

Enhancing Monkeypox Diagnostics: Exploring the Potential of EfficientNet and Big Transfer

Sharia Arfin Tanim¹, Kazi Tanvir², Al Rafi Arnob¹, Md. Hasibur Rahman¹, Tasmia Binte Munir Maisha³, and Kamruddin Nur^{1,*}

¹Department of Computer Science, American International University-Bangladesh (AIUB), Dhaka, Bangladesh

²Department of Mathematics, Vellore Institute of Technology, Tamil Nadu, India

³Department of Computer Science and Engineering, Bangladesh University of Business and Technology (BUBT), Dhaka, Bangladesh

Email: shariaarfin096@gmail.com (S.A.T.); kazitanvir.ai@gmail.com (K.T.); inc.aurnob@gmail.com (A.R.A.)
mhasiburrahman01@gmail.com (M.H.R.); maishatasmia451@gmail.com (T.B.M.M.);

kamruddin@aiub.edu (K.N.)

*Corresponding author

Abstract—The purpose of this research is to investigate the utilization of hybrid models in dermatological diagnostics and to demonstrate the potential of these models to advance medical picture classification capabilities. The study presents BiT-EfficientNet, a novel hybrid model developed specifically for the precise classification of monkeypox lesions in skin images. By combining EfficientNet B6 and Big Transfer (BiT-M-R50x1), the model demonstrates exceptional performance in recognizing patterns and managing visual features. BiT-EfficientNet demonstrates superior performance compared to existing models, achieving a precision of 98.25%, recall of 95.48%, F1-Score of 96.84%, and accuracy of 96.86%. It is positioned as a strong contender through comparative analysis. A highly accurate model is achieved through careful parameter optimization, resulting in a training accuracy of 99.14%. Assessing resilience through empirical means validates it. The findings have a significant impact on increasing diagnostic accuracy for illnesses like monkeypox, which can result in prompt interventions in professional medical professionals' healthcare.

Keywords—monkeypox disease, deep learning, ensemble learning, image processing, skin lesion detection

I. INTRODUCTION

The Monkeypox Virus (MPXV), which is an enveloped DNA virus with double strands and a member of the Orthopoxvirus Poxviridae family, is the infectious disease that causes monkeypox. There are two genetic clades that make up the family of viruses that cause monkeypox. The World Health Organization has concluded that the virus can occasionally be transmitted from animals to humans as well as from person to person. This is despite the fact that its natural reservoir is still unknown. An outbreak of monkeypox happened all across the world in the years 2022–2023, following the elimination of smallpox in 1980 and the invention of the disco [1]. The development of

cutaneous lesions, which typically appear within one to five days of infection with monkeypox, is common. The aforementioned rashes manifest initially on the face before spreading to multiple body parts, including the genitalia, eyes and oral mucosa. The resemblance to rashes observed in varicella cases frequently complicates the diagnostic process. As the rash progresses from fluid-filled lesions to crusty, scaly areas, it transforms into this appearance. The number and severity of these lesions may vary from person to person, as some may show a lot of papules that later join together to form large rashes [2, 3].

There are two main types of learning algorithms, based on a lot of study in machine learning: deep learning and ensemble learning. Deep learning methods can work on large scales, solve difficult problems, and automatically pull out features from unstructured data [4]. But building deep learning models takes a lot of work, and finding the best hyper-parameters takes a lot of trial and error, which is a boring and time-consuming process. Also, the chance of overfitting goes up as the deep neural network is trained to be more complicated [5]. The prevalence of Convolutional Neural Networks (CNNs) in academic research pertaining to image processing and classification is indicative of their extensive application. EfficientNet B6 is a member of the EfficientNet family and embodies a convolutional neural network architecture that has been specifically designed to handle image classification tasks, and its architecture is robust. The model showcases remarkable efficiency and effectiveness by means of precise calibration of its depth, width, and resolution, thereby optimizing computational resources [6–8]. BiTM-R50x1 is a highly advanced image classification model that has achieved the best possible results on various benchmarks, which makes it suitable for large-scale applications. It is very compatible with transfer learning and may be adjusted for specific applications such as

object detection, segmentation, and picture captioning [9, 10]. The main contributions of our research are the following:

- Introduction of BiT-EfficientNet, a customized deep-learning-based ensemble model for early detection and accurate categorization of monkeypox.
- Comprehensive comparison with pre-trained deep-learning models, validating the efficacy of our proposed architecture, which outperformed each of them.
- Demonstration of the applicability of our approach beyond monkeypox classification, showing potential in other medical sectors.

In contrast, ensemble learning is a learning approach that merges multiple baseline models into a larger, more powerful model than the sum of its parts [11]. Because there are different starting models in ensemble learning, the risk of overfitting is also lower. Ensemble learning works better than single models and has been used successfully in many fields [12]. The EfficientNet architecture's EfficientNetB6 variation is known for its model size and computing resource efficiency. Big transfer models like the BiT architecture are for transfer learning. Pre-trained models on several datasets and tasks are very adaptable and resilient. Ensemble learning reduces overfitting and maximizes model strengths by combining predictions. EfficientNetB6 and Big Transfer models in an ensemble learning framework are powerful image classification models. This ensemble technique tries to outperform individual models by using their efficiency, resilience, and diversity. Ensemble learning with EfficientNetB6 and BiT could improve image categorization by letting practitioners combine model strengths.

In the midst of numerous skin diseases, this study presents an innovative hybrid model that has been specifically engineered to classify monkeypox lesions with enhanced accuracy. By seamlessly integrating the capabilities of EfficientNet B6 and Big Transfer (BiT-M-R50x1). Domain generalization and intricate feature extraction are areas in which the model excels. Recognized for its effectiveness in scaling models, EfficientNet B6 detects intricate patterns within skin lesions. Simultaneously, BiT-M-R50x1, renowned for its remarkable domain generalization, effectively manages discrepancies pertaining to signature dimensions, pen pressures, and writing styles. By means of rigorous refinement, the hybrid model attains enhanced levels of accuracy, recall, F1-Score, and precision. This study provides a cutting-edge model for the early detection of monkeypox lesions, which advances dermatological diagnostics. The findings may have significant ramifications, including the ability to implement timely healthcare interventions and enhance patient outcomes.

This paper is organized in the following manner: In Section II, previous research in this area is discussed. The proposed methodology is presented in Section III. Section IV is dedicated to discussing the experimental

results. The conclusion and possible directions for future research are covered in Section V.

II. RELATED WORK

There have been a number of studies that have been carried out to investigate the efficacy of various deep-learning models for the detection of monkeypox. This is done with the intention of classifying monkeypoxes through the use of deep learning. Monkeypox can now be classified alongside other diseases as a result of this. An analysis of comparison and contrast was carried out on a total of thirteen distinct deep-learning models that had been pre-trained to identify the monkeypox virus [13]. A way to identify monkeypox was also suggested: it would be like a binary classification task, and a deep neural network that had already been trained would be used [14]. This method was applied to the identification of monkeypox. The identification of monkeypox was accomplished with the help of this method. Both of them looked into deep transfer learning strategies, with the latter team using a Convolutional Block Attention Module (CBAM) for image-based classification. Deep transfer learning was something that both of them investigated. The examination of skin lesions was supposed to be used in order to identify monkeypox illness [15]. Deep learning strategies were supposed to be utilized in order to accomplish this. Furthermore, we were able to show how important this factor is by looking at bad monkeypox skin imaging datasets [16]. This showed us how important dataset quality is when building models. This allowed us to better understand the significance of this factor.

The identification of monkeypox through the utilization of CNN in conjunction with transfer learning demonstrates the effectiveness of utilizing pre-trained models to achieve better diagnostic results [17]. The range of models offered by EfficientNet extends from B0 to B7. Furthermore, the EfficientNetB0 model serves as the foundational model for the EfficientNet architecture [18]. EfficientNet-B0 will be compared against CNNs. EfficientNet-B0 outperforms MobileNet and InceptionV3 with 85.12% accuracy. With fewer parameters than ResNet-50, it offers good sensitivity (78.46%) and specificity (91.78%) [19].

The COVID-19 epidemic makes monkeypox diagnosis urgent and suggests using machine learning to diagnose it early and accurately. The study uses data augmentation and transfer learning to create a monkeypox diagnosis machine learning model utilizing image processing. The monkeypox classification model PoxNet22 has 100% precision, recall, and accuracy. These findings should help clinicians diagnose and classify monkeypox, according to the study [20]. Monkeypox's milder symptoms and strong similarities to smallpox make diagnosis difficult. The EfficientNet Transfer Learning Method is applied. Photos, training, and testing are ready for analysis after normalization and feature extraction. The dataset is analyzed using Xception, ResNet152, EfficientNetV2L, InceptionV3, MobileNetV2, NASNet-Large, and DenseNet201. Evaluation metrics include accuracy, precision, recall, F1-Score, and training and validation loss [21]. Diagnoses utilize trained standalone deep

learning models like InceptionV3, EfficientNet, VGG16, and SENet Attention. Classifying monkeypox using SENet-based attention models and trunk branches of DL architectures. Experimental results show improved monkeypox diagnosis and classification in accuracy, precision, recall, and F1-Score [22].

A new framework was also proposed that uses Al-Biruni Earth Radius optimization-based Stochastic Fractal Search (BERSFS) to improve the deep Convolutional Neural Network (CNN) layers in order to find monkeypox disease in photos [23]. Utilized the Al-Biruni Earth Radius Optimization Algorithm and transfer learning to enhance deep CNN models for the purpose of Monkeypox image classification [24]. From Table I, we can see that it

presents a novel few-shot learning approach utilizing the deep-learning architecture to accurately categorize photos of monkeypox and other comparable skin diseases, even when only a limited number of samples are available [25].

These studies collectively showcase the capability of deep learning models, transfer learning, and attention mechanisms for precisely identifying and categorizing monkeypox from photos of skin lesions. Additionally, they emphasize the need for high-quality datasets and the creation of innovative frameworks to optimize the fine-tuning of deep Convolutional Neural Network (CNN) layers, hence improving the categorization of monkeypox sickness.

TABLE I. SUMMARY OF LITERATURE REVIEW

Reference	Year	Image	Model	Accuracy	Limitation
[26]	2022	monkeypox (102), others (126)	VGG-16, ResNet50, Inception-V3, ensemble	82.96%	The dataset is limited and demographically homogeneous, which may limit the use of deep learning models for monkeypox lesion identification.
[27]	2022		AlexNet, GoogLeNet, VGG-16, SVM, k-NN, Naive Bayes, DT, RF	91.11%	The Monkeypox-Skin-Lesion-Dataset has few original and augmented pictures. So, making deep CNN model and machine learning classifier comparisons challenging.
[14]	2022	monkeypox (102), others (126)	Modified MobilNetV2	91.1%	The research does not explore pre-trained network biases or restrictions for monkeypox classification.
[28]	2022	monkeypox, chickenpox, smallpox, cowpox, measles, normal	ResNet50, DenseNet121, Inception-V3, SqueezeNet, MobileNet-V2, ShuffleNet-V2, ensemble	83.00%	The study utilized a dataset obtained through web scraping, which may not comprehensively capture the variety of Monkeypox skin images.
[29]	2022	monkeypox, tuberculosis	VGG-16, ResNet50, Inception-V3, ensemble	88.64%	The research does not examine the drawbacks of idea finding approaches for infectious disease classification.
[30]	2022	monkeypox, others (total 2142)	5 deep learning algorithms	88.64%	Replicating the trials is difficult because the paper does not specify the hyperparameters for each deep learning model.
[31]	2022	monkeypox, chickenpox, measles, normal	13 deep learning models.	88.64%	The paper does not discuss the potential challenges or disadvantages of using pre-trained deep learning models.

III. MATERIALS AND METHODS

This section will include a concise summary of the materials and procedures employed in this study (Fig. 1), represents a graphical illustration of the complete overview, showing the sequential phases in data collection, preprocessing, model training, and the classification of monkeypox.

A. Data Gathering and Preprocessing

This study dataset, comprising 228 high-quality photos of monkeypox and others, was obtained from Kaggle [32]. The dataset is available online (<https://www.kaggle.com/datasets/nafin59/monkeypox-skin-lesion-dataset>). There are 3 folders in the dataset. The dataset is divided into two classes, each representing a different type of data at 224×224 pixels. The dataset contains a small number of original images. The total number of original images is 228, with monkeypox accounting for 102 and others accounting for 126. As a

result, using augmented images is extremely beneficial. The use of augmented images can be extremely beneficial. An overview of the dataset is presented in Table II. There are 3,192 augmented images total in the dataset; 1,428 of those images are of monkeypox, and the remaining 1,764 are of other types.

We split the data into training, test and validation sets using an 80:10:10 ratio. The training set is augmented with data augmentation techniques such as rescale, rotation, shift, shear, zoom, and flipping using the Image-DataGenerator class. The `flow_from_dataframe` function creates data batches. The function scales the input pictures to 224×224 pixels and sets the batch size to 32. The class mode is categorical, and we shuffle the training set. The train data and test data are then passed to the train generator and test generator, respectively, as shown in Fig. 2.

TABLE II. DATASET OVERVIEW

Class	Original Images	Augmented Images
Monkeypox	102	1428
Others	126	1764
Total	228	3192

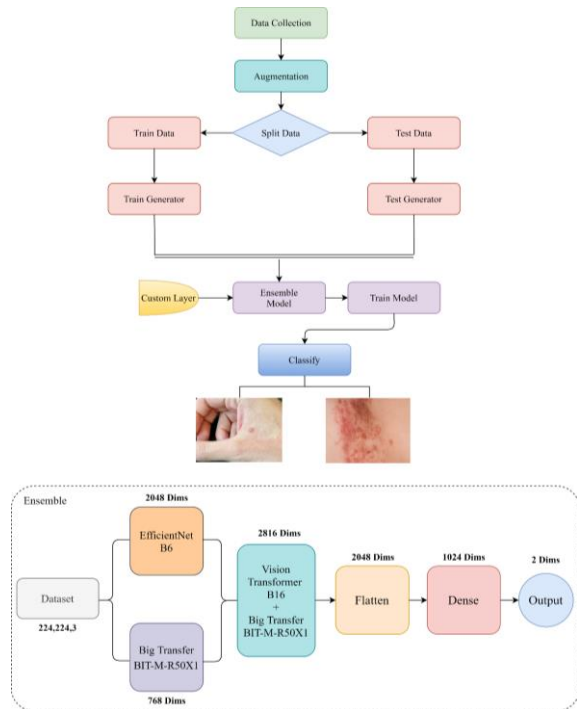


Fig. 1. Comprehensive diagram illustrating the entire sequence of our proposed study.



Fig. 2. Exemplary image used for training purposes.

B. Model Development

1) EfficientNet B6

A convolutional neural network design is featured within the broader EfficientNet model series [18, 33]. This architectural framework is a member of the EfficientNet family of models and is also renowned for its efficacy in deep learning applications. The neural network employs a compound scaling approach to systematically adjust the network’s width, depth, and resolution in a methodical manner. EfficientNet B6, characterized by 43.3 million parameters and 360 layers, exemplifies the application of this scaling method [34]. EfficientNet B6 attains cutting-edge precision across diverse image classification tasks, encompassing benchmarks like ImageNet, CIFAR-100,

and Flowers. Notably, it accomplishes this feat with significantly fewer parameters compared to alternative models of similar complexity [35]. Furthermore, EfficientNet B6 is compatible with popular deep learning frameworks such as TensorFlow and Keras. EfficientNet B6 processes an input image and categorizes it into one of multiple predefined classes. Notably, the model demonstrates a rapid training pace, facilitating efficient learning, and can be easily fine-tuned to accommodate custom datasets [36]. EfficientNet B6 presents a commendable blend of efficiency and performance, rendering it well-suited for a diverse array of image classification tasks. The model’s architecture and scaling strategy have been meticulously crafted to enhance performance by maintaining a harmonious equilibrium among network depth, width, and resolution. Employing the compound scaling method, which uniformly adjusts all dimensions of depth, width, and resolution, contributes to heightened model accuracy and efficiency. Extensive evaluations of the model’s performance across diverse datasets affirm its robust transferability and its ability to achieve state-of-the-art accuracy in various image classification endeavors. It’s worth mentioning that EfficientNet B6 supports transfer learning, which lets you fine-tune it on custom datasets. This makes it useful for a wide range of computer vision applications.

2) BigTransfer

BigTransfer (BiT) stands out as an advanced transfer learning technique in the realm of computer vision [9]. A notable member of the BiT model family, the BiTM-R50x1 model, is seamlessly integrable with popular deep learning frameworks like TensorFlow and Keras [37]. Serving as a robust pre-trained model, it accepts an image as input and proficiently categorizes it among multiple predefined classes. The BiT-M-R50x1 model gets better accuracy and efficiency by using a compound scaling method to make the network width, depth, and resolution all the same. Boasting 23.5 million parameters, this model demonstrates exceptional performance across diverse image classification tasks, rendering it suitable for a broad spectrum of computer vision applications. Its ability to perform fine-tuning on unique datasets, striking a careful balance between efficiency and performance, further emphasizes its adaptability. The intricately designed architecture and scaling method of the model contribute to its superior performance by harmonizing network depth, width, and resolution. Particularly advantageous for applications like image classification, object detection, and image recognition, the BiT-M-R50x1 model emerges as a valuable asset in the landscape of computer vision [38].

3) Supervised ensemble model

Big Transfer is a creation of Google AI, a large-scale vision model trained on extensive datasets of images and text. This broad training equips it with the capability to discern and learn new features, whereas a seasoned AI expert well-versed in the intricacies of the visual domain. In contrast, EfficientNet B6 is renowned for its convolutional neural network design, adept at achieving a delicate balance between accuracy and efficiency. Notably, it delivers state-of-the-art performance while demanding

lower computational resources compared to its counterparts, thereby embodying a streamlined and highly efficient machine. The ensemble model's effectiveness stems from leveraging the strengths of both Big Transfer and EfficientNet B6, combining their complementary features to enhance performance across various tasks. This synergy allows the ensemble model to capitalize on the robustness and versatility of Big Transfer alongside the efficiency and precision of EfficientNet B6, resulting in improved overall performance and effectiveness.

Ensemble learning is a technique that combines multiple distinct models in order to improve generalization performance. At present, deep learning architectures are demonstrating greater efficacy in comparison to pre-trained models. Fig. 3 illustrates an ensemble model architecture designed to process input images with dimensions of $224 \times 224 \times 3$ (224 pixels in width, 224 pixels in height, and 3 color channels). The image undergoes initial processing using a Bit+EfficientNetB6 block as illustrated in (Fig. 4), which consists of a convolutional

neural network pre-trained on a vast collection of photos. This module extracts several characteristics from the image. Subsequently, the characteristics are transmitted through a functional layer, which is a custom layer that executes a particular action on the data. The Keras deep learning packaging provides a standard sort of layer known as a Keras layer, which is used to process the output of the functional layer. In this instance, the Keras layer employed is a concatenate layer, which merges the characteristics extracted from the two preceding layers. Subsequently, the collective attributes are transformed into a unidimensional array through the process of flattening. Then, the compressed characteristics are transmitted through a compact layer, which is a fully connected layer that generates a prediction regarding the class of the image.

Let $f_{\text{EffNet}}(X_{\text{tr}})$ represent the predictions made by the EfficientNet B6 model on the training set X_{tr} , and $f_{\text{BiT}}(X_{\text{tr}})$ represent the predictions made by the BigTransfer model on the same training set.

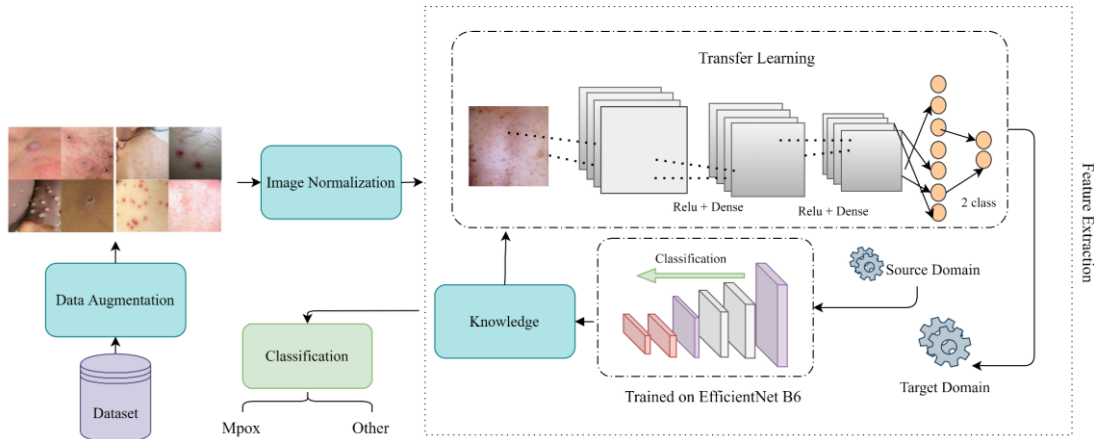


Fig. 3. Architectural diagram of proposed method.

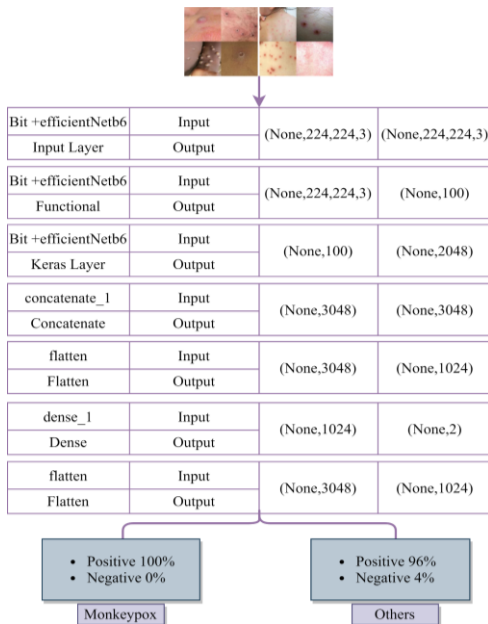


Fig. 4. BiT-EfficientNet model architecture.

Similarly, let $f_{\text{EffNet}}(X_{\text{val}})$ and $f_{\text{BiT}}(X_{\text{val}})$ represent the predictions made by the EfficientNet B6 and BigTransfer models on the validation set X_{val} respectively.

Then, the combined features can be represented as:

$$\text{Combined}_{tr} = \{f_{\text{EffNet}}(X_{tr}) f_{\text{BiT}}(X_{tr})\} \quad (1)$$

$$\text{Combined}_{val} = \{f_{\text{EffNet}}(X_{val}) f_{\text{BiT}}(X_{val})\} \quad (2)$$

$$\text{Model} = \text{Train}(\text{Combined}_{tr}, y_{tr}, \text{Combined}_{val}, y_{val})$$

where y_{tr} and y_{val} are the true labels for the training and validation sets respectively.

Algorithm 1 describes the process of training an ensemble model using features extracted from EfficientNet and BiT models. Initially, the features are extracted from the training and validation datasets using the EXTRACT procedure, followed by concatenation to create combined features. These combined features are then classified using the CLASSIFY procedure, and the resulting outputs are utilized for model training through the TRAIN procedure. The model is initialized with specified loss function and optimizer, compiled, and trained on the training data with validation data for a set number of

epochs. Finally, the trained model is returned for evaluation. Overall, Algorithm 1 presents a succinct framework for ensemble model training leveraging pre-trained EfficientNet and BiT features.

Algorithm 1: Ensemble Model Training

```

1: procedure ENSEMBLE_TRAIN( $X_{tr}, X_{val}, Y_{tr}, Y_{val}$ )
2: FeffNet_tr, FbiT_tr  $\leftarrow$  Extract( $X_{tr}$ )
3: FeffNet_val, FbiT_val  $\leftarrow$  Extract( $X_{val}$ )
4: Combined_tr  $\leftarrow$  Concatenate(FeffNet_tr, FbiT_tr)
5: Combined_val  $\leftarrow$ 
   Concatenate(FeffNet_val, FbiT_val)
6: Output_tr  $\leftarrow$  Classify(Combined_tr)
7: Output_val  $\leftarrow$  Classify(Combined_val)
8: Model  $\leftarrow$  Train(Combined_tr, y_tr, Combined_val, y_val)
9: return Model
10: end procedure
11: procedure EXTRACT(X)
12: FeffNet  $\leftarrow$  EfficientNetFeature(X)
13: FbiT  $\leftarrow$  EbiTFeature(X)
14: return FeffNet, FbiT
15: end procedure
16: procedure CLASSIFY(Features)
17: Flattened_Features  $\leftarrow$  FlattenedFeatures
18: Dense_Output  $\leftarrow$  Dense(FlattenedFeatures)
19: Output  $\leftarrow$  Softmax(DenseOutput)
20: return Output
21: end procedure
22: procedure TRAIN( $X_{tr}, X_{val}, Y_{tr}, Y_{val}$ )
23: Model  $\leftarrow$  InitializeModel()
24: Loss  $\leftarrow$  CategoricalcrossEntropy()
25: Optimizer  $\leftarrow$  AdamOptimizer()
26: Model Compile(loss = Loss, optimizer = Optimizer,
   metrics = ['accuracy'])
27: Model Fit( $X_{tr}, Y_{val},$  validation_data ( $X_{val}, Y_{val}$ ),
   epochs = N)
28: return Model
29: end procedure

```

C. Optimizing Parameter

The selected parameters for the ensemble model's compilation are designed to efficiently optimize the model's weights using the Adam optimizer, reduce the difference between predicted and actual class distributions with categorical cross-entropy loss, and analyze the model's performance mainly through classification accuracy. The parameters are ideal for training and assessing classification models, particularly in situations where accuracy is the main performance measure. We optimized our model by fine-tuning parameters such as input shape, batch size, number of epochs, validation split, optimizer, and activation function, resulting in enhanced model performance and improved overall accuracy Table III.

D. Evaluation Metrics

In the evaluation of the model on the test set, notable performance metrics were observed. The precision (1), indicating the model's ability to correctly identify positive instances, reached 98.25%, calculated as the ratio of true positives to the sum of false positives and true positives.

Similarly, the Recall Eq. (2), representing the model's capability to capture all actual positive instances, achieved a rate of 95.48%. The F1-Score Eq. (3), a balanced measure considering both precision and recall, also marked an impressive 96.84%. Finally, the train accuracy Eq. (4), reflecting the overall correctness of the model, reached 99.14%. These metrics collectively demonstrate the robustness and effectiveness of the model in making accurate predictions on the given test set.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

TABLE III. HYPER-PARAMETERS OF ENSEMBLE MODEL

Parameter	Value
Image Size	224×224
Batch Size	32
Epochs	100
Validation Split	0.1
Class Number	2
Rescaling	1/255
Pooling	Avg
Activation Function	ReLU and Softmax
Optimizer	Adam
Loss Function	Categorical Crossentropy
Dataset Size	228
Total Parameters	69,889,721

IV. RESULT AND DISCUSSION

A. Monkeypox Categorization Result

By completing 100 epochs of training, our model achieved a training accuracy of 99.14%. The model exhibits strong performance on the validation set, with a validation accuracy of 96.86% (Fig. 5). The loss gradually decreased over the epochs, suggesting that the model was learning from the data. The training loss after 100 epochs is 0.0232, and the validation loss is 0.1466 (Fig. 6). The curves demonstrate a consistent improvement in accuracy throughout the training process for both sets of data. However, towards the end of the epoch, the validation accuracy is gradually decreasing. The loss curves for training and validation data illustrate a decrease in loss throughout the training process. The model accurately classifies 141 cases as monkeypox, which represents true positive results, and correctly identifies 169 cases as not being monkeypox, which represents true negative results. It indicates that the false negative number is 2, indicating the cases where the model inaccurately classified a non-monkeypox case as monkeypox. The false positive rate is 8, representing the instances where the model incorrectly classified cases as monkeypox when they were actually attributed to a different condition (Fig. 7).

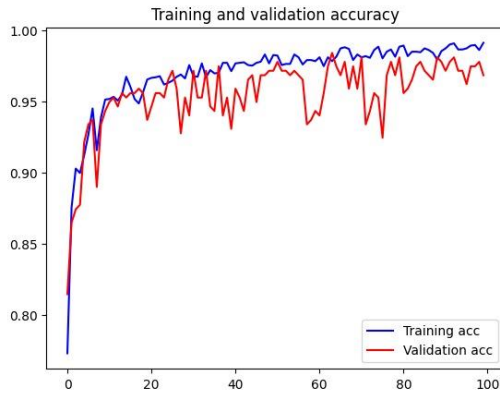


Fig. 5. Training and validation accuracy curve.

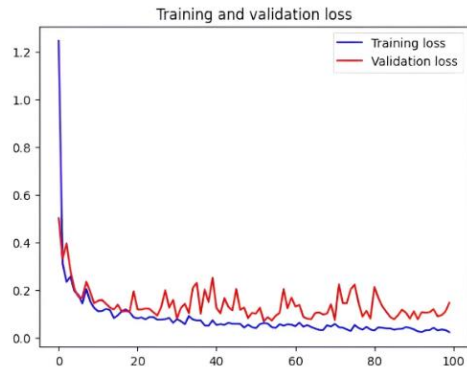


Fig. 6. Training and validation loss curve.

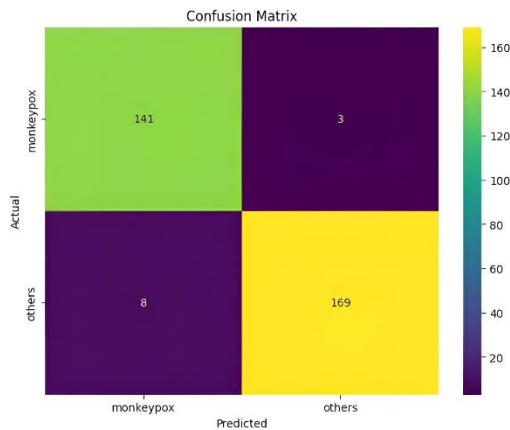


Fig. 7. Confusion matrix of Bit-EfficientNet architecture.

B. Result Model Performance Comparison

In this subsection, we assess and contrast the performance metrics of various classification models. A comprehensive summary of precision, recall, F1-Score, and accuracy for widely used models, including VGG16, VGG19, Inception v3, Ensemble, AlexNet, GoogleLeNet, MobileNet-V2, ShuffleNet-V2, and RESNET50. The proposed Bit-EfficientNet model is provided in Table IV. Significantly superior to its competitors, this model achieves a precision, recall, and F1-Score of 98.25%, 95.48%, 96.84% respectively demonstrating its robustness in accurate classification. Fig. 7, visually depicts the classification outcomes in the confusion matrix. This further underscores the efficacy of the model in increasing the occurrence of both true positives and true negatives. With a remarkable accuracy rate of 100%, Bit-EfficientNet establishes itself as a highly prospective contender for tasks involving the classification of monkeypox. The present analysis illustrates the importance of model selection in the context of image classification, specifically in attaining high-performance results. The model’s exceptional precision, recall, and F1-Score demonstrate its proficiency in accurately detecting instances of monkeypox and differentiating them from other ailments. The precision score of 98.25% signifies that the model accurately predicts cases of monkeypox 98.25% of the time. This indicates that the model has acquired highly efficient capabilities for recognizing the distinct attributes of monkeypox cases. The recall score of 95.48% indicates that the model can correctly detect 95.48% of all real monkeypox cases. This suggests that the model has successfully acquired knowledge from a large volume of data and is capable of applying that knowledge to new, unfamiliar data. The F1-Score, calculated as the harmonic mean of precision and recall, is 96.84%. It quantifies the model’s accuracy by considering both precision and recall in a balanced manner. A high F1-Score indicates that the model effectively balances recall and precision, demonstrating its optimal tuning for both aspects without sacrificing one for the other. After analyzing the performance metrics, it appears that combining the generalization capabilities of BiT with the efficiency and scalability of EfficientNetB6 gives a model that is both accurate and efficient. Findings like these suggest that the model is good at both correctly identifying cases of monkeypox and lowering the number of wrong classifications for those conditions and those in Table IV.

TABLE IV. EVALUATION OF OTHER MODELS’ PERFORMANCE

Study	Classes	Best Perform Model	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
[39]	Monkeypox Chickenpox Measles	Xception CBAM Dense	90.70	89.10	90.11	83.89
[40]	Chickenpox Measles Monkeypox Normal	MobileNetV2	90.50	90.50	88.25	91.37
[41]	Monkeypox Others	EfficientNet- B4	89.60	88.90	89.20	88.89
[42]	Monkeypox Normal	GoogLeNet	86.34	86.48	86.87	86.27
[43]	Chickenpox Measles Monkeypox Normal	EfficientNet-B0	96.57	96.53	95.52	96.53
[14]	Monkeypox	Custom CNN	90.00	90.00	90.00	91.11
This work	Monkeypox Others	BiT-EfficientNet	98.25%	95.48%	96.84%	96.86%

V. CONCLUSION

The utilization of the BiT-EfficientNet model in medical image classification exemplifies the transformative capacity of hybrid deep learning approaches and signifies a substantial progression in dermatological diagnostics. The diagnostic tool's efficacy in identifying monkeypox lesions underscores its significance as a means to accelerate treatment and improve patient outcomes. Moving forward, it is crucial to contemplate the more comprehensive ramifications of incorporating this model into clinical practice, enhancing its functionalities, and cultivating cooperation to facilitate the exchange of data. By effectively tackling these obstacles, we can not only substantiate the effectiveness of the model but also establish a foundation for its wider implementation in the field of dermatology and other disciplines, thereby capitalizing on the capabilities of artificial intelligence to transform the provision of healthcare. Subsequent areas of investigation ought to center on practical implementation and verification, longitudinal assessment of performance progression, investigation into the integration of electronic health records, refinement of models, expansion of applications to encompass additional dermatological conditions, improvement of interpretability, and facilitation of worldwide data exchange. Through the implementation of these strategies, it is possible to validate the efficacy of the model, enhance patient outcomes, and advance the domain of dermatological diagnostics.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

The experiments of this research were devised by Sharia Arfin Tanim. He was involved in the conceptualization, design, and execution of the initial draft of the manuscript. Kazi Tanvir, AL Rafi Arnob, Md. Hasibur Rahman and Tasmia Binte Munir Maisha conducted the experiments and contributed to the initial draft of the manuscript. Kamruddin Nur supervised, edited, and reviewed the manuscript. The final draft of the manuscript has been reviewed and approved by all authors.

REFERENCES

- [1] World Health Organization. (2023). Monkeypox fact sheet. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/monkeypox>
- [2] M. Hraib, S. Jouni, M. M. Albitar, S. Alaidi, and Z. Alshehabi, "The outbreak of monkeypox 2022: An overview," *Annals of Medicine & Surgery*, vol. 79, July 2022.
- [3] D. Kmiec and F. Kirchhoff, "Monkeypox: A new threat?" *International Journal of Molecular Sciences*, vol. 23, no. 14, 2022.
- [4] A. Kamilaris and F. X. Prenafeta-Boldu, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
- [5] A. Mohammed and R. Kora, "A comprehensive review on ensemble deep learning: Opportunities and challenges," *Journal of King Saud University-Computer and Information Sciences*, vol. 35, no. 2, pp. 757–774, 2023.
- [6] G. Marques, D. Agarwal, and I. de la T. Diez, "Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network," *Applied Soft Computing*, vol. 96, 106691, 2020.
- [7] F. Hoorali, H. Khosravi, and B. Moradi, "Irunet for medical image segmentation," *Expert Systems with Applications*, vol. 191, 116399, 2022.
- [8] A. Zhaputri, M. Hayaty, and A. D. Laksito, "Classification of brain tumour MRI images using efficient network," in *Proc. 2021 4th International Conference on Information and Communications Technology (ICOIACT)*, 2021, pp. 108–113.
- [9] A. Kolesnikov, L. Beyer, X. Zhai, J. Puigcerver, J. Yung, S. Gelly, and N. Houlsby, "Big Transfer (BiT): General visual representation learning," arXiv preprint, arXiv: 1912.11370, 2019. doi: <https://doi.org/10.48550/arXiv.1912.11370>
- [10] K. Tanvir, "Signature verification system: Using big transfer (BITM-r50x1) for accurate authentication," *Journal of Image Processing and Intelligent Remote Sensing*, vol. 3, no. 5, 10, 2023.
- [11] V. Kumar, P. S. S. Ayday, and S. Minz, "Multi-view ensemble learning using multi-objective particle swarm optimization for high dimensional data classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 10, pp. 8523–8537, 2022.
- [12] G.-W. Weber, P. Vasant, N. Barsoum, V. N. Dieu, H. Anwar, U. Qamar, and A. W. Muzaffar Qureshi, "Global optimization ensemble model for classification methods," *The Scientific World Journal*, vol. 2014, 313164, 2014.
- [13] C. Sitaula and T. B. Shahi, "Monkeypox virus detection using pretrained deep learning-based approaches," *Journal of Medical Systems*, vol. 46, no. 11, 78, 2022.
- [14] V. H. Sahin, I. Oztel, and G. Y. Oztel, "Human monkeypox classification from skin lesion images with deep pre-trained network using mobile application," *Journal of Medical Systems*, vol. 46, no. 11, 79, 2022.
- [15] M. Altun, H. Guruler, O. Ozkaraca, F. Khan, J. Khan, and Y. Lee, "Monkeypox detection using CNN with transfer learning," *Sensors*, vol. 23, no. 4, 1783, 2023.
- [16] C. Vega, R. Schneider, and V. Satagopam, "Analysis: Flawed datasets of monkeypox skin images," *Journal of Medical Systems*, vol. 47, no. 1, 37, 2023.
- [17] M. Altun, H. Guruler, O. Ozkaraca, F. Khan, J. Khan, and Y. Lee, "Monkeypox detection using CNN with transfer learning," *Sensors*, vol. 23, no. 4, 2023.
- [18] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," arXiv preprint, arXiv: 1905.11946, 2019.
- [19] P. E. N. Taruno, G. S. Nugraha, R. Dwiyanaputra, and F. Bimantoro, "Monkeypox classification based on skin Images using CNN: EfficientNet-B0," in *Proc. 8th International Conference on Industrial, Mechanical, Electrical and Chemical Engineering (ICIMECE 2023), Symposium on Electrical, Information Technology, and Industrial Engineering*, 2023, vol. 465, no. 02031, pp. 1–6. doi: [10.1051/e3sconf/202346502031](https://doi.org/10.1051/e3sconf/202346502031)
- [20] F. Yasmin, M. M. Hassan, M. Hasan, S. Zaman, C. Kaushal, W. ElShafai, and N. F. Soliman, "Poxnet22: A fine-tuned model for the classification of monkeypox disease using transfer learning," *IEEE Access*, vol. 11, pp. 24053–24076, 2023.
- [21] M. S. Devi, J. A. Pandian, M. B. Lonare, and Y. Praveen, "Efficient net transfer learning based early prediction of monkey pox lesion," in *Proc. 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC)*, 2022, pp. 1297–1300.
- [22] S. Surati, H. Trivedi, B. Shrimali, C. Bhatt, and C. M. Travieso Gonzalez, "An enhanced diagnosis of monkeypox disease using deep learning and a novel attention model senet on diversified dataset," *Multimodal Technologies and Interaction*, vol. 7, no. 8, 2023.
- [23] D. S. Khafaga, A. Ibrahim, E.-S. M. El-Kenawy, A. A. Abdelhamid, F. K. Karim, S. Mirjalili, N. Khodadadi, W. H. Lim, M. M. Eid, and M. E. Ghoneim, "An al-biruni earth radius optimization based deep convolutional neural network for classifying monkeypox disease," *Diagnostics*, vol. 12, no. 11, 2022.
- [24] S. Surati, H. Trivedi, and B. Shrimali et al., "An enhanced diagnosis of monkeypox disease using deep learning and a novel attention model senet on diversified dataset," *Multimodal Technologies and Interaction*, vol. 7, no. 8, 2023.
- [25] R. Liu, "Convolutional siamese network-based few-shot learning for monkeypox detection under data scarcity," in *Proc. Second*

- International Conference on Biological Engineering and Medical Science (ICBioMed 2022)*, vol. 12611, 2023, pp. 1459–1465.
- [26] S. N. Ali, M. T. Ahmed, J. Paul, T. Jahan, S. Sani, N. Noor, and T. Hasan, “Monkeypox skin lesion detection using deep learning models: A feasibility study,” arXiv preprint, arXiv: 2207.03342, 2022.
- [27] V. Kumar, “Analysis of CNN features with multiple machine learning classifiers in diagnosis of monkeypox from digital skin images,” *medRxiv*, 2022. <https://doi.org/10.1101/2022.09.11.22278797>
- [28] T. Islam, M. A. Hussain, F. U. H. Chowdhury, and B. R. Islam, “Can artificial intelligence detect monkeypox from digital skin images?” *BioRxiv*, 2022.
- [29] E. Sizikova, J. Vendrow, X. Cao *et al.*, “Automatic infectious disease classification analysis with concept discovery,” arXiv preprint arXiv:2209.02415, 2022.
- [30] M. E. Haque, M. R. Ahmed, R. S. Nila, and S. Islam, “Classification of human monkeypox disease using deep learning models and attention mechanisms,” arXiv preprint, arXiv: 2211.15459, 2022.
- [31] C. Sitaula and T. B. Shahi, “Monkeypox virus detection using pretrained deep learning-based approaches,” *Journal of Medical Systems*, vol. 46, no. 11, 78, 2022.
- [32] S. N. Ali, M. T. Ahmed, T. Jahan, J. Paul, S. M. S. Sani, N. Noor, A. N. Asma, and T. Hasan, “A web-based mpox skin lesion detection system using state-of-the-art deep learning models considering racial diversity,” arXiv preprint, arXiv: 2306.14169, 2023.
- [33] S. A. Tanim, A. R. Aurnob, Z. H. Anik, and M. I. Hossain, “Precise detection of tomato leaf diseases using deep learning approach with EfficientNet,” in *Proc. 2023 26th International Conference on Computer and Information Technology (ICCIT)*, 2023, pp. 1–6.
- [34] C. Lin, P. Yang, Q. Wang, Z. Qiu, W. Lv, and Z. Wang, “Efficient and accurate compound scaling for convolutional neural networks,” *Neural Networks*, vol. 167, pp. 787–797, October 2023.
- [35] P. Foret, A. Kleiner, H. Mobahi, and B. Neyshabur, “Sharpness aware minimization for efficiently improving generalization,” arXiv preprint, arXiv: 2010.01412, 2020.
- [36] S. M. Jaisakthi, P. Mirunalini, C. Aravindan, and R. Appavu, “Classification of skin cancer from dermoscopic images using deep neural network architectures,” *Multimedia Tools and Applications*, vol. 82, no. 10, pp. 15763–15778, 2023.
- [37] J. Yung and J. Puigcerver, “Bigtransfer (BIT): State-of-the-art transfer learning for computer vision,” *TensorFlow Core*, May 20 2020.
- [38] A. Kolesnikov, L. Beyer, and X. Zhai, *et al.*, “BigTransfer (BiT): State-of-the-art transfer learning for computer vision,” in *Proc. the Computer Vision–ECCV 2020: 16th European Conference*, Glasgow, UK, 2020, pp. 491–507.
- [39] M. E. Haque, M. R. Ahmed, R. S. Nila, and S. Islam, “Human monkeypox disease detection using deep learning and attention mechanisms,” in *Proc. 2022 25th International Conference on Computer and Information Technology (ICCIT)*, 2022, pp. 1069–1073.
- [40] M. C. Irmak, T. Aydin, and M. Yaganoglu, “Monkeypox skin lesion detection with mobilenetv2 and VGGNET models,” in *Proc. 2022 Medical Technologies Congress (TIPTEKNO)*, 2022, pp. 1–4.
- [41] M. F. Almufareh, S. Tehsin, M. Humayun, and S. Kausar, “A transfer learning approach for clinical detection support of monkeypox skin lesions,” *Diagnostics*, vol. 13, no. 8, 1503, 2023.
- [42] U. Singh and L. S. Songare, “Analysis and detection of monkeypox using the googlenet model,” in *Proc. 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS)*, 2022, pp. 1000–1008.
- [43] M. R. Hossen, N. Alfaz, A. Sami, S. A. Tanim, T. B. Sarwar, and M. K. Islam, “An EfficientNet to classify monkeypox-comparable skin lesions using transfer learning,” in *Proc. 2023 IEEE International Conference on Omni-layer Intelligent Systems (COINS)*, 2023, pp. 1–6.

Copyright © 2024 by the authors. This is an open access article distributed under the Creative Commons Attribution License ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.