



Robust Deep Learning Model for Black Fungus Detection Based on Gabor Filter and Transfer Learning

Esraa Hassan¹, Fatma M. Talaat¹, Samah Adel², Samir Abdelrazek³, Ahsan Aziz⁴,
Yunyoung Nam^{4,*} and Nora El-Rashidy¹

¹Machine Learning and Information Retrieval Department, Faculty of Artificial Intelligence, Kafrelsheikh University, Kafrelsheikh, Egypt

²Electronics and Communication Engineering Department, Faculty of Engineering Horus University, Damietta, Egypt

³Information Systems Department, Faculty of Computers and Information, Mansoura University, Mansoura, 35516, Egypt

⁴Department of ICT Convergence, Soonchunhyang University, Asan, Korea

*Corresponding Author: Yunyoung Nam. Email: ynam@sch.ac.kr

Received: 06 November 2022; Accepted: 13 April 2023; Published: 28 July 2023

Abstract: Black fungus is a rare and dangerous mycology that usually affects the brain and lungs and could be life-threatening in diabetic cases. Recently, some COVID-19 survivors, especially those with co-morbid diseases, have been susceptible to black fungus. Therefore, recovered COVID-19 patients should seek medical support when they notice mucormycosis symptoms. This paper proposes a novel ensemble deep-learning model that includes three pre-trained models: reset (50), VGG (19), and Inception. Our approach is medically intuitive and efficient compared to the traditional deep learning models. An image dataset was aggregated from various resources and divided into two classes: a black fungus class and a skin infection class. To the best of our knowledge, our study is the first that is concerned with building black fungus detection models based on deep learning algorithms. The proposed approach can significantly improve the performance of the classification task and increase the generalization ability of such a binary classification task. According to the reported results, it has empirically achieved a sensitivity value of 0.9907, a specificity value of 0.9938, a precision value of 0.9938, and a negative predictive value of 0.9907.

Keywords: Black fungus; COVID-19; Transfer learning; pre-trained models; medical image

1 Introduction

Mucormycosis is a fatal infection responsible for 9%–11% of fungal infections [1]. It is formalized, usually caused by direct contact with spores in the surrounding environment (i.e., soil, air, etc.).



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Five types of mucormycosis were identified, including rhino cerebral (45%–50%), cutaneous (9%–15%), pulmonary (10%–11%), and gastrointestinal (2%–11%). Clinical infection with mucormycosis progresses rapidly, causing death in a few days.

It primarily affects people with a weakened immune system, such as those taking medications to fight germs. Others attribute this to the virus's nature as an immunosuppressive factor that aids in spreading the fungus. The black fungus affects the patient's skin, resulting in black spots on the face and body. Unfortunately, this effect is interleaved with other skin infections such as poikiloderma of Civatte, melasma, linea nigra, etc. Therefore, the early diagnosis of black fungus using the traditional visual technique is considered a challenge.

Deep learning is an advanced type of artificial neural network since it has more layers that allow for higher degrees of abstraction and better data predictions. Among deep learning technologies, convolution neural networks (CNNs) have specifically demonstrated to be valuable tools capable of handling a wide range of computer vision tasks [2]. The deep belief convolution network's success is based on its capacity to learn midlevel and high-level abstractions. The deep belief convolution network's success is based on its ability to learn midlevel and high-level abstractions from the studied image. This makes it a highly effective technique for detecting and localizing cancers in natural photos [3].

Image classification is considered one of the main tasks in computer vision and is widely used in various domains such as object detection, age recognition, etc. The concept of Transfer Learning (TL) was developed to resolve this issue by using pre-trained models (i.e., Resnet, Inception, Xception, VGG, alexnet, etc.) that had previously been trained on massive datasets (i.e., Imagenet). Such pre-trained models help improve classification results, especially with limited datasets. Ensemble models improve classification accuracy and build models with lower bias and variance. The key points in developing an effective ensemble model are the current methods for aggregating all models. Recently, ensemble learning has been used in several medical domains [4], for example. Mahmoud et al. [5] propose a stacking ensemble model for predicting mortality inside the ICU. [6,7].

Our paper's main contribution can be summarised as follows: (i) proposing an ensemble model that can be used to predict black fungus and distinguish it from other skin infections; As far as we know, no published research could differentiate between black function and other skin illnesses. (ii) developing an ensemble model for predicting black fungus and distinguishing it from other skin infections. (iii) utilizing the Gabor filter to extract the essential features. (iv) Using different pre-trained models to design our model results in an accurate result with excellent generalization capability. (v) To assure and validate the suggested framework and the findings, further investigation of the impacts of utilizing various optimizers, such as ADAM and SGD, under the strict supervision of medical professionals is required. (vi) In statistical comparisons with other pre-trained models, the suggested model outperformed them all. The EBF model demonstrated 0.9860 accuracy and a 0.9860 F-score, demonstrating that our model resists overfitting. The remaining sections are arranged as follows: The related works are discussed in Section 2. Techniques and introductions are discussed in Section 3. Section 4 provides specifics on the suggested framework. The results and discussion were presented in Sections 5 and 6 concluded the paper.

2 Related Work

DL is a popular area in artificial intelligence with promising results, recently gaining popularity in automatic diagnosis, classification, and segmentation systems in different domains. The existence of the COVID-19 pandemic and its consequences has rapidly increased the need for AI models.

Therefore, it must be fast and accurate to provide timely support to patients and medical experts. Recently, several radiology images have been utilised for disease detection, as illustrated in Table 1. In [8], Mansour et al. presents a method based on established techniques, such as convolutional neural networks (CNN) and support vector machines (SVM). They analyze a dataset of eye photographs of patients with and without black fungus infection, collected from the real-time records of individuals affected by COVID, followed by the black fungus. Khayyat et al. [9] present a combination of traditional therapeutic approaches and deep learning frameworks to detect skin lesions accurately. Their research examines the potential of utilizing image data, handcrafted lesion features, and patient-centric metadata for successful skin cancer diagnosis. The ISIC2018 dataset demonstrated 90.49% sensitivity and 97.76% specificity, while the more extensive and more varied ISIC2019 dataset achieved 85.58% sensitivity and 98.29% specificity. Armstrong et al. [10] utilized two medical datasets to assess their method, including skin lesion diagnosis with the ISIC 2018 challenge and thoracic disease classification with a chest X-ray. Their technique outperforms numerous robust semi-supervised methods in both single-label and multi-label image classification scenarios. Bissoto et al. [11] presented the same in assisting physicians by identifying shoulder pictures acquired from X-ray images.

The performance of 26 deep learning-based pre-trained models on medical dataset VGG (19), Inception (v3), and Resnet were utilized as the pre-trained models employed (50). Ensemble models were used with pre-trained models to achieve a complete accuracy of 0.8455, 0.8472, ROC and area under thr roc curve (AUC) of 0.8862 percent and 0.8695 percent, respectively. Panda et al. [12] apply several transfer learning techniques, such as Visual Geometry Group (VGG19), Xception, Densenet, Inception, MobileNet, NasNetMobile, and Resnet. Based on the ISIC-2019 Challenge for categorizing skin lesions and recent publications, Adegun et al. [13] assess the state-of-the-art in the classification of dermoscopic pictures. Transfer learning and model fine-tuning are used to adapt several deep neural network architectures previously trained on the ImageNet data set to a combined training data set of publically accessible dermoscopic and clinical images of skin lesions [14]. The results demonstrate that their suggested technique can be successfully built as a decision support system for clinical applications such as the classification of wound images. According to one or more pre-trained models by an individual, the prior and traditional procedures share several difficulties. In some methods, the average performance is not satisfactory. The impact on our suggested model and achieving strong accuracy for black fungus infection depends on the ensemble technique.

Table 1: The state of the art of Robust Deep Learning Model for Black Fungus Detection

Year	Author	Model	Task	Metrics	Dataset
2020	[15]	Semi-supervised methods	Skin lesion classification	N/A	ISIC 2018
2020	[16]	VGG (19) Inception (v3)	Skin lesion classification	AUC = 0.8862	ISIC 2018
2020	[17]	CNN	Skin lesion classification	ACC = 94.28	Covid-19
2020	[14]	Data-driven model	Skin lesion classification	Trap = 62.4 ± 3.3	ImageNet ObjectNet ISIC 2018 Task 1

(Continued)

Table 1 (continued)

Year	Author	Model	Task	Metrics	Dataset
2021	[13]	VGG19,	Skin lesion classification transfer learning		ISIC 2018
2021	[12]	Ensemble	Skin lesion classification	Acc+ 0.634	ISIC 2019

3 Techniques and Preliminaries

3.1 Feature Extraction Technique

The Gabor filter is a convolution filter that contains sinusoidal and Gaussian terms. It considers the optimum features that allow multi-resolution analysis to extract the optimal qualities in both the frequency and spatial domains. The Gaussian component generates the weights, whereas the sine component gives the directionality. In this paper, we used a 2D Gabor filter that could be represented with the following equations:

$$g(X, y, \lambda, \theta, \psi, \sigma, \gamma) = \text{Exp}\left(\frac{X^2 + \gamma^2 y^2}{2\sigma^2}\right) \text{Exp}\left(i\left(2\pi \frac{x'}{\lambda} + \psi\right)\right) \quad (1)$$

where $x' = x\cos\theta + y\sin\theta$

And $y' = -x\sin\theta + y\cos\theta$

where λ is the wavelength of the component, θ is the orientation of the filter, ψ is the phase offset, σ is the standard deviation of the gaussian kernel, and γ is the spatial aspect ratio.

3.2 Transfer Learning Approaches

In TL, the model used to train one task is reused for another related task after enhancement curated toward the specific required task. It is used for particular learning tasks where sufficient labeling images are not presented, such as in the medical domain. Transfer learning is considered a specific case of deep neural networks with many parameters that need training. Using TL, the model starts with initial weights that are just a slight modification. A pre-trained model is used for new tasks in two main ways [15]. The classifier is trained on top of the pre-trained model to perform the classification task after the pre-trained model is treated as a feature extractor. The entire network then adjusted the required new task.

As a result, the pre-trained model weights are used as the new model's starting point and updated during training. Transfer learning is frequently employed in the medical field to address the issue of a small amount of data.

3.3 Ensemble Learning

Ensemble learning is a meta-approach used to achieve better performance by combining several models' predication. It proved efficient over traditional learning. Ensemble learning is widely used in deep learning; several works utilized ensemble learning to improve the model's overall performance. This improvement returns to building several models in a way that could reduce variance (i.e., bagging) [16], the process usually used to minimize bias (i.e., boosting). It also contributes to generating a

robust model with high generalization ability. Ensemble learning has two main types of learning. First, homogeneous learning that uses the same algorithms is used in different bagging and boosting) and heterogeneous learning that uses various algorithms on the same dataset (i.e., voting and stacking). The result is combined and voted from all methods in a way that prevents overfitting [17].

4 The Proposed Ensemble Black Fungus (EBF)

We propose the EBF model depends on pre-trained models of the ImageNet dataset for medical imaging tasks, compared with our other proposed model, Accurate Black Fungus (ABF), which depends on three transfer learning pre-trained models. The primary purpose is to offer a classification model that could overcome the lack of training data, especially in the medical image domain [18–20]. The general steps to build the classification model can be summarized as follows. (i) Loading dataset (binary class), which includes 3225 images of black fungus and uses several Gabor filters (i.e., horizontal, vertical, diagonal, etc.). To extract the essential features, (ii) utilizing cross-validation techniques used to choose training and testing data, (iii) integrating three different pre-trained models, including Inception, VGG (19), and resnet (50), are used to generate a binary classification model, (iv) for each pre-trained model. The last layer was removed and replaced with a group of CONV, Maxpooling, and dropout layers. Details of the added layers mentioned in subsection 4.3: (v) implementing an ensemble model that integrates the three pre-trained models, then using relative voting to obtain the result, and (vi) using various metrics to measure the reliability and accuracy of the proposed model. The superiority of the proposed model result ensures the ability to be utilized for any medical imaging task with a small number of labeled images. Fig. 1 shows the general architecture of the proposed model.

4.1 Augmentation Techniques

In this paper, we utilized some data augmentation techniques. We employed random rotation ranges between 45 and 180 degrees, a width shift range with a value of 0.2, and a height shift range of 0.2, applying horizontal flip and vertical flip. As mentioned above, we used the Gabor filter to extract the optimal features. Numerous digital features (kernels) can be identified by varying the Gabor filter parameters [21]. For example, changing the theta value to zero will generate horizontal features while making it cause vertical features will not. One filter is usually insufficient to extract the essential features [22,23]. Therefore, tens of filters are created and passed to the AFB model. The following Table 2 includes some examples of the generated features.

4.2 The Classification Model (AFB)

The AFB model is based on an efficient DCNN model that integrates many advanced elements to address issues, including better feature extraction. In Table 2, a thorough explanation of the suggested model is provided. The proposed architecture's input in the black fungus classification models is (165 154), the changed input size is (80 80). Two conventional convolutional layers that operate in succession are the model's initial stages. The first one has a kernel size of 2×2 and a filter number of 128. The kernel size of the second convolution is 2×2 . ReLu layers come after both convolutional layers. Finally, one dropout layer is used to join fully connected layers. The result is finalized using SoftMax.

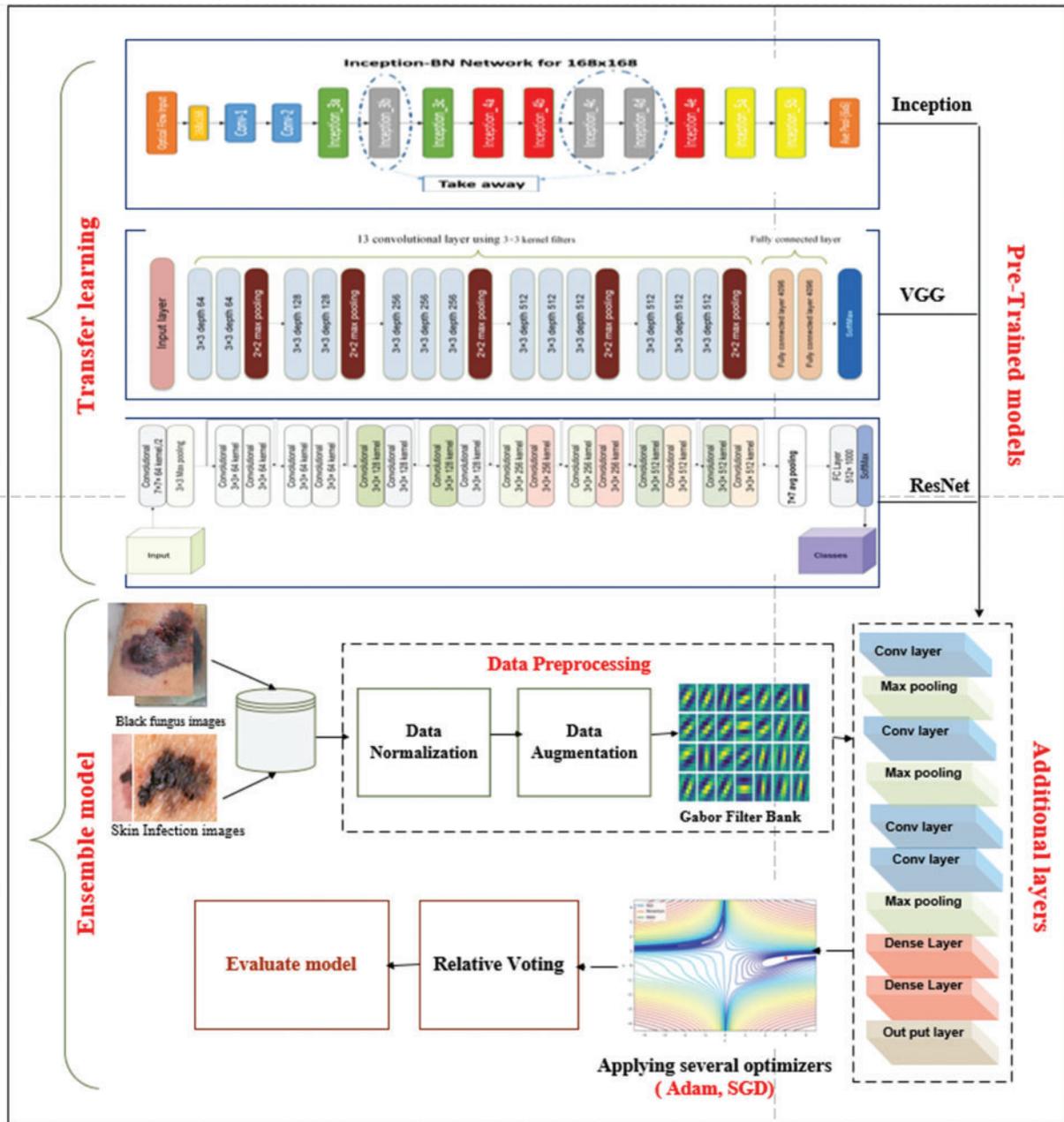


Figure 1: General architecture of proposed ensemble model

5 Results and Discussion

The experimental setup, the preparation procedures, the dataset used, the assessment metrics, and the outcomes of our suggested strategy are all included in this part. The performance of our proposed model was compared to the performance of each pre-trained model separately to ensure its efficacy. The result is finalized using SoftMax. Algorithm 1 shows the steps of the proposed model.

Table 2: The values and sizes of Gabor filter parameters

Num/Parameter	K size	Lamda(λ)	Theta (θ)	Phi(ψ)	Sigma(σ)	gamma (γ)
1	15	1 * np.pi/4	1 * np.pi/4	0.8	5	0.1
2	15	1 * np.pi/4	1 * np.pi/2	0.8	5	0.1
3	15	1 * np.pi/4	1 * np.pi/4	1	5	0
4	9	1 * np.pi/4	1 * np.pi/2	1	10	0.99
5	15	1 * np.pi/4	1 * np.pi/4	0	0	0.1

Algorithm 1: Mathematical equation for the ensemble pre-trained model (Proposed Model)

1 Input	For n training sample images, N Number of pre-trained models, w_j is the adaptive layer in the i -th transferred model, X_i the j -th image sample, y_i the j -th image sample
2	For our binary classification problem, the loss function could be calculated using the following equation $\text{Arg min} = \sum_{j=1}^n \sum_w \{((f(x_i - w_i, p_i), -y_j^2))\}$
3 Decision function	The decision function of this method is expressed as follows $F'_{(x)} = \text{sign}(\sum \text{label}(f((f(x_i - w_i, p_i), -y_i)))$
4	Where Sign(x) is a function described as follows $\text{Sign}(x) = \begin{cases} 1, & \text{if } x > 0 \\ -1, & \text{if } x \leq 0 \end{cases}$
5	Note that the sign (.) values calculate the function's sum of all output labels of a sample in all pre-trained models; it is considered its predicted value when the voting method is used.

5.1 Dataset

The Black Fungus and Melanoma (BFM) dataset [24] contains three compressed folders with images inside. The first file includes 260 images of black fungus after applying data preprocessing, which becomes 1300 images. The second file contains 385 images of melanoma after using data preprocessing, which becomes 1925. The BFM dataset includes images, as shown in Fig. 2, that facilitate training and validation while utilizing deep learning algorithms for skin disease recognition and classification. They were randomly split into 60/20/20 ratios for training and validation and testing the DL models, respectively, to have the following in each class. 1935 for training, 645 for confirmation, and 645 for testing.

5.2 Evaluation Metrics

The evaluation metrics of the black fungus categorization scenario are presented in this section. The following assessment measures were used to assess our model: accuracy, F1-score, sensitivity, specificity, precision, negative predictive value, false positive rate, false discovery rate, false negative rate, and Matthews Correlation Coefficient. These evaluation measures were computed using the TN (true Negative), TP (true positive), FN (false negative), and FP (false positive) calculations. The number of correctly categorized negative and positive instances, respectively, is defined as TN and TP. The quantity of incorrectly identified positive and negative models is also defined as FN and FP, respectively. Table 3 lists each evaluation metric's definition and equation.

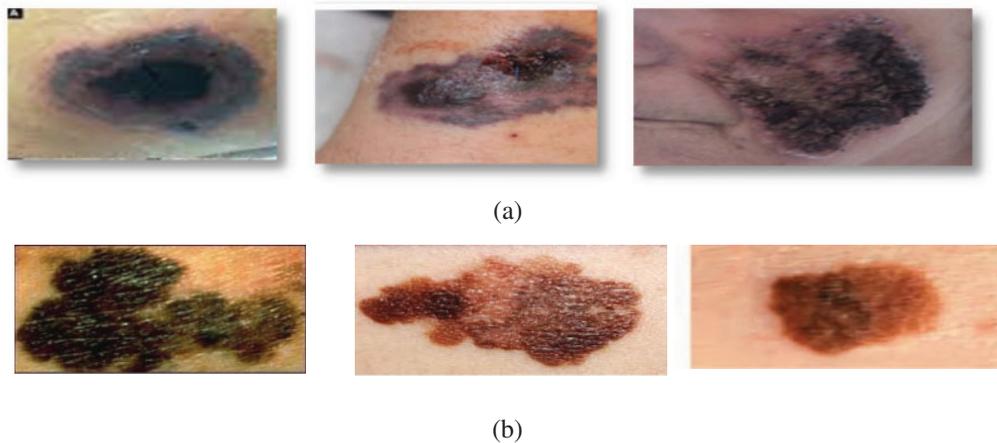


Figure 2: (a) Samples of black fungus cases for skin infection, (b) Samples of Melanoma cases for skin infection

Table 3: Evaluation metrics

Equation	#	Equation	#
Recall = $\frac{TP}{(TP + FN)}$	(2)	FPR = $\frac{FP}{(FP + TN)}$	(7)
Specificity = $\frac{TN}{(FP + TN)}$	(3)	NPV = $\frac{TN}{(TN + FN)}$	(8)
Precision = $\frac{TP}{(TP + FP)}$	(4)	Accuracy = $\frac{TN + TP}{TP + FP + TN + FN}$	(9)
FDR = $\frac{FP}{(FP + TP)}$	(5)	F1-score = $\frac{2TP}{2TP + FP + FN}$	(10)
FNR = $\frac{FN}{(FN + TP)}$	(6)	MCC = $\frac{TP * TN - FP * FN}{\sqrt{((TP + FP) * (TP + FN) * (TN + FP) * (TN + FN))}}$	(11)

Notes: NPV: Negative Predictive Value, FPR: False Positive Rate, FDR: False Discovery Rate, FNR: False Negative Rate, MCC: Matthews Correlation Coefficient.

Several observations concerning hyper-parameters were made throughout the training and testing phase, shown in Table 4. This was also useful for improving the model's performance by changing the values of the variables. The Adam optimizers are briefly summarised using these parameters:

Table 4: The parameters values for Adam and SGD optimizers

Optimizer	Parameter's value
Adam	Lr = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, Decay = 0, amsgrad = false
SGD	Lr = 0.001, Momentum = none, Decay = 0, Nesterov = false

- **Lr**: learning_rate: The learning rate to use in the algorithm. It defaults to a value of 0.001.
- **β_1** : The value for the exponential decay rate for the 1st-moment estimates. It has a default value of 0.9.
- **β_2** : The value for the exponential decay rate for the 2nd-moment estimates. It has a default value of 0.999.
- **amsgrad**: It is a boolean that specifies whether to apply the AMSGrad variant of this algorithm.

5.3 Training Using Several Pre-Trained Architectures

This section compares the AFB model with Inception V3 Residual 50, VGG (16), and VGG (19). Observation: the effect of changing between Adam and SGD optimizers on the final performance of learning curves.

5.3.1 Inception V3

In this paper, we use Inception V3 to concentrate on the image’s critical parts that may vary in size. The classifiers did not significantly contribute until the end of the training period when the Batch Norm in the Auxiliary Classifiers was used to strengthen the network. After the AFB training model was based on the Inception V3 pre-trained model with 500 epochs, the training accuracy value was 0.9341 percent, which differed from the training accuracy value as illustrated in Figs. 3 and 4, where some overfitting occurred between the training and testing values depending on the Adam optimizer that compares with the SGD optimizer in Figs. 5 and 6.

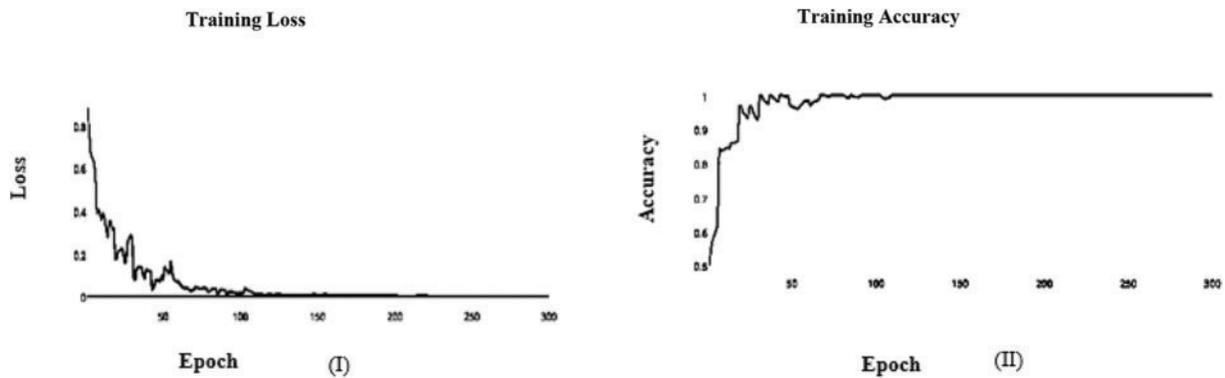


Figure 3: (I) ABF architecture learning accuracy, (II) ABF architecture learning loss (Adam)

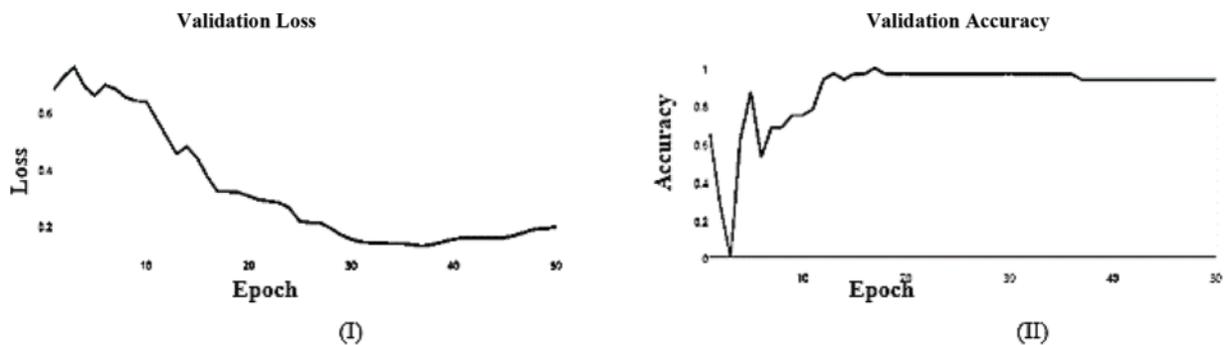


Figure 4: (I) ABF architecture learning accuracy, (II) ABF architecture learning loss (Adam)

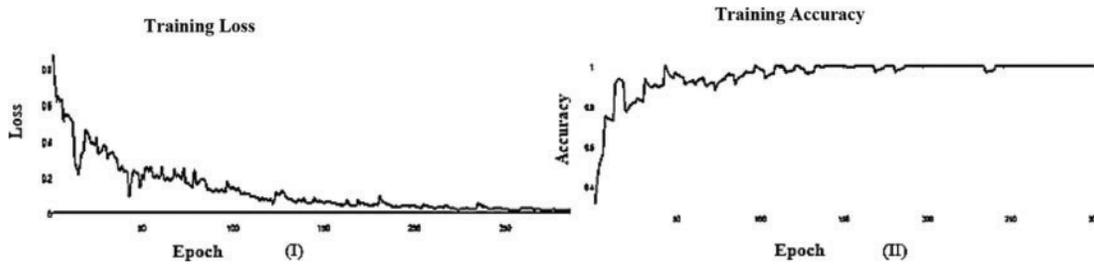


Figure 5: (I) ABF architecture learning accuracy, (II) ABF architecture learning loss (SGD)

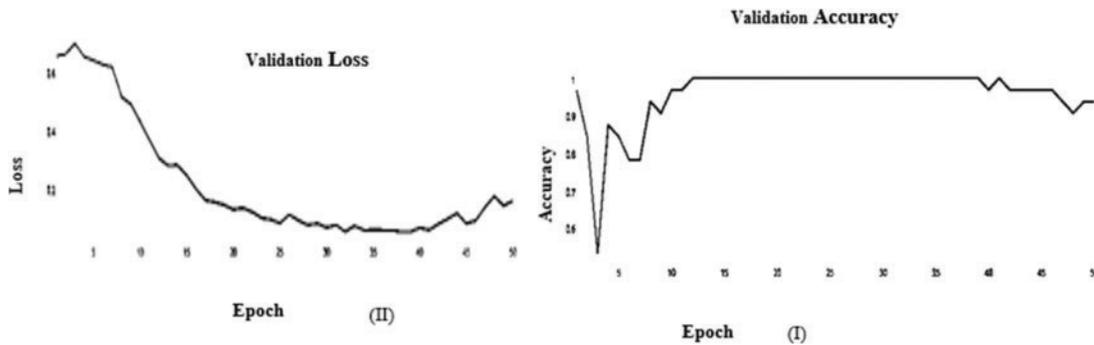


Figure 6: (I) ABF architecture validation accuracy, (II) ABF architecture validation loss (SGD)

The learning curve helps us to check whether a model has learned everything it can about the data. A perfect learning curve that adapts to new data. The learning curves for testing and training converge at similar values with a decreasing gap.

SGD and Adam optimizer achieve a balanced slip for the training and testing stages. We observed that the Adam optimizer with Inception V3 is more suitable in the validation stage than the SGD optimizer. Results of the inception model in the training and testing process are shown in [Tables 5](#) and [6](#), respectively.

Table 5: The overall standard metrics for the training stage of the Inception V3 pre-trained model

Inception V3 (Training)										
Metrics	Recall	Specificity	Precision	NPV	FPR	FDR	FNR	Accuracy	F-score	MCC
Adam	1.0000	0.8889	0.9890	1.0000	0.1111	0.0110	0.0000	0.9899	0.9945	0.9376
SGD	1.0000	0.8889	0.9892	1.0000	0.1111	0.0108	0.0000	0.9901	0.9946	0.9377

Table 6: The overall standard metrics for the testing stage of the Inception V3 pre-trained model

Inception V3 (Testing)										
Metrics	Recall	Specificity	Precision	NPV	FPR	FDR	FNR	Accuracy	F-score	MCC
Adam	0.9130	0.9444	0.9438	0.9140	0.0556	0.0562	0.0870	0.9286	0.9282	0.8576

5.3.2 Residual Neural Network

It's a CNN modification proven to improve image net classification performance empirically. These layers can classify images and perform other deep network activities. After the AFB training model was based on ResNet(50) pre-trained model with 500 epochs, the training accuracy value was 0.9341 percent, which differed from the training accuracy value. As illustrated in Figs. 7 and 8, where some overfitting occurred between the training and testing values depending on the Adam optimizer that compares with the SGD optimizer in Figs. 9 and 10. The proposed model uses the pre-trained resnet50 model after training. The research model's results during the training and testing stages are shown in Tables 7 and 8, respectively.

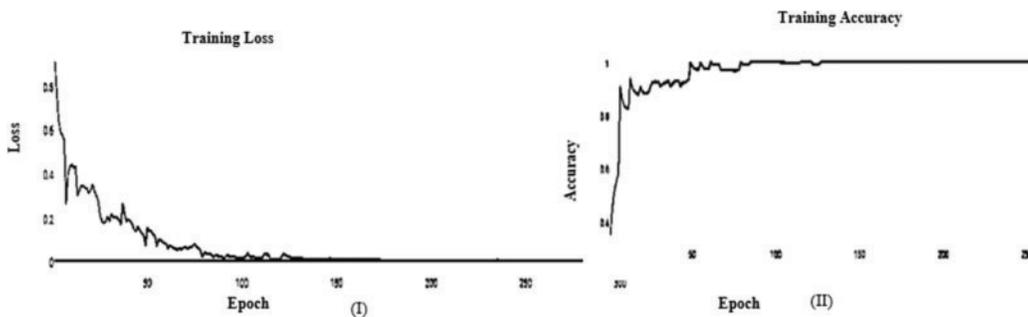


Figure 7: (I) ABF architecture learning accuracy, (II) T ABF architecture learning loss (Adam)

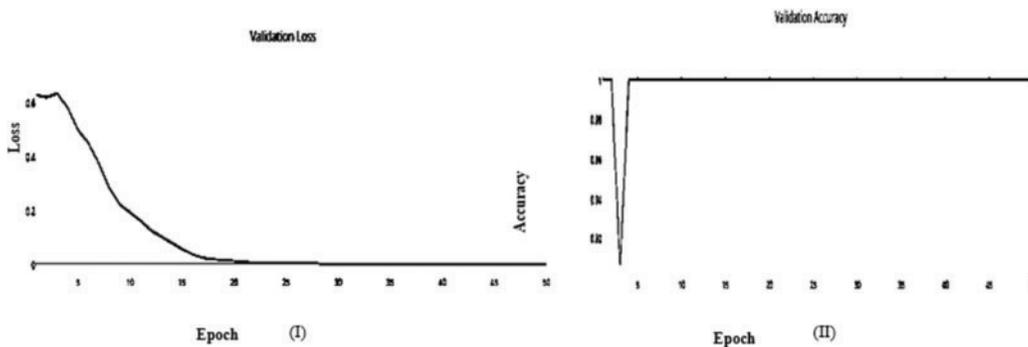


Figure 8: (I) ABF architecture validation accuracy, (II) ABF architecture validation loss (Adam)

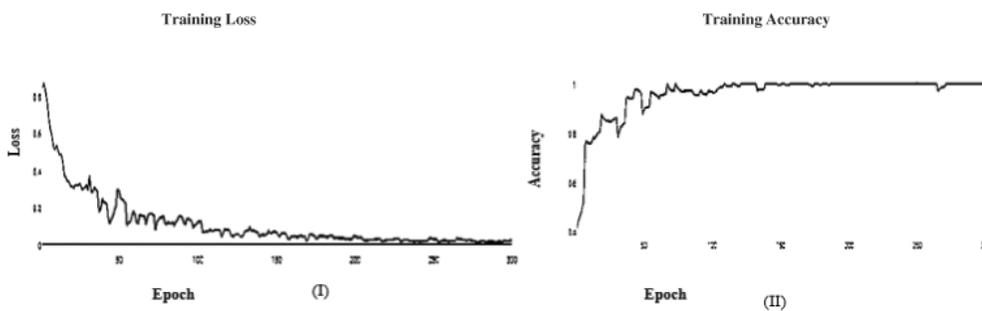


Figure 9: (I) ABF architecture learning accuracy, (II) ABF architecture learning loss (SGD)

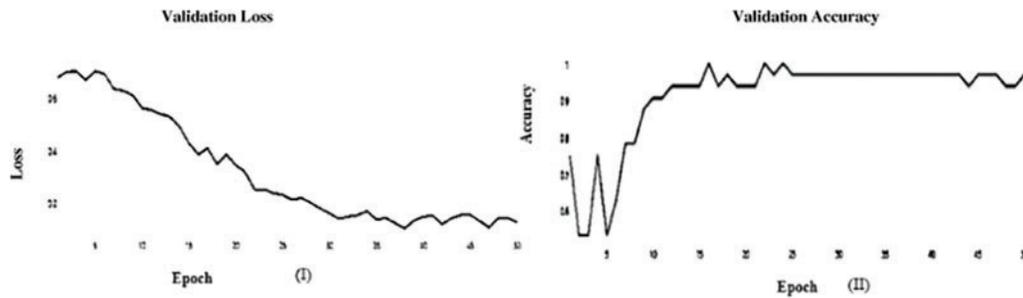


Figure 10: (I) ABF architecture validation accuracy, (II) The Loss value of ABF architecture validation loss (SGD)

Table 7: The overall standard metrics for the training stage of the resnet(50) pre-trained model

VGG (19) (Training)										
Metrics	Recall	Specificity	Precision	NPV	FPR	FDR	FNR	Accuracy	F-score	MCC
Adam	0.9901	0.9877	0.9901	0.9877	0.0123	0.0099	0.0099	0.9890	0.9901	0.9778
SGD	1.0000	0.9877	0.9903	1.0000	0.0123	0.0097	0.0000	0.9945	0.9951	0.9890

Table 8: The overall standard metrics for the testing stage of Resnet (50) pre-trained model

Resnet (50) Testing										
Metrics	Recall	Specificity	Precision	NPV	FPR	FDR	FNR	Accuracy	F-score	MCC
Adam	0.9668	0.9554	0.9581	0.9646	0.0446	0.0419	0.0332	0.9612	0.9624	0.9224
SGD	0.9524	0.9877	0.9901	0.9412	0.0123	0.0099	0.0476	0.9677	0.9709	0.9356

5.3.3 Visual Geometry Group (VGG (19))

In the initial layer, VGG (19) uses the total model mentioned above to focus on short window widths and strides. The optimum performance of the model when using VGG (19) was determined by looking at its sensitivity, specificity, precision, negative predictive value, false positive rate, false discovery rate, accuracy, F-score, and Matthews Correlation Coefficient. Increasing the number of training epochs is an excellent way to improve the model's performance while avoiding overfitting. After the AFB training model was based on the VGG (19) pre-trained model with 500 epochs, the training accuracy value was 0.9341 percent, which differed from the training accuracy value as illustrated in Figs. 11 and 12, where some overfitting occurred between the training and testing values depending on the Adam optimizer that compares with the SGD optimizer in Figs. 13 and 14. Following training, the proposed model employs the pre-trained resnet50 model. The VGG (19) model's training and testing results are shown in Tables 9 and 10.

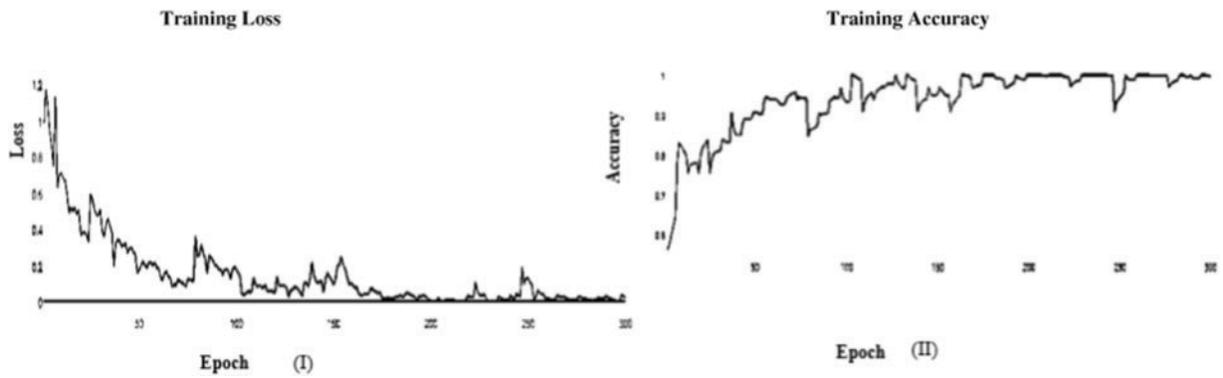


Figure 11: (I) ABF architecture learning accuracy, (II) ABF architecture learning loss (Adam)

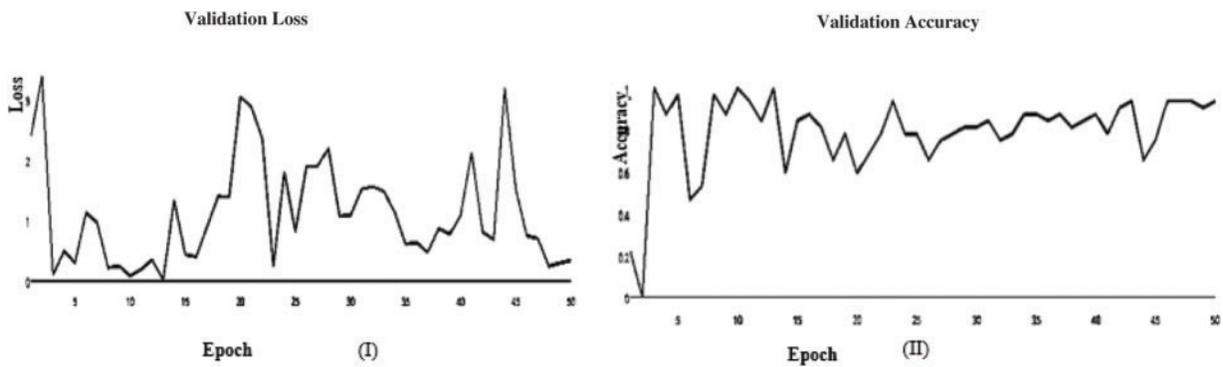


Figure 12: (I) ABF architecture validation accuracy, (II) ABF architecture validation loss (Adam)

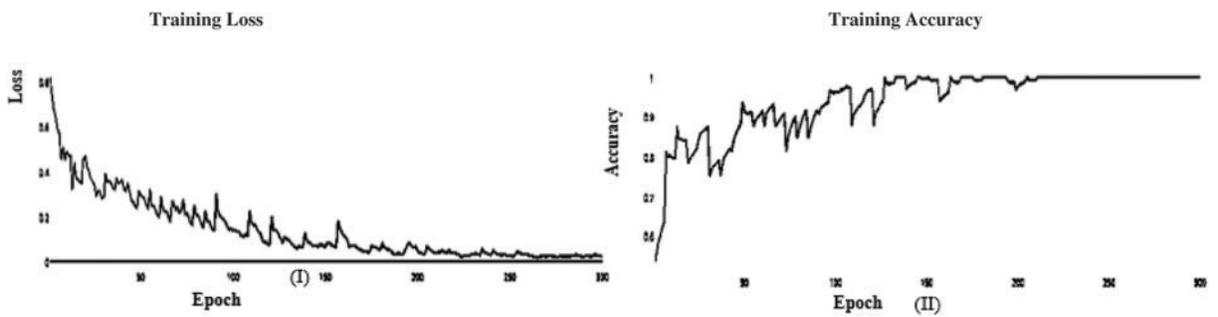


Figure 13: (I) of ABF architecture learning accuracy, (II) ABF architecture learning loss (SGD)

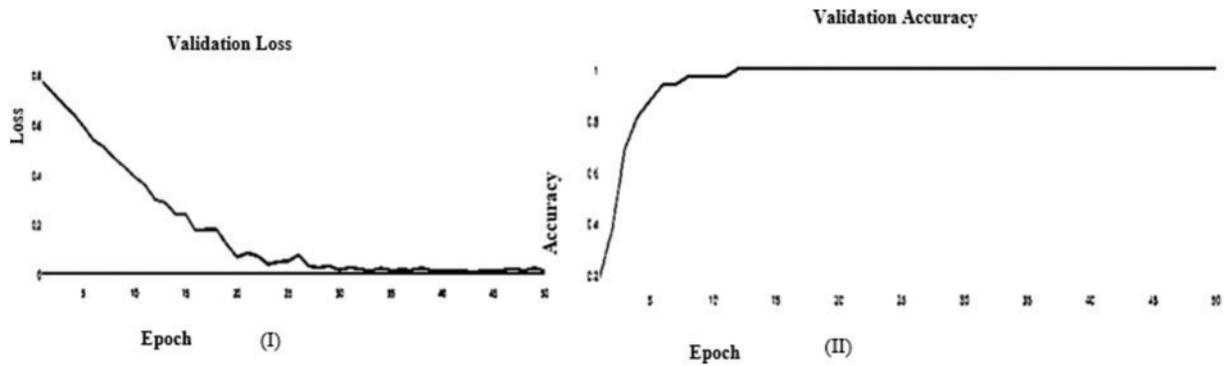


Figure 14: (I) ABF architecture validation accuracy, (II) ABF architecture validation loss (SGD)

Table 9: Comparison between the four different pre-trained models depends on VGG (19)

VGG (19) (Training)										
Metrics	Recall	Specificity	Precision	NPV	FPR	FDR	FNR	Accuracy	F-score	MCC
Adam	0.9901	0.9877	0.9901	0.9877	0.0123	0.0099	0.0099	0.9890	0.9901	0.9778
SGD	1.0000	0.9877	0.9903	1.0000	0.0123	0.0097	0.0000	0.9945	0.9951	0.9890

Table 10: The overall standard metrics for the testing stage of VGG (19) pre-trained model

VGG (19) (Testing)										
Metrics	Recall	Specificity	Precision	NPV	FPR	FDR	FNR	Accuracy	F-score	MCC
Adam	0.9524	0.9091	0.9346	0.9333	0.9333	0.0654	0.0476	0.9341	0.9434	0.8647
SGD	0.9585	0.9554	0.9554	0.9585	0.0446	0.0446	0.0415	0.9569	0.9569	0.9139

From the previous tables, we can deduce the following: (1) All models produced comparable findings. (2) The lowest accuracy was obtained from the inception model in terms of accuracy of 0.9286 and an f-score of 0.9282. (3) With an accuracy of 0.9612 and an F-score of 0.9624, the resnet model outperformed all other pre-trained models. (3) Compared to the SGD optimizer, the Adam optimizer achieves more excellent performance with all models. As a result, the Resnet provides superior performance over the other models. This returns to the following: (1) Resnet utilizes identity mapping to overcome the problem of vanishing gradients. (2) empower the features so that they increase the achieved accuracy.

5.4 Training Using Proposed Ensemble Model

From [Tables 11](#) and [12](#), we make the following observation: (i) The proposed model achieves superior performance over the pre-trained model. The validity of the proposed model was measured by f score, precision, and recall of 0.9860 and 0.9860. (ii) High precision proves the ability of the model to classify other classes. (iii) High specificity indicates that our model has a low false positive rate, ensuring our model’s reliability with the proposed model. [Tables 11](#) and [12](#) show the performance of the proposed model during the training and testing processes. The confusion matrix in [Fig. 15](#) shows

the result of the classification of the proposed ensemble model. The four outcomes often plotted on a confusion matrix are TN: 1580, FP: 20, FN: 1600, and TP: 25 from testing data, achieving total accuracy without data. The proposed model proposes a rapid solution for black fungus detection with a low cost and provides an accurate solution for the problem of data unavailability, which affects the generalization ability. From the previous results, we could summarise the advantages of our model as follows: (1) The black fungus used in this study could be gathered from hospitals. (2) Our proposed system is entirely end-to-end without feature extraction techniques. (3) The proposed ensemble model is built using three pre-trained models, including (resnet, VGG, and Inception). (4) Although the problem is a new problem and the data is limited in size, the results are promising. The main problem in our study is the limited number of images. Therefore, we used transfer learning to overcome this issue. We plan to improve the models if we get more data in the coming days.

Table 11: The overall standard metrics for the Training stage of the Proposed Ensemble Model

Proposed ensemble model (Training)										
Metrics	Recall	Specificity	Precision	NPV	FPR	FDR	FNR	Accuracy	F-score	MCC
Adam	0.9938	0.9907	0.9907	0.9938	0.0093	0.0093	0.0063	0.9922	0.9922	0.9844

Table 12: The overall standard metrics for the Testing stage of the Proposed Ensemble Model

Proposed ensemble model (Testing)										
Metrics	Recall	Specificity	Precision	NPV	FPR	FDR	FNR	Accuracy	F-score	MCC
Adam	0.9907	0.9938	0.9938	0.9907	0.0062	0.0062	0.0093	0.9922	0.9922	0.9845

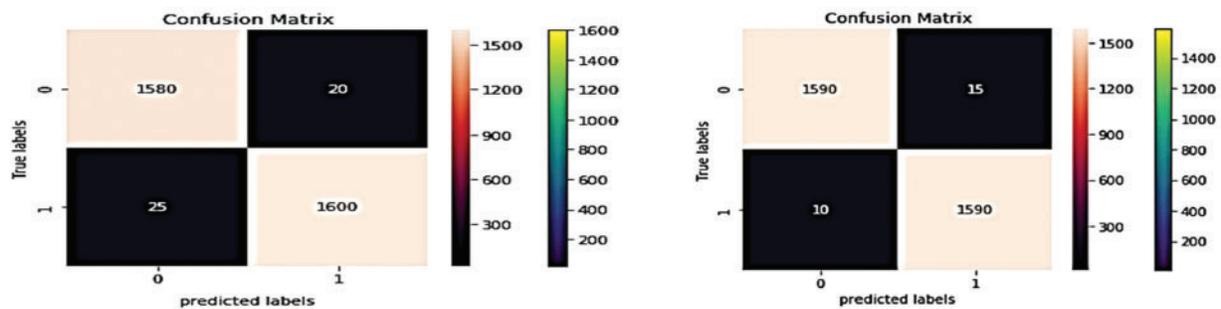


Figure 15: The confusion matrixes for Ensemble proposed model

The outcomes also demonstrate that the ensemble performance of the DL models was superior to that of the individual models. It might extract features performing well in the dataset, making it more flexible, rather than removing features that perform well generally, which results in overfitting rather than extracting features based on intriguing characteristics of the class in the images. Since the classifiers weren't tested on other datasets, it is impossible to confirm or refute if the deep learning models were overfitted on this one. The suggested paradigm should be extrapolated to comparable tasks and datasets with equivalent complexity. Fig. 16 shows the visualization testing for output classes.



Figure 16: The visualization testing results for the output classes

6 Comparison with Other Work

Skin black fungus is a common fungal infection that can be treated with antifungal medications or other laser or ultraviolet light therapy. However, there are limitations to these treatments, and recurrence is possible even after successful treatment. Recent research has focused on developing new treatments for black skin fungus that are more effective than traditional antifungal medications. In [23] focuses on transfer learning-based fault diagnosis under data deficiency. The main limitation of this paper is that it does not discuss how to ensure that the data used for training is accurate and reliable. [24] Urban focuses on environmental and social factors influencing the spread of SARS-CoV-2 in São Paulo, Brazil, during the COVID-19 pandemic. The main limitation of this paper is that it does not address the issue of access to healthcare services in rural areas or how to ensure that those living in rural areas have access to adequate healthcare services during a pandemic such as COVID-19. Additionally, the paper does not discuss the generalization and robustness of their proposed model. This limitation is handled in our manuscript by using optimization techniques and an ensemble DL model that generates robust and generalized models with promising results.

7 Conclusions

Pathologists' knowledge and experience are used to diagnose black fungus; reports from different laboratories doing physical inspections may differ. This paper presents the EBF model for classifying black fungus images, which has proven to be a promising method for detecting and categorizing information from cases. The results are compared to traditional machine learning and CNN-based approaches, proving that deep learning-based image categorization is feasible. To accomplish this, the dataset conversion rate must be increased. Without the requirement for manual feature extraction, it can carry out binary and multi-class classification. The outcomes combine three widely used pre-trained models: VGG (19), Resnet (50), and Inception (v3). The proposed approach can significantly improve the performance of the classification task and increase the generalization ability of such a binary classification task. They have achieved a sensitivity value of 0.9907, a specificity value of 0.9938, a precision value of 0.9938, and a negative predictive value of 0.9907.

For our future work, we intend to make the following enhancements (1) exploration of different structure CNN models, which may help increase learning capacity. (2) aggregating more training data, as training in more datasets will make the model more robust and increase model generalization ability. (3) model expandability will be considered to increase the model interpretation. (4) trying to explore the capabilities of IOT and Cloud in making the real-time diagnosis [23–25]. (5) trying other representative computational intelligence algorithms can be used to solve the problems, like

monarch butterfly optimization (MBO), earthworm optimization algorithm (EWA), elephant herding optimization (EHO), moth search (MS) algorithm, Slime mold algorithm (SMA), hunger games search (HGS), Runge Kutta optimizer (RUN), colony predation algorithm (CPA), and Harris hawk's optimization (HHO) and other fuzzy systems [26–28].

Funding Statement: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) Program (IITP-2023-2020-0-01832) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation) and the Soonchunhyang University Research Fund.

Availability of Data and Materials: Data availability is found at <https://data.mendeley.com/datasets/79hbdrfjg8>.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] Y. Liu, A. Jain, C. Eng, D. Way, P. Bui *et al.*, “A deep learning system for differential diagnosis of skin diseases,” *Nature Medicine*, vol. 26, no. 6, pp. 900–908, 2020. <https://doi.org/10.1038/s41591-020-0842-3>
- [2] S. Xia, Y. Xia, H. Yu, Q. Liu and Y. Xia, “Transferring ensemble representations using deep convolutional neural networks for small-scale image classification,” *IEEE Access*, vol. 7, pp. 168175–168186, 2019. <https://doi.org/10.1109/ACCESS.2019.2912908>
- [3] A. Eisa, N. EL-Rashidy, M. Alsheri, H. El-bakry and S. Abdelrazek, “Incremental learning framework for mining big data stream,” *Computers, Materials & Continua*, vol. 71, no. 2, pp. 2901–2921, 2022. <https://doi.org/10.32604/cmc.2022.021342>
- [4] V. Nirmala and B. Gomathy, “Hybrid deep learning method for diagnosis of cucurbita leaf diseases,” *Computer Systems Science and Engineering*, vol. 44, no. 3, pp. 2585–2601, 2023. <https://doi.org/10.32604/csse.2023.027512>
- [5] N. Mahmoud, S. El-sappagh, T. Abuhmed, S. M. Abdelrazek, S. Abderazik *et al.*, “Intensive care unit mortality prediction: An improved patient-specific stacking ensemble model,” *IEEE Access*, vol. 8, pp. 1, 2020. <https://doi.org/10.1109/ACCESS.2020.3010556>
- [6] V. Gautam, N. Trivedi, A. Anand, R. Anand, A. Zaguia *et al.*, “Early skin disease identification using deep neural network,” *Computer Systems Science and Engineering*, vol. 44, no. 3, pp. 2259–2275, 2023. <https://doi.org/10.32604/csse.2023.026358>
- [7] N. El Rashidy, T. Abyhmed, L. Alarabi, H. El Bakry, S. Abdelrazek *et al.*, “Sepsis prediction in intensive care unit based on genetic feature optimization and stacked deep ensemble learning,” in *Neural Computing and Applications*, vol. 1. London: Springer, 2021. <https://doi.org/10.1007/s00521-021-06631-1>
- [8] R. F. Mansour, “Blockchain assisted clustering with intrusion detection system for industrial internet of things environment,” *Expert Systems with Applications*, vol. 207, pp. 117995, 2022. <https://doi.org/10.1016/j.eswa.2022.117995>
- [9] M. M. Khayyat, M. M. Khayyat, S. Abdel-Khalek and R. F. Mansour, “Blockchain enabled optimal Hopfield Chaotic Neural network based secure encryption technique for industrial internet of things environment,” *Alexandria Engineering Journal*, vol. 61, no. 12, pp. 11377–11389, 2022. <https://doi.org/10.1016/j.aej.2022.05.002>
- [10] S. Armstrong, “COVID-19: Tests on students are highly inaccurate, early findings show,” *British Medical Journal*, vol. 371, pp. m4941, 2020. <https://doi.org/10.1136/bmj.m4941>

- [11] A. Bissoto, E. Valle and S. Avila, “Debiasing skin lesion datasets and models? Not so fast,” in *2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, WA, USA, pp. 3192–3201, 2020. <https://doi.org/10.1109/CVPRW50498.2020.00378>
- [12] S. Panda, A. S. Tiwari and M. R. Prusty, “Comparative study on different Deep Learning models for Skin Lesion Classification using transfer learning approach,” *International Journal of Scientific and Research Publications*, vol. 11, no. 1, pp. 219–232, 2021. <https://doi.org/10.29322/IJSRP.11.01.2021.p10923>
- [13] A. Adegun and S. Viriri, “Deep learning techniques for skin lesion analysis and melanoma cancer detection: A survey of state-of-the-art,” *Artificial Intelligence Review*, vol. 54, pp. 811–841, 2021. <https://doi.org/10.1007/s10462-020-09865-y>
- [14] S. D. Mahalakshmi, “An optimized transfer learning model based kidney stone classification,” *Computer Systems Science and Engineering*, vol. 44, no. 2, pp. 1387–1395, 2023. <https://doi.org/10.32604/csse.2023.027610>
- [15] H. Liu and M. Cocea, “Nature-inspired framework of ensemble learning for collaborative classification in granular computing context,” *Granular Computing*, vol. 4, no. 4, pp. 715–724, 2019. <https://doi.org/10.1007/s41066-018-0122-5>
- [16] H. Saoud, A. Ghadi and M. Ghailani, “Breast cancer diagnosis using machine learning and ensemble methods on large seer database,” *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 3, pp. 594–604, 2021.
- [17] B. Sluban and N. Lavra, “Relating ensemble diversity and performance: A study in class noise detection using neurocomputing,” *Neurocomputing*, vol. 160, pp. 120–131, 2015. <https://doi.org/10.1016/j.neucom.2014.10.086>
- [18] T. E. Mathew, K. S. A. Kumar and K. S. Kumar, “Breast cancer diagnosis using stacking and voting ensemble models with bayesian methods as base classifiers,” *International Journal of Computer Science and Mobile Computing*, vol. IX, no. II, pp. 108–121, 2020.
- [19] T. T. Akano and C. C. James, “An assessment of ensemble learning approaches and single-based machine learning algorithms for the characterization of undersaturated oil viscosity,” *Beni-Suef University Journal of Basic and Applied Sciences*, vol. 11, no. 4, pp. 149–155, 2022.
- [20] A. T. Lopes, E. de Aguiar, A. F. de Souza and T. Oliveira-Santos, “Facial expression recognition with convolutional neural networks: Coping with few data and the training sample order,” *Pattern Recognition*, vol. 61, pp. 610–628, 2017.
- [21] D. Al-Karawi, S. Al-Zaidi, N. Polus and S. Jassim, “Machine learning analysis of chest ct scan images as a complementary digital test of coronavirus (COVID-19) patients,” *medRxiv*, 2020. <https://doi.org/10.1101/2020.04.13.20063479>
- [22] Y. Chen, L. Zhu, P. Ghamisi, X. Jia, G. Li *et al.*, “Hyperspectral images classification with gabor filtering and convolutional neural network,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2355–2359, 2017. <https://doi.org/10.1109/LGRS.2017.2764915>
- [23] E. Hassan, “Black fungus dataset,” 2021. [Online]. Available: <https://data.mendeley.com/datasets/79hbdrfjg8>
- [24] S. H. Cho, S. Kim and J. H. Choi, “Transfer learning-based fault diagnosis under data deficiency,” *Applied Sciences*, vol. 10, no. 21, pp. 7768, 2020. <https://doi.org/10.3390/app10217768>
- [25] L. Y. K. Nakada and R. C. Urban, “COVID-19 pandemic: Environmental and social factors influencing the spread of SARS-CoV-2 in São Paulo Brazil,” *Environmental Science and Pollution Research*, vol. 28, no. 30, pp. 40322–40328, 2021. <https://doi.org/10.1007/s11356-020-10930-w>
- [26] N. Ullah, J. A. Khan, S. El-Sappagh, N. El-Rashidy, M. A. Sohail Khan *et al.*, “A holistic approach to identify and classify COVID-19 from chest radiographs, ecg, and ct-scan images using shufflenet convolutional neural network,” *Diagnostics*, vol. 13, no. 2, pp. 162, 2023. <https://doi.org/10.3390/diagnostics13010162>

- [27] G. G. Wang, D. Gao and W. Pedrycz, “Solving multiobjective fuzzy job-shop scheduling problem by a hybrid adaptive differential evolution algorithm,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 12, pp. 8519–8528, 2022. <https://doi.org/10.1109/TII.2022.3165636>
- [28] D. Gao, G. G. Wang and W. Pedrycz, “Solving fuzzy job-shop scheduling problem using de algorithm improved by a selection mechanism,” *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 12, pp. 3265–3275, 2020. <https://doi.org/10.1109/TFUZZ.2020.3003506>