International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys.14, No.1 (Sep-2023)

http://dx.doi.org/10.12785/ijcds/140160

Neural Network Based Approach to Diagnose and Classify Monkeypox Disease

Sakshi[1], Sachin Lodhi[2] and Vinay Kukreja[1]

[1]*Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India*
[2]*University Institute of Technology, Barkatullah University, Bhopal, Madhya Pradesh, India*

Received 18 Oct. 2022, Revised 22 May 2023, Accepted 31 Jul. 2023, Published 01 Sep. 2023

Abstract: Numerous organizations, including local governments, health, and medical institutions, and even the WHO, have expressed concern about the fatality and transmission rates of this illness due to the rapid increase in monkeypox cases around the world. Monkeypox illness and the COVID-19 virus have a very similar pattern of spread. The number of reported cases worldwide has increased in recent months, continuing the same trend as COVID-19. The characteristics and signs of the monkeypox virus are similar to those of any other viral illness. The main signs and symptoms are fever, chills, fatigue, headache, etc. At this juncture, the diagnosis is a significant challenge. Another challenge is posed by the disease when the obvious symptoms occur like rashes on the skin. The problem with these specific symptoms is that their appearance resembles other diseases like Chickenpox, Cowpox, and so on. Thus, the correct classification and proper diagnosis of the disease are tough and strenuous. Thus, for efficient and correct classification of this severe epidemic, a neural network-based solution is proposed, which classifies the disease at its initial stage with a competent accuracy rate of 94.38%. the proposed solution is excelling the diagnosis problem by performing efficient classification.

Keywords: health, healthy lives, improving mortality, measles, mortality, mortality rate, medical, monkeypox

1. Introduction

The monkeypox virus, which belongs to the genus Orthopoxvirus in the family Poxviridae, causes the disease. Since 1 January 2022, more than 31799 cases have been reported [1], [2]. Though this number may seem very small compared to the timespan, the possibility of unreported cases can not be ignored. Though the number of reported deaths is limited to only 12, as the probability of unreported cases exists, these numbers can not be considered conclusive. The data come from 90 countries, so the list is not exhaustive. The reported cases mostly come from developing and developed countries, bolstering the underreporting of the cases [3], [4].

In the previous outbreak of the novel Coronavirus or COVID-19, health facilities and the medical system were directly hit, and many people succumbed to the disease for various reasons. In the developed countries, the cases and deaths were reported with the best accuracy and for that reason, the countries with an open economy and democratic government seem to be worst hit by the COVID-19 virus [5], [6]. The reported cases of deaths till September 2, 2022, around the globe are depicted in figure 1. The scale in the bottom left corner of the graph depicts the varying magnitude of the reported deaths around the globe. The scale goes from dark green to dark red, which shows the transition from the no death cases reported to the highest death cases reported in the world. The Light green or yellow color represents a moderate number of such cases.

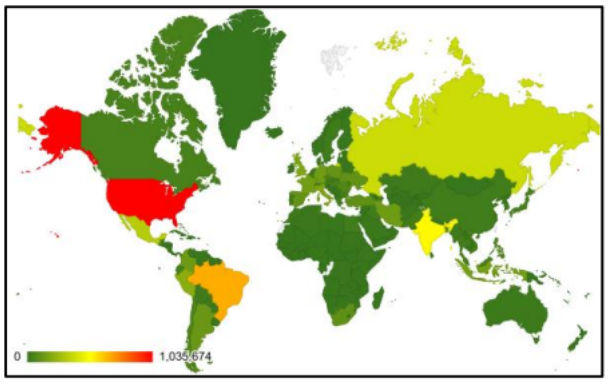


Figure 1. Worldwide reported deaths(COVID-19)

E-mail address: asakshi541@gmail.com, sachinlodhi8614@gmail.com, vinay.kukreja@chitkara.edu.in

From the graph (fig.1), it is quite obvious that the counties with more open and free media reported a higher number of casualties instead of having a lesser population than the countries with less open media. The worldwide cases (reported) of monkeypox are shown in fig. 2. The scale in the bottom left corner of the graph shows the changing number of reported confirmed cases around the world. The scale has a color range from dark green (leftmost) to red (rightmost) and yellowish color in between. The dark green color represents the least to no cases reported and the dark red color represents the most number of cases reported and the color in between both of them represents the moderate number of reported cases.

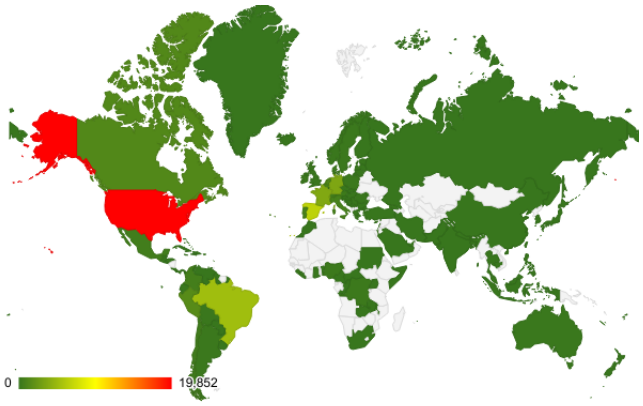


Figure 2. Worldwide reported confirmed cases (Monkeypox)

In fig.2 it can be observed that the more developed countries, in the initial phase of the outbreak, start reporting the most number of cases. This trend, in both graphs(fig1&2), can be attributed to various factors like better medical facilities, a better economy, and independent media houses. On the other hand, the countries with the least number of casualties or reported cases are either undeveloped nations or have a closed economy or state-controlled media [7].

The countries from the African continents are among the least impacted countries, according to the observation from both the graphs for both the diseases but in reality the poor medical and health infrastructure and weak economy can be the main factor [8]. In these nations, either people do not have access to the hospitals like the other developed nations or they do not report the case deliberately to the authorities [9].

Monkeypox spreads in various modes, some with direct contacts like skin-to-skin contact with the monkeypox rashes or body fluids. Another model by using objects of the infected person [10]. Respiratory secretion also can spread. Making intimate contact also can contribute to the spread of the virus. Pregnant women can also spread the virus through the placenta to their foetuses. The prior symptoms appear within three weeks of exposure to the virus. The first symptoms resemble the common flu-like headache, fever, chills, swollen lymph nodes, exhaustion, muscle aches, backache, and respiratory symptoms, e.g., sore throat, nasal congestion, and cough [11], [12]. These prior symptoms are followed by more disease-specific symptoms like rashes on the skin [13], [14]. Some people have these symptoms in swap order, and few may experience only rashes. But the general symptoms occur as discussed in the first case.

The main problem occurs in the phase of the appearance of the secondary symptoms, i.e. when the rashes occur on the skin. The appearance of the rashes is so close to other diseases like Chickenpox, cowpox, measles, smallpox, and so on [15], [16].

This raises the problem of the misdiagnosis of the disease if the diagnosis phase is performed manually. . Thus, the main concerns of the article are

- ***To tackle this problem of misclassification of the disease, based on the visual interpretation of the features present in the symptoms.***
- ***To extend and explore the robustness of neural network - based classification systems in the domain of health and disease detection***

The model classifies the input images to their respective and true class of the present disease in the image. Though there exists an entire class of models that are utilized in the medical and health field as the outbreak of the diseases is very recent there are not many such networks that can deal with such problems as discussed in the prior section. The model achieves an accuracy of 94.38% on the unseen dataset and hence ensures that classification and the diagnosis of the disease are done with the best accuracy percentage rate possible.

2. Motivation for the Work

- **To contribute to enhancing the guard a gainst the deadly virus:**

Though the monkeypox virus is not new and has been detected around 1958 in a lab in Denmark and classified as an illness, the sudden outbreak of this virus around the globe has raised the world's concern [1]. The worst experience witnessed by the world during various waves of the COVID-19 virus is the motivation to produce the piece of work. During COVID-19 entire world was unprepared and under-prepared at its best [17]. Because of this under-preparedness and weak and inefficient medical infrastructure, the world faced the heaviest loss of lives in the wave phase. The work has been produced with the vision of improving the system so the disease can be diagnosed at its early stages.

- **Resolution for the potential misdiagnosed disease:**

As monkeypox disease is a disease that is related to the skin and the most prominent and obvious symptoms are seen on the skin of the patient, there is the possibility that the diagnosis can be quite difficult as there exist many

other diseases like measles, cowpox, Chickenpox and so on which give the similar impression as Monkeypox gives on the skin [18]. The erroneous diagnosis at the primary stage can worsen the situation as the medications and treatment largely depend on the initial phases, like diagnosis and symptoms [19]. This work is proposed to classify the disease with the best accuracy to solve the proximity of the appearance of various diseases like Monkeypox.

- **Providing alternatives and more feasible options to perform diagnosis:**

In the studies, Monkeypox has been attributed to a disease with high virulence and transmissibility. In such a scenario, the traditional walk-in methods to get diagnosed may act as one of the triggering points to an outbreak of the virus in a particular region or sub-region [4], [20]. On the other hand, if the proposed model is mounted on mobile phone devices accessible to most people around the world, it is quite easy and efficient to diagnose the disease, and they can be tagged as infected, and their proper treatment can be suggested [21]. This does not help the patient suffer less because of all those extra precautionary steps required before going anywhere. Still, also it helps the government and higher authorities to tag and classify area(s) based on the number of patients.

- **Introducing self-diagnosis devices to prevent super spreading conditions.**

At the time of the outbreaks, the situation is most likely to be panic which means that the public is in a rush because of the uncertainty. The people prepare on a personal level by stocking the essential needs, getting the medicine, and in the condition of initial symptoms, the public rush to the hospital or medical facilities to get the proper diagnosis of the underlying medical condition [22]. In such cases, when the masses of the public are gathered around the places like hospitals, these places behave like hotspots and super-spreaders [23], [24]. The reason is that medical and health amenities are very few available facilities during the sudden outbreak. And when people rush to such places, they may spread the virus or get it from other infected people [25]. When people have a device that can diagnose and detect the conditions based on such classifications, the patient can order or get the medications prescribed by the doctor. This would save time as there is no need to go anywhere, and the treatment would be available on the go.

3. Research Methodology

The entire research method has been crafted with the best possible and optimal solutions to the sub-phases involved in the entire method. The first step is the collection of the dataset sample, followed by preprocessing the raw images. Then the outcome of this step is the final dataset split into two parts and loaded into the objects. Further phases are the construction of the model, followed by training and testing. The final phase of the whole methodology is analyzing and discussing the results over the results obtained. Every step in the research methodology has been selected and designed among all the possible candidates by considering the associated challenges and the advantages compared to the other steps to achieve the same task. From the collection of the dataset to the parameters over which the model's performance is discussed and every other step involved in between, every minor detail is considered, and other possible methods to achieve the corresponding goal. Consider the very first step of the data collection. There are various methods by employing which the goal can be achieved but the authors finalized the collection by the method of scraping and cleaning the dataset as the employed method (scraping) can accumulate a large amount of data (relevant and irrelevant) in a shorter amount of time compared to any other manual method. Similarly, for model selection and construction, the transfer learning method is employed to save training time. Though the model can be built from scratch and would take a greater amount of time to train such a model as compared to the model which is finalized using the transfer learning method.

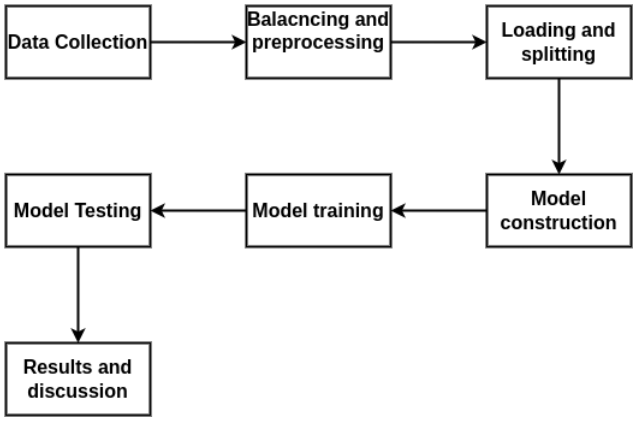


Figure 3. Research Methodology

4. Experimentation

The method to conduct research and find conclusive results have been divided into sub-phases, and each phase is decisive in its merit. By considering the gravity of the task, the sensitivity of the data, and the significance of the results and conclusions that the study will produce, the authors have given each phase time and intense attention. While considering every aspect and potential outcome of such steps, the authors carefully designed the workflow of the research proposed (see fig. 1). As the medical condition is currently getting pace and more and more data is being available, the authors started the process to search suitable data which is compatible and follows all the guidelines and ethos of the medical practices. The preprocessing phase wand loading then follows the data collection. The model is built in the next phase, and then training starts. The output of the training and testing phase is discussed in the end.

The following subsequent sections discuss different

measures that constitute the experimentation phase.

A. Data Collection

Since the global outbreak of Monkeypox in various countries and places, the data about the disease, for instance, the medical conditions and visuals of the symptoms, have been accumulated around the globe. The visual dataset may include images of various features of the symptoms [6]. Other statistical datasets can store information like the readings of the various values in the body of the patient. Though features are rich and have great variety, selecting the correct dataset poses a challenge [26]. Another problem with the process of the collection of the dataset is its authenticity. On the internet, the large number of data available also raises concerns about the authenticity of the dataset.

The third problem is about the consent of the data sharing by the patient. Generally, every country has a law that gives the patient the right to deny sharing data. Considering every such situation in mind, the authors performed the search for the best possible solution. They found the dataset, which is accumulated using web-scarping and PoC, is published on biorxiv.org. The authors consider it authentic as it follows every rule, and the paper has taken care of all the terms and conditions. The individuals involved in collecting data make sure to secure the patient's identity and to achieve that, they block the eyes of the patient. Few more similar measures have been carried out with the intent to preserve the privacy of a patient.

The collected dataset has image samples of Chickenpox, cowpox, measles, monkeypox, and smallpox. The dataset also contains images of healthy skin to differentiate the diseases from each other and that healthy skin. In that manner, six classes of the image samples are in the dataset, and the class names are represented by the folder names inside the main dataset directory. The number of samples in the various sub-directories varies. The subdirectories or classes named Monkeypox, Chickenpox, Smallpox, Cowpox, Measles, and Healthy have the images 1200, 1004, 1102, 935, 1045, and 1156 in order. In total, in the raw and original dataset, there are 6442 image samples collectively. This skewed structure of the main dataset may produce issues and hurt the performance of the model to a great extent by adding unintended bias in the model because of the imbalance in the dataset that is going to be used as training and validation of the model. These issues are addressed and discussed in-depth in the following sections. A few of the samples are shown, in figure 2.

B. Experimentation setup

Experimentation in its entirety has been performed on the instance of the Google Colab environment. Google Colab is a SaaS utility provided by Google which runs the Jupyter notebook instance, and hardware resources are provided in varying capacities according to the selected plan. Google Colab provides two plans: Google Colab paid and Google Colab Free. The authors chose the second

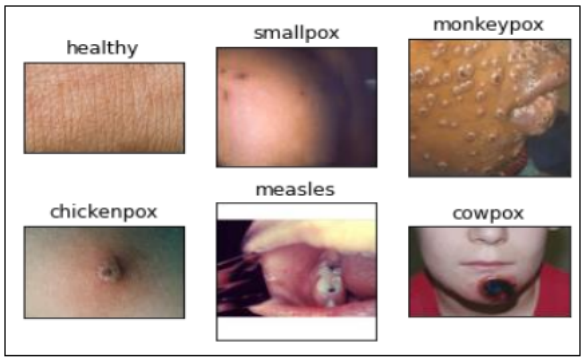


Figure 4. Collected data samples

version, i.e., the Google Colab Free edition. The Free edition is loaded with 13 GB of memory or RAM and 12 GB of Graphical Memory. The GPU is NVIDIA Tesla K80 GPU which is optimized to perform deep learning and machine learning tasks. It also supports CUDA, a library capable of making the most of provided GPU. Apart from GPU and RAM, the collab instance also provides storage space in varying ranges to keep the files and temporary data. The range of this space varies from 40 Gigabytes to 70 Gigabytes. Every Google Colab instance is accessed through a web client. In the experimentation discussed in this study, the Google Chrome web browser is employed to access the Colab. The Chrome browser is installed on the machine running Pop Os 22.04 LTS with the support of 8GB of RAM and an Intel i5 processor clocked at 2GHz.

C. Preprocessing

The dataset, in raw format, has various issues and hence requires preprocessing to enhance and make it compatible to be used as the training and validation dataset. The following sub-sections discuss all the necessary steps in this preprocessing phase.

1) Normalization (Standardization):

As the dataset is collected by scraping online websites there is a strong possibility that the collected samples may not be having the same inherent parameters. For instance, one image sample may have high contrast and the other one may constitute high brightness amount. This disparity of the magnitudes in the properties of the images presents a challenge for the classification as the trained model on non-standardized data may not have acceptable accuracy after the unseen image samples [27].

2) Denoising:

Various samples, taken under varying lighting conditions, using different camera sensors; would eventually have a certain amount of noise in them. The amount and type of the noise may vary depending on the conditions but the presence of the noise participates in the overfitting and hence model starts learning from noise instead of signals. To address the issue of noise presence, a simple technique

named blurring is applied in the preprocessing phase to reduce the noise in the image samples [28].

3) Discarding unusable samples:

Few image samples in the collected dataset have the extreme value of black and white and they are not recognizable. As they serve as outliers and hence thwart the accuracy, they are discarded before the final dataset is prepared. A simple algorithm that sums up the total pixel between two values[whitish and blackish] and then calculates the percentage presence or participation in the image of such pixel values in the range determines whether the image would be kept in the final dataset or not. If the participation overshoots the predefined threshold then the image is discarded on the go [29] [30].

4) Data Augmentation:

As the number of samples is very limited in each class of the dataset, the authors performed the data augmentation and generated new images based on various basic transformations named rotation, noise adding, flipping, shear and zoom. Each method on its own and, if combined with other transformation(s), produces new samples. The augmented dataset has a total of 84465 images which are distributed over the 6 classes as Monkeypox have 15084 samples, Chickenpox has 12847 samples, Smallpox has 17542 samples, Cowpox has 13853 samples, Measles has 12094 samples and Healthy class has 13045 samples respectively The augmentation breaks the initial barrier of insufficient data, which could lead to the underfitting of the model [31].

5) Removal of imbalance in the dataset:

As the augmented dataset has more than 39k images, the model can be fed with this data to start the learning phase. But though the amount of data is sufficient, the distribution of the samples of the dataset over the class is inconsistent. Hence, the current state of the dataset can be considered skewed, which may pose a challenge in the training as it would introduce unintended bias in training for a class with more data samples. This would lead to misclassification, and the post-training model would be partial towards the class with greater samples during the model's training phase. The authors have proposed two solutions to overcome this problem: further augmentation and normalization of the count. While considering all the aspects of each approach, the authors concluded that it would be superior and better to opt for the second approach, normalization, because further data augmentation would lead to the addition of noise. The samples would be repeated, reducing the overall quality of the dataset. On the other hand, normalizing the number of data samples concerning the class with the least samples would not be prone to such issues. Hence the authors prepared the dataset that has 12000 samples in each class (concerning the Measles class having 12094 image samples). The augmented samples from the final dataset have been shown in figure 3.

Figure 5. Augmented data samples

D. Dataset Splitting and Loading

Now as the finalized dataset have 12094 samples in each class, the samples can be split into training split, validation split, and testing split. By utilization of the image_dataset_from_directory function Keras module, the final dataset directory is iterated over and the first object named train_ds is assigned with the dataset's samples. The function image_dataset_from_directory takes arguments like the path of the dataset which is a string indicating the absolute path of the directory containing all the classes(sub-directory). The second parameter is the validation split which is the decimal value indicating the split ratio of the entire dataset. In this case, the value of validation_split is 0.2 which means that 20% i.e. 14400 images would be kept in the validation split, and remaining of the images i.e. 70% of the dataset or 50400 images would be kept in the training split. The test_split contains 10% of the entire data set i.e. 7200 images would be in the testing split and each class would contain 1200 images. The subset parameter is assigned value as training which indicates that the object contains a dataset for training the model. Image size is kept as 229x229 and batch size is kept as 10. In the same manner, the val_ds object is initialized which would contain 20% or 14400 of the total images i.e. 72000 images that would be used for validation. All the other parameters are kept identical.

E. Model Details and Construction

1) Introduction

For the identification and classification of the images containing features of the diseases, a deep neural network model named InceptionV3 is used in the experimentation phase. The InceptionV3 model differs from the traditional CNN in various manners. First, the structure of InceptionV3

is different from the normal CNN. The stacking of the layers in a traditional CNN would follow the sequential structure but in the case of the InceptionV3 model the layers are introduced in a parallel fashion and the blocks(containing layers in a parallel manner) are stacked sequentially. Secondly, the traditional CNN model has some specific layers in them like the convolutional layer, pooling layer, and activation layer but in the case of InceptionV3, there are additional layers like concatenation, both of the pooling layers.

The third difference comes from the parameter numbers and the resource requirements. Deep neural networks are recognized as resource-hungry networks which with the deepening of the model increase the number of parameters and hence also require the resources to train those parameters. Unlike the traditional CNN, the InceptionV3 model, despite being a deep neural network does not require as much memory and resources as its counterpart. The InceptionV3 model is 48 layers deep and very resource efficient as compared to the normal CNN with the same depth. This feature of being resource efficient can be attributed to the implementation of the various layers in various blocks of the InceptionV3 model. Various techniques like factorization of the normal convolutions into the smaller convolutions which reduces the number of parameters by 28% and into the asymmetric convolutions which reduce the number of parameters by 33%.

The introduction of convolutional layers with various filter sizes in a parallel manner allows the model to learn the small and large features in the image together. Another technique named auxiliary classifier fusion is used in which the second auxiliary classifier head is introduced before the main classifier head of the model. This block serves as the module which prevents the overfitting of the model by stopping the vanishing gradient problem. This is done by the error calculation at the auxiliary head and back-propagating weight updation.

All the mentioned operations are performed inside the blocks. Each block has a different number of parameters and the operation it performs also differs. For instance, the factorization, normal and asymmetric, is performed in the Inception block which contains the convolutional and pooling layer in a parallel fashion. The channel reduction is performed in another block named the Reduction block. The auxiliary classifier head is introduced with a different branch that does not impact the main classifier head directly.

2) Construction

The InceptionV3 model is constructed by stacking a first convolutional block CBA, which contains 3 convolutional layers with a filter size of 3x3 and many filters as 32,32, and 64, respectively. The strides used in these layers are 2, 1, and 1 respectively. After the CBA pooling layer is used with the MaxPooling variant having the filter size as 3x3 and stride as 2. The 3rd and 4th layers are convolutional layers having filter sizes of 1x1 and 3x3 and strides of 1 and 1 respectively. After this sequence the first Inception block IBA is introduced and is repeated 3 times. The outcome of this is sent to the reduction block named RA which is followed by 4 inception blocks named IB. The branch after the 4th IB is bifurcated. The output from the previous IB is sent to the auxiliary classifier head (CHAux) and another is sent to the main classifier head (CHMain). The CHAux consists of the 5x5 average pooling followed by 1x1 convolution and 2 fully connected layers. The main classifier head or CHMain is followed by another reduction block named RB which is followed by another inception block named IC repeated twice. The following layer is the global average pooling layer and the last two layers in CHMain are fully connected layers.

The schematic diagram of the InceptionV3 model has been depicted in figure 4.

F. Hyperparameters and Helper functions initialization

Every classification task or task that deals with neural networks require the assertion of a set of parameters that determine the training phase of the model. Various aspects like the convergence of the model, the pace with which the model learns, the required resources, and many such crucial facets depend upon the set of parameters, called hyperparameters, which are needed to be declared before the training commencement. In this study, the set of hyperparameters is the learning rate, number of epochs, optimizer function that would be used for gradient descent, loss function, and the type of metrics that would be monitored to get an insight into the model's performance.

The hyperparameters are the requirement to start the training. Still, another set of parameters is optional but is highly promoted to be used and initialized before the training. This set of parameters is named the callback functions. These helper functions facilitate a few features like post-training visualization and provide immunity against the overfitting case of the training. The first candidate from the list of callbacks is named the tensorboard visualization, which stores the training logs like the number of epochs, the number of iterations, the accuracy and loss achieved during the entire training phase, and so on. These values can be used once the training is finished and one wants to visualize the trends of that phase.

These values can be used once the training is finished and if one wants to visualize the trends of that phase. This callback function does not save only these values but many other values are also stored to visualize every subtle feature of the training phase. For instance, the complex structure and the flow of the data between various blocks of the network can be visualized using the tensorboard callback. The loss pattern in every epoch can be interpreted as the same callback. The second member from this list is the early stopping callback. This function assumes some predefined constraints like the amount of change that is required to be made during each epoch in accuracy, the patience which tells the number of epochs for which even if there is no change in the accuracy the model would

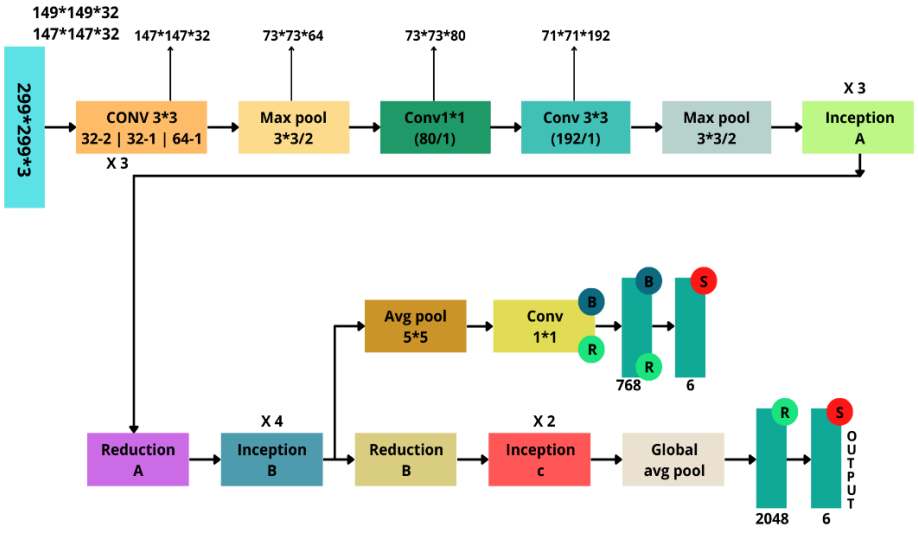


Figure 6. Proposed Model Architecture

continue to train. This callback assures that the model is not going into the training state in which it starts learning noise from the data with the signals which leads to the potential case of overfitting. In the training –utilized but there are 7 total parameters accepted by the early stopping method. The first parameter is named the monitor which determines what value to be monitored during training. It can be loss or accuracy over either set, training set, or validation set. The next parameter is min_delta which takes the floating point number as a value and defines a threshold value and if the model does not show a change in the monitored value according to this parameter training stops. Though there may be some cases in which a certain number of epochs may not show the improvements and training may seem to be of constantly monitored value. Hence not observing the trends according to min_delta, the training would stop which may be a wrong step. So to solve this issue the third parameter named patience is introduced which ensures the training is kept on even if the model does not show any change for the patience number of epochs. The fourth parameter is verbose which determines if the logs are to be printed on the prompt or not. The next parameter is a mode that specifies the continuation of the training based on the parameter passed. It can take one among the 3 parameters named min, max, and auto. The min parameter stops the training if the monitored value stops decreasing. In the case of the max parameter, the training would stop if the monitored value stops increasing. In the case of the auto parameter, the decision is made on the go by considering the features of the monitored quantity. The second to last parameter is the baseline which acts as a threshold and if the model does not show improvement over the baseline then the training stops. The last parameter is restore_best_weigths which gives the feature to save the weight of the epoch in which the model achieved the best magnitude of monitored value.

5. Training Phase

The initialization of the parameters and hyperparameters succeeded during the training stage. The entire training stage consists of 8 trials, and each trial has varying epochs for which the model is trained. For each trial, the number of epochs used is in order of 2,4,6,8,10,12,14,16. The model's accuracy is found to be the best if the model is trained for 12 epochs. The selection of the number of epochs for each trial is determined by the change in the accuracy, which crosses the predefined threshold value. Defining the threshold also keeps checking the training trends to ensure that the model is safe from overfitting. It works in a manner that the model's accuracy improves significantly. The accuracies achieved by the model are then compared to the accuracy reported by the last trial. If there is an improvement in the model's accuracy, then the new number of epochs for the next trial would be calculated as Epochnew = Epochold + 2. Suppose the accuracy does not improve in the current trial. In that case, the training is stopped at the current trial, and the accuracy achieved is interpreted as the final accuracy of the model. The accuracy trend for the finalized trial is shown in figure 5.

It can be interpreted from the graph (see figure 5) that the accuracy trends on both the splits, training split, and testing split improve over time with an increasing number of epochs. The training curve takes the lead in accuracy for half of the starting epochs but superseded the validation curve after epoch 7. The converging nature of both curves indicates that the model does not go into the state where it starts learning noise, i.e., the model does not go into the overfitting state. After epoch 8, both the curves almost superimpose on each other. During further phases, after the 8th epoch, the difference in the accuracy of the model on validation and training split is negligible.

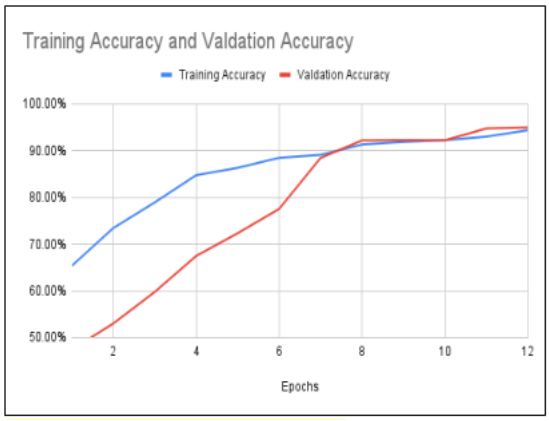


Figure 7. Epochs vs. Accuracy

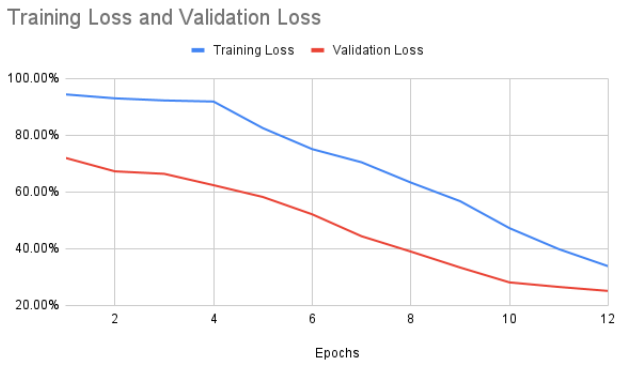


Figure 8. Epochs vs. Loss

The loss trends, for the finalized number of epochs, on the training dataset split and validation dataset split are shown in the graph(see fig 6). The decline in the loss across the increasing number of epochs for both the split also bolsters the achieved accuracy of the model. For the first four epochs, the fall in the loss value is constant on both dataset splits, but after epoch four, the decline is more sharp but moderate and continuous. Around 80% of epochs, the decline in the loss curve takes a gradual turn, and both the training and validation curves can be seen as converging towards each other and downhill overall.

In figure 7, the varying accuracy percentages are shown against the trials performed with varying numbers of epochs. Each bar graph represents the accuracy scored by the model on the corresponding trial with the given number of epochs. The close inspection of the graph gives a very important insight that the model's accuracy increases as the trial progresses, and the number of epochs is stacked in each succeeding trial. The gradual growth in the accuracy is maintained until the 6th trial when the model achieves an accuracy of 94.38%, and then the accuracy decreases in further trials.

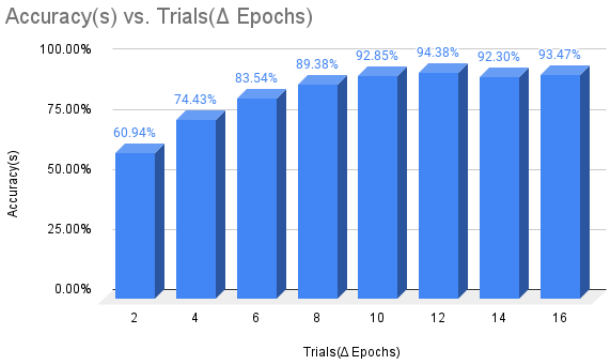


Figure 9. Trials vs. Accuracy

6. Testing on Unseen Data and Accuracy Measures

At the time of dataset splitting, the test split was assigned a total of 7200 images. Even the distribution of the samples in each of the 6 classes would keep 1200 samples in each class. After that, the model sends the split for the prediction by the model, and the results are reported in table 1, which shows the prediction distribution in the confusion matrix.

TABLE I. Confusion Matrix

True/Predicted	MP	ChP	SP	CoP	M	H
MP	1117	23	11	34	15	0
ChP	18	1098	23	21	10	30
SP	5	13	1130	14	18	20
CoP	22	9	39	1084	40	6
M	11	51	35	31	1027	45
H	19	11	20	9	7	1134

Where,

MP = Monkeypox

ChP = Chickenpox

SP = Smallpox

CoP = CowPox

M = Measles

H = Healthy

Calculation of Precision, Recall, and F1 score for various classes is done as follows:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

Using the formulas above the values of Precision, Recall, and F1 score are calculated for each class of dataset and the results are reported in table 2.

TABLE II. Accuracy Measure values distribution

Measure/Class	MP	ChP	SP	CoP	M	H
Precision	0.94	0.91	0.9	0.91	0.92	0.92
Recall(Sensitivity)	0.93	0.92	0.94	0.9	0.85	0.95
F1 Score	0.935	0.915	0.92	0.95	0.883	0.935

As shown in table 2, the F1 score for the Monkeypox class is 0.935, i.e., the achieved accuracy in the Monkeypox class is 93.5%. Similarly, the accuracy for the Chickenpox class is 91.5%, the Smallpox class has an accuracy of 92%, Cowpox has an accuracy of 95%, the Measles class has an accuracy of 88.3%, which is the lowest among all, and the Healthy class reports an accuracy or F1 score as 94%.

7. Results and Discussion

The application of deep learning is not a new concept [29], [32]. Since it began and has grown quickly, many sub-fields of healthcare are being studied, and many new ways are being found to solve the problems that still exist in the medical world. There are neural networks that are built to do specific jobs, like detecting, classifying, recognizing, etc. The deep learning-based neural network model is used to figure out if a person might have monkeypox based on how their skin looks. This makes sure that the diagnosis is made quickly and lowers the chance that monkeypox will be mistaken for another skin disease.

Authorities are worried about what will happen if this dangerous disease keeps spreading and the number of cases of monkeypox keeps going up around the world. The dataset for this model is full of images with lots of details that can be used to teach the model to spot the disease in the image. The accuracy rates (see Figure 5) on the validation dataset or the unknown images show that the trained model can be used to classify diseases based on images.

The efficiency and accuracy of the model are represented by the degree to which the curves, training set, and validation set converge. This means that the model is still learning from the single and with the noise, and has not entered the overfitting state. Overfitting can severely degrade the accuracy metric on new data sets. Also, as training continues, the model learns to predict with less and fewer inaccuracies, as shown by the progressively flatter loss curve (see fig. 6). The optimum number of epochs is also demonstrated by all eight trials, which were conducted with different values. If the model is trained, it outperforms competing options in the evaluation. Figure 7 shows that after 12 iterations of training, the model achieves the highest accuracy reported. On such a boosted dataset, the claimed accuracy of 94.38% over 12 epochs is reasonable. To validate the efficiency of the model, the authors have experimented with existing competent models, and compiled the performance results in table 3.

TABLE III. Comparison of other methods with the proposed TL-based method

Method	Accuracy	Precision	Recall	F1 Score
KNN	0.75	0.72	0.78	0.74
SVM	0.8	0.78	0.82	0.8
Random Forest	0.85	0.84	0.86	0.85
Traditional CNN	0.83	0.84	0.83	0.83
Our Method	**0.914**	**0.916**	**0.91**	**0.925**

In Table 3, the performance of our model to that of other traditional and neural network-based approaches have been compared. Accuracy, Precision, Recall, and F1-score are averaged across six dataset classes. The table clearly demonstrates that our transfer-learning-based InceptionV3 model outperforms other models in every metric. The influence of depth and fine-tuning is evident in the data.

8. Conclusion and Future Recommendations

The proposed neural-network-based architecture does a significant and the most crucial job of identifying the type of skin features and based on the identification, the model can classify the input image of the skin accordingly. This identification and the classification of the input skin images does serve as the primary step, called diagnosis, to provide the best possible treatment plan. The model also solves the problem of confusion that may arise because of the similar appearance of skin features in various types of diseases. This correct classification saves time and decreases the potential risk of the wrong treatment because of the incorrect diagnosis if done in manual mode. Though the model achieves an accuracy of 94.38% on the unseen data this can not be considered the perfect score as the medical domain requires the best possible scores to maintain assurance and reliability about the entire diagnosis of the disease. The accuracy can be improved by increasing the candidate samples in the dataset and the architecture of the network can also be modified to a certain degree to get improved accuracy on the unseen data. Some third-party libraries and modules can be used to deploy this network on handheld devices like smartphones so the end user can perform the diagnosis himself.

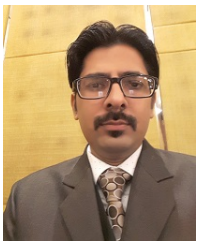
References

[1] E. Who, U. Cdc, M. Cdc, and Health, 2022. [Online]. Available: https://www.cdc.gov/poxvirus/monkeypox/response/2022/world-map.html

[2] E. Who, F. Mathieu, S. Spooner, H. Dattani, M. Ritchie, Roser, and Monkeypox, 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/monkeypox
[3] A. Sherwat, J. T. Brooks, D. Birnkrant, and P. Kim, "Tecovirimat and the Treatment of Monkeypox-Past, Present, and Future Considerations," *N. Engl. J. Med*, 2022.
[4] H. Adler, "Clinical features and management of human monkeypox: a retrospective observational study in the UK," *Lancet Infect. Dis*, 2022.
[5] S. Seang, "Evidence of human-to-dog transmission of monkeypox virus," *Lancet*, vol. 400, no. 10353, pp. 658–659, 2022.
[6] R. A. Farahat, "Monkeypox outbreaks during COVID-19 pandemic: are we looking at an independent phenomenon or an overlapping pandemic?" *Annals of clinical microbiology and antimicrobials*, vol. 21, no. 1, pp. 1–3, 2022.
[7] A. K. Rao, "Monkeypox in a traveler returning from Nigeria-Dallas," *Morb. Mortal. Wkly. Rep*, vol. 71, no. 14, pp. 509–509, 2021.
[8] J. G. Rizk, G. Lippi, B. M. Henry, D. N. Forthal, and Y. Rizk, "Prevention and treatment of monkeypox," *Drugs*, pp. 1–7, 2022.
[9] J. Kaler, A. Hussain, G. Flores, S. Kheiri, and D. Desrosiers, "Monkeypox: a comprehensive review of transmission, pathogenesis, and manifestation," *Cureus*, vol. 14, no. 7, pp. 2022–2022.
[10] E. M. Bunge, "The changing epidemiology of human monkeypox-A potential threat? A systematic review," *PLoS Negl. Trop. Dis*, vol. 16, no. 2, pp. 10 141–10 141, 2022.
[11] J. P. Thornhill, "Monkeypox virus infection in humans across 16 countries," *N. Engl. J. Med*, 2022.
[12] J. S. Bryer, E. E. Freeman, and M. Rosenbach, "Monkeypox emerges on a global scale: a historical review and dermatological primer," *J. Am. Acad. Dermatol*, 2022.
[13] N. L. Bragazzi, "Epidemiological trends and clinical features of the ongoing monkeypox epidemic: A preliminary pooled data analysis and literature review," *J. Med. Virol*, 2022.
[14] N. Z. Alshahrani, "Assessment of knowledge of monkeypox viral infection among the general population in Saudi Arabia," *Pathogens*, vol. 11, no. 8, pp. 904–904, 2022.
[15] D. Mileto, "New challenges in human monkeypox outside Africa: A review and case report from Italy," *Travel Med. Infect. Dis*, pp. 102 386–102 386, 2022.
[16] T. Islam, M. A. Hussain, F. U. H. Chowdhury, and B. M. R. Islam, pp. 2022–2022.
[17] M. A. Onari, S. Yousefi, M. Rabieepour, A. Alizadeh, and M. J. Rezaee, "A medical decision support system for predicting the severity level of COVID-19," *Complex & Intell. Syst*, pp. 1–15, 2021.
[18] F and D. Gennaro, "Human Monkeypox: A Comprehensive Narrative Review and Analysis of the Public Health Implications," *" Microorganisms*, vol. 10, no. 8, pp. 1633–1633, 2022.
[19] A. Patel, "Clinical features and novel presentations of human monkeypox in a central London centre during the 2022 outbreak: descriptive case series," *" bmj*, vol. 378, pp. 2022–2022.
[20] O. Uwishema, "The burden of monkeypox virus amidst the Covid-19 pandemic in Africa: a double battle for Africa," *Ann. Med. Surg*, vol. 80, pp. 104 197–104 197, 2022.
[21] O. J. Peter, S. Kumar, N. Kumari, F. A. Oguntolu, K. Oshinubi, and R. Musa, "Transmission dynamics of Monkeypox virus: a mathematical modelling approach," *Model. Earth Syst. Environ*, vol. 8, no. 3, pp. 3423–3434, 2022.
[22] E. F. Alakunle and M. I. Okeke, "Monkeypox virus: a neglected zoonotic pathogen spreads globally," *Nat. Rev. Microbiol*, pp. 1–2, 2022.
[23] N. Kumar, A. Acharya, H. E. Gendelman, and S. N. Byrareddy, "The 2022 outbreak and the pathobiology of the monkeypox virus," *J. Autoimmun*, pp. 102 855–102 855, 2022.
[24] B. Atkinson, "Infection-competent monkeypox virus contamination identified in domestic settings following an imported case of monkeypox into the UK," *Environ. Microbiol*, 2022.
[25] A. A. Saied, A. A. Metwally, and O. P. Choudhary, "Monkeypox: an extra burden on global health," *Int. J. Surg*, vol. 104, pp. 106 745–106 745, 2022.
[26] D. Daskalakis, R. P. Mcclung, L. Mena, J. Mermin, C. F. Disease, P. M. R. Control, and . Team, "Monkeypox: avoiding the mistakes of past infectious disease epidemics," *Annals of Internal Medicine*, vol. 175, no. 8, pp. 1177–1178, 2022.
[27] V. Sakshi and Kukreja, "A Hybrid SVC-CNN based Classification Model for Handwritten Mathematical Expressions (Numbers and Operators)," *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, pp. 321–325, 2022.
[28] S. Sakshi, C. Gautam, V. Sharma, and O. Kukreja, "Handwritten Mathematical Symbols Classification Using WEKA," *Applications of Artificial Intelligence and Machine Learning*, pp. 33–41, 2021.
[29] V. Sakshi and Kukreja, "Image Segmentation Techniques: Statistical, Comprehensive, Semi-Automated Analysis and an Application Perspective Analysis of Mathematical Expressions," *Arch. Comput. Methods Eng*, pp. 1–39, 2022.
[30] ——, "Segmentation and Contour Detection for handwritten mathematical expressions using OpenCV," *2022 International Conference on Decision Aid Sciences and Applications (DASA), 2022*, pp. 305–310.
[31] V. Kukreja and Sakshi, "Machine learning models for mathematical symbol recognition: A stem to stern literature analysis," *Multimed. Tools Appl*, pp. 1–37, 2022.
[32] P. Sakshi, S. Das, C. Jain, V. Sharma, and Kukreja, "Deep Learning: An Application Perspective," *Cyber*

Intell. Inf. Retr. Proc. CIIR 2021, vol. 291, pp. 323–323, 2021.

Sakshi is currently a full-time Ph.D. student at Chitkara University, Punjab. She has completed her Msc. in Computer Science at Guru Nanak Dev University, Punjab. Her area of interest is pattern recognition, machine learning, and deep learning. She has published more than 10 papers in reputed journals and conferences. She is deliberately attending and presenting papers at several national and international conferences.



Vinay Kukreja earned his Ph.D. Degree in computer science from Chitkara University. He is presently working as a Professor at Chitkara University, Punjab, India. He has more than 16 years of teaching experience; has published more than 55 national/international papers in reputed journals & conferences, 3 books, granted 3 patents and filed more than 20 patents.

Sachin Lodhi has completed his BTech in 2021 from UIT, Barkatullah University, Bhopal. He has been working on browser automation for two years. His core research areas include image processing, image analysis, automation, computer vision, machine learning and IoT. He has in-depth and hands-on knowledge in the field of Artificial intelligence and Image Processing. He has done many projects.