IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 July 2023, accepted 27 July 2023, date of publication 1 August 2023, date of current version 9 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3300793

RESEARCH ARTICLE

Monkeypox Diagnosis With Interpretable Deep Learning

MD. MANJURUL AHSAN[1], MD. SHAHIN ALI[2], MD. MEHEDI HASSAN[3], (Member, IEEE), TAREQUE ABU ABDULLAH[4], KISHOR DATTA GUPTA[5], (Member, IEEE), ULAS BAGCI[6], CHETNA KAUSHAL[7], AND NAGLAA F. SOLIMAN[8]

[1]Department of Radiology, Machine and Hybrid Intelligence Laboratory, Northwestern University, Chicago, IL 60208, USA
[2]Department of Biomedical Engineering, Islamic University, Kushtia 7003, Bangladesh
[3]Computer Science and Engineering Discipline, Khulna University, Khulna 9208, Bangladesh
[4]Department of Computer Science, Lamar University, Beaumont, TX 77705, USA
[5]Department of Computer and Information Science, Clark Atlanta University, Atlanta, GA 30314, USA
[6]Department of Radiology, BME, and ECE, Machine and Hybrid Intelligence Laboratory, Northwestern University, Chicago, IL 60611, USA
[7]Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140401, India
[8]Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia

Corresponding authors: Md. Mehedi Hassan (mehedihassan@ieee.org) and Naglaa F. Soliman (nfsoliman@pnu.edu.sa)

This work was supported by the Deanship of Scientic Research, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, under Project PNURSP2023R66.

ABSTRACT As the world gradually recovers from the impacts of COVID-19, the recent global spread of Monkeypox disease has raised concerns about another potential pandemic, highlighting the urgency of early detection and intervention to curb its transmission. Deep Learning (DL)-based disease prediction presents a promising solution, offering affordable and accessible diagnostic services. In this study, we harnessed Transfer Learning (TL) techniques to tweak and assess the performance of an array of six different DL models, encompassing VGG16, InceptionResNetV2, ResNet50, ResNet101, MobileNetV2, VGG19, and Vision Transformer (ViT). Among this diverse collection, it was the modified versions of the VGG19 and MobileNetV2 models that outshone the others, boasting striking accuracy rates ranging from an impressive 93% to an astounding 99%. Our results echo the findings of recent research endeavors that similarly showcase enhanced performance when developing disease diagnostic models armed with the power of TL. To add to this, we used Local Interpretable Model Agnostic Explanations (LIME) to lend a sense of transparency to our model's predictions and identify the crucial features correlating with the onset of Monkeypox disease. These findings offer significant implications for disease prevention and control efforts, particularly in remote and resource-limited areas.

INDEX TERMS Deep learning, monkeypox, disease diagnosis, transfer learning, image processing, LIME.

I. INTRODUCTION

Monkeypox, also known as Monkeypox Virus (MPXV) disease, is caused by infection with the virus of the same name and is usually found in monkeys [1]. Monkeypox usually occurs in Africa, Central and West Africa, and Asia [2]. Although it can infect any mammal, the virus spreads to humans primarily by biting an infected animal like a bat or primate such as a monkey [3]. Early monkeypox symptoms include muscle pain, headache, fatigue, and fever. The disease almost resembled chickenpox, smallpox, and measles. It may be identified by swollen glands behind the ear, below the jaw, on the neck, or in the groin before the formation of the rash [4]. Even though the virus is not life-threatening, it causes complications in severe cases, including sepsis, pneumonia, and loss of eye vision [5]. Although monkeypox does not affect humans very often, the slight chance of it affecting them should still make people think twice about their exposure to monkeys and rodents, especially in certain parts of the world where outbreaks of this disease happen

The associate editor coordinating the review of this manuscript and approving it for publication was Humaira Nisar.

more often. The Centres for Disease Control and Prevention (CDC) predict that in the next few years, millions of people worldwide will get a new strain of monkeypox [6].

Monkeypox was initially found in 1958 and was rediscovered in the Republic of Congo in West Africa in 2014 [7], [8]. While it may not be as well-known as Ebola or Zika, the monkeypox virus can become just as large a global health threat as those two if enough is not done to stop it from spreading further. Recently, the virus has been slowly spreading, and the number of Monkeypox patients is increasing daily. According to CDC, as of June 06, 2022, the virus had spread among 29 countries, and the number of Confirmed cases was around 1029 [9]. Currently, no appropriate treatments are available for Monkeypox disease [10], [11]. Nevertheless, Brincidofovir and Tecovirimat, two oral medicines primarily used to treat the smallpox virus, are advised for immediate treatment. The best way to protect against this illness is via vaccination [12], [13].

Monkeypox, smallpox, and measles symptoms are almost similar, making it difficult to identify without a proper laboratory test [14]. Testing the skin lesions using electronic microscopy is the gold standard for diagnosing the virus. Polymerase chain reaction (PCR) testing is also widely utilized for COVID-19 testing and may be used to identify the virus [15], [16]. PCR is usually used for laboratory tests. However, during the onset of COVID-19, we have seen that the PCR test kit cannot correctly diagnose COVID-19 patients around 40% of the time [17], [18], which means that multiple tests are required to increase accuracy. As a result, it will be challenging to provide enough tool kits to test both Monkeypox and COVID-19 patients if the disease becomes another worldwide pandemic. The high cost of such a set of tools makes regular usage impractical for many wealthy countries. In the event that Monkeypox or COVID-19 becomes a global pandemic, providing an adequate number of testing kits could prove to be a difficult task. The high cost associated with producing these kits would make it impractical for many affluent countries to utilize them on a regular basis [19].

In the recent past, Machine learning (ML) demonstrated promising results in medical imaging and disease diagnosis [20], which means that diseases such as cancer, pneumonia, and COVID-19 can be detected without doctor intervention [21]. Several studies published this year revealed that deep learning (DL)-based frameworks could be a convenient alternative for detecting Chickenpox and Measles illness, which have symptoms similar to Monkeypox [22], [23], [24]. As an illustration, Chae et al. employed a deep neural network (DNN) and a long-short-term memory (LSTM) model to detect chickenpox, which outperformed the conventional autoregressive integrated moving average (ARIMA) model [22]. In order to detect the potentially blinding varicella-zoster virus, Arias and Mejía [23] designed a deep-learning algorithm. The proposed model has a detection accuracy of up to 97% for the virus. Using CNN, Bhadula et al. [24] were able to identify skin diseases. Acne and lichen planus are detected with 96% and 92% accuracy, respectively, using a CNN model developed by the authors [24]. In 2019, Sriwong et al. (using the CNN method) obtained an accuracy of 79.2% for identifying skin disorders. The authors' primary interest was identifying actinic keratoses, basal cells, and benign keratoses, among other skin conditions [25].

As Monkeypox symptoms significantly affect the human skin, images containing infected human skin can be used to develop an ML-based diagnosis model [10]. Several studies have shown that DL approaches can detect monkeypox skin disease. For instance, Sitaula and Shahi [26] used a pre-trained DNN to detect the monkeypox virus. In their study, the authors tested 13 prospective DL models and evaluated the relevant performance of those models. The ensemble techniques perform the best, with a precision of 85.44%, a recall of 85.47%, an F1-score of 85.40%, and an accuracy of 87.13%, according to their first computational findings [26]. Sahin et al. [27] also used a pre-trained network and incorporated the prediction model with a mobile application. According to the author, their proposed model is 91.11% accurate. However, there are no clear indications of whether the proposed approach is approved for clinical testing or not [27]. Akin et al. [28] used CNN-based models to detect monkeypox skin lesions, and explainable AI was used to assist the models' interpretations. The author reports that their proposed models attained an accuracy of approximately 98.25%, a sensitivity of 96.55%, a specificity of 100%, and an F1-score of 98.25% as per their study [28].

Yasmin et al. [29] present research that uses image analysis approaches based on machine learning to construct a model that can reliably diagnose monkeypox illness. Particularly during the COVID-19 pandemic, the research highlights the significance of early discovery in preventing disease propagation. Using transfer learning and data augmentation approaches, six unique deep learning models were trained using a preprocessed dataset to tackle the challenge. Comparing the models with the highest precision, recall, and accuracy led to the development of the ''PoxNet22'' model. In terms of classifying monkeypox, the proposed model beats current strategies with 100% precision, recall, and accuracy. This discovery makes a substantial contribution to illness identification and detection, opening the door for more precise and reliable diagnostic procedures [29].

In their recent 2022 research, Shah and colleagues present a novel solution for accurately classifying Monkeypox skin lesions using transfer learning-based approaches. The authors tackle the challenge of limited access to PCR tests in remote areas by providing a computer-aided diagnosis method using Deep Learning techniques, which have shown promise in automating skin lesion detection. They train five different transfer learning models on a skin lesion image dataset from various sources and compare their performance to select the best-performing one. Their experiments show that the MobileNetv2 model achieves the highest accuracy of 98.78% in classifying Monkeypox skin lesion images. This proposed

FIGURE 1. Images retrieved from open source data repositories for our proposed research.

approach can provide quick and automated detection of Monkeypox skin lesions, aiding in disease diagnosis and treatment, especially in areas with limited healthcare access [30].

However, none of the studies provides enough reasoning for their higher prediction rates. Moreover, no generalization or regularization approaches have been introduced, which often limits the overfitting issues of DL-based approaches. Based on the referenced literature, it can be inferred that DL techniques have been used extensively to detect various skin diseases caused mainly by various infections. Therefore, traditional CNN is a good approach for developing deep learning-based models to diagnose Monkeypox disease diagnosis. In this work, we proposed and tested six distinct improved deep CNN models by adopting models, namely, VGG16 [31], InceptionResNetV2 [32], ResNet50 [33], [34], [35], ResNet101 [36], and MobileNetV2 [37] using transfer learning approaches. The following is a summary of our contribution:

1) Improvement and assessment of six distinct DL models (VGG16, InceptionResNetV2, ResNet50, ResNet101, MobileNetV2, and VGG19) using generalization and regularization approach for detecting Monkeypox disease.
2) Verify and interpret the model's performance employing Local Interpretable Model Agnostic Explanations (LIME).

II. METHODOLOGY

We leveraged established deep learning (DL) models through transfer learning approaches to extract critical features that are challenging to identify visually. We extract key features that are hard to see with the naked eye because Monkeypox symptoms can look like chickenpox or measles. We then use the model's most complex layer to detect the disease.

A. DATASET

The traditional deep learning approach needs a lot of data for training. However, transfer learning techniques have shown that even a limited dataset can be used to build an efficient CNN model with strong predictive abilities and a pre-trained model [17], [38], [39]. We used a dataset available from Kaggle for our research (See Kaggle repository). We divided the dataset differently for two separate studies. The first study used 76 samples, including 43 Monkeypox and 33 Non-Monkeypox images. For the second study, we used 818 samples, which included 587 Monkeypox and 231 Normal images. Table 1 summarizes the dataset allocation utilized in this investigation. Figure 1 displays a collection of typical images featuring subjects with Monkeypox and Non-Monkeypox conditions derived from the previously mentioned dataset.

TABLE 1. Monkeypox datasets used in this study.

Dataset	Label	Train set	Test set	Total
Study One	Monkeypox	34	9	76
	Normal	26	7	
Study Two	Monkeypox	469	118	818
	Normal	185	46	

B. DATA AUGMENTATION

To mitigate the potential risks of overfitting, we have utilized data augmentation techniques, which encompass stochastic modifications to images such as flipping, rotation, height and weight adjustments, shearing, and zooming as follows:

1) FLIPPING

Flipping is performed on each image inside the training set. Vertical flips capture a distinctive property of medical images: reflection invariance in the vertical direction. Vertical inversions may not always correlate to images that are input. For input images, only horizontal flips of the source images are used. In contrast, a vertical mass flip would still result in a considerable weight.

2) SCALING

Each I is scaled in either the x or y-axis; more precisely, affine transformation is used.

$$A = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix}$$

3) ROTATIONS

Rotations are performed using the affine transformation as well,

$$A = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix},$$

where θ is a value between 0 and 360 degrees.

4) SHEARING

The affine transformation used to represent a shear angle in degrees, which rotates counter-clockwise and is scaled by 0.2,

is expressed as follows:

$$A = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}.$$

s defines the amount that I is sheared, and it is in the range of [0, 0.2].

5) TRANSLATION

Simply moving the image on the X or Y axis is required for translation (or both). For example, it might be assumed that the image seems to have a black background outside its boundaries and translate the image accordingly. This enhancement can discover the vast majority of things anywhere inside the image. Thus, CNN is compelled to examine everywhere.

To perform image augmentation, the *trainAug.flow()* function, a method of the *ImageDataGenerator* class in Keras, was utilized. It generates batches of augmented image data in real time during model training. It applies specified data augmentation techniques like scaling, rotation, and flipping to input images on-the-fly to create new and slightly different images each time the generator is called. During the call to *trainAug.flow*, the training data, batch size, and other parameters were passed to create the generator object, which is then used by the *fit_generator* method to feed augmented data to the model during training [40].

Data augmentation aims to heighten the variety of training data, facilitating the model's ability to generalize more effectively when presented with novel, unfamiliar data. Data augmentation effectively increases the training dataset size without including additional images by generating new and slightly different images each time. This leads to a more robust and accurate model, which can perform well on various input data [45].

Therefore, despite utilizing image augmentation techniques to increase the efficiency of the model's training, it did not affect the training and test data size. Instead, diverse data is created during each batch while keeping the same number of images [46].

C. FINE-TUNING WITH TRANSFER LEARNING

For fine-tuning pre-trained deep learning model(s), we adapted the topmost layers to facilitate the classification process while the initial layers were frozen. Figure 3 displays the flow diagram of the proposed models used during this study. While variations may exist in layer count (typically first 7), specific layer operations, and overall architecture, these models' fundamental principles of feature extraction and hierarchical learning persist.

Our proposed model harnesses the capabilities of several transfer learning models, including VGG16, InceptionResNetV2, ResNet50, ResNet101, MobileNetV2, and VGG19. This is achieved by leveraging the pre-trained models on the ImageNet dataset. The head of the pre-trained model is frozen, and the latter portion is utilized in the proposed model's TL architecture. Using pre-trained models in TL has become a common practice, particularly when working with limited datasets, as it allows the network to leverage knowledge learned from an extensive quantity of data. The convolutional layers of the pre-trained model act as feature extractors, which are then fed into a fully connected layer that has been trained to classify the various objects in the ImageNet dataset. Using this pre-trained model, the proposed model's performance can be significantly improved, particularly when working with limited training data.

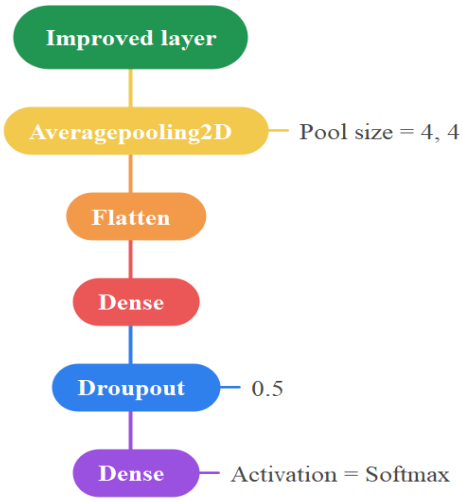


FIGURE 2. Modified layer.

D. PROPOSED MODEL ARCHITECTURE

A pre-existing network refers to one that has undergone training on many datasets and can be implemented for other datasets with similar properties. The pre-trained model is known for its ability to extract features from images. The model has undergone training on the extensive ImageNet dataset, containing over a million images and thousands of object categories. This makes it capable of detecting features in different images, including images of body parts such as faces and hands. Furthermore, the application of transfer learning can enable the utilization of the acquired knowledge from the pre-existing model to enhance the performance of the particular task, despite the relatively limited dataset. The advantages of pre-trained models and their higher performance have also been found in many of the reference literature mentioned in this manuscript [17], [55]. VGG16 networks serve as a prominent example, having been employed in several recent studies to identify COVID-19 cases during the initial stages of the pandemic [17], [56]. An example of a modified architecture proposed in this study for VGG16 is presented in Figure 4. The modified architecture comprises three main steps, as follows:

1) The first step is initializing the model with a pre-trained network that lacks fully connected (FC) layers.
2) To improve the performance of the pre-trained VGG16 model, a new fully connected (FC) layer with the

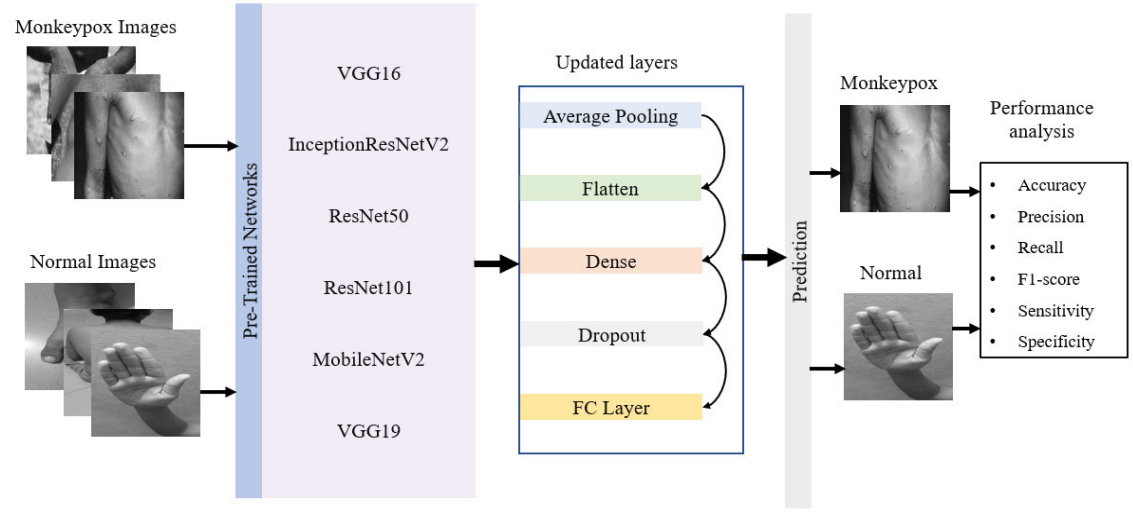


FIGURE 3. Flow diagram of the proposed models.

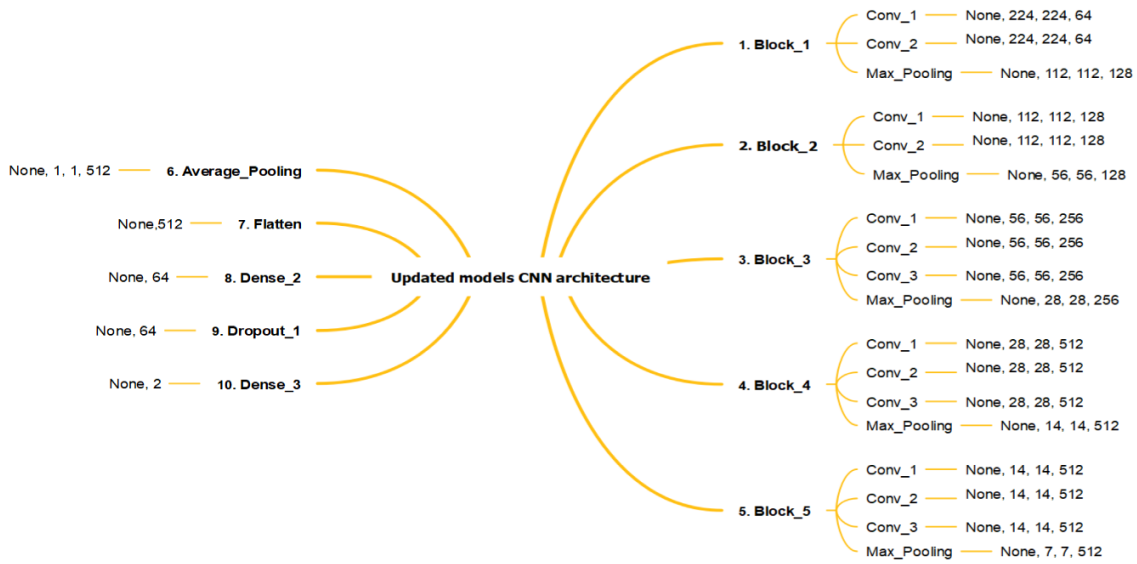


FIGURE 4. Updated VGG16 network implemented during this experiment.

TABLE 2. Parameter settings used during this study.

Parameters	Study One		Study Two	
	Initial parameter	Optimize parameters	Initial parameters	Optimize parameters
Batch size	5, 10, 20, 40	10	10, 20, 30, 40	30
Epochs	20, 30, 40, 50	50	30, 40, 50, 70	50
Learning rate	.001, .01, 0.1	.001	.001, .01, 0.1	.001

"Softmax" activation function is added to the top of the existing model.

3) In the final stage of training, the weights of the convolution layers are frozen, allowing the newly added FC layer to be trained on the specific task.

A similar procedure is followed to develop every transfer learning model used during this study. In this study, the primary modification of the model for all CNN architecture is developed as shown in Figure 2. Although pre-trained methods have been employed previously for image classification

TABLE 3. Monkeypox prediction performance of all models on the training set using various statistical measurements for Study One.

Model	Performance				
	Accuracy	Precision	Recall	F1-Score	Specificity
VGG16	0.98	0.98	0.98	0.98	0.96
InceptionResNetV2	1	1	1	1	1
ResNet50	0.56	0.32	0.57	0.41	0
ResNet101	0.68	0.80	0.68	0.63	0.26
MobileNetV2	1	1	1	1	1
VGG19	0.98	0.98	0.98	0.98	0.96
ViT	0.82	0.82	0.82	0.82	0.76

TABLE 4. Monkeypox prediction performance of all models on the testing set using various statistical measurements for Study One.

Model	Performance				
	Accuracy	Precision	Recall	F1-Score	Specificity
VGG16	0.81	0.86	0.81	0.80	0.42
InceptionResNetV2	0.90	0.88	0.88	0.77	-
ResNet50	0.56	0.32	0.56	0.40	0
ResNet101	0.56	0.32	0.56	0.40	0
MobileNetV2	0.81	0.82	0.81	0.81	0.85
VGG19	0.93	0.94	0.94	0.94	0.85
ViT	0.75	0.76	0.75	0.74	0.57

tasks, our work makes a unique contribution to medical image analysis by focusing on the issue of monkeypox diagnosis, which has not been extensively studied in prior research.

Our proposed study optimized three parameters during the training phase: batch size, epochs, and learning rate. Manually tunning that parameter is time-consuming; therefore, we have used the grid search method to find the best results. Table 2 summarizes the initial and optimal parameters found during the experiment.

E. EXPERIMENT SETUP

The Adam optimizer, known for its robust performance in binary image classification, is utilized to optimize the proposed model. The experimental setup consisted of a standard Windows 10 computer equipped with an Intel Core i7 processor and 16 GB of RAM. The dataset is split into the following ratios, often common in the machine learning domain train set:test set = 80:20 [57]. The overall experimentation is run five times, and the final result is presented by averaging all five results. To evaluate the model's performance following statistical measurement is used [49]:

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{1}$$

$$Precision = \frac{T_p}{T_p + F_p} \tag{2}$$

$$Recall = \frac{T_p}{T_p + F_n} \tag{3}$$

$$F1{-}score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

$$Specificity = \frac{T_n}{T_n + F_p} \tag{5}$$

Here, T_p (True Positive) = Monkeypox infected individual identified as Monkeypox

T_n (True Negative) = Monkeypox infected individual identified as non-Monkeypox

F_p (False Positive) = Non-Monkeypox individual identified as Monkeypox

F_n (False Negative) = Monkeypox individual identified as non-Monkeypox.

III. RESULTS

A. STUDY 1

In Study One, the performance of each CNN model on both the train and test sets was evaluated using Equations 1, 2, 3, 4 and 5. In this particular scenario, InceptionResNetV2 and MobileNetV2 performed far better than any other models regarding the accuracy, precision, recall, F-1 score, sensitivity, and specificity, and the results are presented in Table 3. Notably, ResNet50 presented the poorest performance among all the evaluated metrics.

The results of the performance of all of the models concerning the test set are presented in Table 4. While ResNet50 and ResNet101 performed less, InceptionResNetV2 and VGG19 demonstrated the best performance across all criteria.

1) CONFUSION MATRICES

The total prediction of the model can be represented using confusion matrices. Here, most of the confusion matrix results were found consistent for the repeated measure, and hence only the best-performing outputs are presented. Figure 5 illustrates the confusion matrices acquired for the train set during Study One. Both the InceptionResNetV2 and MobileNetV2 models accurately categorized

TABLE 5. Monkeypox prediction performance of all models on the training set using various statistical measurements for Study Two.

Model	Performance				
	Accuracy	Precision	Recall	F1-Score	Specificity
VGG16	0.97	0.98	0.97	0.98	0.94
InceptionResNetV2	1	1	1	1	1
ResNet50	0.72	0.51	0.72	0.60	0
ResNet101	0.72	0.36	0.50	0.42	0
MobileNetV2	1	1	1	1	1
VGG19	0.94	0.94	0.92	0.93	0.85
ViT	0.89	0.89	0.89	0.89	0.81

TABLE 6. Monkeypox prediction performance of all models on the testing set using various statistical measurements for Study Two.

Model	Performance				
	Accuracy	Precision	Recall	F1-Score	Specificity
VGG16	0.93	0.93	0.90	0.91	0.82
InceptionResNetV2	0.98	0.98	0.97	0.98	0.95
ResNet50	0.72	0.51	0.72	0.60	0
ResNet101	0.72	0.36	0.50	0.42	0
MobileNetV2	0.99	0.99	0.99	0.99	0.97
VGG19	0.90	0.89	0.86	0.87	0.76
ViT	0.87	0.87	0.87	0.87	0.69

TABLE 7. Confidence Interval ($\alpha = 0.05$) of all models used in Study One regarding accuracy.

Study One	Model	Train Accuracy	Methods	
			Wilson Score	Binomial proportion interval
Train set	VGG16	0.98	0.851-0.994	0.914-1.0
	InceptionResNetV2	1	0.93-0.99	1.00-1.00
	ResNet50	0.56	0.44-0.68	0.44-0.692
	ResNet101	0.68	0.55-0.78	0.56-0.80
	MobileNetV2	1	0.93-0.99	1.0-1.0
	VGG19	0.98	0.91-0.99	0.95-1.0
	ViT	0.82	0.72-0.92	0.70-0.90
Test set	VGG16	0.81	0.56-0.93	0.62-1.0
	InceptionResNetV2	0.87	0.63-0.96	0.713-1.0
	ResNet50	0.56	0.33-0.76	0.32-0.81
	ResNet101	0.56	0.33-0.76	0.32-0.81
	MobileNetV2	0.81	0.56-0.93	0.62-1.0
	VGG19	0.93	0.71-0.98	0.81-1.0
	ViT	0.75	0.54-0.96	0.51-0.90

TABLE 8. Confidence Interval ($\alpha = 0.05$) of all models used in Study Two in terms of accuracy.

Study Two	Model	Train Accuracy	Methods	
			Wilson Score	Binomial proportion interval
Train set	VGG16	0.97	0.95-0.98	0.95-0.98
	InceptionResNetV2	1	0.99-1.0	1.0-1.0
	ResNet50	0.72	0.68-0.75	0.68-0.75
	ResNet101	0.72	0.68-0.75	0.68-0.75
	MobileNetV2	1	0.99-1.0	1.0-1.0
	VGG19	0.94	0.91-0.95	0.92-0.95
	ViT	0.89	0.87-0.91	0.86-0.91
Test set	VGG16	0.93	0.88-0.96	0.89-0.97
	InceptionResNetV2	0.98	0.94-0.99	0.961-1.0
	ResNet50	0.72	0.68-0.75	0.68-0.75
	ResNet101	0.72	0.68-0.75	0.68-0.75
	MobileNetV2	0.99	0.96-0.99	0.98-1.0
	VGG19	0.90	0.84-0.93	0.85-0.94
	ViT	0.87	0.82-0.92	0.81-0.91

all patients, as shown in Figure 5. Contrarily, VGG16, ResNet50, ResNet101, and VGG19 misclassified 1, 26, 19, and 1 patients, respectively. As a result, we can conclude that the performance of the ResNet50 model is the worst among all of the models based on assessing the confusion matrices.

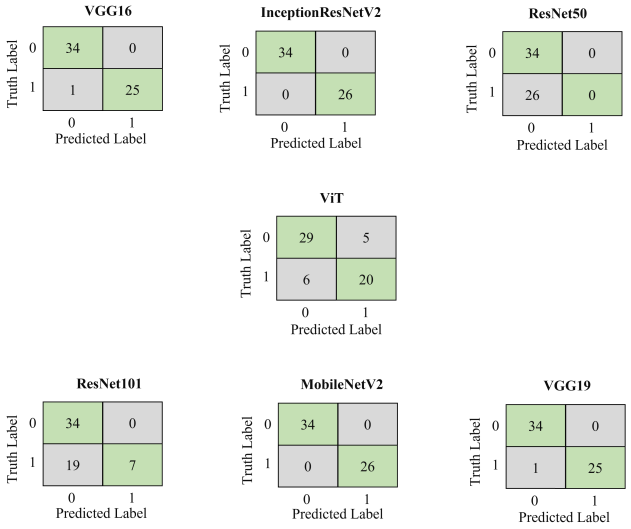


FIGURE 5. Study 1 confusion matrices for the train set.

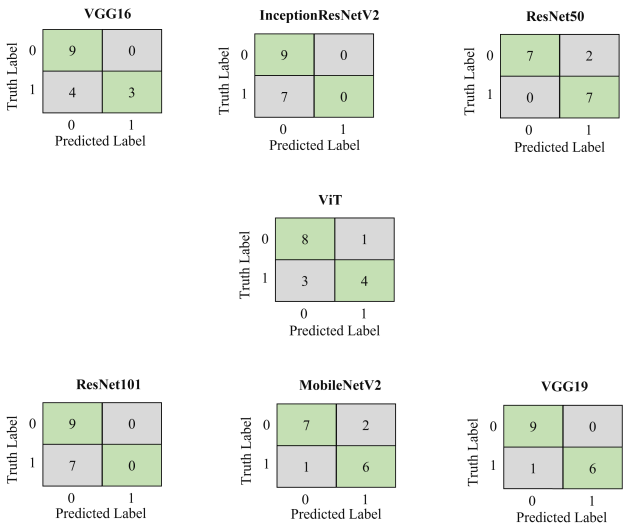


FIGURE 6. Study 1 confusion matrices for the test set.

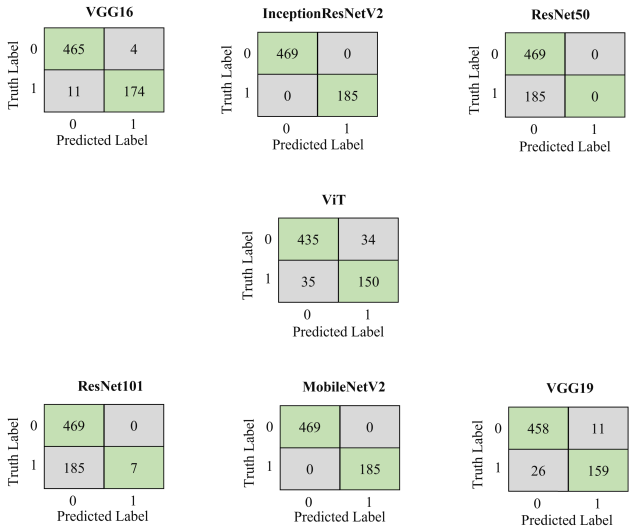


FIGURE 7. Study 2 confusion matrices for the train set.

FIGURE 8. Study 2 confusion matrices for the test set.

Figure 6 displays the confusion matrices for Study One on the test set. Models InceptionResNetV2 displayed the best performance by misclassifying only two patients, while both ResNet50 and ResNet101 demonstrated the worst performance by misclassifying seven patients. Considering the performance on both the train and test set, it can be inferred that the performance of modified InceptionResNetV2 remains constant for both the train and test set, ultimately demonstrating the best performance compared to all other models.

During the experiment, it was observed that the performance of the confusion matrix remained consistent throughout. This consistency in performance can be attributed to the utilization of a robust and proven trained architecture of transfer learning models. The transfer learning technique enabled the utilization of pre-trained models already trained on large datasets. This approach has been proven effective in improving the performance of machine learning models, particularly in cases where the dataset size is limited.

2) MODEL PERFORMANCE DURING TRAINING

Figure 9 shows train and test performance during each epoch for the CNN models used in this study. The figure shows that the model's performance during each epoch improves for InceptionResNetV2 and VGG19 until 50 epochs. On the other hand, ResNet50 and ResNet101's performance fluctuated during each epoch and overfitted at 50 epochs.

3) EXPLAINABLE AI WITH LIME

Our study utilizes Local Interpretable Model-Agnostic Explanations (LIME) to elucidate the predictions of the Monkeypox diagnosis models. LIME provides interpretive insights into which features in the input images are most

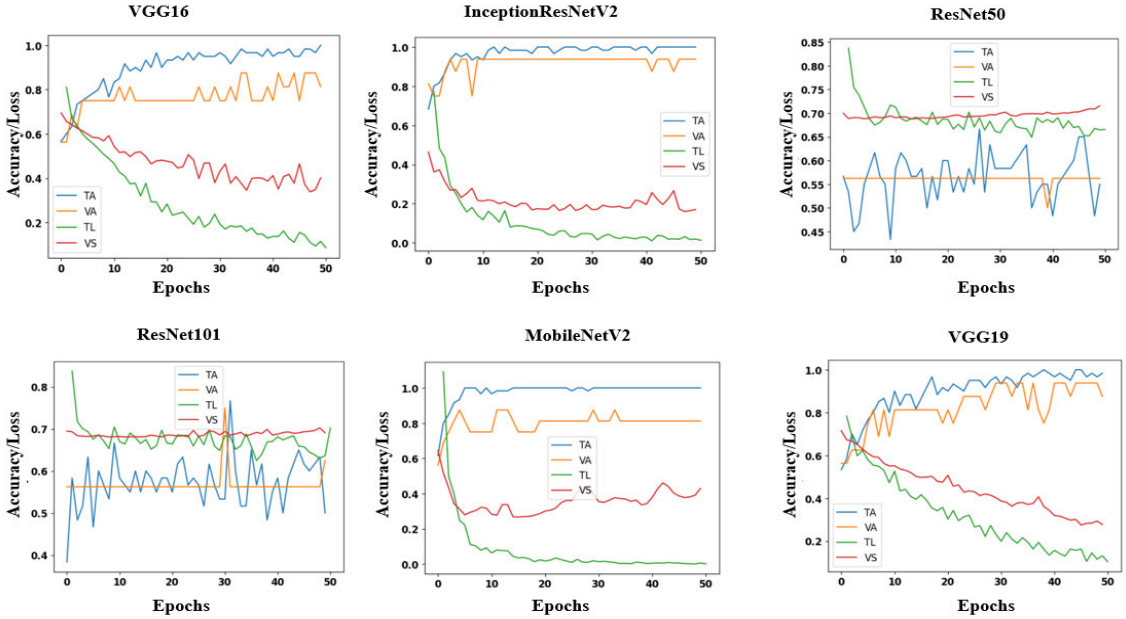


FIGURE 9. Accuracy and loss of the modified models during each epoch for Study One. "TA" represents train accuracy, "VA" represents validation accuracy, "TL" represents train loss, and "VL" represents validation loss.

crucial for the model's predictions. The top four features were identified based on their corresponding LIME scores, encompassing the presence of skin lesions, their size, shape, and texture [58], [59].

To determine the top features, we utilized the LIME algorithm to generate explanations for each prediction made by the model. The features were ranked based on their LIME scores, which quantify the relevance of each feature in the model's decision-making process. Additionally, we computed the frequency of occurrence of each feature in the dataset of Monkeypox images. Additionally, we calculated the correlation between these features and the diagnosis outcome to validate their significance. We found that the top four features highly correlated with the diagnosis outcome and were present in a significant proportion of the dataset. By providing quantitative statistics, we substantiated the importance of the identified features and validated their significance in Monkeypox diagnosis.

The utilization of quantitative parameters to identify superpixels from patients exhibiting symptoms of Monkeypox is a crucial aspect of the proposed methodology, as highlighted in Table 9. The table outlines the optimal values for the function, maximum distance, kernel size, and ratio, which were determined after experimentation with various values. The optimal value for the maximum distance was found to be 200, while the optimal kernel size was 4, and the optimal ratio was 0.2. Identifying superpixels through quantitative parameters is important in accurately identifying the top features and diagnosing Monkeypox symptoms.

Figure 10 explains the model's prediction using LIME. The best performance was observed when the LIME model was used with 150 perturbations. The difference between each perturbation was calculated using a cosine metric with a kernel size of 0.2. A simple linear interpolation is used to explain the models. Furthermore, the coefficient was calculated for each superpixel value. Finally, the top four features potentially affecting the model's prediction are identified.

TABLE 9. The quantitative parameter utilized to identify superpixels from patients exhibiting symptoms of Monkeypox is a crucial aspect of our methodology.

Function	Value	Optimal Value
Maximum distance	150, 200, 250	200
Kernel size	2, 4, 6, 8	4
Ratio	0.1, 0.2, 0.3, 0.4	0.2

B. STUDY 2

In Study Two, on the train set (refer to Table 5), InceptionResNetV2 and MobileNetV2 outperformed all other models across all measures, while ResNet50 and ResNet101 demonstrated the worst performance.

Table 6 shows the test performance of various models applied during study two as a transfer learning approach. The highest performance was observed for MobileNetV2, and the lowest was identified for ResNet101 across all measures.

1) CONFUSION MATRICES

Figure 7 shows the confusion matrices for the train set during Study Two. Both the InceptionResNetV2 and MobileNetV2 accurately categorized all patients, as seen in Figure 7.

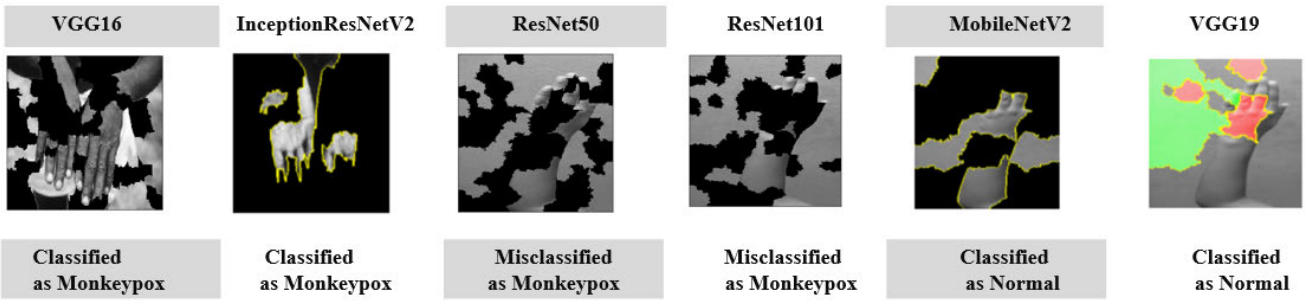


FIGURE 10. The top four features that facilitated the classification and misclassification of Monkeypox patients from infected or non-infected images during Study 1.

In contrast, VGG16, ResNet50, Resnet101, and VGG19 misclassified 15, 185, 185, and 39 patients, respectively. As a result, we can conclude that the performance of the ResNet50 and ResNet101 models is the worst out of all of them based on the analysis of the confusion matrices.

Figure 8 displays the confusion matrices for Study One on the test set. MobileNetV2 displayed the best performance by misclassifying only one patient, while ResNet50 and ResNet101 indicated the worst performance by misclassifying 46 patients. Considering both train and test analysis, it can be inferred that the best performance of modified MobileNetV2 remains constant for both train and test sets.

2) MODELS' PERFORMANCE DURING TRAINING

Figure 11 plotted train and test results during each epoch for all the CNN models for Study Two. The figure shows that the model's performance during each epoch shows better performance for most of the models except ResNet50 and ResNet101. The accuracy and loss for the train and test set fluctuated and overfitted for models ResNet50 and ResNet101 after 25 epochs.

The model's performance and explanation found in Study 1 align with Study Two. As shown in Figure 12, it is clear that ResNet50 and ResNet101 demonstrated the least performance by misclassifying Monkeypox patients.

IV. DISCUSSION

In order to differentiate between patients who have monkeypox and those whose symptoms are not related to monkeypox, the authors of this work propose and evaluate six separate modified deep learning-based models. Our dataset only has a small number of samples, and we have assessed the precision of our model using confidence intervals with a 95% level of certainty, which were derived from research that has been cited in the past [23], [25], [56]. The results of Study One's model correctness are shown in Table 7, along with confidence intervals of 95% for both the train and test sets. Two approaches have been applied to measure the performance with CI: Wilson score and binomial proportional interval. As stated in Table 7, InceptionResNetV2 exceeded all other models in terms of accuracy. In contrast, ResNet50 demonstrates significantly lower performance on the train and test set while measuring the accuracy with CI.

Table 8 manifests the accuracy performance of six different models for Study Two. As stated in the table, MobileNetV2 outperformed all other models by performing better, while ResNet50 and ResNet101 demonstrate the worst. Further, we have used Heatmap as a data visualization technique demonstrating the phenomenon as color in two dimensions for different CNN layers during the prediction. Using a heatmap, it is easy to identify the specific region where CNN establishes more focus during the prediction. The idea is to use the weights from the last dense layers and multiply them with the final CNN layer. In this work, we identify the potential Monkeypox infected areas using global average pooling (GAP). A class activation map (CAM) is initially used to collect each convolutional layer output and combine it in one shot. Figure 13 (a) shows only the heatmap mask on the spotted regions without the accompanying images. Figure 13 (b) demonstrates the identification of different regions of the images that were individually spotted. Finally, Figure 13 (c) presents the entire area of the spotted regions in the complete image. From the overall analysis across various measures, it can be inferred that MobileNetV2 and InceptionResNetV2 are the best models that can detect Monkeypox patients accurately. It is relevant to emphasize that, at the time of the experiment, there was no published or preprint literature where the transfer learning or deep learning approach was utilized to develop an AI-driven model to detect Monkeypox patients. Consequently, we could not evaluate our model's performance by comparing it with others' work. However, our higher performance results align with many referenced research, which established the potentiality of the transfer learning approach to develop image-based disease diagnosis. Examples include the InceptionResNetV2 model proposed by Narin et al. (2020), which trained with only 100 images to detect COVID-19 patients. Fujisawa et al. (2022) developed a deep learning-based model to detect skin tumors. The model was developed with 1842 patient data and achieved an accuracy of 76.5%. Our proposed model predictions are explained with LIME and cross-checked by expert doctors' opinions. Our findings suggest that our modified MobileNetV2 and InceptionResNetV2 demonstrated a significant contribution

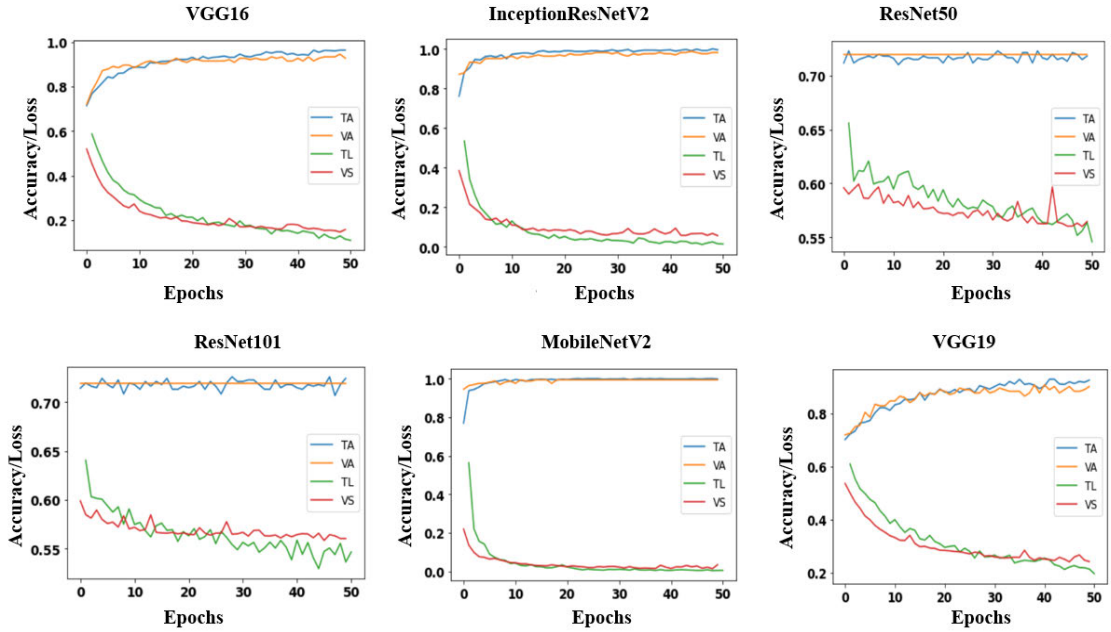


FIGURE 11. Modified models accuracy and loss during each epoch applied to Study Two. TA—train accuracy, VA—validation accuracy, TL—train loss, VS—validation loss.

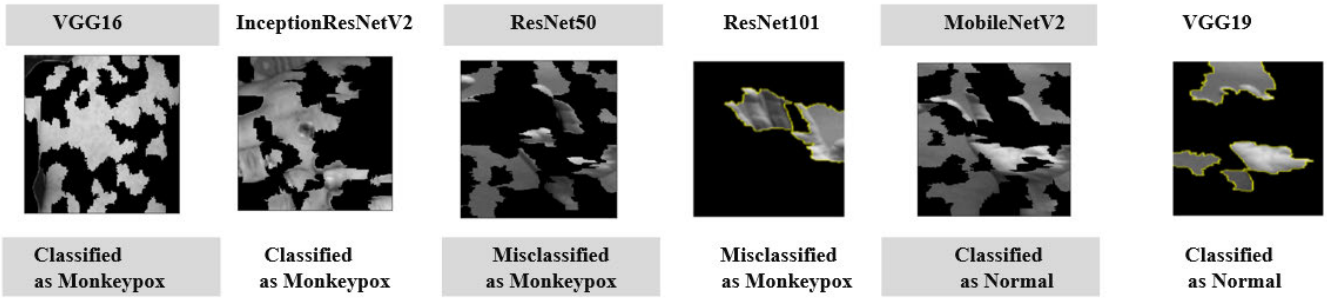


FIGURE 12. The top four features that facilitated the classification and misclassification of monkeypox patients from infected or non-infected hand images during Study 2.

to identifying the top feature that characterizes the onset of Monkeypox by identifying essential features from the images.

To further aid in comparing the model's final prediction with the visual evidence, we have also employed an image segmentation method [60] to highlight highly infected areas. Figure 14 depicts the outcomes of threshold-based segmentation applied to images of hands infected with the Monkeypox virus. The region of interest (ROI) has been defined using the pixel values of the infected areas. Then, a mask is generated for segmenting and generating the final output.

The major purpose of this research was to create a deep learning-based model capable of detecting patients with Monkeypox symptoms using a dataset of images of Monkeypox patients. Due to a lack of data, the scope of the current literature on this topic remains limited at this time. In the near future, it will be fascinating to examine how our suggested model performs on the multiclass and a large dataset.

The DL-based model proposed for Monkeypox disease diagnosis has been compared with other existing literature in Table 10. The table reveals that the accuracy of the proposed models in Studies One and Two is significantly higher than that of the other models considered in the literature. In Study Two, the modified MobileNetV2 model achieved an accuracy of 99%, demonstrating its superior performance. Furthermore, the proposed models' precision, recall, and F1-score exceed those of most of the other models considered.

Additionally, we found that, even though the ViT model is supposed to perform better than traditional CNN and TL-based approaches on big datasets, it does not do as well on smaller and medium-sized datasets as found in this study. One potential reason is that the ViT was designed for large and complex data, can sometimes get too specific, or can not generalize well when there is not enough training data. On the other hand, CNNs, which are good at picking out features and understanding space, continue to work better on smaller and

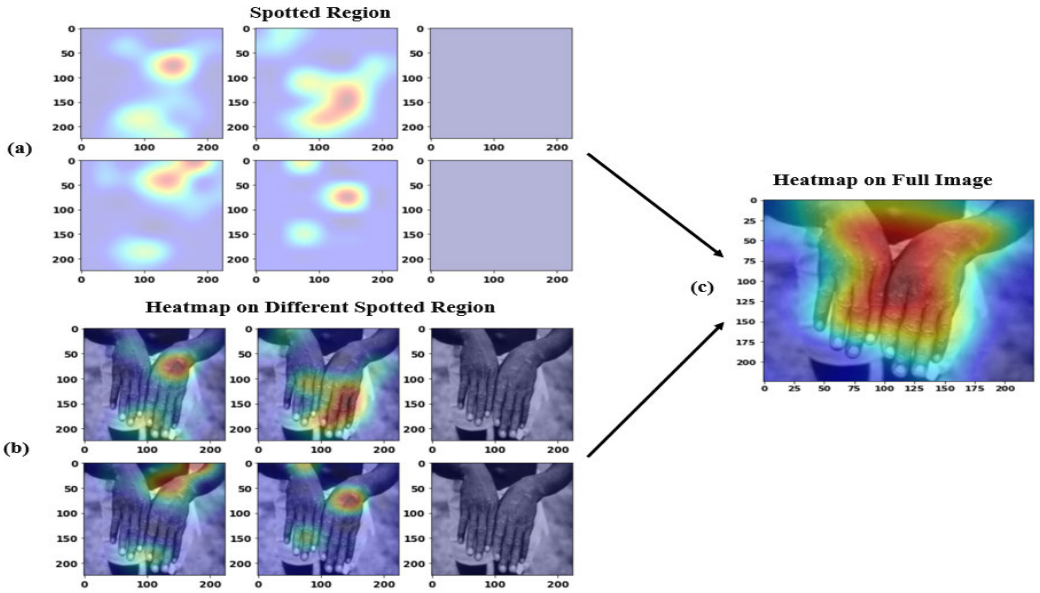


FIGURE 13. Heatmap of class activation on Monkeypox patient image.

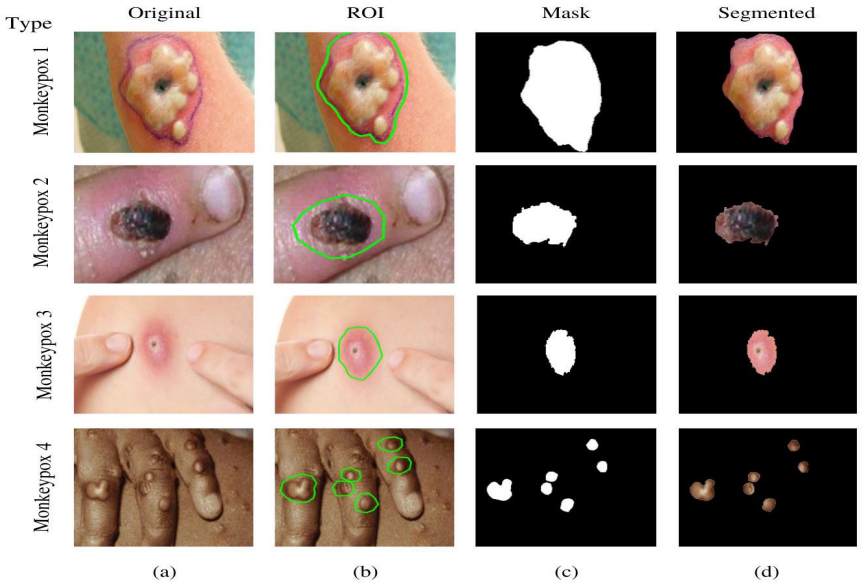


FIGURE 14. Threshold-based segmentation (a) Original, (b) ROI, (c) Mask, (d) Segmented.

medium-sized datasets. This shows how important it is to pick the right model for the amount and complexity of the data, which is something we learned from our research.

1) POTENTIAL REMARKS

- **Application focus:** Our paper specifically focuses on image-based Monkeypox disease classification, aiming to provide a comprehensive overview of this specific domain. We acknowledge that Monkeypox disease can exhibit similar symptoms to other diseases, such as chickenpox and measles. However, our intention was to delve into the specific problem of Monkeypox disease detection using image analysis techniques.
- **Dataset limitations:** Although it could be valuable to broaden the dataset to encapsulate samples from related diseases, such an undertaking was beyond the purview of our current study. Our primary objective was to underline the practicality of deploying TL techniques in detecting Monkeypox, paving the way for the next wave of research in this area.
- **Practical value:** In demonstrating the efficiency of TL within the context of Monkeypox classification, we offer a pragmatic solution that has the potential to lay the groundwork for future investigations and real-world applications. Our work underlines the adaptability and relevance of TL techniques in the field of disease

TABLE 10. Comparison of the proposed transfer learning-based model with the existing state-of-the-art techniques that considered the Monkeypox disease diagnosis model.

Study	Model	Accuracy	Precision	Recall	F1-score
[61]	ResNet50	72%	0.59	0.5	0.54
	InceptionV3	70%	0.71	0.5	0.61
	ShuffleNet-V2	79%	0.79	0.57	0.67
[62]	VGG16	82%	0.86	0.82	0.84
	ResNet50	83%	0.88	0.83	0.84
	InceptionV3	75%	0.74	0.81	0.78
[63]	MobileNetV2	92%	0.91	0.87	0.89
[64]	MobileNetV3-s	96%	-	0.97	0.98
[27]	MobileNetV2	92%	-	-	-
Our best model (Test set)					
Study One	modified VGG19	93%	0.94	0.94	0.94
Study Two	modified MobileNetV2	99%	0.99	0.99	0.99

detection, opening up avenues for further exploration and enhancement. This will, in turn, propel the development of more refined and robust diagnostic tools in healthcare.

- **Innovation and contribution:** Our work carves valuable insights and sets down a strong foundation for future research in this field. We aim not to revolutionize but to add another brick to the edifice of knowledge, focusing on the practical applicability of artificial intelligence in healthcare. We firmly believe that every incremental advance, such as ours, is crucial in using technology to solve real-world health challenges. Hence, our focus on Monkeypox classification and applying TL techniques provides valuable insights and establishes a baseline for further research in this area.

To summarize, our paper hones in on classifying Monkeypox disease based on images, using the power of Transfer Learning (TL) techniques. We fully recognize that our work has room for improvement, particularly in dataset expansion. Nevertheless, the practical significance of our study is undiminished. Our work offers a meaningful contribution to the research community, shedding light on the specialized area of Monkeypox disease diagnosis through the analysis of infected images. This represents a key stride in healthcare, highlighting the potential for machine learning to aid in detecting critical health conditions. Our findings pave the way for future studies to optimize further and refine image-based disease diagnosis methods, ultimately contributing to the global fight against diseases like Monkeypox.

V. CONCLUSION

This study introduces and evaluates six modified deep-learning models, including Vision Transformer (ViT), specifically tuned for detecting Monkeypox disease. Our experiments reveal that modified versions of VGG19 and MobileNetV2 demonstrate outstanding proficiency in distinguishing between Monkeypox and Non-Monkeypox patients, with accuracy levels soaring from 90% to 100%. We further enhance our approach's transparency by applying the LIME (Local Interpretable Model-Agnostic Explanations) method, allowing us to better understand and validate our models' predictions by opening the black box nature of the most deep networks, including what we applied and creating a trustworthy environment for diagnostic results' interpretations. Our experiments affirm the superior capability of our proposed models in detecting Monkeypox, marking a significant stride in the field. With our findings and extensive evaluations, we offer valuable insights into our study to future scholars and practitioners into the potential of leveraging transfer learning models and explainable AI for the development of safe and reliable Monkeypox disease diagnosis models.

DATA AVAILABILITY

The data used to support the findings of this study are available from the corresponding author upon request.

AUTHOR CONTRIBUTION

Md. Manjurul Ahsan: Conceptualization, Methodology, Software, Visualization, Writing-Original draft preparation. **Md. Shahin Ali:** Methodology, Software, Visualization, Writing-reviewing and editing. **Md. Mehedi Hassan:** Data curation, Writing-reviewing and editing. **Tareque Abu Abdullah:** Writing-reviewing and editing. **Kishor Datta Gupta:** Writing-reviewing and editing. **Ulas Bagci:** Writing-reviewing and editing. **Chetna Kaushal:** Writing-reviewing and editing. **Naglaa F. Soliman:** Writing-reviewing and editing.

COMPETING INTEREST

The authors declare that there is no conflict of interest regarding the publication of this article.

ACKNOWLEDGMENT

The authors would like to acknowledge Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R66), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

REFERENCES

[1] L. V. Patrono, K. Pléh, L. Samuni, M. Ulrich, C. Röthemeier, A. Sachse, S. Muschter, A. Nitsche, E. Couacy-Hymann, C. Boesch, R. M. Wittig, S. Calvignac-Spencer, and F. H. Leendertz, "Monkeypox virus emergence in wild chimpanzees reveals distinct clinical outcomes and viral diversity," *Nature Microbiol.*, vol. 5, no. 7, pp. 955–965, Apr. 2020.

[2] S. Parker, A. Nuara, R. M. L. Buller, and D. A. Schultz, "Human monkeypox: An emerging zoonotic disease," *Future Microbiol.*, vol. 2, no. 1, pp. 17–34, Feb. 2007.

[3] S. Wong, S. Lau, P. Woo, and K.-Y. Yuen, "Bats as a continuing source of emerging infections in humans," *Rev. Med. Virol.*, vol. 17, no. 2, pp. 67–91, Mar. 2007.

[4] Centres for Disease Control and Prevention. (2022). *Monkeypox Signs and Symptoms*. [Online]. Available: https://www.cdc.gov/poxvirus/mpox/symptoms/index.html

[5] M. Reynolds, A. McCollum, B. Nguete, R. S. Lushima, and B. Petersen, "Improving the care and treatment of monkeypox patients in low-resource settings: Applying evidence from contemporary biomedical and smallpox biodefense research," *Viruses*, vol. 9, no. 12, p. 380, Dec. 2017.

[6] A. S. Fauci, "Emerging and reemerging infectious diseases: The perpetual challenge," *Academic Med.*, vol. 80, no. 12, pp. 1079–1085, Dec. 2005.

[7] Michaeleen Doucleff. (2022). *Scientists Warned us About Monkeypox in 1988. Here's Why they were Right*. [Online]. Available: https://t.ly/QbTJ

[8] M. Dwivedi, R. G. Tiwari, and N. Ujjwal, "Deep learning methods for early detection of monkeypox skin lesion," in *Proc. 8th Int. Conf. Signal Process. Commun. (ICSC)*, Dec. 2022, pp. 343–348.

[9] Centres for Disease Control and Prevention. (2022). *2022 Monkeypox and Orthopoxvirus Outbreak Global Map*. [Online]. Available: https://www.who.int/emergencies/situations/monkeypox-oubreak-2022

[10] M. M. Ahsan, M. R. Uddin, M. Farjana, A. N. Sakib, K. Al Momin, and S. A. Luna, "Image data collection and implementation of deep learning-based model in detecting monkeypox disease using modified VGG16," 2022, *arXiv:2206.01862*.

[11] M. M. Ahsan, M. R. Uddin, and S. A. Luna, "Monkeypox image data collection," 2022, *arXiv:2206.01774*.

[12] Centres for Disease Control and Prevention. (2022). *Monkeypox and Smallpox Vaccine*. [Online]. Available: https://t.ly/e3b5

[13] F. Yasmin, M. M. Hassan, S. Zaman, S. T. Aung, A. Karim, and S. Azam, "A forecasting prognosis of the monkeypox outbreak based on a comprehensive statistical and regression analysis," *Computation*, vol. 10, no. 10, p. 177, Oct. 2022.

[14] D. L. Heymann, M. Szczeniowski, and K. Esteves, "Re-emergence of monkeypox in africa: A review of the past six years," *Brit. Med. Bull.*, vol. 54, no. 3, pp. 693–702, Jan. 1998.

[15] T. Wawina-Bokalanga, N. Sklenovska, B. Vanmechelen, M. Bloemen, V. Vergote, L. Laenen, E. Andre, M. Van Ranst, J.-J. T. Muyembe, and P. Maes, "An accurate and rapid real-time PCR approach for human monkeypox virus diagnosis," *medRxiv*, 2022.

[16] Centres for Disease Control and Prevention. (2022). *Diagnostic Tests*. [Online]. Available: https://www.cdc.gov/dpdx/diagnosticprocedures/index.html

[17] M. M. Ahsan, M. T. Ahad, F. A. Soma, S. Paul, A. Chowdhury, S. A. Luna, M. M. S. Yazdan, A. Rahman, Z. Siddique, and P. Huebner, "Detecting SARS-CoV-2 from chest X-ray using artificial intelligence," *IEEE Access*, vol. 9, pp. 35501–35513, 2021.

[18] S. Upadhayay, R. Arthur, D. Soni, P. Yadav, U. Navik, R. Singh, T. G. Singh, and P. Kumar, "Monkeypox infection: The past, present, and future," *Int. Immunopharmacology*, vol. 113, Dec. 2022, Art. no. 109382.

[19] M. M. Ahsan, T. E. Alam, T. Trafalis, and P. Huebner, "Deep MLP-CNN model using mixed-data to distinguish between COVID-19 and non-COVID-19 patients," *Symmetry*, vol. 12, no. 9, p. 1526, Sep. 2020.

[20] M. M. Ahsan and Z. Siddique, "Machine learning-based heart disease diagnosis: A systematic literature review," *Artif. Intell. Med.*, vol. 128, Jun. 2022, Art. no. 102289.

[21] M. M. Ahsan, S. A. Luna, and Z. Siddique, "Machine-learning-based disease diagnosis: A comprehensive review," *Healthcare*, vol. 10, no. 3, p. 541, Mar. 2022.

[22] S. Chae, S. Kwon, and D. Lee, "Predicting infectious disease using deep learning and big data," *Int. J. Environ. Res. Public Health*, vol. 15, no. 8, p. 1596, Jul. 2018.

[23] R. Arias and J. Mejía, "Varicella zoster early detection with deep learning," in *Proc. IEEE Eng. Int. Res. Conf. (EIRCON)*, Oct. 2020, pp. 1–4.

[24] S. Bhadula, S. Sharma, P. Juyal, and C. Kulshrestha, "Machine learning algorithms based skin disease detection," *Int. J. Innov. Technol. Exploring Eng.*, vol. 9, no. 2, pp. 4044–4049, 2019.

[25] K. Sriwong, S. Bunrit, K. Kerdprasop, and N. Kerdprasop, "Dermatological classification using deep learning of skin image and patient background knowledge," *Int. J. Mach. Learn. Comput.*, vol. 9, no. 6, pp. 862–867, Dec. 2019.

[26] C. Sitaula and T. B. Shahi, "Monkeypox virus detection using pre-trained deep learning-based approaches," *J. Med. Syst.*, vol. 46, no. 11, pp. 1–9, Oct. 2022.

[27] V. H. Sahin, I. Oztel, and G. Y. Oztel, "Human monkeypox classification from skin lesion images with deep pre-trained network using mobile application," *J. Med. Syst.*, vol. 46, no. 11, pp. 1–10, Oct. 2022.

[28] K. D. Akin, C. Gurkan, A. Budak, and H. Karataş, "Classification of monkeypox skin lesion using the explainable artificial intelligence assisted convolutional neural networks," *Avrupa Bilim ve Teknoloji Dergisi*, vol. 40, pp. 106–110, Sep. 2022.

[29] F. Yasmin, M. M. Hassan, M. Hasan, S. Zaman, C. Kaushal, W. El-Shafai, 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.

[30] A. Shah, "Monkeypox skin lesion classification using transfer learning approach," in *Proc. IEEE Bombay Sect. Signature Conf. (IBSSC)*, Dec. 2022, pp. 1–5.

[31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[32] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 4278–4284.

[33] T. Akiba, S. Suzuki, and K. Fukuda, "Extremely large minibatch SGD: Training ResNet-50 on ImageNet in 15 minutes," 2017, *arXiv:1711.04325*.

[34] W. Islam, M. Jones, R. Faiz, N. Sadeghipour, Y. Qiu, and B. Zheng, "Improving performance of breast lesion classification using a ResNet50 model optimized with a novel attention mechanism," *Tomography*, vol. 8, no. 5, pp. 2411–2425, Sep. 2022.

[35] W. Islam, G. Danala, H. Pham, and B. Zheng, "Improving the performance of computer-aided classification of breast lesions using a new feature fusion method," *Proc. SPIE*, vol. 12033, pp. 84–91, Apr. 2022.

[36] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[37] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.

[38] M. K. Islam, M. S. Ali, M. S. Miah, M. M. Rahman, M. S. Alam, and M. A. Hossain, "Brain tumor detection in MR image using superpixels, principal component analysis and template based K-means clustering algorithm," *Mach. Learn. With Appl.*, vol. 5, Sep. 2021, Art. no. 100044.

[39] M. S. Ali, M. K. Islam, J. Haque, A. A. Das, D. S. Duranta, and M. A. Islam, "Alzheimer's disease detection using m-Random forest algorithm with optimum features extraction," in *Proc. 1st Int. Conf. Artif. Intell. Data Anal. (CAIDA)*, Apr. 2021, pp. 1–6.

[40] S. K. M S. Islam, M. A. Al Nasim, I. Hossain, D. M. A. Ullah, D. K. D. Gupta, and M. M. H. Bhuiyan, "Introduction of medical imaging modalities," 2023, *arXiv:2306.01022*.

[41] M. S. Ali, M. S. Miah, J. Haque, M. M. Rahman, and M. K. Islam, "An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models," *Mach. Learn. with Appl.*, vol. 5, Sep. 2021, Art. no. 100036.

[42] M. Hassan, S. Zaman, S. Mollick, M. M. Hassan, M. Raihan, C. Kaushal, and R. Bhardwaj, "An efficient Apriori algorithm for frequent pattern in human intoxication data," *Innov. Syst. Softw. Eng.*, vol. 19, pp. 1–9, Oct. 2023.

[43] A. Kwasigroch, A. Mikolajczyk, and M. Grochowski, "Deep neural networks approach to skin lesions classification—A comparative analysis," in *Proc. 22nd Int. Conf. Methods Models Autom. Robot. (MMAR)*, Aug. 2017, pp. 1069–1074.

[44] M. I. Hasan, M. S. Ali, M. H. Rahman, and M. K. Islam, "Automated detection and characterization of colon cancer with deep convolutional neural networks," *J. Healthcare Eng.*, vol. 2022, pp. 1–12, Aug. 2022.

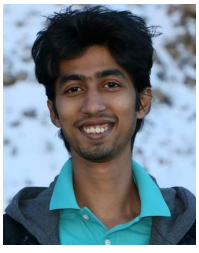
[45] F. Chollet, "Building powerful image classification models using very little data," *Keras Blog*, vol. 5, pp. 90–95, Jun. 2016.

[46] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," 2017, *arXiv:1712.04621*.
[47] M. M. Ahsan, M. R. Uddin, M. S. Ali, M. K. Islam, M. Farjana, A. N. Sakib, K. A. Momin, and S. A. Luna, "Deep transfer learning approaches for monkeypox disease diagnosis," *Expert Syst. Appl.*, vol. 216, Apr. 2023, Art. no. 119483.
[48] M. K. Islam, M. S. Ali, M. M. Ali, M. F. Haque, A. A. Das, M. M. Hossain, D. S. Duranta, and M. A. Rahman, "Melanoma skin lesions classification using deep convolutional neural network with transfer learning," in *Proc. 1st Int. Conf. Artif. Intell. Data Anal. (CAIDA)*, Apr. 2021, pp. 48–53.
[49] M. M. Ahsan, Y. Li, J. Zhang, M. T. Ahad, and M. M. S. Yazdan, "Face recognition in an unconstrained and real-time environment using novel BMC-LBPH methods incorporates with DJI vision sensor," *J. Sensor Actuator Netw.*, vol. 9, no. 4, p. 54, Nov. 2020.
[50] M. M. Ahsan, Y. Li, J. Zhang, M. T. Ahad, and K. D. Gupta, "Evaluating the performance of eigenface, fisherface, and local binary pattern histogram-based facial recognition methods under various weather conditions," *Technologies*, vol. 9, no. 2, p. 31, Apr. 2021.
[51] N. A. Baghdadi, A. Malki, S. F. Abdelaliem, H. M. Balaha, M. Badawy, and M. Elhosseini, "An automated diagnosis and classification of COVID-19 from chest CT images using a transfer learning-based convolutional neural network," *Comput. Biol. Med.*, vol. 144, May 2022, Art. no. 105383.
[52] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, pp. 1–21, 2015.
[53] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.
[54] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," *ACM Comput. Surv.*, vol. 54, no. 10s, pp. 1–41, Jan. 2022.
[55] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
[56] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Anal. Appl.*, vol. 24, no. 3, pp. 1207–1220, Aug. 2021.
[57] K. Jankowsky and U. Schroeders, "Validation and generalizability of machine learning prediction models on attrition in longitudinal studies," *Int. J. Behav. Develop.*, vol. 46, no. 2, pp. 169–176, Mar. 2022.
[58] S. Khedkar et al., "Explainable AI in healthcare," in *Proc. 2nd Int. Conf. Adv. Sci. Technol. (ICAST)*, Apr. 2019.
[59] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 1135–1144.
[60] C. Kaushal, M. K. Islam, S. A. Althubiti, F. Alenezi, and R. F. Mansour, "A framework for interactive medical image segmentation using optimized swarm intelligence with convolutional neural networks," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–21, Aug. 2022.
[61] M. A. Hussain, T. Islam, F. U. H. Chowdhury, and B. M. R. Islam, "Can artificial intelligence detect monkeypox from digital skin images?" *bioRxiv*, 2022.
[62] S. N. Ali, M. T. Ahmed, J. Paul, T. Jahan, S. M. S. Sani, N. Noor, and T. Hasan, "Monkeypox skin lesion detection using deep learning models: A feasibility study," 2022, *arXiv:2207.03342*.
[63] M. C. Irmak, T. Aydin, and M. Yağanoğlu, "Monkeypox skin lesion detection with MobileNetV2 and VGGNet models," in *Proc. Med. Technol. Congr. (TIPTEKNO)*, 2022, pp. 1–4.
[64] M. Altun, H. Gürüler, O. Özkaraca, F. Khan, J. Khan, and Y. Lee, "Monkeypox detection using CNN with transfer learning," *Sensors*, vol. 23, no. 4, p. 1783, Feb. 2023.

MD. MANJURUL AHSAN received the M.S. degree in industrial engineering from Lamar University, USA, in 2018, and the Ph.D. degree in industrial and systems engineering from The University of Oklahoma, Norman, OK, USA, in 2023. He is currently a Research Scholar with the Machine and Hybrid Intelligence Laboratory, School of Medicine, Northwestern University. His research interests include image processing, computer vision, deep learning, machine learning, and model optimization as applied to high-risk applications, such as medical imaging and diagnosis.

MD. SHAHIN ALI received the B.Sc. (Eng.) degree in biomedical engineering from Islamic University, Kushtia, Bangladesh. During the studies, he was a Research Assistant with the Biomedical Image Processing Laboratory. His academic performance throughout the B.Sc. studies was exceptional, earning him recognition and respect among his peers and professors. His research focus during the B.Sc. studies was on detecting various human diseases. His exceptional research contributions have earned him recognition as a crucial reviewer for several reputable journals in the same sector. His dedication and passion for the field of biomedical engineering have led him to enroll in the master's program in the same department, where he plans to expand his knowledge and expertise further. With his exceptional academic achievements, research contributions, and enthusiasm, he is well-positioned to make significant contributions to the field of biomedical engineering. His research interests include medical image processing, signal processing, computer vision, machine learning, deep learning, cognitive radio networks, and health informatics.

MD. MEHEDI HASSAN (Member, IEEE) received the B.Sc. degree in computer science and engineering from North Western University, Khulna, Bangladesh, in 2022, where he excelled in his studies and demonstrated a strong aptitude for research. He is currently pursuing the M.Sc. degree in computer science and engineering with Khulna University, Khulna. He is a dedicated and accomplished Researcher. As the Founder and the CEO of The Virtual BD IT Firm and the VRD Research Laboratory, Bangladesh, he has established himself as a highly respected leader in the fields of biomedical engineering, data science, and expert systems. He is a member of the prestigious Institute of Electrical and Electronics Engineers (IEEE). He is highly skilled in association rule mining, predictive analysis, machine learning, and data analysis, with a particular focus on the biomedical sciences. As a Young Researcher, he has published 28 articles in various international top journals and conferences, which is a remarkable achievement. His work has been well-received by the research community and has significantly contributed to the advancement of knowledge in his field. He is a highly motivated and a skilled researcher with a strong commitment to improving human health and well-being through cutting-edge scientific research. His accomplishments to date are impressive, and his potential for future contributions to his field is very promising. He has filed more than three patents out of which two are granted to his name. His research interests include broad and include important human diseases, such as oncology, cancer, hepatitis, human behavior analysis, and mental health. He serves as a reviewer for 20 prestigious journals.



TAREQUE ABU ABDULLAH received the M.S. degree in computer science from Lamar University, in 2022. He is currently pursuing the Ph.D. degree with the Department of Computer Science, Clark Atlanta University, in January 2023. His research interests include machine learning, computer vision, and medical imaging.

KISHOR DATTA GUPTA (Member, IEEE) is currently pursuing the Ph.D. degree with the Department of Computer Science, The University of Memphis. He is a Research Assistant with the Department of Computer Science, The University of Memphis. His research interests include image processing, computer vision, machine learning, deep learning, and big data analysis.

ULAS BAGCI is currently an Associate Professor (with tenure) with the Radiology and Biomedical Engineering Department and the Electronics and Computer Engineering Department, Northwestern University, Chicago. He was a Staff Scientist and the Laboratory Co-Manager with the Radiology and Imaging Sciences Department, Center for Infectious Disease Imaging, National Institutes of Health, and an Assistant Professor with the Center for Research in Computer Vision, UCF. He serves as a Steering Committee Member for Artificial Intelligence Resource (AIR), NIH. He teaches machine learning, advanced deep learning methods, computer and robot vision, and medical imaging courses. He holds several NIH grants (R01s, U01s, R15, and R03), as a Principal Investigator. He has more than 290 peer-reviewed articles on these topics. His research interests include artificial intelligence, machine learning and their applications in biomedical, and clinical imaging. He has several international and national recognition, including best paper and reviewer awards. He serves as the Area Chair for MICCAI, for several years. He is an Associate Editor of top-tier journals in the fields, such as IEEE TRANSACTIONS ON MEDICAL IMAGING, *Journal of Medical Physics*, and *Medical Image Analysis*.

CHETNA KAUSHAL received the B.Tech. degree in IT from Punjab Technical University, the M.Tech. degree in CSE from DAV University, Punjab, and the Ph.D. degree in CSE from Chitkara University, Punjab. She is currently an Assistant Professor with Chitkara University. She has published numerous research papers in various international/national journals, books, and conferences. She has filed more than 50 patents out of which 20 are granted to her name. She has around ten years' experience in research, training, and academics. She is an exceptionally motivated and talented researcher deeply dedicated to advancing human health and wellbeing through pioneering scientific investigations. Her remarkable achievements thus far are a testament to her capabilities, and her potential for making significant future contributions to her field is undeniably bright. Her research interests include machine learning, soft computing, pattern recognition, image processing, and artificial intelligence. She is a reviewer of many prestigious journals.

NAGLAA F. SOLIMAN received the B.Sc., M.Sc., and Ph.D. degrees from the Faculty of Engineering, Zagazig University, Egypt, in 1999, 2004, and 2011, respectively. She was with the Faculty of Computer Science, PNU, Saudi Arabia. Since 2015, she has been a Teaching Staff Member with the Department of Electronics and Communications Engineering, Faculty of Engineering, Zagazig University. Her research interests include digital image processing, information security, multimedia communications, medical image processing, optical signal processing, big data, and cloud computing.