Monkeypox Detection from Skin Images using Multi-Layer CNN Model

Rapolu Sai Kumar, Swathi Chigurlapalli

Department of Computer Science and Engineering

Sree Dattha Group of Institutions, Hyderabad, Telangana, India.

ABSTARCT

The recent monkeypox outbreak has become a public health concern due to its rapid spread in more than 40 countries outside Africa. Clinical diagnosis of monkeypox in an early stage is challenging due to its similarity with chickenpox and measles. In cases where the confirmatory Polymerase Chain Reaction (PCR) tests are not readily available, computer-assisted detection of monkeypox lesions could be beneficial for surveillance and rapid identification of suspected cases. Deep learning methods have been found effective in the automated detection of skin lesions, provided that sufficient training examples are available. However, as of now, such datasets are not available for the monkeypox disease. This project designs transfer learning based modified VGG16 and Custom CNN algorithm to predict Monkeypox disease as this disease is not deadly but spreading very fastly. To deal with this disease for timely detection doctors can use this algorithm for detection. Normal SKIN and Monkeypox images are uses to train both algorithms. The existing VGG16 gives low accuracy, and the proposed custom CNN gives high accuracy.

Keywords: Monkeypox detection, Polymerase Chain Reaction (PCR), Skin images.

1. INTRODUCTION

As the world recovers from the COVID-19 pandemic, the recent multi-country outbreak of monkeypox has raised concerns in global communities. The World Health Organization (WHO) declared that the outbreak poses a moderate risk to global public health and has stopped short of declaring it a public health emergency. However, healthcare organizations such as World Health Network (WHN) expressed a heightened concern and highlighted the need for immediate and concerted global action against the disease. Monkeypox is a zoonotic disease from the genus Orthopox virus. It closely resembles chickenpox, measles, and smallpox regarding clinical features. The minor differences in the skin rash of these diseases, coupled with the relative rarity of monkeypox have made the early diagnosis of this condition very challenging for healthcare professionals. On the contrary, the confirmatory PCR test is also not widely available.

Although the case fatality ratio has been reported to be 3–6% for the recent outbreak, early detection of monkeypox, corresponding contact tracing, and isolation are essential to limit the community transmission of the virus. In this scenario, AI-based automated computer-aided systems may substantially limit its global spread. In recent years, the multi-faceted applications of deep learning (DL), particularly the variations of Convolutional Neural Networks (CNNs), have revolutionized different domains of medical science due to their superior learning capability. When trained with a large number of data, these deep networks can process images in different layers, automatically extracting salient features and learning to identify the optimal representations for specified tasks. However, the requirement for large amounts of data and time-consuming training with dedicated computational resources limits the applicability of DL-based frameworks [6]. While using accelerators (e.g., GPU, TPU) resolves the time and resource-related issues, the dataset-related concerns persist due to the difficulty of obtaining unbiased, homogeneous medical data. Data augmentation [7] is a well-known method of increasing the dataset size by generating additional samples through slight modifications of existing data. In case of scarcity of data, transfer learning [6] is also a commonly used technique. This method utilizes a CNN model pre-trained on a large dataset (e.g., ImageNet) and transfers its knowledge for context-specific learning on a different, comparatively smaller dataset. Currently, there is no publicly available monkeypox skin lesion dataset for developing automated detection algorithms. There are impediments considering privacy and validity concerns. Moreover, the high prevalence of monkeypox in the under-developed African regions may introduce a bias in the dataset because of very high inter-class similarity and intraclass variability. In this paper, we first introduce the "Monkeypox Skin Lesion Dataset (MSLD)"1, an openly accessible dataset containing web-scrapped images of different body parts (face, neck, hand, arm, leg) of patients with monkeypox and non-monkeypox (measles, chickenpox) cases. We also present a DL-based preliminary feasibility study leveraging transfer learning involving VGG16 [8], Hybrid50 [9] and InceptionV3 [10] architectures to explore the potential of deep learning models for the early detection of the monkeypox disease.

2. LITERATURE SURVEY

Sitaula, et al. proposed [1] developed and evaluated are based on Convolutional Neural Networks models and some ensembles composed of a combination of those models, obtaining automatic classification results between healthy, monkeypox and other skin damages, given a close skin tissue image. The results show a system accuracy greater than 93% when using a unique CNN model (VGG-19 and Hybrid50), and greater than 98% when using a CNN ensemble formed by Hybrid50, EffiscientNet-B0 and MobileNet-V2.

Joseph Paul Cohen, et al. [2] determined that there is a need to establish a dataset including images of patients with Monkeypox disease, which will enable many researchers and practitioners to immediately begin work on developing and proposing a unique AI-assisted strategy. Therefore, the author anticipate that the dataset will serve the same function and assist researchers and practitioners who are eagerly waiting to get access to the dataset in order to construct a model for diagnosing Monkeypox disease.

Luna-Perej´on, F. et al. [3] proposed the diagnosis procedure of the Monkeypox includes initial observations of the unusual characteristics of skin lesions present and the existing history of exposure. However, the definitive way to diagnose the virus is to test skin lesions using electron microscopy. In addition, the Monkeypox virus can be confirmed using polymerase chain reaction (PCR), which is currently being used extensively in diagnosing the COVID-19 patients.

Ali, et al. [4] proposed a binary classification of monkeypox and other skin diseases is performed using skin images taken by users. The authors tested several convolutional neural network models such as VGG16, Hybrid50, and Inception-V3. The dataset uses 102 monkeypox images and 126 images of other skin diseases, but it does not include images of healthy skin tissue. The best classifier has an accuracy greater than 82%.

Ahsan M, et al. [5] uses a custom dataset formed by the four classes "healthy", "measles", "chickenpox" and "monkeypox", containing 54, 17, 47 and 43 images for each class, respectively. Although the authors use a data augmentation process, the dataset has very few images for some classes, and it is quite unbalanced. However, the developed classifiers are trained only for two classes ("monkeypox" versus "others", and "monkeypox" versus "chickenpox"). Using a VGG-16 CNN for each implemented system, authors obtain a 83% accuracy for the "monkeypox" vs "others" study for the training subset, and a 78% accuracy for the second experiment using the training subset.

Taking these potentials into account, we determined that there is a need to establish a dataset including images of patients with Monkeypox disease, which will enable many researchers and practitioners to immediately begin work on developing and proposing a unique AI-assisted strategy.

Md Manjurul Ahsan, et al. [7] Local interpretable model-agonistic explanations (LIME). It is one of the powerful tools that can help to analyse the model's true prediction and offer the opportunity to understand the Blackbox behind any CNN model's final predictions. LIME's impressive performance in describing the complexities of image classification has led to its extensive application in recent years. In the case of image classification, LIME uses super pixel. When an image is over-segmented, super pixels are produced. Super pixels store much data and help to identify essential features of the images during the primary prediction.

Tim Menzies, et al. [8] aims to address the ongoing data scarcity related to Monkeypox virus-infected patient images. The dataset is developed by collecting the images from open source and is publicly available to use without any privacy restrictions, ultimately allowing individuals to share and use that data for experiments and even for commercial purposes. Finally, they have used LIME to present the proper explanation of the reason behind our model's prediction, which is one of the current demands in deploying ML models for clinical trials.

Linda, et al. [9] proposed a low complex convolutional neural network (CNN) to detect skin diseases such as Psoriasis, Melanoma, Lupus, and chickenpox. Many experts in the medical domain believes that artificial intelligence (AI) systems could reduce the burden on clinical diagnosis with the outbreaks by processing image data.

Pan Pan, et al. proposed [10] the LIME parameters that have been used in this study to calculate the superpixel values. Note that the parameters are proven to be useful in many images' prediction analyses, as referred to in many existing literatures. Some of the works have been referenced previously, but other works include monitoring physiological signals, biomechanical gait studies, fall detectors, even aid systems for the diagnosis of cancer and other diseases.

Zhang, et al. [11] the main problem of this dataset is the image classes. It has only two classes, and, moreover, there is no "Healthy" class, and this absence can lead to failures when evaluating the classifier with healthy tissue images; since there is no class containing it, any image that it classifies is labelled as damaged tissueFor correct training of the classifier, it would be necessary to discard a significant part of the images in these datasets due to deficiencies such as those indicated, or even cases where two different images overlap.

Thomas, et al. [12] proposed for both datasets, the images used do not follow a similarity pattern, including full-body images or images of single body parts (even with images in which more than one person appears). For correct training of the classifier, it would be necessary to discard a significant part of the images in these datasets due to deficiencies such as those indicated, or even cases where two different images overlap.

3. PROPOSED SYSTEM

A Convolutional Neural Network is a type of artificial neural network that is primarily used for image recognition applications. It consists of multiple layers called perceptron's for learning features present in images with utmost detail. "Convolution" here represents the understanding of features within an image. Extracting these features entails the entire process of convolution. To extract these features, we use filters and mention the number of filters to be used (Kernels). Pooling is another concept in CNN. It is used to reduce the number of features that we get after we have run our filters over the images.

The filters cause us to get many features compared to the image itself. So pooling is used to get a better generalized representation of the same. For example, a 25x25 image might end up with 100x100 features due to the number of filters (kernels) and their shape. Pooling can reduce these features either using max pooling or mean pooling. Padding is another CNN concept where we add zeros to the edges of an image.

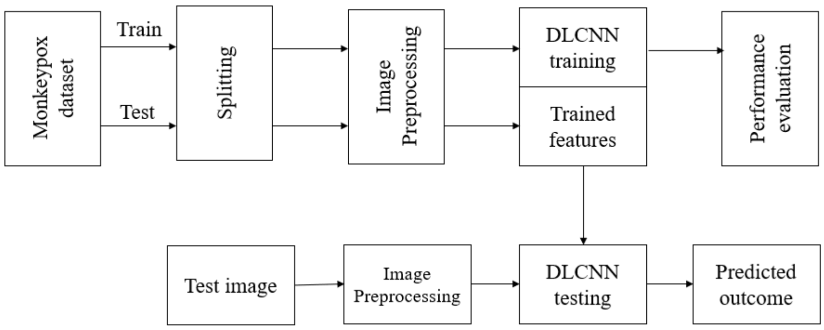


Fig. 1: Proposed Method.

Figure 1 shows the block diagram of proposed method. To better read the edge features and to get similar output as the input image. Lastly the output of the CNN is flattened and sent into a fully connected layer. This is our Neural Network part of CNN. It will enable the machine to learn from the extracted features and create a generalized model. Convolutional Neural Networks are Deep Learning algorithms that process images, assign importance to objects in the image using learnable weights and biases, and can differentiate images from each other. A convolutional neural network is essentially a neural network that uses a convolution layer and pooling layer. The convolutional layer convolves into a smaller area to extract features, while the pooling layer picks the data with the highest value within an area. They require less pre-processing in comparison to other classification algorithms and are able to learn filters and characteristics. The architecture of convolutional neural networks was based on the organization of the visual cortex. They use computer vision, natural language processing, and recommender systems to perform generative and descriptive tasks.

3.1 Data collection

In the time of the rapidly emerging Monkeypox disease among many nations, it is imperative to diagnosis patients with Monkeypox symptoms. Many experts in the medical domain believes that artificial intelligence (AI) systems could reduce the burden on clinical diagnosis with the outbreaks by processing image data [13]. During the onset of COVID-19, we observed that, hospitals in China and Italy, deployed AI-based and image processing-based interpreters to improve the hospitals efficiency in handling COVID-19 patients [14, 16, 24]. However, at the time of writing, unable to find any publicly available Monkeypox dataset hinders taking advantage of deploying an AI-based approach to diagnose and prevent the Monkeypox disease efficiently. As an effect, many researchers and practitioners cannot contribute to detecting Monkeypox disease using advanced AI techniques. Considering these limitations, we collected patients' images with Monkeypox images in this work. Our initial dataset contains very limited samples which will not be an issue for the initial experimentation, as supported by many referenced literatures that previously considered limited dataset in developing AI-based model during the onset of COVID-19 diseases. However, the database will be regularly updated with data contributed by numerous global entities. We followed following procedure to collect the data samples.

1. As there is no established shared dataset is available by the authorized and designated hospital, clinic, or viable source, therefore, to establish a preliminary dataset, the Monkeypox image data is collected from various sources such as websites, newspapers, and online portals and publicly shared samples. To do so, the google search engine is used for the initial searching procedure. Figure 1 displays the procedure used to search the data.
2. To develop the non-Monkeypox samples, a similar procedure is used in collecting the data sample, which contains search terms such as “Chickenpox,” “Measles,” and normal images (i.e., photos of both hands, legs, and faces) without any symptoms of the designated disease.
3. To increase the data sample size, additional Normal images are collected manually from various participants with their consent who do not have any skin disease symptoms. A consent form is used to get approval from all the participants.

3.2 Image pre-processing

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subfield of digital signal processing, digital image processing has many advantages over analogue image processing. It allows a much wider range of algorithms to be applied to the input data — the aim of digital image processing is to improve the image data (features) by suppressing unwanted distortions and/or enhancement of some important image features so that our AI-Computer Vision models can benefit from this improved data to work on. To train a network and make predictions on new data, our images must match the input size of the network. If we need to adjust the size of images to match the network, then we can rescale or crop data to the required size.

we can effectively increase the amount of training data by applying randomized augmentation to data. Augmentation also enables to train networks to be invariant to distortions in image data. For example, we can add randomized rotations to input images so that a network is invariant to the presence of rotation in input images. An augmented Image Datastore provides a convenient way to apply a limited set of augmentations to 2-D images for classification problems.

we can store image data as a numeric array, an ImageDatastore object, or a table. An ImageDatastore enables to import data in batches from image collections that are too large to fit in memory. we can use an augmented image datastore or a resized 4-D array for training, prediction, and classification. We can use a resized 3-D array for prediction and classification only.

There are two ways to resize image data to match the input size of a network. Rescaling multiplies the height and width of the image by a scaling factor. If the scaling factor is not identical in the vertical and horizontal directions, then rescaling changes the spatial extents of the pixels and the aspect ratio.

Cropping extracts a subregion of the image and preserves the spatial extent of each pixel. We can crop images from the center or from random positions in the image. An image is nothing more than a two-dimensional array of numbers (or pixels) ranging between 0 and 255. It is defined by the mathematical function f(x,y) where x and y are the two co-ordinates horizontally and vertically.

Resize image: In this step-in order to visualize the change, we are going to create two functions to display the images the first being a one to display one image and the second for two images. After that, we then create a function called processing that just receives the images as a parameter.

Need of resize image during the pre-processing phase, some images captured by a camera and fed to our AI algorithm vary in size, therefore, we should establish a base size for all images fed into our AI algorithms.

ISSN- 2394-5125 VOL 08, ISSUE 05, 2021

3.4 Proposed Hybrid-ML-CNN

Deep neural network is gradually applied to the identification of Monkeypox diseases and insect pests. Deep neural network is designed by imitating the structure of biological neural network, an artificial neural network to imitate the brain, using learnable parameters to replace the links between neurons. Convolutional neural network is one of the most widely used deep neural network structures, which is a branch of feed forward neural network. The success of AlexNet network model also confirms the importance of convolutional neural network model. Since then, convolutional neural networks have developed vigorously and have been widely used in financial supervision, text and speech recognition, smart home, medical diagnosis, and other fields.

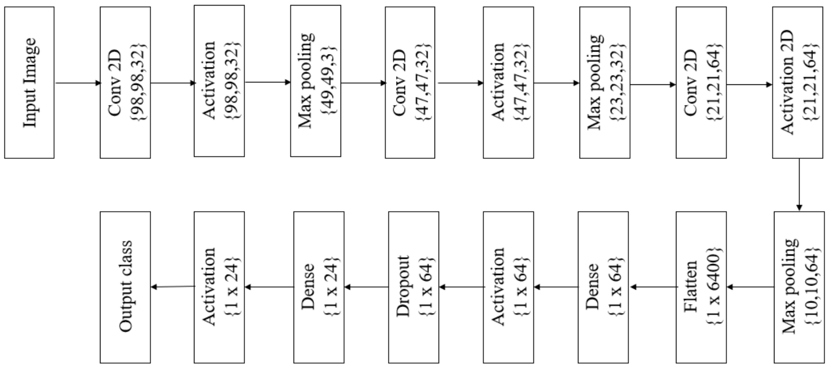


Fig. 2: Proposed Hybrid-ML-CNN.

Table.1: Layers description.

Layer Name	No. of filters	Feature size	Parameters
Conv 2D	32	98 x 98	896
Activation	32	98 x 98	0
Max pooling 2D	32	49 x 49	0
Conv 2D	32	47 x 47	9248
Activation	32	47 x 47	0
Max pooling 2D	32	23 x 23	0
Conv 2D	64	21 x 21	18496
Activation	64	21 x 21	0

ISSN- 2394-5125 VOL 08, ISSUE 05, 2021

Max pooling 2D	64	10 x 10	0
Flatten		1 x 6400	0
Dense		1 x 64	409664
Activation		1 x 64	0
Dropout		1 x 64	0
Dense		1 x 24	1560
Activation		1 x 24	0

Convolutional neural networks are generally composed of three parts. Convolution layer for feature extraction. The convergence layer, also known as the pooling layer, is mainly used for feature selection. The number of parameters is reduced by reducing the number of features. The full connection layer carries out the summary and output of the characteristics. A convolution layer is consisting of a convolution process and a nonlinear activation function ReLU. A typical architecture of CNN model for Monkeypox disease recognition is shown in Figure 2.

The leftmost image is the input layer, which the computer understands as the input of several matrices. Next is the convolution layer, the activation function of which uses ReLU. The pooling layer has no activation function. The combination of convolution and pooling layers can be constructed many times. The combination of convolution layer and convolution layer or convolution layer and pool layer can be very flexibly, which is not limited when constructing the model. But the most common CNN is a combination of several convolution layers and pooling layers. Finally, there is a full connection layer, which acts as a classifier and maps the learned feature representation to the sample label space.

Convolutional neural network mainly solves the following two problems.

1) Problem of too many parameters: It is assumed that the size of the input picture is $50 * 50 * 3$. If placed in a fully connected feedforward network, there are 7500 mutually independent links to the hidden layer. And each link also corresponds to its unique weight parameter. With the increase of the number of layers, the size of the parameters also increases significantly. On the one hand, it will easily lead to the occurrence of over-fitting phenomenon. On the other hand, the neural network is too complex, which will seriously affect the training efficiency. In convolutional neural networks, the parameter sharing mechanism makes the same parameters used in multiple functions of a model, and each element of the convolutional kernel will act on a specific position of each local input. The neural network only needs to learn a set of parameters and does not need to optimize learning for each parameter of each position.

2) Image stability: Image stability is the local invariant feature, which means that the natural image will not be affected by the scaling, translation, and rotation of the image size. Because in deep learning, data enhancement is generally needed to improve performance, and fully connected feedforward neural is difficult to ensure the local invariance of the image. This problem can be solved by convolution operation in convolutional neural network.

3.4 Hybrid-ML-CNN

According to the facts, training and testing of Hybrid-ML-CNN involves in allowing every source image via a succession of convolution layers by a kernel or filter, rectified linear unit (ReLU), max pooling, fully connected layer and utilize SoftMax layer with classification layer to categorize the objects with probabilistic values ranging from [0,1]. Figure 1 discloses the architecture of Hybrid-ML-CNN that is utilized in proposed methodology for CBIR system for enhanced feature representation of word image over conventional retrieval systems.

Convolution layer as depicted in Figure 4.3 is the primary layer to extract the features from a source image and maintains the relationship between pixels by learning the features of image by employing tiny blocks of source data. It's a mathematical function which considers two inputs like source image $I(x, y, d)$ where x and y denotes the spatial coordinates i.e., number of rows and columns. d is denoted as dimension of an image (here $d = 3$, since the source image is RGB) and a filter or kernel with similar size of input image and can be denoted as $F(k_x, k_y, d)$.

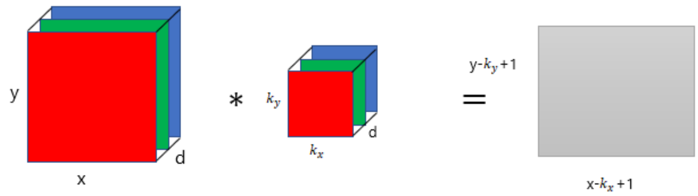


Fig. 3: Representation of convolution layer process.

The output obtained from convolution process of input image and filter has a size of $C\left((x - k_x + 1), (y - k_y + 1), 1\right)$, which is referred as feature map. An example of convolution procedure is demonstrated in Figure 5.2. Let us assume an input image with a size of 5×5 and the filter having the size of 3×3. The feature map of input image is obtained by multiplying the input image values with the filter values as given in Figure 3.

1	1	1	0	0
0	0	1	1	1
1	1	0	0	1
0	0	0	1	1
1	1	1	0	0

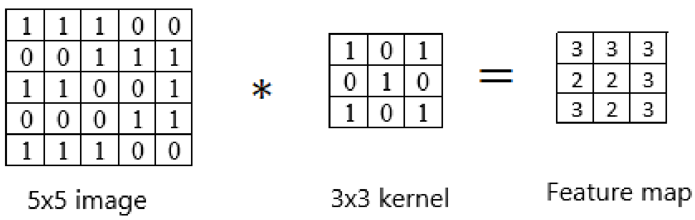
5x5 image

*

1	0	1
0	1	0
1	0	1

3x3 kernel

(a)



(b)

Fig. 4: Example of convolution layer process (a) an image with size $\mathbf{5} \times \mathbf{5}$ is convolving with $\mathbf{3} \times \mathbf{3}$ kernel (b) Convolved feature map

3.4.1 ReLU layer

Networks those utilizes the rectifier operation for the hidden layers are cited as rectified linear unit (ReLU). This ReLU function $\mathcal{G}(\cdot)$ is a simple computation that returns the value given as input directly if the value of input is greater than zero else returns zero. This can be represented as mathematically using the function $max(\cdot)$ over the set of 0 and the input x as follows:

$$\mathcal{G}(x) = \max\{0, x\}$$

3.4.2 Max pooing layer

This layer mitigates the number of parameters when there are larger size images. This can be called as subsampling or down sampling that mitigates the dimensionality of every feature map by preserving the important information. Max pooling considers the maximum element form the rectified feature map.

3.5 Softmax classifier

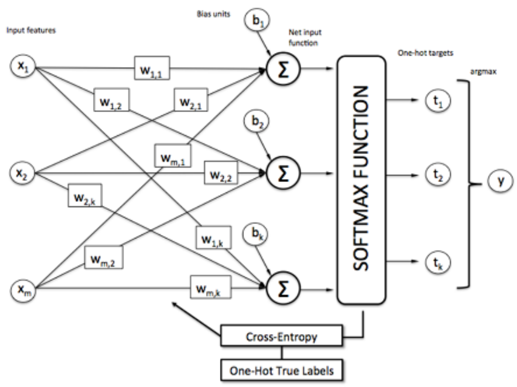


Fig.5: Monkeypox disease prediction using SoftMax classifier.

Generally, as seen in the above picture softmax function is added at the end of the output since it is the place where the nodes are meet finally and thus, they can be classified. Here, X is the input of all the models and the layers between X and Y are the hidden layers and the data is passed from X to all the layers and Received by Y. Suppose, we have 10 classes, and we predict for which class the given input belongs to. So, for this what we do is allot each class with a particular predicted output. Which means that we have 10 outputs corresponding to 10 different class and predict the class by the highest probability it has.

4. RESULTS

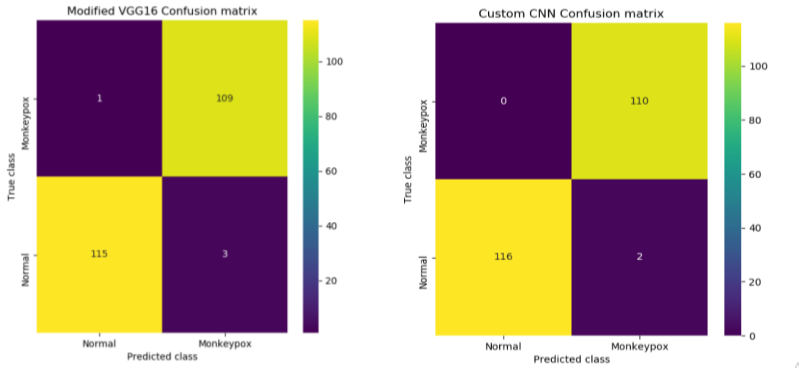


Fig.8: Confusion matrixes of VGG16 and CNN.

Table.1: Performance comparison.

Method	Accuracy	Precision	Recall	FSCORE
Modified VGG16	98.2	98.2	98.2	98.2
Proposed ML CNN	99.1	99.1	99.1	99.1

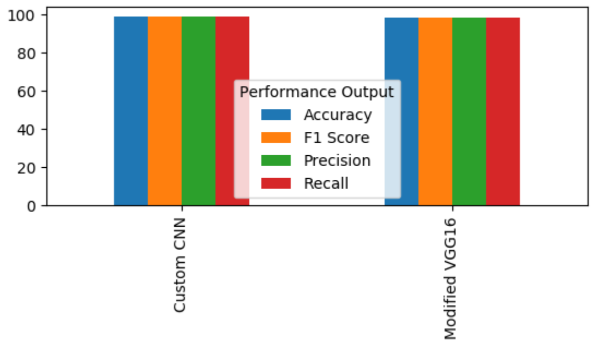


Fig.9: Graphical representation of performance metrics.

Fig. 10: Classified results as proposed method.

5. CONCLUSION

The study aims to address the ongoing data scarcity related to Monkeypox virus-infected patient images. The dataset is developed by collecting the images from open source and is publicly available to use without any privacy restrictions, ultimately allowing individuals to share and use that data for experiments and even for commercial purposes. Additionally, we conducted two studies considering small and moderate datasets wherein a modified ML-CNN model is implemented. Our findings suggest that using transfer learning approaches, the proposed modified ML-CNN can distinguish patients with Monekypox symptoms from others in both Study One and Two with accuracy ranging from 78% to 97%. Finally, we have used ML-CNN to present the proper explanation of the reason behind our model's prediction, which is one of the current demands in deploying ML models for clinical trials. Our model's predictions were crosschecked by doctors to emphasize that the results could be validated. We intend to emphasize the possibilities of artificial intelligence-based approaches, which might play an essential role in diagnosing and preventing the contamination of the onset of the Monkeypox virus. We hope our publicly available dataset will play an important role and provide the opportunity to the ML researcher who cannot develop an AI-based model and conduct the experiment due to data scarcity. As our proposed model is supported by many previously published literatures that uses the transfer learning approach in developing an AI-based diagnosis model, it will also encourage future research and practitioners to take advantage of the transfer learning approach and apply it in clinical diagnosis. Some of the constraints connected with our work can be overcome by updating the dataset by continuously collecting new Monkeypox infected patients' images, evaluating the proposed ML-CNN model's performance on highly imbalanced data, comparing the performance of our model with other researchers' findings (once available), and deploy our proposed model in developing mobile-based diagnosis tool.

REFERENCES

[1] Sitaula, Chiranjibi, and Tej Bahadur Shahi. "Monkeypox virus detection using pre-trained deep learning-based approaches." arXiv preprint arXiv:2209.04444 (2022).

[2] Joseph Paul Cohen, Paul Morrison, and Lan Dao. Covid-19 image data collection. arXiv preprint arXiv:2003.11597, 2020.

[3] Luna-Perej´on, F. et al. (2020). Low-power embedded system for gait classification using neural networks. Journal of Low Power Electronics and Applications, 10 , 14.

[4] Ali, S. N., Ahmed, M., Paul, J., Jahan, T., Sani, S., Noor, N., Hasan, T. et al. (2022). Monkeypox skin lesion detection using deep learning models: A feasibility study. ArXiv preprint arXiv:2207.03342.

[5] Ahsan, M. M., Uddin, M. R., & Luna, S. A. (2022b). Monkeypox image data collection. arXiv preprint arXiv:2206.01774.

[6] Phi-Yen Nguyen, Whenayon Simeon Ajisegiri, Valentina Costantino, Abrar A Chughtai, and C Raina MacIntyre. Reemergence of human monkeypox and declining population immunity in the context of urbanization, nigeria, 2017–2020. Emerging Infectious Diseases, 27(4):1007, 2021.

[7] Md Manjurul Ahsan, Redwan Nazim, Zahed Siddique, and Pedro Huebner. Detection of covid-19 patients from ct scan and chest x-ray data using modified mobilenetv2 and lime. In Healthcare, volume 9, page 1099. Multidisciplinary Digital Publishing Institute, 2021.

[8] Tim Menzies, Jeremy Greenwald, and Art Frank. Data mining static code attributes to learn defect predictors. IEEE transactions on software engineering, 33(1):2–13, 2006.

[9] Linda Wang, Zhong Qiu Lin, and Alexander Wong. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Scientific Reports, 10(1):1–12, 2020.

[10] Pan Pan, Yichao Li, Yongjiu Xiao, Bingchao Han, Longxiang Su, Mingliang Su, Yansheng Li, Siqi Zhang, Dapeng Jiang, Xia Chen, et al. Prognostic assessment of covid-19 in the intensive care unit by machine learning methods: model development and validation. Journal of medical Internet research, 22(11):e23128, 2020.

[11] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770–778).

[12] Thomas, S. M. et al. (2021). Interpretable deep learning systems for multiclass segmentation and classification of non-melanoma skin cancer. Medical Image Analysis, 68 , 101915.

[13] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014. [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE CVPR, 2016, pp. 770–778.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE CVPR, 2016, pp. 770–778.

[15] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," J. of Big Data, vol. 6, no. 1, pp. 1–48, 2019.

[16] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Classification of skin lesions using transfer learning and augmentation with Alex-net," PloSone, vol. 14, no. 5, p. e0217293, 2019.