands, the United Kingdom, and Belgium, with an average of 34,870 million tons throughout the 2017–2019 period [3]. Specifically, Germany is currently the largest potato producer in the EU, with an average yearly potato production of slightly more than 10.4 million tonnes. France (i.e., 8.3 million tonnes) comes next, followed by the Netherlands (i.e., 6.8 million tonnes), the United Kingdom (i.e., 5.5 million tonnes), and Belgium (i.e., 3.8 million tonnes) between 2017 and 2019 [3]. However, potatoes face challenges, particularly from climate change, pollution, and diseases. Climate change affects potato growth with higher temperatures and irregular rainfall patterns [4]. Pollution from farming practices harms soil and water quality, impacting potato cultivation [5]. Additionally, diseases caused by viruses and bacteria threaten potato crops, reducing yields and farmer incomes [6].

Consequently, advancements in technology and machinery have revolutionized potato cultivation in agriculture, optimizing efficiency and yield while minimizing labor and environmental impact. Precision planting equipment ensures accurate spacing and depth, enhancing seedling establishment and uniformity across fields [7]. Additionally, state-of-the-art harvesting machinery, such as mechanical diggers and conveyor systems, streamline the process, minimizing damage to tubers and reducing post-harvest losses. Modern sorting and grading machines utilize advanced sensors and algorithms to identify and segregate potatoes based on size, shape, and quality criteria, improving marketability and reducing manual labor. Besides, Genetic modification of crops has also shown good effects in increasing productivity, nutritional quality, and disease resistance [8] These advancements not only increase productivity and profitability but also contribute to sustainable agriculture by minimizing inputs and environmental footprint in potato production.

Besides mechanical and genetic technologies, several studies on computer technology were developed for potatoes. Thus, artificial intelligence (AI) is creating and making significant strides in the potato farming sector, particularly in planting and harvesting processes. AI uses smart algorithms and sensors to improve various stages of potato cultivation. When it comes to planting, AI helps in determining the best spacing and depth for potato seeds, ensuring efficient resource utilization and better yields. AI-powered machinery, equipped with sensors, reduces wastage and increases productivity. During harvesting, AI aids in identifying ripe potatoes using computer vision, distinguishing them from soil and foliage, thus optimizing the harvesting process and minimizing crop damage. Moreover, AI analyzes environmental factors like soil moisture, temperature, and weather conditions in real-time, providing valuable insights to farmers for informed decision-making and better outcomes. Consequently, AI not only enhances efficiency and productivity in potato farming but also promotes sustainable agricultural practices by minimizing resource usage and environmental impact.

Deep learning is a subset of artificial intelligence and it has revolutionized various fields, including agriculture and industry. Besides, computer vision appeared as a new way for classification and segmentation of a lot of aspects of images [9] [10] [11] [12] [13], through techniques like transfer learning and fine-tuning. For example, image technologies have come out as invaluable tools in potato farming, offering correct solutions for classification, segmentation, and detection tasks. Classification algorithms are used to assess potato images, categorizing them based on diverse attributes including size, shape, and quality, thereby facilitating farmers in optimizing sorting procedures and ensuring consistency across their yield [14]. Segmentation techniques are employed to segment potato disease regions or potato growing areas to detect weeds [15], facilitating precise attribute measurements such as size and color distribution, which in turn assists in grading and quality evaluation. Detection algorithms play a pivotal role in identifying diseases, pests, and other anomalies in potato crops [16], enabling prompt intervention and mitigation measures to curtail yield losses.

In this study, various algorithms have used the progressions in machine learning and computer vision, and large datasets of potato plant images afflicted with various diseases for classification and segmentation images, deep learning models can accurately classify these images. Transfer learning has been used in this process, it is a technique where a pre-trained model developed for one task is adapted for another [17] [18] [19], allowing researchers to leverage the knowledge gained from training models on massive datasets to enhance the performance of models in potato disease classification and segmentation. Furthermore, fine-tuning, a process of adjusting the parameters of a pre-trained model to better fit the specific characteristics of a new dataset, has been crucial in refining the accuracy and efficiency of disease segmentation and detection algorithms [20] [21] [22]. By fine-tuning pre-existing deep learning architectures such as convolutional neural networks (CNNs), researchers can customize these models to effectively identify and delineate regions of potato plants affected by diseases, enabling early intervention and targeted treatment.

Overall, the integration of deep learning techniques, including transfer learning and fine-tuning, has significantly advanced the field of potato disease management by providing accurate, efficient, and contributing to improved crop yield and food security. This study advances agricultural technology by using ConvNeXtSmall in the Keras library to categorize potato disease photos with very high accuracy. Additionally, utilizing k-mean clustering for the segmentation of anomalous positions offers a thorough solution that supports agricultural efforts in the early diagnosis and treatment of potato illnesses by assisting in the accurate identification of abnormal zones.

The contributions of this paper are as follows:

- A customized CNN model based on ConvNeXtSmall is presented by the research for the purpose of classifying and segmenting potato diseases into seven groups: pink rot, dry rot, blackleg, black scurf, common scab, miscellaneous, and healthy. Thus, it might offer a quick and easy way for the farmer to boost profitability and productivity depending on the amount of potatoes they produce.
- In the scenario of seven classes classification, our model achieved excellent validation accuracy, test accuracy, and F1 score (i.e., 99.49%, 98.97%, and 98.97%). As a result, a table and confusion matrix were also made to illustrate how successful the model was in terms of training and testing duration.
- The article proposes K-means clustering in image segmentation for identifying potato diseases. By grouping pixels with similar characteristics, it enables the differentiation of healthy areas from diseased ones based on color and texture. This method aids in precise disease mapping, facilitating targeted treatment strategies for agricultural management.
- Gradient-weighted Class Activation Mapping (Grad-CAM) was applied and assisted in the visual explanation of potato diseases by pinpointing regions in images where the model concentrates its attention. This technique aids in the precise identification and comprehension of various potato diseases, enhancing diagnostic accuracy and agricultural management strategies.
- This research gathered photos of both sick and healthy potatoes, as confirmed by experts at the Bangladesh Agricultural Research Institute. This dataset can be used to teach agriculture students and is validated for the creation of automated machine learning and deep learning algorithms for the classification, segmentation, and detection of potato illnesses.

The organization of our research paper is structured into six principal sections. Firstly, Section I serves as a comprehensive overview providing a general introduction to the article. Following this, Section II extensively explores related research, offering a thorough examination of the existing literature upon which our work is based. Subsequently, Section III delineates the methodology utilized, furnishing detailed insights into the methods employed throughout the article. Section IV then elaborates on the experiments conducted, encompassing the procedures for their execution and the evaluation of each scenario. Furthermore, Section V presents the results of the best experiment and conducts a comparative analysis with existing scenarios. Finally, the article encapsulates the key findings and analyzes the fundamental domains associated with our research in Section VI.

II. Related Works

Exploring potato diseases led to groundbreaking research in transfer learning and fine-tuning and various research was published to promote disease segmentation and classification in potatoes. For example, A deep convolutional neural network was trained to distinguish between four types of potatoes (i.e., Red, Red Washed, Sweet, and White) using a public dataset of 2400 photos by Abeer A. Elsharif et al. The trained model attained an accuracy of 99.5% of the test accuracy [23]. Besides, Sofia Marino et al. have presented an effective unsupervised adversarial domain adaptation method to classify six potato classes in two different scenarios and reach an average F1-score of 84% [24]. Furthermore, Qinghua Su et al. used a convolutional neural network model and achieved a high success rate in size classification at 94.5% and appearance classification at 91.6% [25].

One of the most important tasks in post-harvest quality control during potato manufacturing is identifying bad surfaces for potatoes. Therefore, Chenglong Wang et al. used transfer learning to refine a basic model using three DCNN modes. As a result, RFCN ResNet101 had the best overall performance, achieving accuracy levels of 92.5%, 95.6%, and 98.7%, respectively [26]. Moreover, Kaili Zhang et al. improved U-Net and showed that the accuracy of the potato surface evaluation method proposed in the study was greater than 97.55% [27], In addition, Black scurf, common scab, black leg, pink rot, and healthy are the five categories of potato diseases that Khalid Hamza et al. detect and classify. In multiple classes, the accuracy can reach 98% and 100% [28].

To classify potato problems early, machine learning technology has been widely used as an affordable and nondestructive diagnostic tool. For instance, Ali Arshaghi et al. use image processing and convolution neural networks to identify and classify surface potatoes from a collection of 5,000 photos of potatoes that have been split into five types. As a result, the results show that the accuracy of the deep learning proposed was 98% and 100% accuracy in some of the classes [29].To classify potatoes, Hyeon-Seung Lee et al. used Mask R-CNN, one of the object identification technologies utilizing deep learning, and were surprised with the result of 93.0% [30]. Furthermore, Israa Mohammed Hassoon et al. proposed a PDCNN framework that is very effective in classifying four types of potato tuber diseases including black dot, common scab, potato virus y, and ring rot with 91.3% accuracy [31].

To expand the author's limited knowledge about transfer learning and fine-tuning in CNN. Hence, several research about potato leaf disease classification have been investigated. With the help of deep learning and the VGG16 and VGG19 convolutional neural network architectural model, Rizqi Amaliatus Sholihati et al. developed a system that can classify the four different types of diseases in potato plants based on leaf conditions. As a result. the model achieved an average accuracy of 91% [32]. Moreover, Aditi Singh et al. presented a model that can reach a 95.99% accuracy rate when using the K-means approach for feature segmentation and the multi-class support vector machine methodology for classification [33]. To identify potato leaf disease, Rabbia Mahum et al. employed an additional transition layer in DenseNet-201 along with a pre-trained Efficient DenseNet model. As a result, the performance was evaluated and gave an accuracy of 97.2% [34].

The quantity and quality of potatoes are greatly impacted by many diseases. Because manually explaining these leaf diseases is labor-intensive and time-consuming. As a result, Divyansh Tiwari et al utilized a pre-trained model VGG19 for fine-tuning the dataset and reached 97.8% in classification accuracy over the test dataset[35]. Moreover, Asif Iqbal et al. proposed an image processing and machine learning-based automatic system that will identify and classify over 450 images of healthy and diseased potato leaves with an accuracy of 97% [36]. Kulendu Kashyap Chakraborty et al proposed a methodology using four deep learning models such as VGG16, VGG19, MobileNet, and ResNet50. Hence, it achieved 97.89% accuracy for classification between late and early blight syndromes as compared to healthy potato leaf [37].

Besides using computer vision to recognize and classify potatoes, other fruits and vegetables also apply this method for having the advantages of speed, and high accuracy in dividing a final product. For instance, Alper Taner et al. used popular seven CNN architectures (i.e., VGG16, VGG19, InceptionV3, MobileNet, Xception, ResNet150V2, and DenseNet201) and it was found that DenseNet201 had the highest classification accuracy of 97.48% in classifying apple varieties [38]. Moreover, Dhiya Mahdi Asriny et al. proposed the classification model to classify orange images using CNN and shows an accuracy of 96% [39]. In conclusion, Table I shows that related studies have been compiled for easier evaluation and synthesis.

TABLE I. RELATED PAPERS IN AGRICULTURE

Product	Method	Accuracy	Year	Author
Potato	CNN	99.5%	2020	Abeer A. Elsharif et al [23]
Potato	FCN	F1 score = 84%	2020	Sofia Marino et al [24]
Potato	CNN	91.6%	2020	Qinghua Su et al [25]
Potato	RFCN ResNet101	98.7%	2021	Chenglong Wang et al [26]
Potato	VGG and U-Net	97.55%	2023	Kaili Zhang et al [27]
Potato	CNN	98%˜100%	2022	Khalid Hamza et al
Potato	CNN	98%˜100%	2020	Ali Arshaghi et al [29]
Potato	Mask R-CNN	93%	2020	Hyeon-Seung Lee et al [30]
Potato	PDCNN	91.3%	2021	Israa Mohammed Hassoon et al [31]
Potato leaves	VGG16 and VGG19	91%	2020	Rizqi Amaliatus Sholihati et al [32]
Potato leaves	CNN	95.99%	2021	Aditi Singh et al [33]
Potato leaves	DenseNet-201 and Efficient DenseNet	97.2%	2022	Rabbia Mahumet et al [34]
Potato leaves	VGG19	97.8%	2020	Divyansh Tiwariet et al[35]
Potato leaves	D-CNN	97%	2020	Asif Iqbal et al [36]
Potato leaves	VGG16	97.89%	2022	Kulendu Kashyap Chakraborty et al[37]
Apple	DenseNet201	97.48%	2024	Alper Taner et al[38]
Orange	CNN	96%	2020	Dhiya Mahdi Asriny et al [39]

III. METHODOLOGY

A. The Research Implementation Procedure

This study proposed a method including 12 steps from input to output shown in Fig. 1. The roles of the steps are shown as follows:

1) Collecting dataset: the dataset selected from various sources and rigorously verified by the Bangladesh Agricultural Research Institute (BARI), boasts 451 images capturing diverse potato illnesses such as Common scab, Blackleg, Dry rot, Pink rot, Black scurf, Miscellaneous, and Healthy Potatoes. It is a valuable resource for academic research, offering comprehensive insights into potato disease management and cultivation practices.

2) Pre-processing image and data augmentation: image pre-processing techniques such as resizing and normalization are crucial for standardizing input data, and ensuring consistency in potato disease classification models. Furthermore, using data augmentation methods such as rotation, flipping, and contrast enhancement enhances the variety of the dataset, and assists model training. These algorithms provide a foundation for developing accuracy and prediction to classify and segment potato diseases, thus safeguarding agricultural productivity and income.

3) Dividing the dataset into three categories train, validation, and test: after increasing data augmentation by 451 default participants, which were chosen at random for the training, validation, and testing phases, the pictures dataset contains 5833 subjects. The datasets are randomly selected using an 8-1-1 scale and are then placed into 8 training, 1 validation, and 1 testing folder. This guarantees a balanced distribution, which is essential for trustworthy model development and evaluation.

4) Dividing dataset for scenarios: the dataset was divided into four scenarios. In the first scenario, four classes healthy, black scurf, common scab, and pink rot were selected because they can be classified by surface observation. Next to that, four classes healthy, blackleg, dry rot, and miscellaneous because it is an internal harm. Finally, the next two scenarios employed all classes for the experiment.

5) Building the model: our work employed transfer learning to a pre-trained model based on the CNN architecture prototype to conduct tests. External layers were employed during fine-tuning to adapt the pre-trained model to the specific data of the intended task. For our training test, the ConvNeXtSmall model thus yields an excellent result.

6) Applying transfer learning: using pre-trained deep learning models, this approach aims to transfer knowledge from related domains to enhance the classification accuracy of potato disease. By adapting neural networks to recognize patterns indicative of various potato diseases.

7) Validating and collecting accuracy score: after the model had completed training, its efficiency was assessed using its training accuracy as well as its other scores. Next, The validity of the test was then determined using the initially divided testing set.

8) Applying fine-tuning: fine-tuning involves adjusting the parameters, and additional layers of a pre-trained neural network particularly in the latter layers, typically focusing on refining performance. This process allows the model to borrow knowledge acquired from a broader scope while customizing it to the particular requirements.

9) Validating, collecting, and explaining results with Grad-CAM: Grad-Cam was used for the analysis of heat maps generated by the model to highlight regions of interest. By correlating these areas with known symptoms, researchers can validate the model's accuracy, collect valuable data for further analysis, and elucidate its decision-making process.

10) Image segmentation by k-means clustering: this step includes partitioning the image into distinct clusters based on pixel intensities. By iterative assigning pixels to clusters with similar characteristics, this method effectively separates different disease regions within the potato image, facilitating targeted analysis and diagnosis of specific ailments.

11) Reconstructing and comparing the cycles with other models: after one phase, the procedure was reworked and compared with another model including, EffecientNetB3, ResNet50, MobileNet, Inception V3, Xception, ConvNeXtSmall, ConvNeXtTiny, ConvNeXtLarge to create the final result

12) Showing the result: the data will be presented as tables and graphs after procedures to enable pertinent comparisons.

B. Pre-processing Image and Data Augmentation

Pre-processing plays an important role in boosting the quality of images before subjecting them to potato disease classification algorithms. Resizing (1) and normalization (2) are two fundamental techniques used in this process. Resizing (1) connection to change the dimensions of images to a uniform size, which aids in reducing computational complexity and ensuring consistency across the dataset. Mathematically, resizing can be represented as follows:

$$Resize\,image = resize(original\,image, target\,size) \quad (1)$$

Algorithm 1 Resizing Algorithm

Require: Original Image, target_size
Ensure: Resized Image

Load the Original Image:
Define the target_size = (224,224)
Resize the Original Image to the target_size using the resize function:
  $ResizedImage = \text{resize}(OriginalImage, (224, 224))$
return Resized Image

Here, original image (1) represents the raw input image, and target size (1) indicates the desired dimensions of the output image. Besides, Algorithm 1 is a pseudo-code that presents flow on how it works in coding. Furthermore, Normalization (2) fixes standard pixel values within a certain range. It can be expressed as:

$$Normalized\, image = \frac{original\, image - mean}{std} \quad (2)$$

Algorithm 2 Normalization Algorithm

Require: Original Image
Ensure: Normalized Image
1: Compute the mean and standard deviation of pixel intensities:
2: $mean = \frac{1}{n}\sum_{i=1}^{n} pixel_i$
3: $std = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(pixel_i - mean)^2}$
4: Normalize the Original Image by subtracting the mean and dividing by the standard deviation:
5: $NormalizedImage = \frac{OriginalImage - mean}{std}$
6: **return** Normalized Image

Specifically, mean (2) and std (2) represent the mean and standard deviation of pixel intensities across the entire dataset, respectively. This normalization process ensures that pixel values are centered around 0 with a standard deviation of 1, to facilitate convergence during training and mitigate the influence of illumination variations, thereby increasing the stability of the training process. In addition, To clarify the process a pseudo-code has been provided at Algorithm 2.

Data augmentation techniques are tools for improving the diversity of training samples, thereby improving the generalization ability of the classifier. Rotation (3), flipping (4) (5), and contrast enhancement (6) are commonly employed augmentation strategies in the context of potato surface disease classification. Rotation (3) involves rotating the image by a certain angle to simulate variations in orientation. Mathematically, rotation can be represented as:

$$Rotated\, image = rotate(original\, image, \theta) \quad (3)$$

Algorithm 3 Rotation Algorithm

Require: Original Image, angle
Ensure: Rotated Image
1: Load the Original Image
2: Specify the angle of rotation (θ)
3: Rotate the Original Image by the specified angle using the formula:
4: Rotated image = rotate(original image, θ)
5: **return** Rotated Image

In the equation, θ (3) strand for the angle of rotation and Algorithm 3 presents the detail of code flow which is provided in pseudo-code for overview. Moreover, flipping (4) (5) entails flipping the image horizontally or vertically to introduce variations in perspective. It can be mathematically expressed as:

$$P(x, y) = (width - x - 1, y) \quad (4)$$

$$P(x, y) = (x, height - y - 1) \quad (5)$$

Algorithm 4 Flipping Algorithm

Require: Original Image, axis
Ensure: Flipped Image
1: Load the Original Image
2: Specify the axis along which flipping should occur: horizontal (H) or vertical (V)
3: **if** axis is H **then**
4: Flip the Original Image horizontally using the formula: $P(x, y) = (width - x - 1, y)$
5: **else if** axis is V **then**
6: Flip the Original Image vertically using the formula: $P(x, y) = (x, height - y - 1)$
7: **end if**
8: **return** Flipped Image

In the context of image flipping, (x, y) 4 5 represents the pixel coordinates of the Original Image. In Algorithm 4 flipping horizontally (H), the pixel x-coordinate is reversed concerning the image width, and when flipping vertically (V), the pixel y-coordinate is reversed concerning the image height.

$$P_{\text{enhanced}}(x, y) = CDF(P_{\text{original}}(x, y)) \times (L - 1) \quad (6)$$

Algorithm 5 Contrast Enhancement Algorithm

Require: Original Image
Ensure: Enhanced Image
1: Load the Original Image
2: Compute the cumulative distribution function (CDF) of pixel intensities
3: Apply histogram equalization to map pixel intensities to a new range using the formula:
4: $P_{\text{enhanced}}(x, y) = \text{CDF}(P_{\text{original}}(x, y)) \times (L - 1)$
5: where $P_{\text{enhanced}}(x, y)$ is the pixel intensity in the Enhanced Image,
6: $P_{\text{original}}(x, y)$ is the pixel intensity in the Original Image,
7: and L is the number of intensity levels
8: **return** Enhanced Image

The contrast enhancement algorithm aims to improve the contrast of an image through histogram equalization. It begins by loading the original image and computing its histogram, which represents the frequency distribution of pixel intensities. Subsequently, Algorithm 5 calculates the cumulative distribution function (CDF) from the histogram. This CDF provides a mapping between original pixel intensities and their corresponding enhanced values. The enhancement is achieved by applying the formula (6). where L (6) illustrates the number of intensity levels. Each pixel in the original image is mapped to a new intensity level based on its CDF value, resulting in an image with improved contrast. Finally, These data augmentation techniques collectively contribute to the robustness of the classification model by exposing it to a diverse range of image variations.

C. Transfer Learning and Fine-tuning of ConvNeXtSmall

Transfer learning is a technique in machine learning where knowledge gained from solving one problem is applied to a different but related problem [17] [18] [19]. In the context of image classification, it involves leveraging pre-trained models, which have been trained on large datasets, and adapting them to new classification tasks with relatively smaller datasets. This approach is particularly useful when the target dataset is not large enough to train a model from scratch effectively.

In the classification of potato surface disease, transfer learning can be employed by utilizing a pre-trained Convolutional Neural Network (ConvNet), such as ConvNeXtSmall, which has been trained on a large dataset. The initial layers of ConvNeXtSmall have learned to extract low-level features like edges and textures, which are generally applicable to various image recognition tasks. By reusing these learned features and adjusting the later layers to suit the specifics of potato surface disease classification, this algorithm can expedite the training process and improve performance.

Fine-tuning is a key aspect of transfer learning, where the parameters of the pre-trained model are further adjusted to better fit the new dataset [20] [21] [22]. In the case of ConvNeXtSmall, fine-tuning includes unfreezing some of the later layers and retraining them using the new dataset. This allows the model to learn higher-level representations more adapted to the characteristics of potato surface disease, refining its ability to distinguish between disease states or healthy potato surfaces.

Moreover, adding extra layers to ConvNeXtSmall in Fig. 2 can enhance its accuracy and overall performance. These additional layers can capture more complex patterns and relationships within the data, providing the model with a deeper understanding of the distinguishing features of different disease conditions. However, care must be taken to prevent overfitting, where the model becomes too specialized to the training dataset and performs poorly on unseen data. Regularization techniques such as dropout and weight decay can be employed to mitigate overfitting and ensure the generalization ability of the model.

In summary, transfer learning and fine-tuning of ConvNeXtSmall offer effective strategies for the classification of potato surface disease by borrowing pre-existing knowledge and adapting it to the specific characteristics of the target dataset. By incorporating additional layers and carefully fine-tuning the model, this research improves its accuracy and prediction in classifying different disease states, ultimately aiding in the early detection and management of potato crop diseases.

D. Visual Explanation with Gradcam

Visual explanations through techniques like Grad-CAM (Gradient-weighted Class Activation Mapping) offer valuable insights into the classification process, aiding researchers and farmers in making informed decisions. Potato surface diseases encompass a range of fungal, bacterial, and viral infections that affect the external appearance of the potato tubers. Timely detection and classification of these diseases are essential for maintaining crop health and yield.

Grad-CAM is a technique used in computer vision to understand the decision-making process of deep learning models. It highlights regions of an image that contribute most significantly to the model's classification decision. In the context of potato surface disease classification, Grad-CAM helps elucidate which features or regions of the potato surface are indicative of particular diseases. in the case of the black scurf and black leg of Fig. 3, Grad-CAM may highlight regions of the potato surface where characteristic lesions or discolorations are present.

The Grad-CAM method functions by generating a heat map that delineates the significance of various regions within the input image for prediction. This heat map is derived by computing the gradient of the target class score concerning the final convolutional layer feature maps. Mathematically:

$$H^c_{(i,j)} = \text{ReLU}\left(\sum_k \alpha^c_k \cdot A^k_{(i,j)}\right) \tag{7}$$

Specifically, $H^c_{(i,j)}$ (7) represents the heat map value at position (i, j) for class c. α^c_k (7) denotes the importance of the kth feature map for class c, and $A^k_{(i,j)}$ (7) is the activation of the kth feature map at position (i, j). Equation (7) essentially encapsulates the importance of each feature map activation, weighted by its corresponding importance score. The ReLU function is employed to ensure that only positive contributions are considered.

In conclusion, visual explanation techniques like Grad-CAM furnish a valuable means of interpreting deep learning models in the classification of potato surface diseases. By accentuating crucial regions within input images, Grad-CAM facilitates the understanding of model predictions and offers insights for refining disease management strategies.

E. Image Segmentation by k-means Clustering

Image segmentation is a main task in classifying potato surface diseases, assisting in identifying and analyzing abnormal areas on the potato surface. Among various segmentation techniques, k-means clustering is an efficient method for partitioning images into distinct clusters based on pixel intensity values. In this context, k-means clustering (8) facilitates the categorization of pixels into groups, differentiating healthy potato regions from those affected by diseases.

Mathematically, the k-means algorithm aims to minimize the within-cluster sum of squares, defined as:

$$W(C_k) = \sum_{i=1}^{n} \sum_{x_j \in C_k} ||x_j - \mu_k||^2 \tag{8}$$

In every detail, C_k (8) represents the cluster k, x_j (8) show the jth (8) data point (pixel), and μ_k (8) is the centroid of cluster k. The goal is to assign each pixel to the cluster whose centroid is nearest to it in terms of Euclidean distance.

Specifically, Algorithm 6 outlines the k-means clustering approach for segmenting images. it begins with an initialization phase, during which k initial centroids are randomly selected.

Following this, the assignment step assigns each pixel to the nearest centroid, effectively partitioning the image into k clusters. Subsequently, in the update phase, the centroids are recalculated based on the mean of all pixels assigned to each cluster. This iterative process continues until convergence is achieved, typically indicated by a condition such as centroids no longer undergoing significant changes between iterations. This methodical approach enables the algorithm to effectively segment images by distinguishing regions based on pixel similarities, making it particularly valuable for applications such as disease detection on potato surfaces.

Algorithm 6 K-Means Clustering

Require: Image I, Number of clusters k = 3 (Background, healthy surface and disease surface)
Ensure: Segmented image I_{seg}, Centroids $\{\mu_1, \mu_2, ..., \mu_k\}$
Initialization:
Randomly select k initial centroids: $\mu_1, \mu_2, ..., \mu_k$
$I_{\text{prev}} \leftarrow$ Copy of I
repeat
  Assignment Step:
  for each pixel p in I **do**
    Assign p to the nearest centroid μ_i
  end for
  Update Step:
  for each cluster i **do**
    Recalculate centroid μ_i as the mean of all pixels in cluster i
  end for
  $I_{\text{seg}} \leftarrow$ Image formed by assigning pixels to their respective clusters
  $I_{\text{prev}} \leftarrow$ Copy of I_{seg}
until Convergence criteria are met (e.g., centroids do not change significantly)
return Segmented image I_{seg}, Centroids $\{\mu_1, \mu_2, ..., \mu_k\}$

By applying k-means clustering to potato surface images, Fig. 4 effectively segments the image into regions of similar pixel intensities, thereby distinguishing between healthy and diseased areas. The centroids obtained represent characteristic color values associated with each cluster, aiding in the identification of disease patterns based on pixel color.

Furthermore, the simplicity and efficiency of k-means clustering make it a suitable choice for real-time or large-scale image processing tasks, contributing to the rapid and accurate classification of potato surface diseases. Overall, leveraging k-means clustering in image segmentation enhances the precision and scalability of disease detection systems, facilitating timely interventions to mitigate agricultural losses.

IV. Experiments

A. Dataset and Performance Metrics

The research used a single dataset for both the training, validation, and testing phases in this analysis. The data was selected from various sources and rigorously verified by the Bangladesh Agricultural Research Institute (BARI). 451 images constitute the comprehensive dataset in Fig. 5 including 62 Common scabs, 60 Blackleg, 60 Dry rot, 57 Pink rot, 58 Black scurf, 74 Miscellaneous, and 80 Healthy Potato images.

In the classification of potato surface disease, various performance metrics are utilized to evaluate the effectiveness of the classification model. These metrics help in understanding the ability to correctly classify instances of potato surface disease and its performance in terms of both precision and recall.

One of the fundamental metrics used is the Accuracy (i.e., ACC), which measures the proportion of correctly classified instances out of the total instances. Mathematically, it can be expressed as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

Where TP (9) indicates true positives (correctly classified instances of potato surface disease), TN (9) represents true negatives (correctly classified instances of absence of potato surface disease), FP (9) stands for false positives (instances incorrectly classified as having potato surface disease), and FN (9) represents false negatives (instances incorrectly classified as not having potato surface disease).

Another important metric is the Recall in equation (10), also known as sensitivity or true positive rate. It measures the proportion of actual positive instances that are correctly identified by the model. Moreover, This metric is crucial in scenarios where the consequences of false negatives (misclassifying actual instances of potato surface disease as negative) are severe. Mathematically, it can be defined as:

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

The Precision metric in equation (11) evaluates the proportion of true positive instances among all instances classified as positive by the model. It can be calculated as:

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

The F1 Score in equation (12) is a metric that combines precision and recall into a single value. It is the harmonic mean of precision and recall, and it provides a balance between the two metrics. Mathematically, it is represented as:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{12}$$

B. Scenario 1: the Result of Classifying Potato Diseases into four Classes: Healthy, Black Scurf, Common Scab, Pink Rot

Table II presents a comparative analysis of various deep learning models in classifying potato disease images into four classes: healthy, black scurf, common scab, and pink rot, using both transfer learning and fine-tuning techniques. Among the models evaluated, EfficientNet B3 emerges as the top performer, exhibiting remarkable accuracy in both transfer learning (i.e., 99.70% validation accuracy, 99.40% test accuracy) and fine-tuning (i.e., 100.00% for both validation and test accuracy). In contrast, Xception shows comparatively lower accuracy levels across both transfer learning and fine-tuning

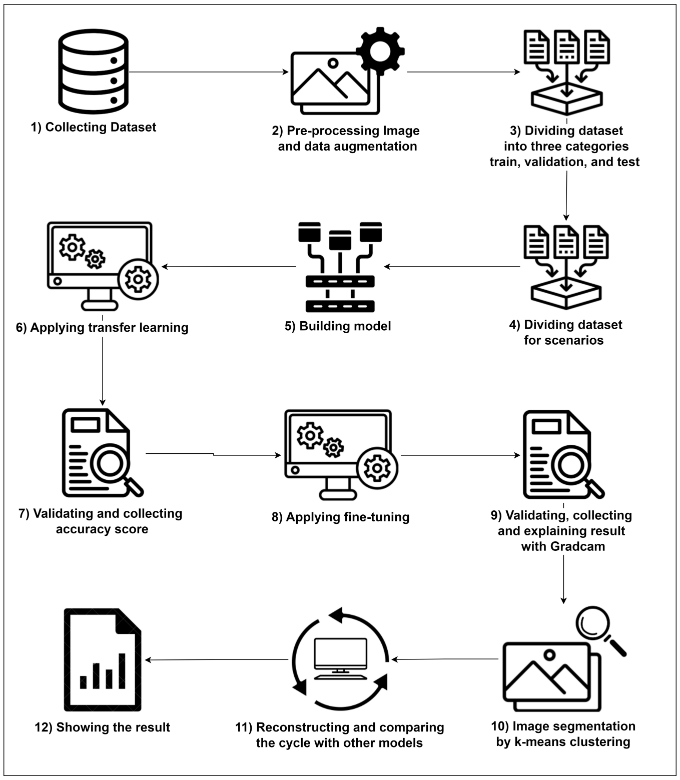

Fig. 1. The implementing procedure flowchart.

TABLE II. THE ACCURACY OF CLASSIFYING POTATO DISEASE IMAGES INTO FOUR CLASSES: HEALTHY, BLACK SCURF, COMMON SCAB, PINK ROT IN TRANSFER LEARNING AND FINE-TUNING, FOR EACH DEEP LEARNING MODEL

Model	Phase	Valid acc	Test acc	Precision	Recall	F1
EfficientNet B3	Transfer Learning	99.70%	99.40%	97.44%	97.32%	97.30%
	Fine Tuning	100.00%	100.00%	100.00%	100.00%	100.00%
ResNet50	Transfer Learning	100.00%	99.83%	99.71%	99.70%	99.70%
	Fine Tuning					
MobileNet	Transfer Learning	97.31%	95.04%	94.96%	94.94%	94.94%
	Fine Tuning	97.31%	97.46%	96.74%	96.73%	96.73%
InceptionV3	Transfer Learning	89.55%	91.91%	90.92%	90.77%	90.74%
	Fine Tuning	91.64%	94.99%	92.69%	92.56%	92.58%
Xception	Transfer Learning	84.48%	84.52%	84.91%	84.52%	84.57%
	Fine Tuning	91.34%	89.66%	89.49%	89.29%	89.33%
Our Model	**Transfer Learning**	**98.51%**	**93.75%**	**97.93%**	**97.92%**	**97.92%**
	Fine Tuning	**99.40%**	**96.88%**	**99.11%**	**99.11%**	**99.11%**

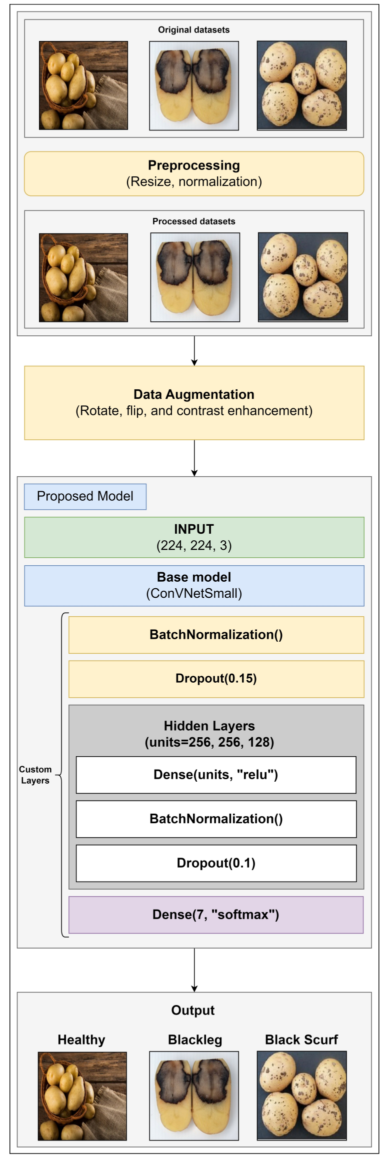

Fig. 2. Procedure of transfer learning and fine-tuning in our model with custom layers.

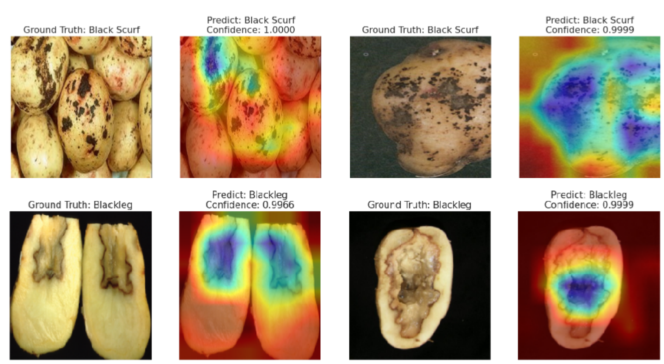

Fig. 3. Visual explanation by gradcam of potato diseases.

phases, with validation accuracy of 84.48% and 91.34%, and test accuracy of 84.52% and 89.66%, respectively. However, Our model showcases competitive performance, particularly in fine-tuning (i.e., an increase of 3.13 % with transfer learning), where it achieves a validation accuracy of 99.40% and a test accuracy of 96.88%, closely aligning with EfficientNetB3. This suggests that the proposed model demonstrates robustness comparable to the state-of-the-art EfficientNetB3 while outperforming the least effective model, Xception, underscoring its potential as a viable alternative in potato disease classification tasks.

Fig. 6 and Fig. 7 indicate the accuracy and loss scores in the two training and validation phases. With this line chart, the accuracy and loss scores are presented intuitively and simply to help with general assessment through the training epoch. Besides, the confusion matrix in Fig. 8 assesses the performance of deep learning models for potato disease classification. It evaluates accuracy, identifies misclassifications, and guides parameter optimization. Additionally, it detects class imbalances, aids in error analysis, and facilitates model comparison, enabling improvement in accuracy and robustness.

C. Scenario 2: the Result of Classifying Potato Diseases into Four Classes: Healthy, Blackleg, Dry Rot, Miscellaneous

TABLE III. The Accuracy of Classifying Potato Disease Images Into Four Classes: Healthy, Blackleg, Dry Rot, Miscellaneous in Transfer Learning and Fine-tuning, for Each Deep Learning Model

Model	Phase	Valid acc	Test acc	Precision	Recall	F1
EfficientNet B3	Transfer Learning	99.40%	96.60%	96.18%	96.13%	96.14%
	Fine Tuning	99.11%	99.24%	99.42%	99.40%	99.40%
ResNet50	Transfer Learning	99.11%	98.59%	98.82%	98.81%	98.81%
	Fine Tuning	99.70%	98.54%	98.54%	98.51%	98.51%
MobileNet	Transfer Learning	90.18%	90.09%	89.36%	89.29%	89.20%
	Fine Tuning	91.67%	89.93%	91.40%	91.37%	91.36%
InceptionV3	Transfer Learning	79.76%	77.33%	77.03%	76.79%	76.68%
	Fine Tuning	84.23%	83.23%	82.97%	82.74%	82.73%
Xception	Transfer Learning	76.19%	78.77%	76.34%	76.19%	76.01%
	Fine Tuning	83.93%	85.58%	85.11%	85.12%	85.08%
Our Model	**Transfer Learning**	**98.21%**	**96.88%**	**94.35%**	**94.35%**	**94.31%**
	Fine Tuning	**99.11%**	**96.88%**	**96.17%**	**96.13%**	**96.12%**

Compared with Table II, Table III shows that the performance of our model remained stable at 96.68% in the classification four classes healthy, blackleg, dry rot, miscellaneous although it is internal damage. However, EfficientNetB3 witnessed a slight decrease (i.e., a decrease of 0.76%) when compared with scenario 1. Moreover. Xception presents inef-

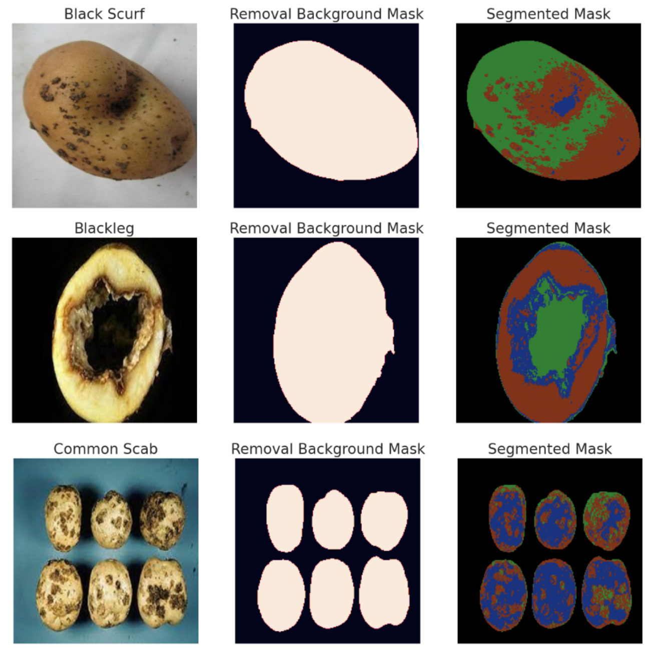

Fig. 4. Image segmentation in potato diseases by k-mean clustering.

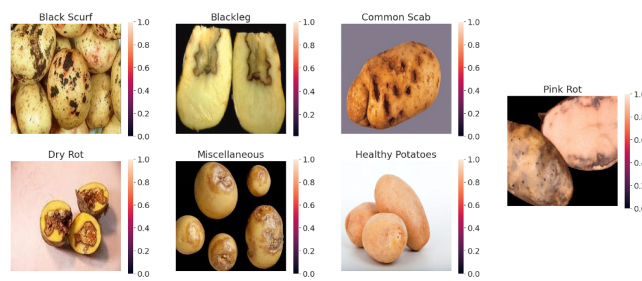

Fig. 5. The dataset of potato images.

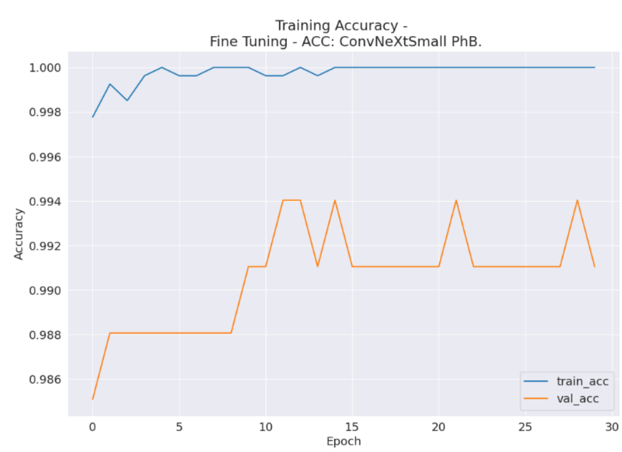

Fig. 6. Training accuracy and validation accuracy by fine-tuning of our model at scenario 1.

fectiveness in evaluating the performance of classifying potato diseases.

Furthermore, Training and validation on both accuracy and loss scores are presented in Fig. 9 and Fig. 10. Following the figures, the evaluation performance of our model presents the balance when the dataset is changed. Moreover, Fig. 11 is provided for evaluating, optimizing, and understanding the performance of deep learning models in classifying potato diseases, providing insights that can lead to improved accuracy and reliability.

D. Scenario 3: the Result of Classifying Potato Diseases Into Seven Classes: Healthy, Black Scurf, Common Scab, Pink Rot, Blackleg, Dry Rot, Miscellaneous

TABLE IV. THE ACCURACY OF CLASSIFYING POTATO DISEASE IMAGES INTO SEVEN CLASSES: HEALTHY, BLACK SCURF, COMMON SCAB, PINK ROT, BLACKLEG, DRY ROT, MISCELLANEOUS IN TRANSFER LEARNING AND FINE-TUNING, FOR EACH DEEP LEARNING MODEL

Model	Phase	Valid acc	Test acc	Precision	Recall	F1
EfficientNet B3	Transfer Learning	98.00%	98.22%	98.23%	98.22%	98.22%
	Fine Tuning	99.11%	98.00%	98.08%	98.00%	98.00%
ResNet50	Transfer Learning	98.67%	97.56%	97.63%	97.56%	97.54%
	Fine Tuning	99.56%	98.44%	98.49%	98.44%	98.45%
MobileNet	Transfer Learning	86.22%	80.89%	81.05%	80.89%	80.77%
	Fine Tuning	91.56%	92.22%	92.27%	92.22%	92.21%
InceptionV3	Transfer Learning	68.44%	68.67%	68.38%	68.67%	68.18%
	Fine Tuning	99.33%	98.44%	98.47%	98.44%	98.44%
Xception	Transfer Learning	65.78%	61.11%	62.50%	61.11%	61.10%
	Fine Tuning	98.67%	97.33%	97.38%	97.33%	97.33%
Our Model	**Transfer Learning**	**94.85%**	**94.69%**	**94.72%**	**94.69%**	**94.67%**
	Fine Tuning	**99.49%**	**98.97%**	**98.98%**	**98.97%**	**98.97%**

Table IV illustrates a successful classification when our model experiences a dramatic rise in test accuracy of 98.97% of fine-tuning (i.e., a growth of 2.09%). Moreover, other scores such as prediction, recall, and f1 reached a high point. Thus, our model successfully demonstrated that the performance in classifying images in multiple classes (i.e., seven classes) is better than other models. However, EfficientNetB3 presents a decline in performance when working with seven classes. Despite ResNet50, MobileNet, InceptionV3, and Xception climb significantly. In particular, Xception has an increase of 11.75% when compared with Table III.

In addition, Fig. 12 and Fig. 13 show progress with nearly reached the highest point in a surprise outcome of validation accuracy = 99.49%. Moreover, training and validation loss decreased significantly and was achieved at 0.07. To describe more detail, Fig. 14 represents a confusion matrix for the

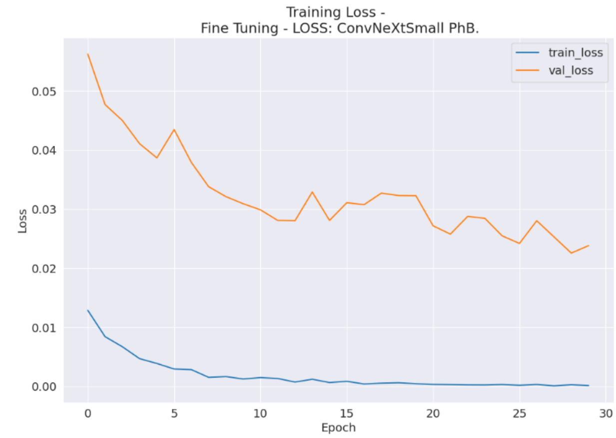

Fig. 7. Training loss and validation loss by fine-tuning of our model at scenario 1.

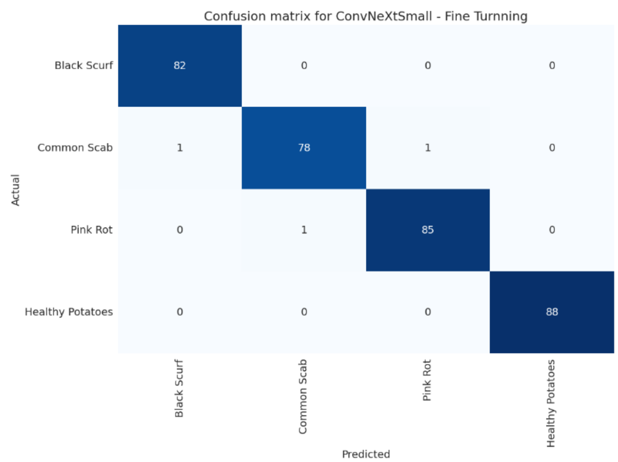

Fig. 8. Confusion matrix in fine-tuning for our model at scenario 1.

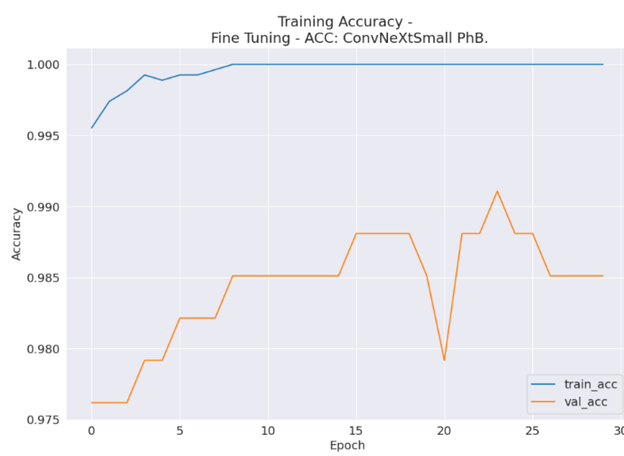

Fig. 9. Training accuracy and validation accuracy by fine-tuning of our model at scenario 2.

final result and helps the research article to have an overall assessment.

E. Scenario 4: the Result of Classifying Potato Diseases Into Seven Classes with ConvNeXt Model: Small, Tiny, and Large

TABLE V. THE ACCURACY OF CLASSIFYING POTATO DISEASE IMAGES INTO SEVEN CLASSES: HEALTHY, BLACK SCURF, COMMON SCAB, PINK ROT, BLACKLEG, DRY ROT, MISCELLANEOUS IN TRANSFER LEARNING AND FINE-TUNING, FOR EACH DEEP LEARNING MODEL

Model	Phase	Valid acc	Test acc	Precision	Recall	F1
ConvNeXt Large	Transfer Learning	100.00%	99.14%	99.18%	99.14%	99.14%
	Fine Tuning	-	-	-	-	-
ConvNeXtTiny	Transfer Learning	96.57%	97.09%	97.12%	97.09%	97.10%
	Fine Tuning	97.43%	97.60%	97.60%	97.60%	97.60%
Our model	**Transfer Learning**	**94.85%**	**94.69%**	**94.72%**	**94.69%**	**94.67%**
	Fine Tuning	**99.49%**	**98.97%**	**98.98%**	**98.97%**	**98.97%**

ConvNeXtLarger reaches the highest score in validation and test accuracy in Table V. Because it causes hardware overload and is too heavy for training small and medium datasets. Thus, this scenario points out that ConvNeXtLarger is not necessary for classifying this dataset and it can lead to a waste of resources and time to train the model. The customized ConvNeXtSmall model presents an effective and suitable classification although it is marginally lower than ConvNeXtLarger in transfer learning (I.e., lower 0.17% in test accuracy). In conclusion, the choice of models in the ConvNeXt family should be carefully considered because their performance may be very little different, although there are different requirements in terms of time and resources for training.

V. RESULTS AND COMPARISON

A. Results Explanation

Throughout scenarios 1, 2, and 3, our customized model presents effectiveness and sustainability when classifying many classes in Fig 15. Scenario 3 shows that test accuracy, precision, recall, and F1 score reached a surprise point (I.e., 99.49% in validation accuracy and 98.97% in test accuracy) in the seven potato disease classes classification. However, Scenarios 1 and 2 point out that in classifying a few of the classes the customized model worked unsuccessfully with the researcher's desire although it still reached a high score (i.e., 99.11% in validation accuracy and 96.88% in test accuracy). These issues will be explored and improved in subsequent studies.

In addition, Fig. 16 illustrates ConvNeXt family performance between ConvNeXtLarge, ConvNeXtTiny, and Customized ConvNeXtSmall. Specifically, Our model used fewer resources and a shorter time for training the model in classifying seven classes which reached a surprise result. However, ConvNext Large achieved slightly higher results than can be expected but it used a higher resource and time of computer

Fig. 10. Training loss and validation loss by fine-tuning of our model at scenario 2.

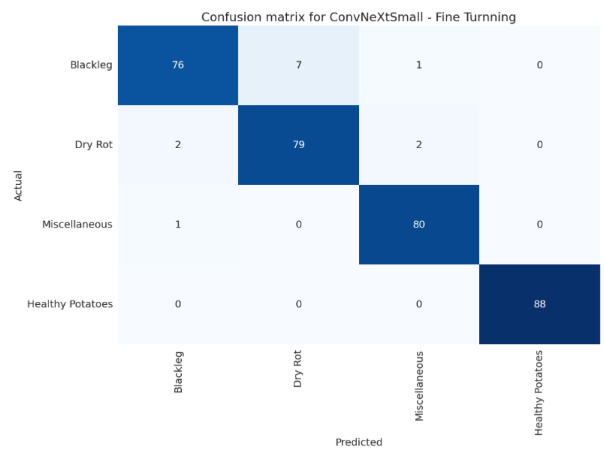

Fig. 11. Confusion matrix in fine-tuning for our model at scenario 2.

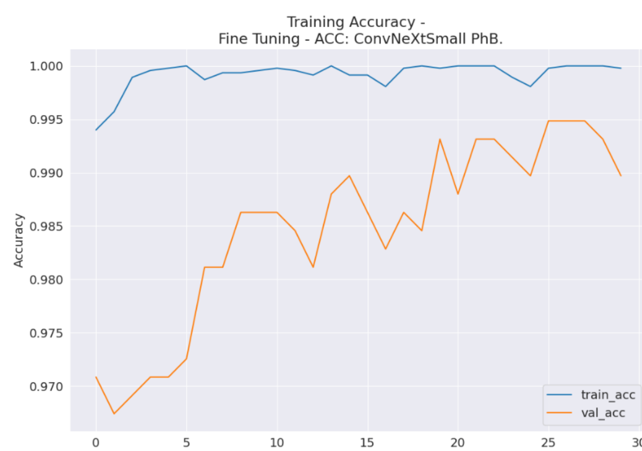

Fig. 12. Training accuracy and validation accuracy by fine-tuning of our model at scenario 3.

Fig. 13. Training loss and validation loss by fine-tuning of our model at scenario 3.

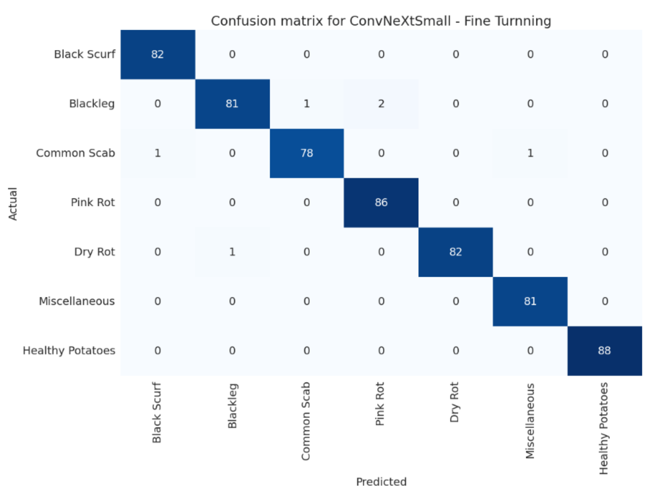

Fig. 14. Confusion matrix in fine-tuning for our model at scenario 3.

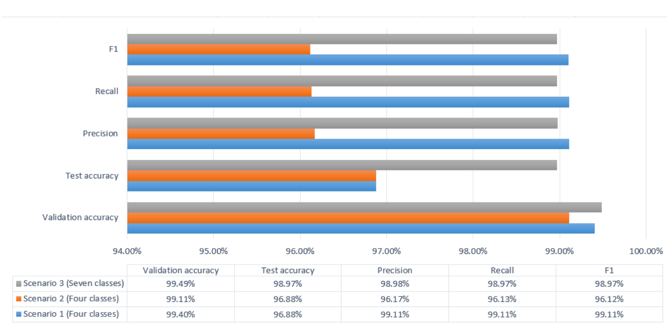

Fig. 15. The result of fine-tuning in our model.

for training and it can be crashed. As a result. Our model shows that it is an efficient and economical model for the classification of potato diseases.

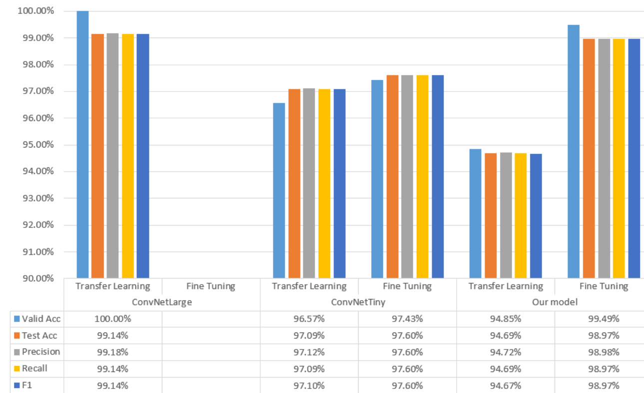

Fig. 16. The result of comparing in ConvNeXt family.

Besides, Grad-CAM for visual explanation and k-means clustering for image segmentation can significantly enhance the accuracy and interpretability of the models. Grad-CAM offers insights into the decision-making process of deep learning models by highlighting regions crucial for classification in Fig. 3, aiding in model validation and refinement. Meanwhile, in Fig. 4 k-means clustering was used for segmented potato images based on features like color and texture, enabling finer analysis and feature extraction for improved classification. Together, these techniques provide a deeper understanding of disease patterns and facilitate more precise identification and classification of potato diseases, crucial for effective agricultural management.

B. Comparison with others State-of-the-art Methods

To examine the accuracy of the proposed model that our article has just given out in the previous subsection, this subsection compares the accuracy score of the proposed model with other architectures. The result of getting the value of accuracy on the test set is illustrated in Table VI.

TABLE VI. Comparison with Others State-of-the-art Methods

Ref.	Proposed	Classes	Accuracy
Abeer A. Elsharif et al	CNN	4 classes	99.5%
Sofia Marino et al	FCN	6 classes	F1 score = 84%
Qinghua Su et al	CNN	5 classes	91.6%
Chenglong Wang et al	RFCN ResNet101	3 classes	98.7%
Kaili Zhang et al	VGG and U-Net	5 classes	97.55%
Khalid Hamza et al	CNN	5 classes	98% -100%
Ali Arshaghi et al	CNN	5 classes	98% -100%
Hyeon-Seung Lee et al	Mask R-CNN	-	93%
Israa Mohammed Hassoon et al	PDCNN	4 classes	91.3%
Proposed model (7 classes)			98.97%

Evaluating trade-offs according to task-specific priorities is necessary when comparing these measures. Precision and recall address more subtle elements, while accuracy offers a wide overview. The F1 score balances their interplay to ensure a well-informed assessment of classification models within the parameters of specific objectives.

VI. Conclusion

In agricultural management, accurately identifying and classifying potato diseases is crucial for maintaining crop health and yield. Recently, a study showcased significant advancements in this domain, introducing a novel model for disease classification with impressive accuracy rates. This model effectively categorizes potato images into various disease classes, including healthy specimens and those affected by ailments like black scurf, common scab, pink rot, blackleg, dry rot, and miscellaneous conditions. Our new model achieved remarkable performance metrics, boasting a validation accuracy of 99.49%, test accuracy of 98.97%, and an F1 score of 98.97% across seven disease classes. The success of this model lies in its utilization of the ConvNeXt family, a type of deep-learning architecture specifically designed for image analysis. Notably, the study employed transfer learning, a technique where a pre-trained model (i.e., such as ConvNeXtSmall) is fine-tuned to adapt to a new dataset. By adding dense and dropout layers and adjusting certain parameters, researchers were able to enhance performance.

To improve the decision-making process, the study utilized GradCam. Specifically, GradCam generates heatmaps to highlight regions of the image that are influential in the model's classification decision, aiding in the interpretability and trustworthiness of the results. Moreover, the research incorporated k-means clustering, a popular unsupervised machine learning algorithm, to segment images. K-means clustering partitions data into clusters based on similarity, enabling researchers to identify distinct regions within potato images corresponding to different disease manifestations. This segmentation facilitates more granular analysis and targeted interventions for disease management.

However, there are limitations to be solved. One such limitation is the reliance on the quality and diversity of the dataset. Improving data collection processes and expanding the dataset to encompass a wider range of potato diseases and variations in environmental conditions will be paramount for enhancing model robustness and generalizability. Looking ahead, future work will focus on further refining the model through improved data preparation techniques and leveraging advanced visualization methods. Additionally, expanding the dataset will be essential for accommodating the complexities and nuances inherent in real-world agricultural settings. By continually refining and enhancing the model, this research aims to contribute significantly to the advancement of potato disease detection and agricultural management practices.

In conclusion, the integration of fine-tuning, GradCam, and k-means clustering has propelled the efficacy of potato disease classification models. With ongoing efforts to overcome limitations and refine methodologies, this research holds promise for revolutionizing disease management strategies in agriculture.

Acknowledgment

Luong Hoang Huong was funded by the Vingroup Innovation Foundation (VINIF) 's Master, Ph.D. Scholarship Programme, code VINIF.2023.TS.049.

We would like to extend our heartfelt gratitude to Hao Van Tran, and Phuc Tan Huynh for their invaluable contributions to this project. Their dedication, expertise, and unwavering support have been instrumental in its success.

References

[1] M. E. Calıskan, A. Bakhsh, and K. Jabran, *Potato production worldwide*. Academic Press, 2022.

[2] E. Soare and I.-A. Chiurciu, "Study on the dynamics of potato production and worldwide trading during the period 2012-2019," *Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development*, vol. 21, 2021.

[3] J.-P. Goffart, A. Haverkort, M. Storey, N. Haase, M. Martin, P. Lebrun, D. Ryckmans, D. Florins, and K. Demeulemeester, "Potato production in northwestern europe (germany, france, the netherlands, united kingdom, belgium): characteristics, issues, challenges and opportunities," *Potato Research*, vol. 65, no. 3, pp. 503–547, 2022.

[4] A. Rana, V. Dua, S. Chauhan, and J. Sharma, "Climate change and potato productivity in punjab—impacts and adaptation," *Potato Research*, vol. 63, pp. 597–613, 2020.

[5] M. Abdolmaleky, K. N. Mahdei, and P. Nejatian, "Environmental sustainability assessment: Potato production in western iran," *Process Integration and Optimization for Sustainability*, vol. 6, no. 4, pp. 1063–1073, 2022.

[6] M. Kolychikhina, O. Beloshapkina, and C. Phiri, "Change in potato productivity under the impact of viral diseases," vol. 663, no. 1, p. 012035, 2021.

[7] M. Salimzyanov, V. Pervushin, R. Shakirov, and M. Kalimullin, "Improvement of technology and machines for growing potatoes in agriculture," *technology*, vol. 4, p. 5, 2020.

[8] M. del Mar Martínez-Prada, S. J. Curtin, and J. J. Gutiérrez-González, "Potato improvement through genetic engineering," *GM Crops & Food*, vol. 12, no. 1, pp. 479–496, 2021.

[9] H. L. Duc, T. T. Minh, K. V. Hong, and H. L. Hoang, "84 birds classification using transfer learning and efficientnetb2," in *International Conference on Future Data and Security Engineering*. Springer, 2022, pp. 698–705.

[10] K. V. Hong, T. T. Minh, H. L. Duc, N. T. Nhat, and H. L. Hoang, "104 fruits classification using transfer learning and densenet201 fine-tuning," in *Computational Intelligence in Security for Information Systems Conference*. Springer, 2022, pp. 160–170.

[11] K. D. D. Le, H. H. Luong, and H. T. Nguyen, "Patient classification based on symptoms using machine learning algorithms supporting hospital admission," in *Nature of Computation and Communication: 7th EAI International Conference, ICTCC 2021, Virtual Event, October 28–29, 2021, Proceedings 7*. Springer, 2021, pp. 40–50.

[12] H. T. Nguyen, N. K. T. Nguyen, C. L. H. Tran, and H. H. Luong, "Effects evaluation of data augmentation techniques on common seafood types classification tasks," in *Biomedical and Other Applications of Soft Computing*. Springer, 2022, pp. 213–223.

[13] H. T. Nguyen, Q. T. Quach, C. L. H. Tran, and H. H. Luong, "Deep learning architectures extended from transfer learning for classification of rice leaf diseases," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, 2022, pp. 785–796.

[14] A. Khorramifar, M. Rasekh, H. Karami, U. Malaga-Toboła, and M. Gancarz, "A machine learning method for classification and identification of potato cultivars based on the reaction of mos type sensor-array," *Sensors*, vol. 21, no. 17, p. 5836, 2021.

[15] S. Sabzi, Y. Abbaspour-Gilandeh, and J. I. Arribas, "An automatic visible-range video weed detection, segmentation and classification prototype in potato field," *Heliyon*, vol. 6, no. 5, 2020.

[16] Y. Yang, X. Zhao, M. Huang, X. Wang, and Q. Zhu, "Multispectral image based germination detection of potato by using supervised multiple threshold segmentation model and canny edge detector," *Computers and Electronics in Agriculture*, vol. 182, p. 106041, 2021.

[17] S. Yao, Q. Kang, M. Zhou, M. J. Rawa, and A. Abusorrah, "A survey of transfer learning for machinery diagnostics and prognostics," *Artificial Intelligence Review*, vol. 56, no. 4, pp. 2871–2922, 2023.

[18] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[19] Z. Zhu, K. Lin, A. K. Jain, and J. Zhou, "Transfer learning in deep reinforcement learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[20] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, "Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation," pp. 22 500–22 510, 2023.

[21] W. Chen, Y. Liu, W. Wang, E. M. Bakker, T. Georgiou, P. Fieguth, L. Liu, and M. S. Lew, "Deep learning for instance retrieval: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[22] H. Rasheed, M. U. Khattak, M. Maaz, S. Khan, and F. S. Khan, "Fine-tuned clip models are efficient video learners," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 6545–6554.

[23] A. A. Elsharif, I. M. Dheir, A. S. A. Mettleq, and S. S. Abu-Naser, "Potato classification using deep learning," *International Journal of Academic Pedagogical Research (IJAPR)*, vol. 3, no. 12, pp. 1–8, 2020.

[24] S. Marino, P. Beauseroy, and A. Smolarz, "Unsupervised adversarial deep domain adaptation method for potato defects classification," *Computers and Electronics in Agriculture*, vol. 174, p. 105501, 2020.

[25] Q. Su, N. Kondo, D. F. Al Riza, and H. Habaragamuwa, "Potato quality grading based on depth imaging and convolutional neural network," *Journal of Food Quality*, vol. 2020, pp. 1–9, 2020.

[26] C. Wang and Z. Xiao, "Potato surface defect detection based on deep transfer learning," *Agriculture*, vol. 11, no. 9, 2021. [Online]. Available: https://www.mdpi.com/2077-0472/11/9/863

[27] K. Zhang, S. Wang, Y. Hu, H. Yang, T. Guo, and X. Yi, "Evaluation method of potato storage external defects based on improved u-net," *Agronomy*, vol. 13, no. 10, p. 2503, 2023.

[28] K. Hamza, S. un Nisa, and G. Irshad, "A review on potato disease detection and classification by exploiting deep learning techniques," *Journal of Agriculture and Veterinary Science*, vol. 1, no. 2, pp. 81–88, 2022.

[29] A. Arshaghi, M. Ashourian, and L. Ghabeli, "Potato diseases detection and classification using deep learning methods," *Multimedia Tools and Applications*, vol. 82, 07 2022.

[30] H.-S. Lee and B.-S. Shin, "Potato detection and segmentation based on mask r-cnn," *Journal of Biosystems Engineering*, vol. 45, pp. 233–238, 2020.

[31] I. Hassoon, S. Kassir, and M. Riyadh, "Pdcnn: Framework for potato diseases classification based on feed foreword neural network," *Baghdad Science Journal*, vol. 18, p. 1012, 06 2021.

[32] R. A. Sholihati, I. A. Sulistijono, A. Risnumawan, and E. Kusumawati, "Potato leaf disease classification using deep learning approach," pp. 392–397, 2020.

[33] A. Singh and H. Kaur, "Potato plant leaves disease detection and classification using machine learning methodologies," vol. 1022, no. 1, p. 012121, 2021.

[34] R. Mahum, H. Munir, Z.-U.-N. Mughal, M. Awais, F. Sher Khan, M. Saqlain, S. Mahamad, and I. Tlili, "A novel framework for potato leaf disease detection using an efficient deep learning model," *Human and Ecological Risk Assessment: An International Journal*, vol. 29, no. 2, pp. 303–326, 2023.

[35] D. Tiwari, M. Ashish, N. Gangwar, A. Sharma, S. Patel, and S. Bhardwaj, "Potato leaf diseases detection using deep learning," pp. 461–466, 2020.

[36] M. A. Iqbal and K. H. Talukder, "Detection of potato disease using image segmentation and machine learning," pp. 43–47, 2020.

[37] K. K. Chakraborty, R. Mukherjee, C. Chakroborty, and K. Bora, "Automated recognition of optical image based potato leaf blight diseases using deep learning," *Physiological and Molecular Plant Pathology*, vol. 117, p. 101781, 2022.

[38] A. Taner, M. Mengstu, K. Selvi, H. Duran, I. Gür, and N. Ungureanu, "Apple varieties classification using deep features and machine learning," *Agriculture*, vol. 14, p. 252, 02 2024.

[39] D. Asriny, S. Rani, and A. F. Hidayatullah, "Orange fruit images classification using convolutional neural networks," *IOP Conference Series: Materials Science and Engineering*, vol. 803, p. 012020, 05 2020.