



Detection of Monkeypox from skin lesion images using deep learning networks and explainable artificial intelligence

Tushar Nayak, Krishnaraj Chadaga, Niranjana Sampathila, Hilda Mayrose, G. Muralidhar Bairy, Srikanth Prabhu, Swathi S. Katta & Shashikiran Umakanth

To cite this article: Tushar Nayak, Krishnaraj Chadaga, Niranjana Sampathila, Hilda Mayrose, G. Muralidhar Bairy, Srikanth Prabhu, Swathi S. Katta & Shashikiran Umakanth (2023) Detection of Monkeypox from skin lesion images using deep learning networks and explainable artificial intelligence, *Applied Mathematics in Science and Engineering*, 31:1, 2225698, DOI: [10.1080/27690911.2023.2225698](https://doi.org/10.1080/27690911.2023.2225698)

To link to this article: <https://doi.org/10.1080/27690911.2023.2225698>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 27 Jun 2023.



Submit your article to this journal



Article views: 1755



View related articles



View Crossmark data



Citing articles: 2 View citing articles

Detection of Monkeypox from skin lesion images using deep learning networks and explainable artificial intelligence

Tushar Nayak^a, Krishnaraj Chadaga^b, Niranjana Sampathila^a, Hilda Mayrose^a, G. Muralidhar Bairy^a, Srikanth Prabhu^b, Swathi S. Katta^c and Shashikiran Umakanth^d

^aDepartment of Biomedical Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India; ^bDepartment of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India; ^cPrasanna School of Public Health, Manipal Academy of Higher Education, Manipal, India; ^dDepartment of Medicine, Dr. T.M.A. Pai Hospital, Manipal Academy of Higher Education, Udupi, India

ABSTRACT

Monkeypox (Mpox) resurfaced in January 2022 as a rare zoonotic disease that spreads to many countries. Though the virus is not as dangerous as COVID-19, it has still caused many fatalities worldwide. The Mpox virus spreads when people are in close contact with infected individuals. Among many symptoms, the disease also causes skin rashes, and medical imaging can be used to diagnose the virus successfully. However, other diseases such as smallpox, chickenpox, and measles also cause similar skin rashes. Hence, artificial intelligence (AI) and machine learning (ML) can be highly beneficial in diagnosing Mpox from other similar diseases. After extensive model training, it is advantageous to use a standard camera to capture skin images of an infected patient and run it against deep learning (DL) models. In this research, we have used transfer learning models such as residual networks and SqueezeNet to diagnose Mpox from measles, chickenpox and healthy patients. An average accuracy of 91.19% and an F1-score of 92.55% were obtained for the Mpox class. The findings show that the models can be useful in detecting the contagious virus. Since the classifiers are easily deployable, they can be used on camera-ready devices such as phones and laptops.

ARTICLE HISTORY

Received 10 February 2023

Accepted 10 June 2023

KEYWORDS

Monkeypox; chickenpox; measles; skin lesion; deep learning; explainable artificial intelligence (XAI)

MATHEMATICS SUBJECT CLASSIFICATION

68T09

1. Introduction

A multitude of reasons, including physical injury, overexposure to ultraviolet light, infections and tumours, can cause skin lesions [1]. Due to the rashes being on the skin, non-invasive techniques such as imaging can be used as a diagnostic modality for certain diseases that cause skin lesions using machine learning and deep learning [2]. These algorithms pick up unique features of skin lesions that might be difficult for humans to identify. The usage of deep learning-based classifiers has been very popular due to the increased

CONTACT Niranjana Sampathila  niranjana.s@manipal.edu  Department of Biomedical Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India; Srikanth Prabhu

 srikanth.prabhu@manipal.edu  Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India

computational power of computers and the public availability of data sets and information to train convolutional neural networks. Fatal diseases such as COVID-19 prompted the development of alternate diagnostic modalities due to the high transmission rate of the disease [3]. Common symptoms of COVID-19 include fever, cough, and fatigue, among others. These symptoms were similar to Influenza (flu), which further complicated and increased the number of people that needed to be tested. Digital images of X-Ray and Computer Tomography (CT) scans were used to non-invasively detect coronavirus patients showing symptoms of general respiratory diseases, alongside RT-PCR tests to identify the same infection using nose and throat swabs [4].

A similar problem looms at large over the ongoing Monkeypox (Mpox) outbreak of 2022. Mpox is an infectious viral zoonosis that was first detected nearly five decades ago in 1970 but has re-emerged recently [5]. The Mpox virus is a member of the Poxviridae family of viruses. A human infected by the virus will develop symptoms between five to twenty-one days of infection. Initial symptoms include muscle aches, headache along with swollen lymph nodes, shivering, fever, and skin lesions. The skin lesions (rashes) appear quickly after the infection, with the onset of the lesions taking about 3 days [6].

The current outbreak of the Mpox virus that began in January 2022 has infected nearly 83,000 people and caused the death of over 200 people in 111 countries [7]. A likely factor that has helped it not become the next major global pandemic such as COVID-19 is its direct contact-based transmission. Nevertheless, steps must be taken to ensure that the infected people know about their diagnosis at the earliest to curb the spread of the disease [8].

While the usage of polymerase chain reaction tests (PCR Tests) is the recommended modality of Monkeypox diagnosis, the infrastructure required in the testing of Monkeypox could be a hindering factor in rural areas [9]. They also take a considerable amount of time and are also prone to false negative results. Therefore, differential diagnosis can be used with the help of the skin lesions that form the basis of the work discussed below and our proposed modality. Table 1 compares the onset of skin lesion images in monkeypox, chickenpox and measles patients. Combining the differential diagnosis could be cost-effective and time-efficient.

Because of recent advancements in the semiconductor industry, deep learning-based classifiers are increasingly widespread. These networks have been successfully implemented in distinctly non-medical applications such as in [10] where deep neural networks have been successfully implemented in detecting fault diagnosis of the shaft of Brushless Direct Current Electric motors. The recent COVID-19 pandemic also saw the widespread usage of deep learning bases models that use CT-Scan images [11–13] and X-Ray images of the chest for a non-invasive detection of infection of SARS-CoV-2. These networks have also been used beyond the detection of infectious diseases such as the work detailed in [14] in which residual deep networks have been employed in the recognition of human

Table 1. Monkeypox skin lesions.

Disease	Stages	Development	Distribution
Monkeypox	Single stage	Slow	Face, palm, sole
Chickenpox	Multiple stage	Fast	Trunk
Measles	Multiple stage	Fast	Face then hands/feet



emotions from EEG signals. As stated in [15], deep learning has also been employed for the detection of glaucoma. Skin lesion images have also been used by deep learning-based models for non-invasively detecting disease [16,17].

The process of using deep convolutional neural networks for computer vision tasks involves using data sets used for the training and validation of a proposed model [18]. The usage of stacked residual blocks fits layers in the neural network using residual mapping [19]. Due to these factors, these residual neural networks (ResNets) scale their depth and are inherently much easier to optimize.

The usage of networks trained on competitive or large publicly available data sets with slight modifications to one's data set is known as transfer learning [20]. The main advantage of transfer learning is that the neural network weights have been pre-trained. This can then be further used on a smaller data set, helping improve the accuracy of the final proposed model.

In our proposed image classification modality, we have made use of deep residual networks and SqueezeNet to achieve reliable performance in a four-way classification of skin lesion images which include Mpox), chickenpox, measles and normal/healthy patients. Validation accuracies of all the deep residual networks consistently were above 90%. This automated diagnostic system can be used in healthcare settings to perform initial screening and ease the burden on medical personnel. The rest of the article is as follows. Section 2 reviews existing literature on Mpox diagnosis. The transfer learning methodology is described in Section 3. Section 4 discusses the results obtained by the deep learning classifiers. Conclusion and future directions are described in Section 5.

2. Related work

Various researchers have used deep learning to diagnose the Mpox virus. Ahsan et al. [21] used VGG-16 to diagnose Mpox from skin lesion images. They also created a data set called 'Monkeypox2022'. The accuracy, sensitivity, recall and f1-score obtained by the model were 97%, 97%, 97% and 97%, respectively. To interpret the results the local interpretable model explainer (LIME) was utilized. Abdelhamid et al. [22] used the AI-Biruni Earth Radius Optimization algorithm to diagnose Mpox. GoogLeNet was utilized for feature extraction. Various deep learning algorithms were used and a maximum accuracy of 98.8% was obtained. The sensitivity, recall and f1-score obtained were 62.5%, 99.8% and 76%, respectively.

A mobile application was developed to diagnose Mpox from skin lesion images [23]. Java and android were used to develop the application. A maximum accuracy of 91.11% was obtained. The sensitivity, recall and f1-score obtained were 85%, 94% and 89%, respectively. Islam et al. [24] used deep learning to diagnose the Mpox virus. The data set contained images of Mpox, chickenpox, smallpox, cowpox and measles. Seven different classifiers were utilized in this research. The accuracy, sensitivity, recall and f1-score were 83%, 85%, 94% and 89%. Sitaula et al. [25] used pre-trained DL classifiers to diagnose the Mpox. Eight different classifiers were used to distinguish the four classes. A maximum accuracy, sensitivity, recall, and f1-score of 87.13%, 85%, 85% and 85% were obtained. The various studies are summarized in Table 2.

Our proposed modality is based on the usage of deep residual convolutional neural networks which are trained on four classes: Monkeypox, Measles, Chickenpox and

Table 2. Comparison of existing work.

Classes	Train/validation test split	Accuracy	Sensitivity	Recall	F1-Score	
[21]	2	80/0/20	$97\% \pm 0.018$	$97\% \pm .018$	$97\% \pm 0.018$	$97\% \pm 0.018$
[22]	2	–	98.8%	62.5%	99.8%	76%
[23]	2	70/10/20	91.11%	90%	90%	90%
[24]	6	80/20	83%	85%	94%	89%
[25]	4	70/0/30	87.13%	85%	85%	85%

Healthy/Normal. On average, the 50 residual layers deep ResNet-18 performed the best at batch sizes 32 and 16, which our proposed modality is based on. Individual peak performance (validation accuracy) of the best-performing trial is 95.42%.

While not the diagnosis of monkeypox or other viral diseases using skin lesion images, the work presented in ref. [26] employs a ShuffleNet and EfficientNet to classify across seven diseases: Actinic Keratoses and Intraepithelial Carcinoma, Basal Cell Carcinoma, Benign Keratosis-like Lesions, Dermatofibroma, Melanoma, Melanocytic Nevi and Vascular Lesions that net a highest accuracy of 87.23% on the HAM10000 data set.

3. Methodology

As detailed in Figure 1, our classifiers were trained and validated on the four pre-trained convolution neural networks. The results were averaged and have been tabulated in Section 3. This was mainly to ensure that the obtained results are more trustworthy. The four models used were ResNet-18, ResNet-50, ResNet-101 and SqueezeNet.

3.1. Data set

The data set used in the training and validation of our proposed model is a four-class quaternary classification of the classes ‘Monkeypox’ ‘Chickenpox’ ‘Measles’ and ‘Normal’. The data set has been created using images collected from the internet by researchers at the Department of Computer Science and Engineering at Islamic University, Kushtia, Bangladesh. The data set is publicly available on the online community of data scientists and machine learning practitioners, Kaggle [27]. Samples of the images from the data set have been represented in Figure 2.

The images are all RGB images of resolution 224×224 . The four-class data set has 107 images of class ‘Chickenpox’, 91 images of class ‘Measles’, 279 images of class Monkeypox and 293 images of class ‘Normal’. The total size of the data set is 57 MB.

3.2. Deep residual learning (RESNETS)

Three deep convolutional neural networks of the residual network family available in the Deep Learning Toolbox: ResNet18, ResNet50 and ResNet101 were used for the transfer learning training and validation. Additionally, the lightweight network SqueezeNet was used to draw a baseline result due to its impressive performance in performance-limited devices.

A general assumption in the deep learning space that has been vastly exploited is the relationship between the depth of a neural network and accuracy. Many recent and

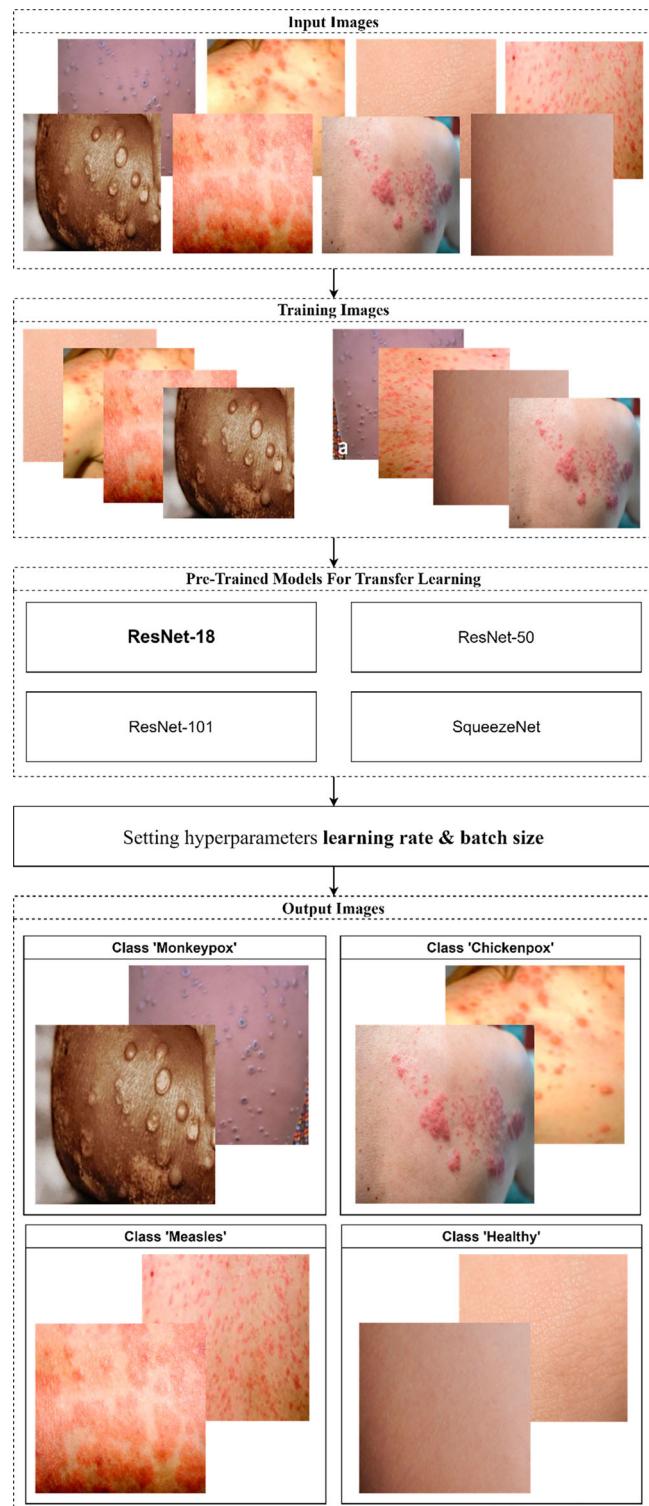


Figure 1. Methodology diagram.

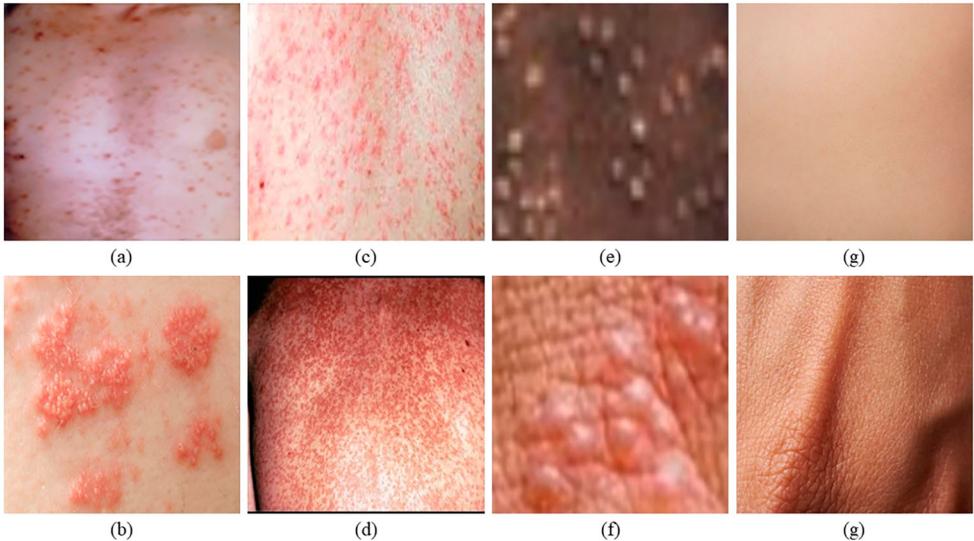


Figure 2. Samples from the data set for the class ‘Chickenpox’ (a, b), ‘Measles’ (c, d), ‘Monkeypox’ (e, f), ‘Healthy’ (g, h).

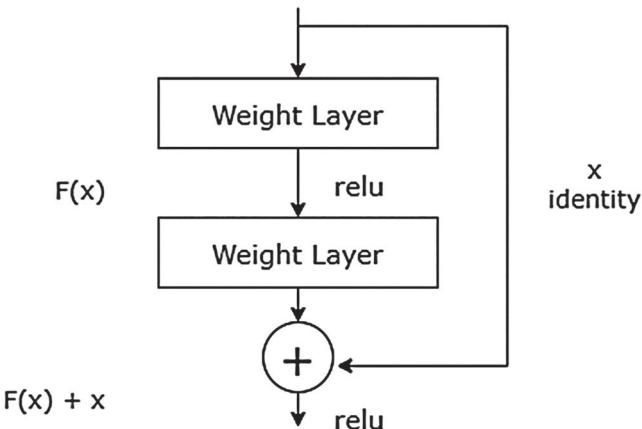


Figure 3. Residual network skip connection.

popular deep networks use 16–30-layer deep models for image classifications. However, in these *very deep* networks, the convergence can be limited by a problem called vanishing/exploiting gradients. To address this, normalized initialization and intermediate initialization are used to allow many layers to converge for stochastic gradient descent with backpropagation [28].

However, this causes a problem called *degradation* where the accuracy gets saturated and begins degrading with the increase in network depth. To combat *degradation* issues, ‘skip connections’ are employed. These ‘skip connections’ are a fundamental element of residual networks, and they work by skipping connections and causing different outputs at the given layer. This is described in Figure 3.

In networks without the skip connection, an input x can be multiplied by the layer's weights and the bias addition. Then, it goes through the activation function $f(x)$ which gives us the output $H(x)$ as represented in Equation 1 which can also be written as Equation 2.

$$H(x) = f(wx + b) \quad (1)$$

$$H(x) = f(x) \quad (2)$$

When the skip connections are introduced, the output transforms to Equation 3. But in this case, the pooling and convolutional layers are a hindrance surface where the input dimensions vary from the output dimensions.

$$H(x) = f(x) + x \quad (3)$$

To solve the problem, the dimensions can be matched by adding 1×1 convolution layers to the input. This method is called the projection method, and it is given by Equation 4. A second approach can also be taken by padding the skip connection with extra zero entries to increase its dimensions where the addition of the $w1$ can be avoided.

$$H(x) = f(x) + w1.x \quad (4)$$

Another problem occurring in deep convolution neural networks is the vanishing gradient problem. Here, the addition of more layers using activation function causes the gradient of the loss function to tend to zero. This in turn makes the neural network much more difficult to train. The skip connections in residual networks also help avoid this problem since these skip connections are an alternate path for the gradient. Additionally, the skip connections enable learning of the identified functions. This allows the neural network's higher layers to have performance comparable to the lower layers. Due to these characteristics, residual networks (ResNets) have been widely used in deep learning-based image classification modalities. The prevalence of ResNets in biomedical applications is also very vast, with it being used in many research works ranging from COVID-19 detection on X-Ray images with high accuracy to digital pathology deep learning-based classifiers [29,30]. It has also been used in the analysis of one-dimensional physiological signals [13]. During the experimentation phase of our proposed model, three variants of ResNet were used, each having a different number of layers: ResNet-18 (18 layers), ResNet-50 (50 layers) and ResNet-101 (101 layers). The architectures of ResNet-18, ResNet-50 and ResNet-101 have been detailed in Figure 4, Figure 5, and Figure 6, respectively.

Due to the ability of the Residual Convolutional Neural Networks (ResNets) to solve the vanishing gradient problem and scale without suffering any performance degradation, our proposed models have primarily been tested on Residual Networks of varying depths. To serve as a frame of comparison, SqueezeNet was also tested to compare their performance to that of the three variants of ResNet.

3.3. SqueezeNet

The network used to serve as a frame of reference for performance evaluation is SqueezeNet. It is a deep neural network for computer vision developed by researchers at the University of California, Berkely and Stanford along with DeepScale [31]. Figure 7

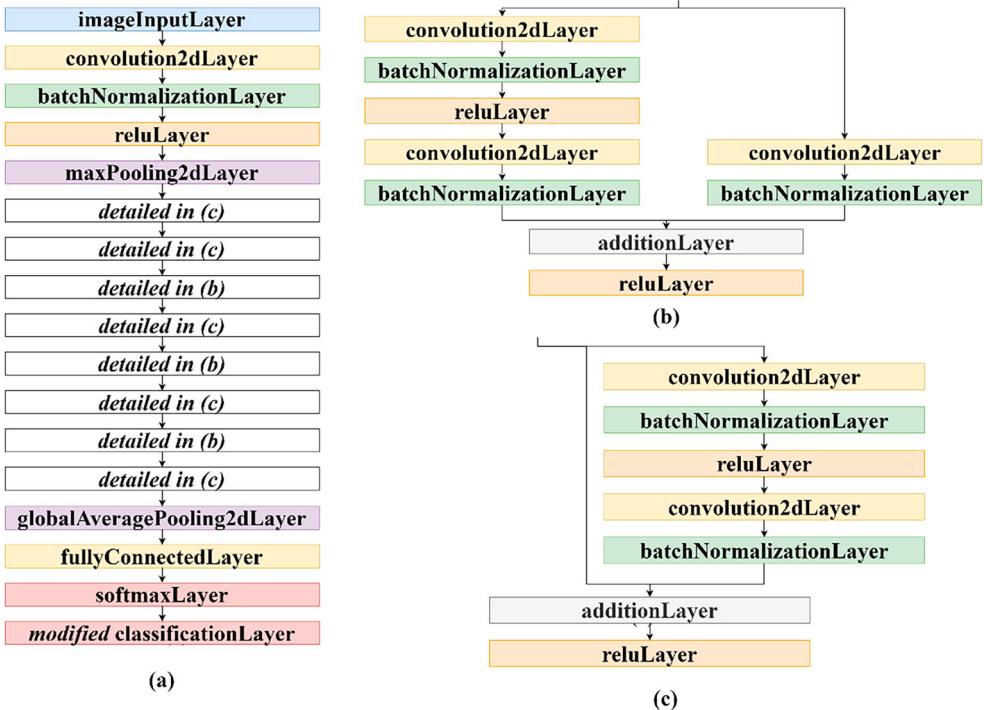


Figure 4. Architecture of ResNet-18: (a) main architecture, (b) and (c) repeating unit.

describes the architecture of SqueezeNet. It is based on making use of a collection of layers called the Fire Module which consists of 1×1 convolution filters that Squeeze, which after a ReLU block goes through 1×1 and 3×3 convolution filters. Nine of these fire blocks make up the core of the SqueezeNet architecture, along with the absence of the fully connected layer. The functionality of which is taken up by a last convolution 2D layer with the filter size set to that of the number of classes that the SqueezeNet-based deep learning classifier is being trained and validated on.

3.4. Experimental set-up

The entire experimentation has been conducted on a laptop running an Intel Core i7 10750H with 16 GB memory. A Nvidia GeForce GTX 1650 GPU with 4 GB VRAM had been set as the execution environment for the training and validation of the models. All the experimental trials were conducted on MathWorks MATLAB 2022b. For fair results, all the trials were conducted three times, with the average scores across all trials being reported in the results section below.

The Deep Learning Toolbox was installed on MATLAB 2022b, giving us access to install the pre-trained networks ResNet-18, ResNet-50, ResNet-101, SqueezeNet and GoogLeNet. Experiment Manager on MATLAB was used to run all the trials for the transfer learning experiment, with hyperparameters assigned, as demonstrated in Table 3. Due to the lower number of epochs used during training, the validation frequency was set to one. Each trial was repeated thrice, and the average result has been tabulated.



Figure 5. Architecture of ResNet-50: (a) main architecture, (b) and (c) repeating unit.

4. Results and discussion

4.1. Performance evaluation metrics

To evaluate the performance of the three residual neural networks and SqueezeNet, metrics such as confusion matrix, accuracy, precision, sensitivity and f1-score have been utilized.

- **Confusion Matrix:** For a quad classifier, a 4×4 confusion matrix is presented with true positives, true negatives, false positive and false negative values for each class. True positive values are the validation images that have been correctly identified and likewise, false positives and false negatives are validation images that have been incorrectly identified for a particular class. An ideal classifier must have a large number of true positive and true negative results. Figure 8 is the confusion matrix of a random trial. The cells shaded blue represent the true positive for the given class.

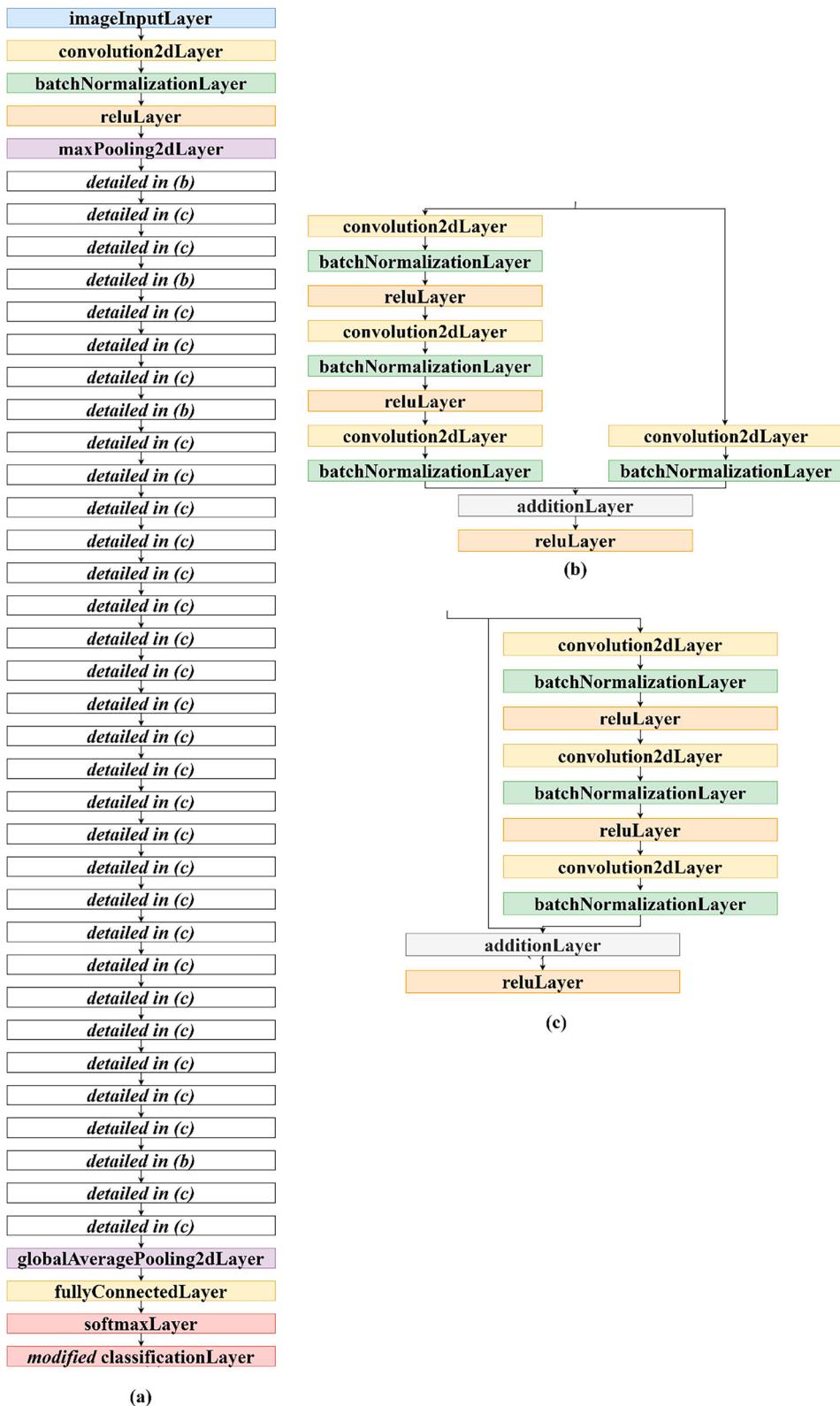


Figure 6. Architecture of ResNet-101: (a) main architecture, (b) and (c) repeating unit.

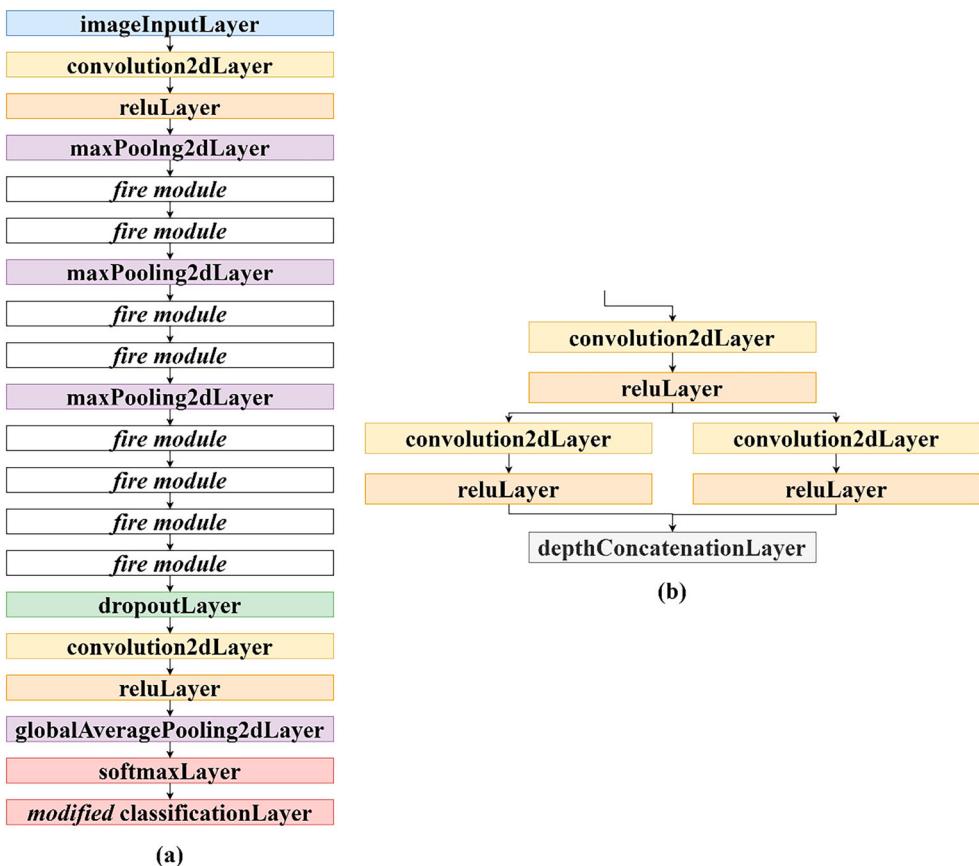


Figure 7. Architecture of SqueezeNet: (a) main architecture, (b) fire module.

Table 3. Hyperparameters for training and validation.

Hyperparameter	Values
Network name	['resnet18', 'resnet50', 'resnet101', 'squeezenet']
Mini batch size	[8, 16, 32, 64]
Learning rate	[0.001, 0.0001]
Epochs	[20]

- **Accuracy:** The accuracy of a classifier is the ratio of true positives to true negatives across all classes. It is calculated using Equation 5,

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}} \quad (5)$$

- **Precision:** The precision of a classifier for a given class is the ratio of the true positives to the total positives identified by the classifier. This value is inversely proportional to the number of false positives for the given class. It is calculated using Equation 6,

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positive} + \text{False Positives}} \quad (6)$$

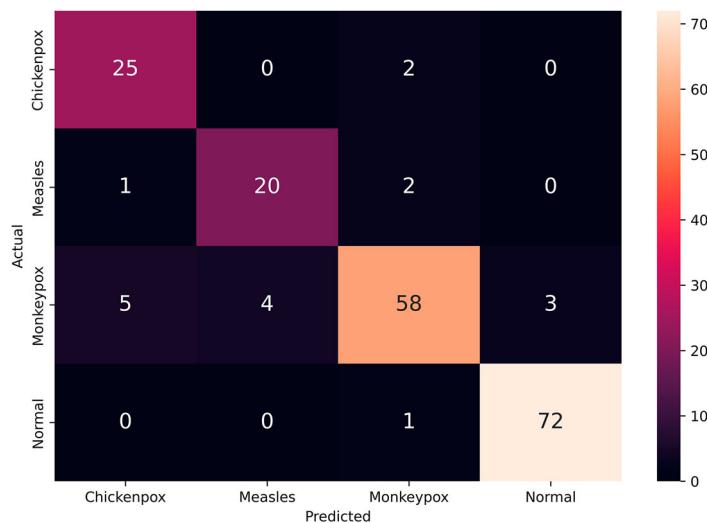


Figure 8. Confusion matrix for a random trial.

- **Sensitivity:** It is also known as recall and is calculated using Equation 7, the sensitivity of a classifier for a given class is the ratio of the true positives to the sum of the true positives and false negatives for the given class. This value is inversely proportional to the number of false negative cases for the given class.

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (7)$$

- **F1-Score:** The F1-Score for a given class considers both precision and recall for a more accurate evaluation of the performance of the classifier. It is calculated using Equation 8.

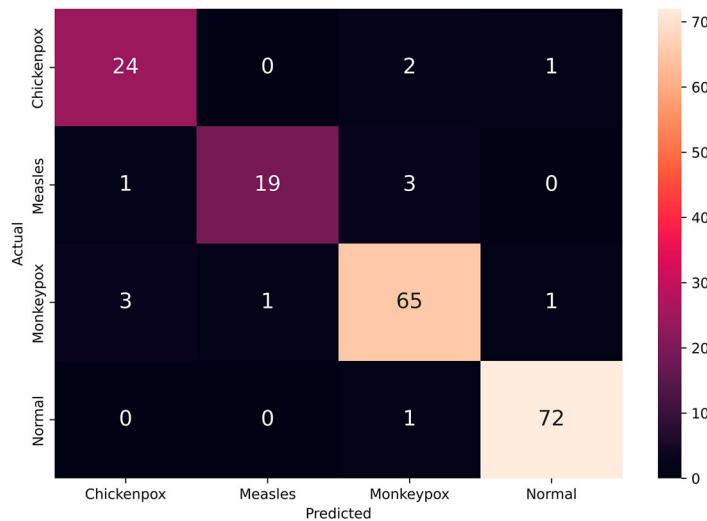
$$F1 - \text{Score} = 2 \frac{\text{Precision} * \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (8)$$

4.2. Model evaluation

In this research, four transfer models ResNet-18, ResNet-50, ResNet-101 and SqueezeNet were utilized to classify skin lesion images of patients diagnosed with Mpox, chickenpox and measles. The results obtained from the models are described in Table 4. Among all, the validation accuracy was the highest for ResNet-18 with 91%. The validation loss was 0.45. The batch size used and the learning rate were 16 and 0.001, respectively. ResNet-50 was the next best-performing model. It was able to obtain an accuracy of 91.01% and a validation loss of 0.47. The batch size and the learning rate used were 16 and 0.001, respectively. The ResNet-101 obtained a good accuracy and validation loss of 90.08% and 0.56, respectively. The residual network models outperformed the SqueezeNet model. The average accuracy obtained by the SqueezeNet model was 86.87%. The confusion matrix for the top-performing trial for ResNet18 is described in Figure 9.

Table 4. Experimental results for the transfer learning experiments.

Network	Batch size	Learning rate	Average training		average validation	
			Accuracy	Loss	Accuracy	Loss
SqueezeNet	8	0.0001	1.0	0.001733333	0.868739333	0.6443
ResNet-50	16	0.001	1.0	0.000566667	0.910189	0.4729
ResNet-18	16	0.001	1.0	7.26937E-05	0.911917	0.450033333
ResNet-101	32	0.001	1.0	0.000833333	0.908462667	0.561072333

**Figure 9.** Confusion matrix for top-performing trial ResNet-18.**Table 5.** Comparison of precision values for all classes.

	Chickenpox	Measles	Monkeypox	Healthy
SQUEEZENET	0.667	0.840667	0.909667	0.913333
RESNET50	0.839667	0.768333	0.933333	0.963667
RESNET18	0.778	0.782667	0.947667	0.968
RESNET101	0.788667	0.811667	0.938333	0.954667

Precision is an important metric in model classification. The precision values are described in Table 5. From the figure, it can be seen that Mpox and healthy patients' skin lesion images were classified better. The models were not able to accurately classify chickenpox cases. For Mpox cases, ResNet-18 obtained the highest precision of 94.7%. The precision obtained by SqueezeNet, ResNet50 and ResNet101 are 90.9%, 93.33% and 93.8%.

Recall is an important machine learning metric that emphasizes on false negative cases. The recall was comparatively high for all four classes. ResNet-18 obtained the highest recall of 90.4% for Mpox classification. The recall obtained by SqueezeNet, ResNet50 and ResNet101 are 88.5%, 88.06% and 87.2%, respectively. The recall values are summarized in Table 6.

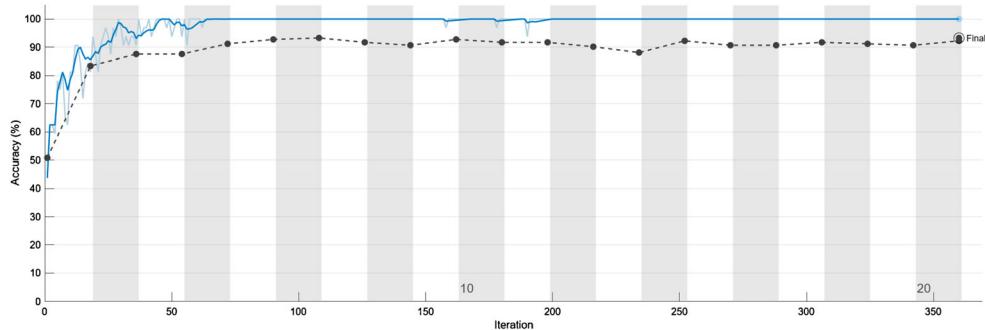
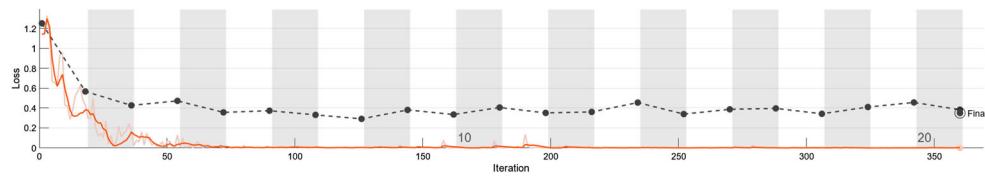
The F1-score obtained by the four models is described in Table 7. Among all the algorithms, ResNet-18 obtained the highest f1-score of 92.54% for Mpox classification. The

Table 6. Comparison of Recall values for all classes.

	Chickenpox	Measles	Monkeypox	Healthy
SQUEEZENET	0.739333	0.814	0.885	0.913
RESNET50	0.871333	0.891	0.880667	0.96766667
RESNET18	0.888	0.904333	0.904333	0.933
RESNET101	0.910667	0.889	0.872667	0.95333333

Table 7. Comparison of F1-Scores for all classes.

	Chickenpox	Measles	Monkeypox	Healthy
SQUEEZENET	0.701306	0.827118	0.897164	0.913167
RESNET50	0.855207	0.825133	0.906235	0.965663
RESNET18	0.829369	0.839113	0.925493	0.950178
RESNET101	0.845287	0.848575	0.904309	0.954

**Figure 10.** Accuracy plots – training (blue) and validation (black) (top-performing trial ResNet-18).**Figure 11.** Loss plots – training (red) and validation (black) (top-performing trial ResNet-18).

F1-scores obtained by SqueezeNet, ResNet50 and ResNet101 are 89.7%, 92.5% and 90.4%, respectively.

Among all the four transfer learning models, ResNet-18 was the best-performing Mpox classifier in this study. All four models were able to obtain good results in classifying the Mpox virus. The accuracy and loss curve for the best-performing trial (ResNet-18) are described in Figures 10 and 11.

The residual deep networks tested here can be used to form a diagnostic modality that can run on moderately performing computers and smartphones to form a non-invasive diagnostic modality. A picture of a skin lesion can be captured and cropped using a camera. The cropped image can be used as an input for the deep learning model. The classifier then categorizes the image into one of the four classes: ‘Monkeypox’, ‘Chickenpox’, ‘Measles’ and ‘Healthy’. This diagnostic modality can be used to complement the existing PCR testing

modality. The high-performing deep learning model can be used for the initial screening of Mpox patients.

Our research was compared with other studies which use deep learning for Mpox diagnosis (multi-classification). In [24], six classes were present, and a maximum accuracy of 83% was obtained. The sensitivity, recall, and f1-score obtained were 85%, 94% and 89%. Researchers in [25] used a similar data set to the one used in this study. The maximum accuracy, sensitivity, recall and f1-score obtained by the model were 87.13%, 85.44%, 85.47% and 85.40%. Our model outperforms [25] with an accuracy, sensitivity, recall and f1-score of 91.11%, 94.26%, 90.43% and 92.55%. The network used is an ensemble of DenseNet-169 & Xception. While this could add to the increased robustness of the network, the resulting network nets lower performance evaluation metrics, and the usage of computationally heavy neural networks could hamper its ability to run efficiently on performance-limited hardware in rural areas. Using the Stochastic Gradient Descent with Moment solver for the training and validation of the ResNet-18-based model, coupled with a learning rate that is appropriate given the features of the data set and batch size fitting the size of the data set, our proposed model is able to outperform others works. These optimizations enable the single network ResNet-18-based deep learning model to perform on par with similar studies using ensemble convolution neural networks. The advantage in performance does not have an efficiency penalty, as the single ResNet-18 model is more computationally efficient than CNNs used in other studies.

4.3. Explainable artificial intelligence (XAI)

Due to the black box approach taken in deep learning classifier modalities, it is important to use visualization techniques to understand the features learned by the convolutional neural network during the training phase. Using these insights, further work can be conducted such as further optimization and alternate models to combat problems such as

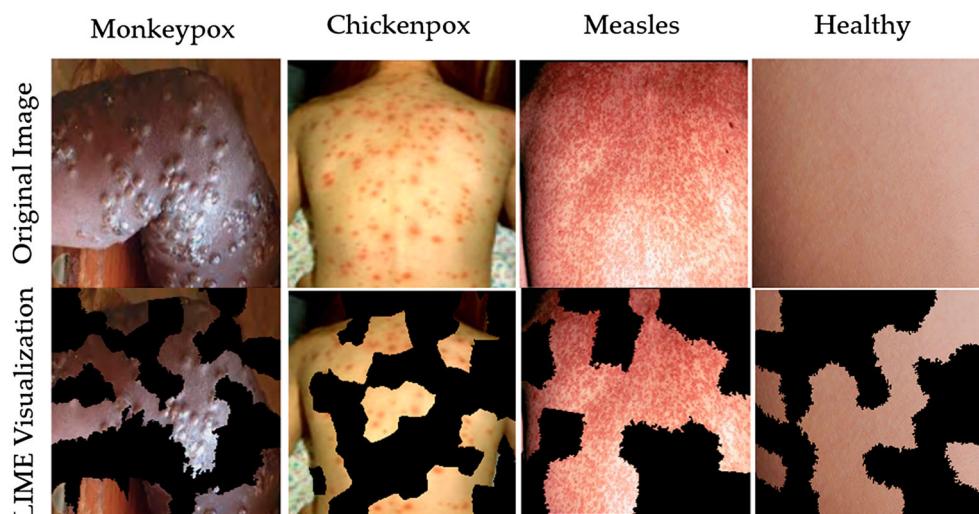


Figure 12. Training explanations for top-performing ResNet50 trial using LIME.

incorrect feature learning and overfitting. The visualization technique used; LIME (Local Interpretable Model-Agnostic Explanations) is heavily used for visualizing the predictions made by the convolutional neural network [32]. LIME works by taking the input image data and randomly perturbing it to observe the workings behind the predictions made.

The XAI visualization of the best-performing model (ResNet50) is shown in Figure 12 for the skin lesion data consisting of ‘Normal’, ‘Measles’, ‘Chickenpox’ and ‘Monkeypox’.

5. Conclusion and future scope

Mpox is a deadly disease which has spread to many countries recently. Mpox causes skin rashes which can be easily diagnosed using imaging. Hence, in this research, we use four transfer learning models (ResNet-50, ResNet-18, ResNet-10 and SqueezeNet) to diagnose Mpox from other similar diseases which cause skin rashes (measles and chickenpox). The data set was obtained from Kaggle which consisted of 107 chickenpox, 91 measles, 279 mpox and 293 healthy skin lesion images. All the models fared relatively well with ResNet-18 obtaining the best results with an accuracy, precision, recall and f1-score of 91.11%, 94.72%, 90.43% and 92.55%. LIME two techniques to explain the predictions made by the model have also been employed with the relevant portions of the image being highlighted either as segments or with a heatmap. The models can be deployed in real time to help healthcare workers with a fast and reliable diagnosis with minimal infrastructure costs due to the ability of the models to run efficiently on pre-existing hardware. The Explainable AI (XAI) techniques can also aid medical health professionals confirm the diagnosis made by the classifier by better understanding the choices made during the prediction.

Various other deep learning algorithms can be used in the future. Medical validation can be performed by doctors to verify the efficacy of the model. More data sets from various countries can be collected since DL models thrive on data. The models can also be deployed on various cloud platforms for easy accessibility and scalability. Advents in the deep learning space such as attention mechanisms and robust ensemble networks can also be implemented on the data set to increase the performance of the deep learning-based classifier.

All authors have read and agreed to the published version of the manuscript. Please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- [1] Ortiz-Martínez Y, Rodríguez-Morales AJ, Franco-Paredes C, et al. Monkeypox—a description of the clinical progression of skin lesions: a case report from Colorado, USA. Ther Adv Infect Dis. 2022 Jul;9:20499361221117726. DOI: [10.1177/20499361221117726](https://doi.org/10.1177/20499361221117726).
- [2] Benyahia S, Meftah B, Lézoray O. Multi-features extraction based on deep learning for skin lesion classification. Tissue Cell. 2022 Feb 1;74:101701. doi:[10.1016/j.tice.2021.101701](https://doi.org/10.1016/j.tice.2021.101701)
- [3] Chemnad K, Alshakhsi S, Almourad MB, et al. Smartphone usage before and during COVID-19: a comparative study based on objective recording of usage data. Informatics. 2022;9:98. doi:[10.3390/informatics9040098](https://doi.org/10.3390/informatics9040098)

- [4] Khanna VV, Chadaga K, Sampathila N, et al. Diagnosing COVID-19 using artificial intelligence: a comprehensive review. *Network Model Anal Health Informat Bioinformat.* **2022 Dec**;11(1):1–23. doi:[10.1007/s13721-022-00367-1](https://doi.org/10.1007/s13721-022-00367-1)
- [5] Gessain A, Nakoune E, Yazdanpanah Y. Monkeypox. *N Engl J Med.* **2022 Nov 10**;387(19):1783–1793. doi:[10.1056/NEJMra2208860](https://doi.org/10.1056/NEJMra2208860)
- [6] Luo Q, Han J. Preparedness for a Monkeypox outbreak. *Infect Med.* **2022 Jul 19**. doi:[10.1016/j.imj.2022.07.001](https://doi.org/10.1016/j.imj.2022.07.001)
- [7] WHO. Monkeypox outbreak 2022. 2022. Available from: <https://www.who.int/news-room/fact-sheets/detail/monkeypox>.
- [8] Rizk JG, Lippi G, Henry BM, et al. Prevention and treatment of monkeypox. *Drugs.* **2022 Jun 28**:1–7. doi:[10.1007/s40265-022-01742-y](https://doi.org/10.1007/s40265-022-01742-y)
- [9] Fomenko A, Weibel S, Moezi H, et al. Assessing severe acute respiratory syndrome coronavirus 2 infectivity by reverse-transcription polymerase chain reaction: a systematic review and meta-analysis. *Rev Med Virol.* **2022 Apr 2**:e2342. doi:[10.1002/rmv.2342](https://doi.org/10.1002/rmv.2342)
- [10] Glowacz A. Thermographic fault diagnosis of shaft of BLDC motor. *Sensors.* **2022**;22(21):8537.
- [11] Irfan M, Iftikhar MA, Yasin S, et al. Role of hybrid deep neural networks (HDNNs), computed tomography, and chest X-rays for the detection of COVID-19. *Int J Environ Res Public Health.* **2021**;18(6):3056.
- [12] Almalki YE, Qayyum A, Irfan M, et al. A novel method for COVID-19 diagnosis using artificial intelligence in chest X-ray images. *Healthcare.* **2021, April**;9(5):522. MDPI.
- [13] Joshi AM, Nayak DR. MFL-Net: An efficient lightweight multi-scale feature learning CNN for COVID-19 diagnosis from CT images. *IEEE J Biomed Health Inform.* **2022**;26(11):5355–5363.
- [14] Nayak T, Chadaga K, Sampathila N, et al. Deep learning based detection of monkeypox virus using skin lesion images. *Med Novel Technol Devices.* **2023**;18:100243. doi:[10.1016/j.medntd.2023.100243](https://doi.org/10.1016/j.medntd.2023.100243)
- [15] An G, Akiba M, Omodaka K, et al. Hierarchical deep learning models using transfer learning for disease detection and classification based on small number of medical images. *Sci Rep.* **2021**;11(1):4250.
- [16] Hoang L, Lee SH, Lee EJ, et al. Multiclass skin lesion classification using a novel lightweight deep learning framework for smart healthcare. *Appl Sci.* **2022**;12(5):2677.
- [17] Glowacz A, Glowacz Z. Recognition of images of finger skin with application of histogram, image filtration and K-NN classifier. *Biocybernet Biomed Eng.* **2016**;36(1):95–101.
- [18] Naseer I, Akram S, Masood T, et al. Performance analysis of state-of-the-art CNN architectures for luna16. *Sensors.* **2022 Jun 11**;22(12):4426. doi:[10.3390/s22124426](https://doi.org/10.3390/s22124426)
- [19] Fan Z, Lin H, Li C, et al. Use of parallel ResNet for high-performance pavement crack detection and measurement. *Sustainability.* **2022 Feb 5**;14(3):1825. doi:[10.3390/su14031825](https://doi.org/10.3390/su14031825)
- [20] Chowdhury D, Das A, Dey A, et al. ABCAndroid: a cloud integrated android app for noninvasive early breast cancer detection using transfer learning. *Sensors.* **2022 Jan 22**;22(3):832. doi:[10.3390/s22030832](https://doi.org/10.3390/s22030832)
- [21] Ahsan MM, et al. Image data collection and implementation of deep learning-based model in detecting Monkeypox disease using modified VGG16. *arXiv preprint arXiv:2206.01862.* 2022. doi:[10.48550/arXiv.2206.01862](https://doi.org/10.48550/arXiv.2206.01862)
- [22] Abdelhamid AA, et al. Classification of Monkeypox images based on transfer learning and the Al-Biruni Earth radius optimization algorithm. *Mathematics.* **2022**;10(19):3614. doi:[10.3390/math10193614](https://doi.org/10.3390/math10193614)
- [23] Sahin VH, Oztel I, Yolcu Oztel G. Human Monkeypox classification from skin lesion images with deep pre-trained network using mobile application. *J Med Syst.* **2022 Nov**;46(11):1–0. doi:[10.1007/s10916-022-01863-7](https://doi.org/10.1007/s10916-022-01863-7)
- [24] Islam T, Hussain MA, Chowdhury FU, et al. Can artificial intelligence detect Monkeypox from digital skin images? *bioRxiv.* 2022 Jan 1. doi:[10.1101/2022.08.08.503193](https://doi.org/10.1101/2022.08.08.503193)
- [25] Sitaula C, Shahi TB. Monkeypox virus detection using pre-trained deep learning-based approaches. *J Med Syst.* **2022 Nov**;46(11):1–9. doi:[10.1007/s10916-022-01868-2](https://doi.org/10.1007/s10916-022-01868-2)
- [26] Gajera HK, Nayak DR, Zaveri MA. A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features. *Biomed Signal Process Control.* **2023**;79:104186.

- [27] Images dataset (MSID) a new multiclass skin-based image datatset for Monkeypox disease detection – Kaggle [cited December 1, 2022]. Available from: <https://www.kaggle.com/datasets/dipuiucse/monkeypoxskinimagedataset>.
- [28] Karthik R, Vaichole TS, Kulkarni SK, et al. Eff2Net: An efficient channel attention-based convolutional neural network for skin disease classification. *Biomed Signal Process Control*. 2022 Mar 1;73:103406. doi:[10.1016/j.bspc.2021.103406](https://doi.org/10.1016/j.bspc.2021.103406)
- [29] Almuhammadi WS, Agu E, King J, et al. OA-pain-sense: machine learning prediction of hip and knee osteoarthritis pain from IMU data. *Informatics*. 2022;9:97. doi:[10.3390/informatics9040097](https://doi.org/10.3390/informatics9040097)
- [30] Krishnadas P, Chadaga K, Sampathila N, et al. Classification of malaria using object detection models. *Informatics*. 2022;9:76. doi:[10.3390/informatics9040076](https://doi.org/10.3390/informatics9040076)
- [31] Clement D, Agu E, Obayemi J, et al. Breast cancer tumor classification using a Bag of deep multi-resolution convolutional features. In *Informatics*. 2022 Oct 28;9(4):91. MDPI. doi:[10.3390/informatics9040091](https://doi.org/10.3390/informatics9040091)
- [32] Ribeiro MT, Singh S, Guestrin C. “Why should I trust you?” Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016, August; pp. 1135–1144.